

وزارة التعليم العالي والبحث العلمي

BADJI MOKHTAR- ANNABA UNIVERSITY
UNIVERSITE BADJI MOKHTAR ANNABA



جامعة باجي مختار- عنابة

Faculté: Sciences de l'ingénieur
Département: Electronique

Année 2008/2009

THESE

Présentée en vue de l'obtention du diplôme de DOCTORAT

Option : Automatique

Intitulé

EXTRACTION DES PARAMETRES POUR L'ANALYSE DES IMAGES
APPLICATION A L'ECRITURE ARABE MANUSCRITE

Par : Mr Bouchareb Faouzi

Directeur De Thèse : Hamdi Rachid : MC UNIVERSITE ANNABA

DEVANT LE JURY

Président : Mr. Hadj Ahmed Abbassi : Pr U.ANNABA

Examineurs :

Mr. Boukrouche Abdelhani	:	Pr U.GUELMA
Mr. Redjimi Mohammed	:	Mc U.Skikda
Mr. Khellil Khaled	:	Mc C.U Souk Ahras
Mr. Messadeg Djemil	:	Mc U.ANNA

خلاصة

يندرج هذا العمل في مجال تحليل الصور و تطبيقاتها على الكتابة العربية اليدوية. نظرا للطبيعة المتصلة للكتابة العربية نقترح استخلاص عناصر جديدة لمختلف مراحل التحليل. سرد طريقة جديدة لتنقية الصورة التي تعتمد على حساب تقديري لسمك خط الكتابة الذي يتعلق مباشرة بالقلم المستعمل, و هيكل الصورة. طريقة حذف التشويش تتمثل في حذف المركبات المتصلة الذي سمك خطها أقل من السمك المقدر, هذه الطريقة تمكن من ابقاء الشكل سوى الرئيسي و نقاطه . تطبق هذه الطريقة على الكلمات و الحروف المعزولة. من الخصائص المهمة للكتابة العربية أن الحروف المتصلة في الكلمة تكتب متماسة في خط أفقي الذي يسمى الخط القاعدي الذي يمكن تقريبه بخط مستقيم أفقي, في أغلب الأحيان الكتابة اليدوية للكلمات تكون مائلة لذي نقترح تصحيح هذا الميل باستعمال محولة هوف, كما نعرض طريقة رياضية تعتبر محاولة لتقطيع الكلمة و ذلك باستعمال الخط القاعدي. لتقطيع الحروف (فصل الشكل الرئيسي عن النقاط) نقترح طريقة تستند على المركبات المتلاصقة و تتبع بتحديد موضع النقاط بالنسبة للشكل الرئيسي.

أما فيما يخص العناصر الوصفية للأشكال استعملنا العزوم, واصفات فوريه للجوانب و الأبعاد المركزية, هذه العناصر ثابتة بالنسبة للسلم, الانسحاب و الدوران, نقوم بعدها بمعالجة شعاع هذه العناصر باستعمال تحليل المركبات الأساسية لتصغير الأبعاد. الشعاع المتحصل عليه يستعمل كمدخل لمصنف آلة ذات أشعة سند الذي يعتبر أحسن مصنف في الوقت الحاضر, يستعمل هذا الأخير لتصنيف الحرف بدون نقاطه كما نستعمل بعدها قواعد تتعلق بعدد و مواضع النقاط لتحديد الحرف النهائي.

ثم تجريب الطرق المعدة لتحليل الكلمات على قاعدة بيانات حقيقية تحتوي على أكثر من أربعة عشر ألف كلمة.

الطرق الخاصة بتحليل و التصنيف للحرف المعزول فقد جربت على قاعدة بيانات تحتوي على ألف حرف يدوي.

النتائج المحصل عليها تعتبر مشجعة.

الكلمات المفتاحية: نص عربي مكتوب بخط اليد ، تحويل هوغ ، تحليل المكون الرئيسي ، آلة متجه الداعم،

واصفات فورييه ، مسافات النقطة الوسطى.

Abstract

This work deals with image analysis applied to Arabic handwriting. Due a cursive nature of Arabic handwriting, we propose an extraction of new parameters for different analysis step: in the smoothing image we propose a novel algorithm based on width writing estimated which depend of the pen and skeleton character.

The mechanism of removing the noise is to extract all of the connected components in the image and we remove all component which has width inferior to width estimated for global image, this method keep only a principal shape and dots this smoothing algorithm can be used for word image and character image

In Arabic writing the baseline has an interesting particularity that can be approached by a straight line, so we propose a new algorithm of base line correction using the Hough Transform and we describe a mathematic method to try to segment a cursive word using a baseline.

In character segmentation task we propose a new method based on connected component analysis to separate a principal shape from dots and we detect the position of these points compared to principal shape

The shape descriptor parameters used in this work are moment, Fouries descriptor of projection and centred distance; these parameters are invariant in scale, rotation and translation we apply Principal component Analysis (PCA) as data processing algorithm to features vector in order to reduce dimension. The classifier proposed in this work is based on Support Vector Machines (SVM) which consider an recent optimal classifier up to now. this classifier used to classify a principal shape of character without dots

We use the position and a number of the dots in order to classify the final character among 28 characters

The methods developed for word analysis are test on image data base which belongs more then 14000 words.

The methods developed for characters are test on image data base contain 1000 isolated Arabic character. The results show that these methods are very powerful.

Key Words: *Handwritten Arabic script, Hough transform, PCA, SVM, Fourier descriptors, Centroid distances.*

Résumé

Ces travaux s'inscrivent dans le cadre de l'analyse des images appliquée à l'écriture arabe manuscrite. Vu la nature cursive de l'écriture arabe nous proposons l'extraction des nouveaux paramètres pour les différentes étapes d'analyse : Dans la partie filtrage nous exposons une nouvelle méthode basée sur l'estimation de l'épaisseur d'écriture qui dépend du stylo utilisé et du squelette de l'image originale ; le mécanisme de suppression du bruit consiste à éliminer que les composantes connexes qui ont une épaisseur plus petit que l'épaisseur globale estimée, cette méthode permet de garder que la forme principale ainsi que ses points. Ce type de filtrage s'applique pour les mots ainsi que pour les caractères isolés.

Comme la ligne de base présente une particularité importante, le fait que cette dernière peut être bien approximée par une ligne horizontale, nous décrivons une méthode de correction de la ligne de base en utilisant la transformée de Hough, et nous exposons une méthode mathématique exploitant la ligne de base qui peut être considérée comme un essai de segmentation du mot.

Pour la partie segmentation du caractère (la séparation des points de la forme principale), nous proposons un algorithme basé sur l'extraction des composantes connexes ; cette étape est suivie par l'extraction de positions des points par rapport à la forme principale. Pour les paramètres de description de la forme nous avons utilisé les moments, les descripteurs de Fourier des profils et des distances centroides ; ces paramètres sont invariants par échelle, translation et rotation, le vecteur paramètre extrait est traité en utilisant l'analyse de la composante principale (ACP), le vecteurs issu de cette étape sera utilisé comme vecteur d'entrée pour un classifieur SMV (machine à vecteurs support) qui est considéré comme étant le meilleur classifieur développé jusqu'à nos jours, afin de classifier la forme principale, le nombre des points et leurs positions sont utilisés par la suite pour déterminer le caractère final parmi les 28 caractères que compte l'alphabet arabe .

Les méthodes développées pour l'analyse du mot sont validées expérimentalement sur des données réelles issues d'une base d'images qui contient plus que 14000 mots arabes, alors que les méthodes qui ont été développées pour le caractère isolé ont été appliquées sur une base d'images qui contient 1000 caractères arabes isolés. Les résultats obtenus sont encourageants.

Mots Clés : *l'écriture arabe manuscrite, la transformée de Hough, ACP, SVM, descripteurs de Fourier, distances centroides.*

DEDICACE

A

Mes Parents

A Ma femme et nos enfants

A Mes Frères et Sœurs

A Tous Ceux que j' Aime

REMERCIEMENTS

Je tiens tout d'abord à remercier à remercier le Dr Hamdi Rachid, qui a dirigé ce travail, pour ses nombreux conseils et suggestions ainsi que pour sa disponibilité.

Je remercie le Pr Mouldi Bedda pour ses conseils.

Je tiens à remercier le Pr Abassi Haj pour avoir accepté de présider le jury de cette thèse.

Je remercie le Pr . Boukrouche Abdelhani, le Dr. Messadeg Djemil, le Dr Redjimi Mohamed et le Dr. Khellil Khaled, pour avoir bien voulu accepter d'être membres du jury de cette thèse et de l'intérêt qu'ils ont porté à ce travail.

Listes des Tableaux

Tableau 2.1: les 28 caractères Arabes	7
Tableau 2.2: les lettres et leurs formes selon leur position.....	8
Tableau 2.3: positions des points sur les caractères	9
Tableau 2.4: différents types des mots.....	10
Tableau 8.1: Coordonnées des points en 2 dimensions.....	71
Tableau 8.2: Coordonnées des points 3 dimensions.....	71
Tableau 10.1 : Résultats obtenus.....	108

Listes des figures

Figure 2.1: Une écriture arabe manuscrite d'un nom d'une ville	6
Figure 2.2: lettre avec voyelles arabes.....	9
Figure 2.3: Exemples d'écritures arabes reflétant ses caractéristiques morphologiques	10
Figure.2.4: Illustrations de certaines variations morphologiques de L'Arabe en fonction de l'identité de la fonte.....	13
Figure. 2.5: Exemples d'écritures arabes coufiques.....	13
Figure.2.6 : Ecriture manuscrite Naskhi.....	14
Figure 3.1: Principe de l'érosion d'un ensemble.	19
Figure 3.2: Exemple d'une dilatation par un élément structurant de forme circulaire..	19
Figure 3.3: Illustration de l'ouverture morphologique.....	20
Figure 3.4: Illustration de l'ouverture et de la fermeture morphologiques...	21
Figure 3.5: Squelettisation.....	22
Figure.3.6: Image avec deux ligne.....	23
Figure.3.7: Deux Crêtes de l'image a TH.....	23
Figure.3.8: L'espace de (x, y)	25
Figure.3.9: L'espace de Hough (ρ, θ)	25
Figure. 3.10: Exemple d'un mot manuscrit.....	25
Figure. 3.11 : La TH d'un mot manuscrit.....	25
Figure.3.12 : squelette d'un mot.....	
Figure.3.13 : TH du squelette du mot.....	26
Figure 4.1: exemple de binarisation.....	30
Figure 4.2: Squelettisation.....	31
Figure.4.3: Filtrage d'un caractère arabe.....	32
Figure.4.4: Image original.....	32
Figure.4.5: squelette d'image.	32
Figure.4.6: image originale d'un nom de willaya.	32
Figure.4.7: image filtrée.....	32
Figure 4.8: points d'intersections.....	33
Figure.4.9: points d'extrémités.	33
Figure.4.10: la courbure d'une fonction.....	34
Figure.4.11: l'angle α de la courbure.....	34
Figure.4.12: points des coins.....	34
Figure.4.13: image d'entrée.....	35
Figure .4.14: squelette de l'image.....	35

Figure.4.15: Image originale.....	36
Figure.4.16: TH d'une image.....	36
Figure.4.17: Correction de la ligne de base et projection vertical.....	36
Figure.4.18: Correction de mot renversé.....	37
Figure 5.1: Exemple de ligature (chauvauchement).	38
Figure 5.2: la lettre sin.	41
Figure.5.3: La courbe d'une Fonction $y=f(x)$	41
Figure.5.4: La courbe d'une Fonction $y=x^n$	42
Figure.5.5: Image originale«سطينف»	43
Figure.5.6: Image sans les éléments secondaires (les points).	43
Figure.5.7: la ligne de base.	44
Figure.5.8: mot segmenté	44
Figure.5.9: La forme de la courbe $Y = f(x)$	44
Figure.5.10: les dérivés de la fonction.....	45
Figure.5.11: Segmentation d'un caractère arabe.....	45
Figure.6.1: exemple des cinq cavités.....	51
Figure 6.2: Application de la distance centroide.....	55
Figure 7.1: Espace d'observation R avec deux classes « gaussiennes».....	59
Figure 7.2: exemple d'un classifieur de Bayes.....	60
Figure 7.3: densité de probabilité à posteriori des classes.....	62
Figure 7.4: kppv $n=1$ pour deux classes	63
Figure 7.5: kppv $n=8$ pour deux classes.....	64
Figure 8.1: transformation d'espace	68
Figure 8.2: Espace d'entrée.....	69
Figure 8.3: Espace de redescription	69
Figure 8.4: Transformation d'un espace non linéairement séparable à un espace linéairement séparable	70
Figure 8.5: espace description de deux dimensions.....	71
Figure.8.6: espace de description de 3 dimentsions.....	72
Figure 8.7: représentation dans \mathfrak{R}^2 de l'hyperplan correspond à la fonction de la décision d'un classifieur linéaire.....	73
Figure 8.8: marge d'un exemple	75
Figure 8.9: Hyperplans canoniques.....	76
Figure 8.10: intérêt de la marge maximale.....	77
Figure 8.11: Un mapping ϕ rendant les exemples linéairement séparable.....	80
Figure 8.12: Un exemple d'une marge souple.....	85
Figure 8.13: la distance séparant un outlier et l'hyperplan canonique.....	86

Figure 9.1: Exemple de l'ACP	93
Figure 10.1 : différentes étapes d'analyse d'un caractère	97
Figure-10.2: exemple de la base de donné	98
Figure10.3: La squelettisation d'images	99
Figure 10.4 : exemple de filtrage	99
Figure-10.5: segmentation d'un caractère	100
Figure10.6: Exemple de cavités du squelette du caractères (kaf)	101
Figure-10.7 :Exemple de cavités du squelette du caractères (taa).....	101
Figure 10.8 : caractère taa	102
Figure 10.9: les quatre profils du caractère taa	102
Figure 10.10 : les quatre profils du caractère taa en même vecteur	102
Figure 10.11 : profils normalisées du caractère taa	103
Figure 10.12: Descripteurs de Fourier des profils du caractère taa	103
Figure 10.13 : Descripteurs de Fourier des 7 moment invariantdu caractère taa en rouge et ba en bleu.....	104
Figure10.14 : Descripteur de Fourier des distances centroides de les lettres (baa) و et (jim)ج	104
Figure 10.15 : Descripteurs de Fourier des profils du caractère baa en bleu et HAA en rouge.....	105
Figure 10.16 : Descripteurs de Fourier des profils du caractère sine en bleu et saad en rouge.....	105
Figure 10.17 : Descripteurs de Fourier des profils de deux écriture différente du caractère ain.....	106
Figure 10.19 : Moments invariants du caractère Mim en bleu et Noun en rouge	106
Figure 10.120 : Descripteur de Fourier des distances centroides de 4 caractères mim	107

Listes des Symboles

$O^B(X)$	Ouverture d'un ensemble X par un élément structurant B
$F_B(X)$	Fermeture d'un ensemble X par un élément structurant B
$Y=E_B(X)$	Erosion d'un ensemble X par un élément structurant B
$Y=D_B(X)$	Dilatation d'un ensemble X par un élément structurant B
$I \oplus E$	dilatation de l'image I par la demi-droite E
\bar{I}	Négation inversion des 0 et des 1
(ρ, θ)	Coordonnées polaires
im	Matrice d'image originale
sk	Matrice d'image squelette
wd	Largeur d'écriture
m_{pq}	Moments d'ordre (p+q)
(u_{pq})	moments centré
$P(w_i)$	probabilités à priori
$p(w_i / x)$	probabilités à posteriori
$\Phi()$	Transformation non linéaire d'espace
\mathfrak{R}^n	Espace de dimension n
$sign$	Fonction signe
ψ	Marge géométrique
L	Lagrange
C	contrainte
$card$	cardinal
S_w	matrices de variance intra classes
s_b	matrice de dispersion inter-classes

Table des matières

خلاصة	I
Abstract	II
Résumé	III
Dédicace	IV
Remerciement	V
Liste Des Tableaux.....	VI
Liste Des Figures	VII
Liste Des Symboles	X
Chapitre 1.....	1
Introduction Générale	16
1.1 Introduction.....	16
1.2 Problématique	18
1.3 Contributions Majeures.....	18
1.4 Organisation de la thèse:.....	19
Chapitre 2.....	21
Caractéristiques de l'écriture Arabe.....	21
2.1 Introduction.....	21
2.2 Particularité de la langue arabe.....	22
2.2.1 l'alphabet de la langue arabe	22
2.2.2 Position des lettres arabes	22
2.2.3 Propriété de l'écriture arabe.....	23
2.3 Caractéristiques morphologiques de l'écriture arabe	25
2.3.1 Structure d'un mot	25
2.3.2 Catégories des mots	26
2.4 Typologie des fontes Arabes.....	26
2.4.1 L'écriture coufique	28
2.4.2 L'écriture cursive ou Naskhi.....	29
2.5 Conclusion	29
Chapitre 3.....	30
Outils Mathématiques	30
3.1 Morphologie Mathématique	30
3.1.1 Introduction.....	30
3.1.2 Définitions et application.....	32
3.1.3. Images binaires	32
3.1.4 L'analyse des images et La morphologie mathématique.....	33
3.1.5 Différentes opérations de la morphologie mathématique	33
a. L'érosion	33
b. Dilatation	34
c. Ouverture morphologique.....	35
d. Fermeture morphologique.....	35

e. La squelettisation	36
3.2 Transformée de Hough	37
3.2.1. Introduction.....	37
3.2.2 Transformée de Hough Standard (THS).....	38
3.2.3 Définition	39
3.2.4. Propriétés de La THS.....	40
3.2.5 La transformé de Hough généralisée	42
3.2.6 Conclusion	43
Chapitre 4.....	44
Paramètres de prétraitement.....	44
4.1 Introduction.....	44
4.2 Binarisation :	44
4.3 Squelettisation :	45
4.4 filtrage basé sur le squelette :.....	46
4.5 Points caractéristiques.....	48
4.5.1 Point d'intersection.....	48
4.5.2 Point d'extrémité.....	48
4.5.3. Point de coin	49
4.6 Ligne de base et correction des mots renversés.....	50
4.6.1. Correction de la linge de base.....	50
4.6.2 Correction du mot inversée.....	51
4.7. Conclusion	52
Chapitre 5.....	53
La Segmentation	53
5.1. Introduction.....	53
5.2 Les méthodes classiques de segmentation.....	54
5.2.1. Méthode de segmentation basée sur la projection verticale	54
5.2.2. Méthode de segmentation basée sur la fonction de distance supérieure.....	54
5.2.3. Méthode de segmentation basée sur le squelette	55
5.2.4. Méthode de segmentation basée sur la reconnaissance	55
5.3. Méthode de segmentation proposée pour un mot	55
5.3.1. Fondements mathématiques.....	56
5.3.2. Extraction de la ligne de base	58
5.3.3. Fonction de la ligne de base.....	59
5.3.4. Calcul de la première et la deuxième dérivé.....	59
5.4 Segmentation du caractère :	60
5.4.1 La position des points	60
5.4.2 l'algorithme de position	61
5.5 Conclusion	61
Chapitre 6.....	62
Paramètres de classification.....	62
6.1 Introduction.....	62
6.2 Méthode des moments invariants	62
6.3 Méthode des moments centrés.....	63
6.4 détection des cavités	64
6.5 Descripteur de Fourier	66

6.5.1 Descripteur par tangente :	66
6.5.2 Représentation complexe	67
6.5.3 Les descripteurs de Fourier pour contours fermés	68
6.5.3.1 Echantillonnage :	68
6.5.3.2 Signature angulaire :	68
6.5.3.3 Descripteur de Fourier de distance centroides.	69
6.6 Conclusion	70
Chapitre 7	71
classification	71
7.1 Introduction	71
7.2 Notion de Classificateur	71
7.3 Apprentissage	72
7.4 Test	72
7.5 Les méthodes de classification	73
7.5.1 Les méthodes paramétriques	73
a. Règle de Bayes et décision Bayésienne	73
b. Probabilité d'erreur et fonctions discriminantes bayésiennes	75
7.5.2 Méthodes non paramétriques	76
a. Principe des méthodes non paramétriques :	76
b. Règle de décision du plus proche voisin	78
c. Les réseaux connexionnistes	79
7.5.3 Les méthodes structurelles	80
7.6 Conclusion	81
Chapitre 8	82
Machines à Vecteurs Support (SVM)	82
8.1 Introduction	82
8.2 Principe	83
8.3 classifieur	86
8.3.1 Définition:	86
8.3.2 Marge de l'hyperplan:	88
a. marge fonctionnelle d'un exemple:	89
b -marge géométrique d'un exemple:	90
c- marge géométrique de l'ensemble d'apprentissage	90
8.4 Hyperplans canoniques:	90
8.5 Relation avec l'apprentissage statistique:	91
8.6 Classificateur à marge maximale	92
8.7-Les fonctions Kernels:	94
8.7.1 Généralités:	94
8.7.2 Exemple de kernels:	95
8.8 Formulation de SVM:	98
8.8.1 Cas linéairement séparables:	98
8.8.2 Cas multi-classe	101
8.9 Conclusion:	102
Chapitre 9	104
Analyse en composante principale (ACP)	104
9.1 Introduction	104

9.2 But :	104
9.3 Réduction de l'espace de représentation :	104
9.4 Description de l'Analyse en composantes principales :	105
9.5 Sélection de paramètres :	108
9.6 Conclusion	111
Chapitre 10.....	112
Application sur les Caractères Arabes Manuscrits Isolés	112
10.1 Introduction.....	112
10.2 Base de donnée	113
10.3-Prétraitement.....	114
10.3.1 La squelettetisation	114
10.3.2 Filtrage.....	114
10.4 Segmentation :	115
10.5 La position des points	115
10.6 Extraction de caractéristique.....	115
10.6.1-Les cavités	115
10.6.2-Les profils.....	116
10.6.3Exemples:.....	118
10.7Classification	122
10.8 Resultats obtenus:	123
10.8 Conclusion :	123
Conclusion Générale et Perspectives	124
Références Bibliographiques	126

Chapitre 1

Introduction Générale

1.1 Introduction

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

Si le fait de lire est un acte relativement simple pour un humain. Concernant la machine cette activité met en oeuvre des processus complexes. Durant le processus de lecture, un humain exploite implicitement des connaissances contextuelles et linguistiques, qui sont difficiles à transmettre et à faire reproduire à une machine. Faire lire une machine nécessite le recours à l'application de plusieurs algorithmes mathématiques.

L'apparition des premières écritures, il y a près de cinq mille ans, a révolutionné l'humanité, qui décida ultérieurement que son histoire débiterait à ce moment. L'importance de l'écriture en tant que moyen de communication pour l'homme n'a fait que croître depuis. Aujourd'hui, le développement constant de l'outil informatique rend nécessaire une communication toujours plus étroite entre l'homme et la machine. C'est à celle-ci de s'adapter aux moyens de communication que l'homme maîtrise à présent, et qu'il se complaît à utiliser, telle l'écriture.

La reconnaissance automatique de caractères manuscrits a pour objet de convertir des images qui sont compréhensibles par l'homme, en un ensemble de codes interprétables par un ordinateur. Deux méthodes peuvent être dégagées: la reconnaissance en-ligne et la reconnaissance hors-ligne. Dans la reconnaissance en-ligne, les caractères sont reconnus au moment même où ils sont écrits. L'information est dynamique et monodimensionnelle, et

consiste en une séquence de traits qui suivent une échelle temporelle. Dans la reconnaissance hors-ligne, au contraire, le processus d'écriture n'est pas accessible. L'information ne possède pas de nature temporelle; elle est bidimensionnelle, et consiste en l'image digitalisée des caractères.

La reconnaissance hors-ligne peut elle-même être scindée en deux domaines, en fonction du nombre de personnes dont l'écriture doit être reconnue. Lorsque ce nombre est limité, le système de reconnaissance peut subir un entraînement spécifique sur la base de l'écriture propre aux personnes concernées. On parle alors de reconnaissance monoscripteur. Lorsqu'un tel entraînement n'est pas possible, parce que les écritures peuvent provenir d'un très grand nombre de personnes distinctes, on parle de reconnaissance multiscripteur. Le style d'écriture lui-même peut en outre être réparti en deux catégories; écriture manuscrite et imprimée. Dans la pratique pour la lecture des chèques, tri postal, et formulaire c'est l'écriture manuscrite qui doit être analysée qui présente des difficultés importante par rapport à l'imprimé vu la variabilité d'une personne à une autre et pour la même personne.

Cette contrainte est alors nécessaire pour limiter la dispersion d'aspect des caractères, et rendre ainsi ces derniers plus aisément exploitables par l'homme. les méthodes de classification sont divisées en deux approches, une approche globale et une approche par segmentation. L'approche globale consiste à reconnaître le mot entier sans l'avoir segmenté; c'est une approche utilisée pour un vocabulaire limité. Par contre dans l'approche par segmentation, elle consiste à segmenter le mot en lettres ou en graphèmes puis en fait la reconnaissance; c'est une approche qui est adaptée pour un vocabulaire non limité.

Peu de travaux ont été consacrés à l'écriture arabe manuscrite, un tour d'horizon des techniques développées pour l'arabe est donné par Adnan.A [1]. Vue la nature cursive de l'écriture arabe manuscrite, quelques chercheurs ont essayé d'appliquer des algorithmes développés pour le Latin mais généralement les résultats de ces travaux sont loin d'être prometteurs, la langue Arabe, est utilisée comme langue officielle dans tous les pays arabes elle est également employée dans la plupart des écrits et à l'oral, dans les situations officielles ou formelles (discours religieux, politiques, journaux). Notons également que le Farsi utilisé principalement en Iran , partage un grand nombre de points commun avec l'écriture arabe.

1.2 Problématique

L'alphabet arabe comporte 28 lettres. La forme des lettres change de forme selon sa position dans le mot. Quelques lettres ne se différencient que par la présence ou non de points.

Une des principales difficultés de ce domaine est liée à la très grande variabilité existant entre les différents styles d'écriture. De plus il existe une forte incertitude, non seulement en raison de la grande variété dans la forme des caractères, mais également à cause des éventuels recouvrement et des liaisons pouvant exister entre caractères. Dans un mot, les lettres peuvent être entièrement isolées ou réunies sous forme d'un ou plusieurs groupes de lettres, ou encore entièrement connectées.

L'écriture arabe manuscrite présente des inclinaisons vue sa nature cursive, cette inclinaison est bien remarquable dans la ligne de base.

La présence des points comme composante essentielle d'un caractère rend la tâche de filtrage difficile.

1.3 Contributions Majeures

L'objectif de ce travail est de proposer de nouvelles méthodes et d'extraire de nouveaux paramètres pour l'analyse de l'écriture arabe manuscrite. Pour cela nous avons développé des nouveaux paramètres pour la phase de prétraitement des mots [2], ainsi que des méthodes de prétraitement et de reconnaissance des caractères isolés décrites dans [3] et [4].

Les principales contributions:

- Estimation de l'épaisseur de trait d'écriture qui dépend du point de stylo utilisé: ce paramètre est utilisé pour la suppression de bruit sans suppression des point du caractère. Ce paramètre est utilisé pour l'analyse du mot ainsi que pour le caractère isolé.
- Extraction de l'angle d'inclinaison d'un mot et correction de la ligne de base du mot qui présente une grande importance car tous les caractères sont connectés sur cette ligne, cette méthode est basée sur une technique mathématique utilisant la transformée de Hough, pour rendre ces tâches rapides nous avons proposé d'utiliser le squelette qui préserve la forme du mot.
- Pour la segmentation du caractère isolé; c'est à dire l'extraction du caractère et ses

points nous avons utilisé une analyse statistique des surfaces des composantes connexes de l'image originale et du squelette.

- Extraction des positions des points (haut, bas ou milieu) selon le caractère.
- Pour l'extraction des paramètres qui sont des informations élémentaires pour la classification, nous avons utilisé des paramètres classiques mais appliqués au squelette de la partie primaire du caractère le caractère sans les points.
- Utilisation un classifieur SVM après un traitement de données par l'ACP pour la classification de forme principale du caractère puis utilisation du nombre de points et de leurs positions pour la classification finale du caractère.

1.4 Organisation de la thèse:

Nous présentons dans le deuxième chapitre les caractéristiques de l'écritures arabe, et les propriétés morphologiques des différents caractères arabes, nous poserons le problème de la variabilité de leurs formes et les différentes fontes existantes ainsi que la présence des points qui forment la partie secondaire de la lettre.

Les outils mathématiques utilisés dans notre travail telle que la morphologie mathématique et la transformée de Hough seront décrites dans le troisième chapitre.

Dans le quatrième chapitre nous exposons les méthodes développées pour le prétraitement en premier lieu c'est la méthode d'estimation de la largeur d'écriture pour le filtrage et nous décrivons une méthode basée sur la transformée de Hough pour la correction de la ligne de base.

Dans le cinquième chapitre nous décrivons une méthode de segmentation du mot en exploitant la particularité qui existe dans l'écriture arabe, le fait que la majorité des lettres dans un mot sont attachées sur la ligne de base et nous utilisons l'extraction des composantes connexes pour la segmentation du caractère.

Le sixième chapitre est consacré à l'extraction des paramètres de classification.

Une revue des méthodes de classification sera décrite dans le septième chapitre.

Dans le huitième chapitre nous exposons en détail les fondements mathématiques de la méthode de classification basée sur SVM.

Le neuvième chapitre est consacré à l'ACP qui est utilisée comme outil de traitement et

de réduction de la dimension du vecteur de données présentée par ses paramètres, ce qui facilite la tâche de classification.

Nous présentons dans le dixième chapitre. L'application des différentes méthodes développées pour l'analyse et la reconnaissance des caractères arabes isolés et les résultats obtenus.

Nous terminons ce travail par une conclusion et des perspectives envisagées pour l'analyse et la reconnaissance de l'écriture Arabe.

Chapitre 2

Caractéristiques de l'Écriture Arabe

2.1 Introduction

L'arabe est écrit par plus de 256 millions de personnes, dans plus de 20 pays Différents. Le manuscrit arabe a évolué d'un type d'araméen, la langue aramaic a moins de consonnes que l'arabe, ainsi de nouvelles lettres ont été créés autour du 7ème siècle en ajoutant des points aux lettres existantes [60]. D'autres petites marques (diacritiques) sont employées pour indiquer des voyelles courtes, mais ne sont pas toujours employées.

L'écriture arabe est une écriture consonantique qui utilise un alphabet composé de vingt-huit lettres. C'est une écriture plutôt semi- cursive dans le sens où le mot peut être composé d'un ou de plusieurs pseudo- mot le lettre change de forme en fonction de sa position, la plupart des lettres arabes possèdent la même forme de caractère et ne se différencient que par la présence et/ou le nombre de points se situant au-dessus ou bien au-dessous de corps du caractère.

Le texte arabe, manuscrit et imprimé, est cursif. Les lettres sont jointes ensemble suivant une ligne d'écriture (ligne de base) Figure 2.1.ceci ressemble au latin cursif. L'arabe s'écrit de la droite vers la gauche.

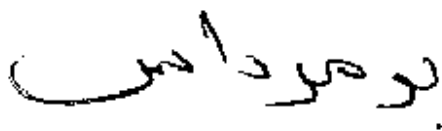


Figure 2.1: écriture arabe manuscrite du nom d'une ville .

2.2 Particularité de la langue arabe

2.2.1 l'alphabet de la langue arabe

Lettre arabe	Correspondant français	Prononciation	Lettre arabe	Correspondant français	Prononciation
أ	a	Alif	ض	D	Dad'
ب	b	Ba'	ط	T	Tah'
ت	T	Ta'	ظ	Z	Zah'
ث	th	Tha'	ع	''	Ayan'
ج	j	Jim'	غ	Gh	Ghayn'
ح	h	Hha	ف	F	Fa'
خ	kh	Kha'	ق	Q	Qaf'
د	d	Del'	ك	K	Kaf'
ذ	d	Thal'	ل	L	Lam'
ر	R	Ra'	م	M	Mim'
ز	Z	Zayn'	ن	N	Nun'
س	s	Sin'	ه	H	ha'
ش	Sh	Shin'	و	W	Waw'
ص	s	Sad'	ي	Y	Ya'

Tableau 2.1: les 28 caractères Arabes (letters isolées).

2.2.2 Position des lettres arabes

La plupart des lettres changent de forme, selon leur position dans le mot au début, au milieu ou à la fin du mot, si elles sont jointes à une autre lettre, ou bien si elles sont isolées. Comme représenté sur Le Tableau 2.1.

- ❖ Au début du mot (DM),
- ❖ Au milieu du mot (MM),
- ❖ la fin du mot mais liée à une lettre (FM),
- ❖ A la fin du mot sans être liée à la lettre (FS).

FS	FM	MM	DM	FS	FM	MM	DM	FS	FM	MM	DM
ع	ع	ع	ع	ر	ر	ر	ر	ا	ا	ا	ا
ق	ق	ق	ق	ز	ز	ز	ز	ب	ب	ب	ب
ك	ك	ك	ك	س	س	س	س	ت	ت	ت	ت
ل	ل	ل	ل	ش	ش	ش	ش	ث	ث	ث	ث
م	م	م	م	ص	ص	ص	ص	ج	ج	ج	ج
ن	ن	ن	ن	ض	ض	ض	ض	ح	ح	ح	ح
ه	ه	ه	ه	ظ	ظ	ظ	ظ	خ	خ	خ	خ
و	و	و	و	غ	غ	غ	غ	د	د	د	د
ي	ي	ي	ي	ف	ف	ف	ف	ذ	ذ	ذ	ذ

Tableau 2.2: les lettres et leurs formes selon leur position

2.2.3 Propriété de l'écriture arabe

Les premières remarques concernant la forme des lettres viennent du fait que la forme de plusieurs lettres ne se distinguent les unes des autres que par des points diacritiques.

- La forme des lettres : C'est ainsi que la lettre **ح** peut se lire **خ** (Kha'), **ح** (Hha'), **ج** (Ja'), compliquant de fait la clarté de la lecture. Cette confusion est renforcée par le problème lié au fait que certaines lettres comme le Da (**د**) ou le Ra' (**ر**), ne peuvent être liées à la lettre qui les suit. Cette situation complique singulièrement la tâche du lecteur qui souvent ne sait où commence un mot, ni même où il finit; et ceci d'autant plus que les espacements entre les mots, même en imprimerie, n'est pas normalisé.
- Position des lettres dans les mots : On peut également reprocher à l'écriture arabe la diversité des formes des lettres, formes qui varient selon la place de la lettre dans le mot, isolée, finale, médiane et initiale.

- Un mot arabe s'écrit avec des consonnes et des voyelles. Les voyelles sont ajoutées au-dessus ou au-dessous des lettres (َ , ِ , ُ) Elles sont nécessaires à la lecture et à la compréhension correcte d'un texte, elles permettent de différencier des mots ayant la même représentation. Cependant, les voyelles ne sont utilisées que pour des textes sacrés et didactiques. Les textes courants rencontrés dans les journaux et les livres ne comportent habituellement pas de voyelle. Pour les voyelles nous distinguons de 3 lettres alif (ا), yaa (ي), waaw (و) et voyelles courtes comme illustré sur la figure suivante :

بُ بَ بِي بُ

Figure 2.2: lettre avec voyelles arabes

- De plus certaines lettres comme Alif peuvent être symbolisé par $\bar{ا}$, $\acute{ا}$ ou $\grave{ا}$ de même que pour les lettres $ي$ et $و$ qui symbolisent respectivement $\bar{ي}$ et $\acute{و}$.

Les points de diacritiques attachés aux lettres qui peuvent être 1, 2 ou 3 points, sont placés au-dessus ou au-dessous de la lettre

un point au-dessous	ب	b
deux points au-dessous	ي	y
un point au-dessus	ن	n
deux points au-dessus	ت	t
trois points au-dessus	ث	th

Tableau 2.3: positions des points sur les caractères

2.3 Caractéristiques morphologiques de l'écriture arabe

L'écriture arabe est semi cursive aussi bien dans sa forme consacrée que manuscrite. Les caractères d'une même chaîne (ou pseudo mot) sont ligaturés horizontalement et parfois verticalement (dans certaines fontes deux, trois et même quatre caractères peuvent être ligaturés verticalement), occultant ainsi toute tentative de segmentation en caractères. Le mot arabe n'a pas de longueur fixe, il peut comprendre un ou plusieurs pseudo mots incluant un nombre souvent différent de caractères. L'étude de la morphologie des pseudo mots montre que l'écriture arabe présente des variations dans des bandes horizontales plus ou moins complexes en fonction de la calligraphie des caractères contenus dans le pseudo mot. La bande centrale est généralement la plus chargée au point de vue densité d'informations en pixels. Elle correspond aux lieux des ligatures horizontales, aux caractères centrés (sans extensions), aux boucles. La Figure (2.3) donne des exemples d'écritures arabes imprimées et manuscrites [5].



Figure 2.3: Exemples d'écritures arabes reflétant ses caractéristiques morphologiques

2.3.1 Structure d'un mot

En arabe un mot peut signifier toute une phrase grâce à sa structure composée qui est une agglutination d'éléments de la grammaire, la représentation suivante schématise la structure possible d'un mot. Notons que la lecture et l'écriture d'un mot se fait de droite vers la gauche [6].

Post fixe	Suffixe	Corps schématique	Préfixe	Antéfixes
-----------	---------	-------------------	---------	-----------

Tableau 2.4 : différents type des mots

- Antéfixes sont des prépositions ou des conjonctions.
- Préfixes et suffixes expriment les traits grammaticaux et indiquent les fonctions : cas du nom, mode du verbe et les modalités (nombre, genre, personne,...)
- Post fixes sont des pronoms personnels.

2.3.2 Catégories des mots

L'arabe considère 3 catégories de mots

- Le verbe : entité exprimant un sens dépendant du temps, c'est un élément fondamental auquel se rattachent directement ou indirectement les divers mots qui constituent l'ensemble.
- Le nom : l'élément désignant un être ou un objet qui exprime un sens indépendant du temps.
- Les particules : entités qui servent à situer les événements et les objets par rapport au temps et l'espace, et permettent un enchaînement cohérent du texte.

2.4 Typologie des fontes Arabes

L'écriture Arabe est connue pour sa richesse en fontes et styles [5]. Il existe environ 450 fontes d'écritures différentes dont seulement quelques-unes sont couramment utilisées dans le monde arabo-musulman, citons à titre d'exemple : le Neskhi, Thoulthi, Roqa, Diwani, Koufi, Farsi...etc.). Le Neskhi demeure aujourd'hui la fonte la plus utilisée pour l'écriture imprimée. Chaque style arabe est régi par des lois particulières. D'un style à un autre, les proportions d'une même lettre et son dessin peuvent changer considérablement. L'examen des dessins des caractères montre des différences considérables inter et intra-fontes à différents degrés[79]. Parmi les variations les plus importantes nous retenons les différences aux niveaux:

- du dessin et/ou du nombre de boucles d'un caractère Figure. 2.4.(a).
- des dimensions ainsi que la forme aussi bien des extensions hautes (hampe) que basses (jambage) Figure. 2.4 -(b).
- du dessin des points diacritiques Figure2.4.-(c) et ceux des signes diacritiques tels que Shadda “ ّ ” et Mada “ ٓ ” Figure. 2.4-(b), 2.4.-(c), ainsi que leur position par rapport au

tracé du caractère,

- des accolements entre Pseudo mots Figure. 2.4 -(d).

- de la disposition spatiale du caractère par rapport à la ligne d'écriture (naturellement horizontale) qui varie dans certaines fontes Figure.2.4-(e).



Figure 2.4.(a): Exemples de dessins de boucle dans des fontes différentes.



Figure. 2.4.(b): Exemples de dessins de hampe et de jambage dans des fontes différentes.

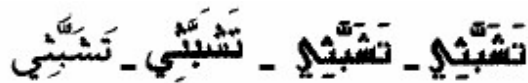


Figure. 2.4.(c): Variation du dessin et de la position des diacritiques en fonction de la fonte.



Figure. 2.4.(d): Différents accolements de pseudo mots dans des fontes différentes.



Figure.2.4.(e): Disposition des caractères par rapport à la ligne d'écriture.



Figure. 2.4.(f): Exemples de dessins de caractères pour une même position contextuelle.



Figure. 2.4.(g): Variation de la ligne d'écriture en fonction de la fonte

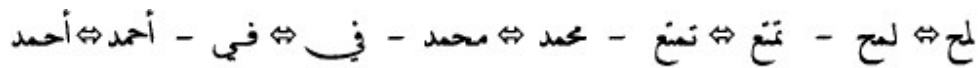


Figure. 2.4 -(h) : Pseudo mots avec et sans ligatures (à gauche et à droite).

Figure.2.4: Illustrations de certaines variations morphologiques de L'Arabe en fonction de l'identité de la fonte.

2.4.1 L'écriture coufique

L'écriture coufique tire son nom de la ville de Kufa, en Mésopotamie, où elle aurait été employée pour la première fois. Elle présente des formes anguleuses et rectilignes. C'est une écriture « penchée » et rigide. Elle est d'abord employée dans les décors architecturaux. Puis les écritures du Coran de la fin du X^e siècle présentant des lettres de plus en plus décoratives.



Figure. 2.5: Exemples d'écritures arabes coufiques

2.4.2 L'écriture cursive ou Naskhi

Le Naskhi est l'écriture courante tracée au calame avec des formes souples et rondes. Chaque caractère peut prendre quatre formes, selon sa position: initiale, médiane, finale ou isolée. L'écriture cursive vient remplacer l'écriture coufique vers le XII^e siècle. La cursive correspond au geste naturel de la main. C'est un univers de courbes souples et denses.



Figure.2.6 : Écriture manuscrite Naskhi.

2.5 Conclusion

Les critiques faites à l'écriture arabe portent principalement sur les défauts de l'écriture manuscrite et les difficultés de l'imprimée, c'est que l'écriture arabe, à la différence des écritures grecques, latines ou cyrilliques, ne se présente que sous une seule forme. Alors qu'en typographie latine, les caractères utilisés se distinguent sensiblement de leurs homologues manuscrits du fait qu'ils ne sont pas liés entre eux par une ligature, en typographie arabe, les caractères sont les mêmes que ceux de l'écriture manuscrite.

Chapitre 3

Outils Mathématiques

Le choix des outils mathématiques joue un rôle très important pour tout système d'analyse des images, cela peut influencer la qualité des résultats envisagés ainsi que la vitesse d'exécution. Nous décrivons dans ce chapitre les outils mathématiques utilisés dans notre travail, telle que la morphologie mathématique et la transformée de Hough.

3.1 Morphologie Mathématique

3.1.1 Introduction

La morphologie mathématique est une théorie non linéaire, utilisée en particulier en analyse d'images dont le but est l'étude des objets en fonction de leur forme, de leur taille, des relations avec leur voisinage (en particulier topologiques), de leur texture, et de leurs niveaux de gris ou de leur couleur. Par les transformations qu'elle propose, elle se situe à différents niveaux du traitement d'images (filtrage, segmentation, mesures, analyse de texture) et fournit ainsi des outils pour la reconnaissance des formes. Elle trouve ses applications dans de nombreux domaines du traitement d'images en robotique et vision par ordinateur. Dans [27] l'auteur montre l'utilité des opérations morphologiques pour l'écriture arabe.

Cette méthode s'appuie sur la théorie des ensembles, des treillis, de la topologie des formes et des probabilités. Elle s'applique ainsi à des structures algébriques variées (ensembles, fonctions, mais également ensembles flous ou propositions logiques), pouvant avoir un caractère aléatoire (pour l'analyse de textures par exemple).

Dans [6] F.S .Douzidia définit à la base des transformations de morphologie mathématique quatre principes fondamentaux qui guident leur construction et qui sont vérifiés pour la plupart des opérations. Ils sont exprimés ci-dessous pour une opération Ψ quelconque, agissant sur un ensemble ou une fonction définie sur \mathbb{R}^n :

- **Compatibilité avec les translations** : Ce principe exprime l'indépendance des

transformations par rapport à l'origine de l'espace :

$$\forall t \in \mathbb{R}^n, \psi(f+t) = [\psi(f)+t] \quad (3.1)$$

Compatibilité avec les homothéties : Ce principe assure l'indépendance des transformations par rapport à un paramètre d'échelle

$$\forall \lambda \in \mathbb{R}, \psi(\lambda f) = [\lambda \psi(f)] \quad (3.2)$$

Connaissance locale : Pour connaître $\Psi(f)$ dans un domaine Z' de \mathbb{R}^n il suffit de connaître f dans un domaine Z de \mathbb{R}^n

$$\forall z' \subset \mathbb{R}^n, z' \text{ borné}, \exists z \subset \mathbb{R}^n, z \text{ borné}, [\psi(f \cap z)] \cap z' = \psi(f) \cap z' \quad (3-3)$$

Où $f \cap z$ désigne la restriction de f au domaine Z

➤ **Semi -continuité** : ce principe assure la robustesse des transformations. Outre ces propriétés fondamentales, les opérateurs de morphologie mathématique peuvent avoir des propriétés algébriques dont les principales sont définies ci-dessous.

❖ **Définition 1. Croissance**: Une transformation Ψ sur des ensembles de \mathbb{R}^n ou des Fonctions de \mathbb{R}^n dans \mathbb{R}^n est croissante si :

$$\forall X, Y / X \subset Y \Rightarrow \psi(X) \subset \psi(Y) \quad (3-4)$$

$$\forall f, g / f \leq g \Rightarrow \psi(f) \leq \psi(g) \quad (3-5)$$

❖ **Définition 2. Extensivité et anti-extensivité** : Une transformation Ψ sur des ensembles de \mathbb{R}^n ou des fonctions de \mathbb{R}^n dans \mathbb{R}^n est extensive si :

$$\forall X, X \subset \psi(X), \quad (3-6)$$

$$\forall f, f \leq \psi(f). \quad (3-7)$$

Ψ est anti-extensive si :

$$\forall X, \psi(X) \subset X \quad (3-8)$$

❖ **Définition 3. Idempotence** : Une transformation Ψ sur des ensembles de \mathbb{R}^n ou des fonctions de \mathbb{R}^n dans $\overline{(\mathbb{R}^n)}$ est idempotente si :

$$\forall X, \psi[\psi(X)] = \psi(X) \quad (3-9)$$

$$\forall f, \psi[\psi(f)] = \psi(f) \quad (3-10)$$

❖ **Définition 4. Dualité** : Deux transformations Φ et Ψ sur des ensembles de \mathbb{R}^n ou des

fonctions de R^n dans (R^n) sont duales par rapport à la complémentation si :

$$\forall X, \Psi(X^C) = [\Phi(X)]^C \quad (3-11)$$

Où X^C désigne le complémentaire de X dans R^n (c'est-à-dire $R^n \setminus X$),

$$\forall f, \Psi(-f) = -\Psi(f) \quad (3-12)$$

3.1.2 Définitions et application

La morphologie mathématique est un ensemble de techniques pouvant être utilisées pour traiter les images binaires qui Permettent d'améliorer le résultat de la segmentation, elle est basée sur des opérateurs mathématiques et ensemblistes, Certains opérations sont étendues sur les niveaux de gris telles que.

- La morphologie mathématique Basée sur l'utilisation de la théorie des ensembles.
- Les images (binaires et au niveau de gris) sont des ensembles sur lesquels les opérations s'effectueront.
- Les transformations morphologiques peuvent s'effectuer en N dimensions Euclidiennes, mais on travaille fréquemment sur des images binaires, dans un espace bidimensionnel.
- L'utilisation des opérateurs algébriques tels que ceux de la morphologie mathématique est utile, parce que l'on peut éliminer des parties sans importance dans des images complexes, tout en gardant des éléments significatifs.
- Ces méthodes forment un groupe important d'opérateurs qui sont utilisés en vision par ordinateur.

3.1.3. Images binaires

Une image binaire est une image pour laquelle chaque pixel ne peut avoir pour valeur que 0 ou 1. La manipulation de telles images regorge d'outils spécialisés ainsi que de théories mathématiques pour plusieurs raisons :

- Les débuts du traitement des images numériques ne permettaient pas le traitement d'images complexes (problème de temps de calcul, d'espace mémoire disponible et de qualité des périphériques de sortie).
- Les images binaires sont un contexte simple permettant une formalisation mathématique des problèmes par des outils tels que la topologie.

3.1.4 L'analyse des images et La morphologie mathématique

Un opérateur de morphologie mathématique reçoit une image en entrée et fournit une image en sortie. A la base, il s'agit d'images binaires, bien que l'on puisse étendre ces opérations Aux images en niveaux de gris.

On peut définir un opérateur de morphologie mathématique de deux manières : par l'intermédiaire d'un élément structurant (c'est la définition la plus proche de l'idée intuitive que l'on a de ces opérateurs) ou par l'intermédiaire d'une table de correspondance (définition moins intuitive, mais plus systématique) [9].

3.1.5 Différentes opérations de la morphologie mathématique

Un élément structurant est une forme qui définit une opération morphologique. On a deux opérations morphologiques de base : l'érosion et la dilatation. A partir de ces deux opérations de base, on peut définir des opérations plus complexes, telles que l'ouverture et la fermeture, le squelettisation et la détection de cavités [10].

a. L'érosion

Un espace R^2 partiellement occupé par un ensemble X . Prenons un élément structurant B représentant une figure géométrique simple, un cercle par exemple. Cet élément B^x est repéré par son centre et placé en X dans l'espace R^2 (Figure 2.1). Il est ensuite déplacé de telle sorte que son centre occupe successivement toutes les positions x de l'espace.

L'ensemble des positions x correspondant à une réponse positive forme un nouvel ensemble Y Appelé érodé de X par B . Cet ensemble satisfait l'équation : $Y = \{X / B_x \subset X\}$ (3.13)

L'érosion Y d'un ensemble X par un élément structurant B est notée habituellement :

$$Y = E_B(X) \quad (3.14)$$



Élément
structurant

الجزائر

image originale

الجزائر

image érodé

Figure 3.1: Principe de l'érosion d'un ensemble.

b. Dilatation

L'opération de dilatation se définit de façon analogue à l'opération d'érosion. En prenant le même élément structurant B_x , on se pose pour chaque point X de R^2 de savoir si B_x touche l'ensemble X , c'est à dire si B_x et X ont une intersection non vide : $B_x \cap X \neq \emptyset$. L'ensemble des points X de R^2 correspondant aux réponses positives forme un nouvel ensemble Y dont la frontière est le lieu des centres géométriques de B_x lorsque B_x touche X . L'ensemble Y satisfait l'équation :

$$Y = \{X / B_x \cap X \neq \emptyset\} \quad (3.15)$$

La dilatation d'un ensemble X par un élément structurant Y est habituellement notée $Y = D_B(X)$. On en trouve un exemple Figure (3.2).

الجزائر

Image originale

الجزائر

image dilaté

Figure 3.2: Exemple d'une dilatation par un élément structurant De forme circulaire

c. Ouverture morphologique

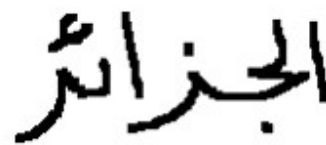
L'ouverture est l'application de l'opérateur érosion puis de l'opérateur dilatation avec le même élément structurant. On appelle ouverture d'un ensemble X par un élément structurant B et on note $O^B(X)$:

$$O_B(X) = D_X(E_B(X)) \quad (3-16)$$

En général, on ne retrouve pas l'ensemble de départ car une partie de la forme éliminée par l'érosion ne peut être recréée par une dilatation. L'ensemble $O^B(X)$ est plus régulier (moins de Détails au niveau du contour) que l'ensemble initial X . En termes "géographiques" ou morphologiques, on dit que l'ouverture adoucit les contours, coupe les isthmes étroits, supprime les petites îles et les caps étroits. On peut considérer que l'ouverture réalise une sorte de filtrage morphologique de la forme de l'ensemble X . La Figure (3.3) illustre ce comportement.



Image originale



ouverture morphologique

Figure 3.3: Illustration de l'ouverture morphologique

d. Fermeture morphologique

La fermeture est l'opération "inverse" de l'ouverture, c'est à dire que l'on applique tout d'abord une dilatation puis une érosion (toujours en gardant le même élément structurant):

$$F^B(X) = E^B(D^B(X)) \quad (3-17)$$

Un ensemble fermé est également moins riche en détail que l'ensemble initial. La transformation par fermeture bouche les canaux étroits, supprime les petits lacs et les golfes étroits (Figure3.4).



Figure 3.4: Illustration de l'ouverture et de la fermeture morphologiques

Sur des ensembles non connexes

Remarque: Les deux opérations d'ouverture et de fermeture sont idempotentes, c'est à dire que le résultat est invariant après itérations:

$$\begin{aligned} O^X(X) &= O^X(O^X(X)) \\ F^B(X) &= F^B(F^B(X)) \end{aligned} \quad (3.18)$$

Ces transformations ne sont pas homotopiques, c'est à dire qu'elles ne préservent pas la connexité car elles peuvent scinder une forme en deux (érosion) ou fusionner deux formes (dilatation).

Elles ne présentent donc pas de propriétés topologiques intéressantes. Ce ne sont que des transformations d'aspect, de simplification ou de filtrage.

e. La squelettisation

La squelettisation est une étape essentielle dans la reconnaissance de forme. Elle a pour but de décrire chaque objet par un ensemble de lignes infiniment fines (analogues à une ossature de l'objet), réduisant sensiblement le volume d'information à manipuler. Le squelette est généralement défini comme étant l'ensemble des lignes médianes, c'est-à-dire l'ensemble des points équidistants à deux points de la frontière. Plusieurs chercheurs ont été intéressés par la squelettisation, T.Zhang propose dans [88] un algorithme parallèle rapide. Dans [89] X.David et H.Yan proposent une amélioration de squelette.

Pour notre application nous avons utilisé l'amincissement qui peut donner une squelette qui garde la forme et la topologie des caractères car l'écriture en général est formée d'un ensemble de traits.

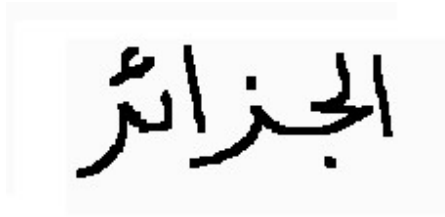


Figure.3.5.a Image originale



Figure3.5.b Squelette d'image

Figure 3.5: Squelettisation

3.2 Transformée de Hough

L'objectif envisagé avec la TH est de définir un changement de représentation de l'image pour que l'ensemble des points appartenant à une forme géométrique spécifique soit identifié par un seul point dans l'espace transformé. Pour cela, on exploite la représentation paramétrique des formes géométriques à détecter.

La transformée de Hough est utilisée et adaptée en tolérant un angle d'inclinaison de 0° à 180° . C'est l'algorithme standard de l'analyse d'image digitale pour détecter des formes qu'on peut paramétriser, par exemple des lignes, courbes ou cercles.

3.2.1. Introduction

La TH a été développée par Paul Hough en 1962 [14], la TH est devenue un outil standard dans le domaine de la vision artificielle. Mais des modifications peuvent être apportées pour reconnaître n'importe quelle forme : c'est la transformée de Hough généralisée est développée par Richard Duda et Peter Harte en 1972 [15].

La TH permet de détecter dans l'image la présence de courbes paramétriques appartenant à une famille connue (droites, cercles, ellipses, etc.) à partir d'un ensemble de points sélectionnés. Elle peut aussi être étendue à des cas de description d'objets plus complexes. La TH permet d'avoir une bonne robustesse contre les discontinuités et les éléments manquants d'une structure. Le principe général de la transformée de Hough est d'établir une projection

entre l'espace de l'image et un espace de paramètres représentatifs de la forme recherchée appelée espace de Hough.

3.2.2 Transformée de Hough Standard (THS)

La transformée de Hough Standard (THS) est une méthode classique de détection de lignes droites dans une image. Elle a été utilisée dans plusieurs applications : la détection des routes dans les images, la lecture robuste de code-barre [12], et dans d'autres applications. Dans ce mémoire nous allons utiliser la THS pour la détection de l'angle d'inclinaison de la ligne de base d'un mot arabe manuscrit.

Partant de la dualité point-sinusoïde et point droite la THS propose un Détecteur de ligne droite dans les images. Pour détecter les lignes droites dans une image, chaque point de l'image binaire est transformé en une sinusoïde dans l'espace des paramètres (discrétisé) et les points parcourus par le plus de sinusoïdes représentent les droites passant par le plus de points dans l'image. La recherche des lignes droites dans une image se réduit donc à la détection des maximums locaux (pics) dans l'espace des paramètres.



Figure.3.6: Image avec deux lignes

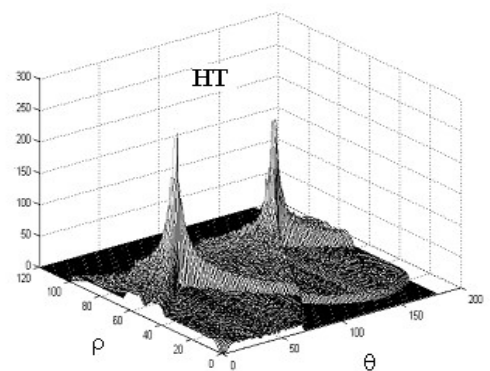


Figure.3.7: Deux Crêtes de l'image a TH

3.2.3 Définition

La transformée de Hough Standard [16] est basée sur la représentation normale d'une droite :

$$y=mx+c \quad (3-19)$$

Chaque point dans l'espace de Hough correspond une ligne un angle et une distance par rapport à l'origine. La valeur est en ce moment une expression de combien des points de donnée se trouvent sur cette ligne. La transformée de Hough peut être employée comme manière de trouver les lignes droites dans un ensemble de données. Les points dans l'espace de Hough avec une valeur élevée signifient beaucoup des points de données se trouvant le long de la ligne Correspondante dans l'espace cartésien. Ceci peut être employé pour trouver la ligne de base parce qu'il est caractérisé par beaucoup de Pixel de premier plan se trouvant suivant une ligne presque droite.

$$\rho =X \cos\theta+Y \sin\theta \quad (3-20)$$

$$D(\rho, \theta) = \{ (X, Y), | X \cos \theta + Y \sin \theta | = \rho \} \quad (3-21)$$

Où $(\cos \theta, \sin \theta)$ est le vecteur normal à la droite et $|\rho|$ représente la distance de la droite à l'origine. (figure.3.6). Soit un point P0 de coordonnées cartésiennes (X_0, Y_0) . Et de coordonnées polaires (θ_0, ρ_0) , de telle sorte que: $(X_0, Y_0) = \rho_0 (\cos \theta_0, \sin \theta_0)$ Ce point peut être identifié à l'intersection des droites passant par P0 (θ_0, ρ_0) . Comme le point P0 (θ_0, ρ_0) doit être sur chacune des droites D (θ, ρ) passant par ce point, nous avons

$$X_0 \cos\theta+Y_0\sin\theta = \rho \quad (3.22)$$

Et en utilisant ses coordonnées polaires, on a

$$\rho_0(\cos\theta_0\cos\theta+\sin\theta_0\sin\theta) = \rho \quad (3.23)$$

D'où

$$\rho_0 \cos(\theta-\theta_0) = \rho \quad (3.24)$$

Ce qui correspond à une sinusoïde de paramètres (θ_0, ρ_0) dans le système d'axes (θ, ρ) Que nous noterons $s(\theta_0, \rho_0)$ Figure.3.9.

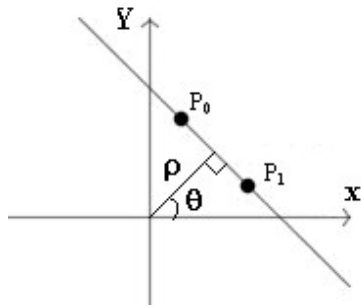


Figure.3.8: L'espace (x, y)

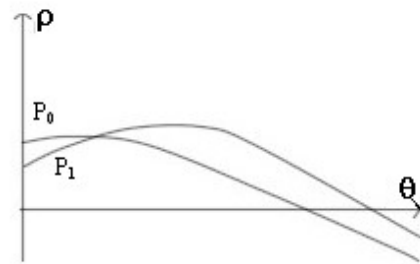


Figure.3.9: L'espace de Hough (ρ, θ)



Figure. 3.10 : Exemple d'un mot manuscrit

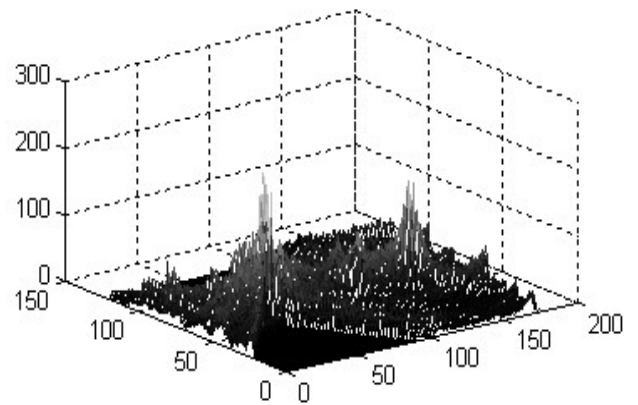


Figure. 3.11 : La TH du mot manuscrit

3.2.4. Propriétés de La THS

Une image I sera décrite à l'aide des coordonnées cartésiennes de ses points. Nous noterons par H l'espace des coordonnées (ρ, θ) associées aux droites. D'après sa définition, la Transformée de Hough Standard possède les propriétés suivantes.

- Propriété 1 : Un point dans l'espace d'observation (image I), correspond à une sinusoïde dans l'espace des paramètres H .
- Propriété 2 : Des points appartenant à la même sinusoïde dans l'espace des

paramètres H correspondent à des lignes qui se coupent en un même point dans l'espace d'observation I .

- Propriété 3 : Un point de l'espace des paramètres correspond à une droite dans l'espace d'observation I .
- Propriété 4 : Des points appartenant à une même droite D dans l'espace d'observation I correspondent à des sinusôides qui se coupent en un même point $P(\rho_0; \theta_0)$ dans l'espace des paramètres H .

- Propriété 5 : En utilisant les notations précédentes nous avons :

$$si(\theta, \rho) \in S(\theta_0, \rho_0) \text{ alors } (\theta + l\pi, (-1)^l \rho) \in S(\theta_0, \rho_0) \text{ Pour tout } l \in \mathbb{Z} \dots\dots (3.25)$$

$$D(\theta, \rho) = D(\theta + l\pi, (-1)^l \rho) \forall l \in \mathbb{Z} \dots\dots\dots (3.26)$$

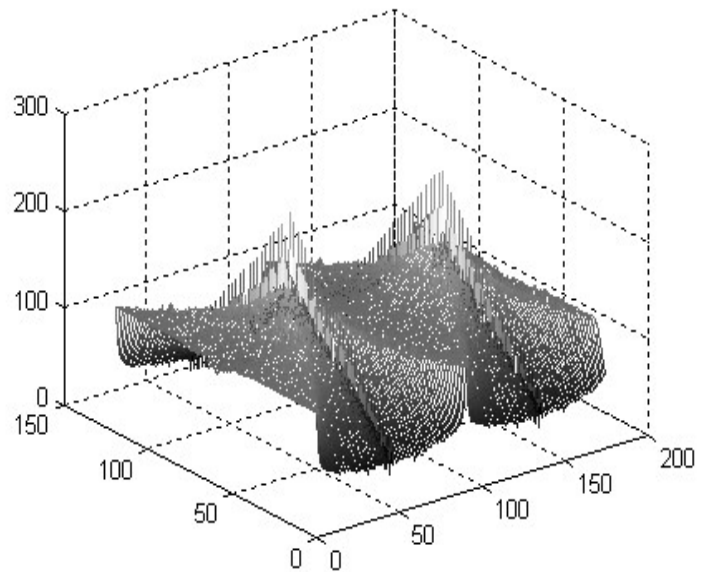


Figure.3.12.a : squelette d'un mot

Figure.3.12.b : TH du squelette du mot

3.2.5 La transformé de Hough généralisée

La technique de la TH a pu être généralisée pour la détection de formes arbitraires contenues dans l'image. Les formes recherchées par la transformée de Hough Généralisée (THG) ne sont pas obligatoirement définies de manière analytique mais plutôt par une silhouette particulière [81]. Un tableau de référence nommé R- table définit donc la correspondance entre l'espace de définition de la forme recherchée (espace image) et l'espace paramétrique [16]. La THG agit donc sur des points caractéristiques de l'image, généralement le contour. Si des points de contrastes de l'image sont alignés, les droites correspondantes de l'espace de Hough passent toutes par un même point (m, c). la THG permet de détecter des formes paramétriques (cercle, droite, ellipse, etc.) ou une forme quelconque.

- Quelconques : $y=f(x,a,b,c,\dots)$
- Ellipses : $[(x-x_0)^2/a^2 + (y-y_0)^2/b^2] = 1$ (4paramètres+éventuellement orientation)
- Cercles : $(x-a)^2 + (y-b)^2 = r^2$ (3paramètres)

Pour les cercles :

On considère l'espace de Hough h (a, b, r). Chaque point (x, y) de l'image correspond à un cône de l'espace (a, b, r). Pour un rayon fixé, chaque point (x, y) correspond à un cercle de l'espace (a, b,r). Idée de l'algorithme :

Pour chaque rayon $r > 0$ on trace les cercles de l'espace de Hough correspondant aux points de l'image. Lorsque tous les cercles se coupent en un même point, on a trouvé le bon rayon, les coordonnées (a, b) de ce point correspondent au centre du cercle. Dans [80] S.Touj et N.B.Aessoukri utilisent la THG pour la description des formes pour la reconnaissance des caractères arabes.

3.2.6 Conclusion

La première partie de ce chapitre décrit la théorie de la morphologie mathématique ainsi que ses différentes opérations sur les images binaires. Dans la suite de notre travail nous exploiterons ses méthodes pour le calcul des cavités et nous utiliserons le squelette de la forme qui à l'avantage de préserver la totalité de l'information forme avec un minimum de points. Dans la deuxième partie de ce chapitre nous nous avons décrit les fondements mathématiques de la transformée de Hough qui permet d'avoir une bonne robustesse contre les discontinuités et les éléments manquants d'une structure. Le principe général de la transformée de Hough est d'établir une projection entre l'espace de l'image et un espace de paramètres représentatif de la forme recherchée appelé espace de Hough. Cette transformée est un outil très intéressant pour la détection de la ligne de base ainsi que l'inclinaison des mots qui sera corrigé par la suite par la rotation dans le sens opposé de l'inclinaison, ces méthodes seront détaillées dans le chapitre suivant.

Chapitre 4

Paramètres de prétraitement

4.1 Introduction

On regroupe souvent sous le terme de prétraitement toutes les opérations qui sont appliquées aux images, indépendamment de leur usage futur pour assurer une bonne qualité. Elles concernent donc essentiellement les corrections de contraste et la suppression du bruit. Cette pratique est souvent critiquée car l'expérience montre qu'il est toujours important d'adapter les traitements plus élaborés et donc aux objectifs à long terme du projet, mais les prétraitements sont souvent appliqués dans des circonstances où l'on destine les images à de nombreuses applications différentes dont on ignore souvent les besoins exacts.

Le rôle du prétraitement est de préparer les données reçues du capteur à la phase suivante d'analyse consacrée à l'extraction des paramètres [34]. Cette phase n'est possible et surtout fiable que si les données du capteur sont dénuées du bruit, corrigées de leurs erreurs éventuelles, homogénéisées, normalisées et réduites à l'essentiel. Toutes les techniques élaborées dans ce sens se gardent de modifier les propriétés essentielles des formes, ce qui pourrait conduire, dans le cas contraire, à de graves erreurs d'analyse et plus tard de reconnaissance.

La plupart des systèmes de reconnaissance comportent une étape de prétraitement après que l'acquisition soit faite, son but est améliorer les résultats et les performances du module de reconnaissance [76].

Nous avons employé les opérations de prétraitements suivantes:

4.2 Binarisation :

Un image binaire est représentée par une matrice rectangulaire dont les valeurs sont 0 ou 1. La binarisation est une transformation de l'image à plusieurs niveaux de gris en une image à deux niveaux de gris noir et blanc, cette méthode consiste à faire un seuillage des niveaux de gris, on distingue deux approches globale par calcul d'histogramme [67], et adaptatif où le seuil n'est pas le même pour tout les points de l'image une méthode de seuillage adaptatif est décrite dans [22].

Une étude comparative des méthodes de binarisation est décrite dans [68], [82] et [74].

Dans [83] et [84] une binarisation spécifique pour les documents est décrite. Dans notre cas nous avons utilisé la méthode globale car le nombre d'objets dans notre cas n'est pas important car le caractère est ses points ont presque le même niveau de gris si le même stylo est utilisé.

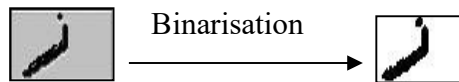


Figure 4.1: exemple de binarisation

4.3 Squelettisation :

Pour une représentation structurelle de l'objet, les caractéristiques décrivant la topologie de l'objet sont préférables. Le contour et le squelette sont les caractéristiques intrinsèques de l'objet. La notion de squelette est décrite dans [42] comme étant le résultat d'une transformation de l'axe médian. La transformation de l'axe médian détermine les points de contour les plus proche de chaque point dans une région S donnée. Un point à l'intérieur de S appartient au squelette si au moins deux de ses points les plus proches sont des points de contour. Le squelette a plusieurs propriétés intéressantes.

Il conserve les propriétés topologiques de la forme d'origine ainsi que ses propriétés géométriques [72].

Le squelette connaît plusieurs applications telles que la reconnaissance des formes, la modélisation de solides pour la conception et la manipulation de formes, l'organisation de nuages de points. Des applications ont été trouvées dans l'indexation d'images dans les bases de données et en compression. Il existe également quelques applications en architecture et en urbanisme, dans le cadre d'analyse morphologique [89].

La squelettisation sert à obtenir une épaisseur égale à un pixel du trait d'écriture et de se

ramener ainsi à une écriture linéaire [54]. Le squelette doit préserver la forme, connexité, topologie et extrémités du tracé, et ne doit pas introduire d'éléments parasites Figure 4.2. Dans notre travail nous avons utilisé la méthode d'amincissement.

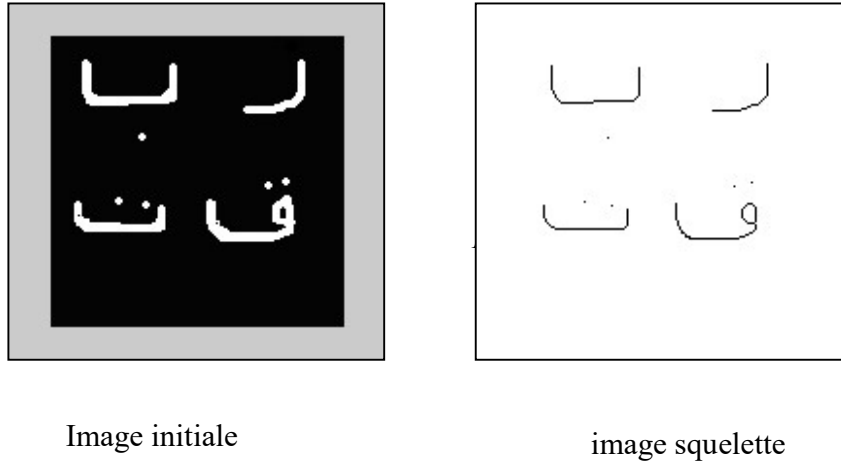


Figure 4.2: La squelettisation

4.4 filtrage basé sur le squelette :

La méthode de filtrage employée dans ce travail est à base du squelette. Premièrement nous estimons l'épaisseur de trait d'écriture, qui dépend du stylo. On prend **im** comme image initiale, **sk** est le squelette de l'image, **maxx** est la largeur de l'image et **maxy** est la longueur de l'image.

L'épaisseur d'écriture **wd** peut être estimé par (4.1).

$$wd = \frac{\sum_{x=1}^{Maxx} \sum_{y=1}^{Maxy} Im(x,y)}{\sum_{x=1}^{Maxx} \sum_{y=1}^{Maxy} Sk(x,y)} \quad (4.1)$$

La deuxième étape consiste à calculer la surface de chaque composant connexe dans l'image

La dernière étape permet de garder seulement les composants qui ont une surface plus grande que $wd^2 / 2$

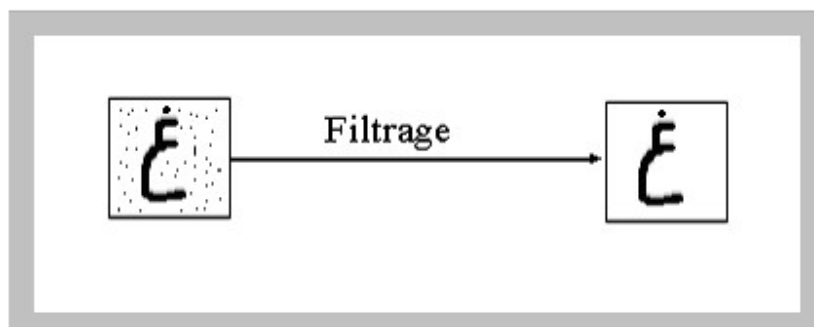


Figure 4.3: Filtrage d'un caractère arabe

Ce traitement a été appliqué pour éviter l'élimination de la forme significative du mot comme les points.

Figure.4.4: Image originale.

Figure.4.5: squelette d'image.

Figure.4.6: image originale d'un nom de willaya.

Figure.4.7: image filtrée.

4.5 Points caractéristiques

Les points caractéristiques sont employés pour une description morphologiques des mots ou des caractère. Il existe trois types de point caractéristique (points d'intersection, point final et point de coin).

4.5.1 Point d'intersection

Concerne le point sur un segment qui a plus de deux branches reliées.

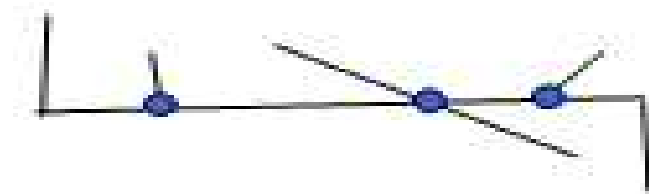


Figure 4.8: points d'intersections.

4.5.2 Point d'extrémité

concerne le point sur un segment qui a seulement un seul points parmi ses huit voisins.

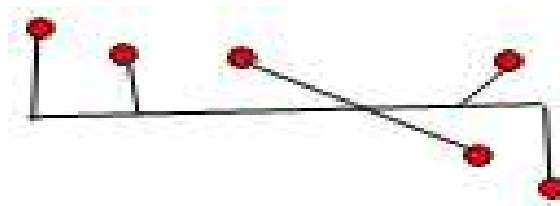


Figure.4.9 : points d'extrémités.

4.5.3. Point de coin

Le point sur un segment où le point n'est pas un point d'intersection et la courbure du segment change brusquement.

Un des éléments qui caractérisent la forme d'une courbe est son degré de flexion, d'incurvation. Soit donnée une courbe qui n'a pas de points double et qui a une tangente déterminée en chaque point. Menons les tangentes à la courbe en deux points quelconques A et B et désignons par α l'angle formé par ces tangentes ou, plus exactement, l'angle de rotation de la tangente quand on passe du point A au point B (Figure.4.7).on appelle cet angle de contingence de l'arc AB. De deux arcs de même longueur .le plus incurvé est celui dont l'angle de contingence est le plus grand angle.

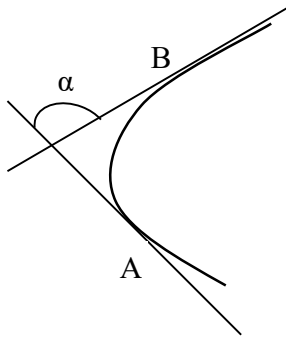


Figure.4.10: la courbure d'une tracé.

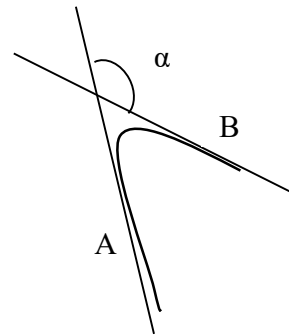


Figure.4.11:l'angle α de la courbure.

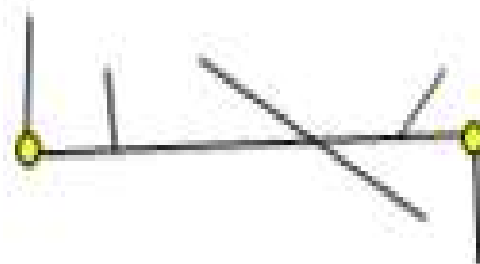


Figure.4.12 : points des coins.

4.6 Ligne de base et correction des mots renversés

4.6.1. Correction de la ligne de base

La ligne de base a une très grande importance dans l'écriture cursive soit pour l'approche globale et l'approche analytique

Dans [69] on utilise des mesures statistiques pour la détection et la correction de la ligne de base. Une méthode basée sur les extrema locaux est détaillée dans [85].

La ligne de base dans l'écriture arabe a des particularités plus importantes que celle du latin car cette ligne pour l'arabe est bien marquée et peut être bien rapprochée par une ligne droite horizontale d'où la nécessité de corriger cette ligne pour faciliter la tâche de reconnaissance pour le cas où on utilise une approche globale ainsi que pour la segmentation du mot, si on utilise l'approche analytique. Dans notre travail nous avons développé que des algorithmes de prétraitement et de segmentation pour le mot [69].

Dans notre travail la correction de la ligne de base est basée sur la TH qui estime l'orientation globale de la ligne de base du mot en recherchant le maximum de TH dans le voisinage horizontal. Ensuite nous corrigeons l'inclinaison du mot en utilisant ces deux équations :

$$x' = x \cos(\theta) + y \sin(\theta) \quad (4-2)$$

$$y' = -x \sin(\theta) + y \cos(\theta) \quad (4-3)$$

(X, Y) est la coordonnée du Pixel dans l'image l'initiale.

(X', Y') est la coordonnée du Pixel dans l'image corrigée. θ est l'orientation estimée pour la ligne de base.



Figure.4.13 : image d'entrée.



Figure .4.14: squelette de l'image.

A chaque point dans l'espace de Hough correspond une ligne et un angle et une distance de l'origine. La valeur est en ce moment une expression de combien des points de donnée se trouvent sur cette ligne. La TH peut être employé comme manière de trouver les lignes droites dans un ensemble de données. Les points dans l'espace de hough avec une valeur élevée signifient beaucoup du point de données se trouvant le long de la ligne correspondant dans l'espace cartésien. Ceci peut être employé pour trouver la ligne de base parce qu'il est caractérisé par beaucoup de pixel de premier plan se trouvant suivant une ligne droite

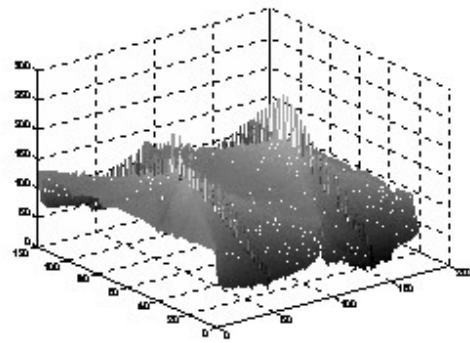


figure.4.15: Image originale

figure.4.16: TH d'une image

4.6.2 Correction du mot inversée

Après correction de la ligne de base nous calculons la projection horizontale du mot et nous appliquons des analyses statistiques formulées par (4.4) et (4.5) qui nous permettent de savoir si le mot est inversé. Nous tournons le mot de 180° s'il est inversé.



Figure.4.17 : Correction de la ligne de base et projection verticale

$$g(y_0) = \max(g(y)) \quad (4.4)$$

$$\sum_{y=y_0-2 \cdot wd}^{y_0-1} g(y) > \sum_{y=y_0}^{y_0+2 \cdot wd} g(y) \quad (4.5)$$

Le $g(y)$ est la projection horizontale du mot y_0 est la coordonnée du maximum de g , wd est la largeur de l'écriture, estimée en lissant la section d'image.

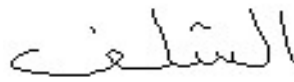


Figure.4.18: Correction de mot renversée.

4.7. Conclusion

Dans ce chapitre nous avons présenté le prétraitement des mots arabes manuscrits et des caractères isolés pour ces mots : nous utilisons une nouvelle méthode de filtrage, et la transformée de Hough pour estimer et corriger la ligne de base et nous avons employé la projection horizontale afin de détecter n'importe quel mot renversé. Dans notre travail nous avons utilisé le squelette d'image pour le filtrage, et les points caractéristiques (point final, intersection et point de coin) pour la correction des segments verticaux et inclinés.

Pour les caractères isolés nous avons utilisé le même algorithme de filtrage utilisé pour les mots.

Chapitre 5

La Segmentation

5.1. Introduction

Plusieurs efforts ont été consacrés pour la reconnaissance de l'écriture cursive mais jusqu'à présent le problème est loin d'être résolu, beaucoup de méthodes ont été développées pour l'écriture cursive pour le texte latin imprimé, bien que certaines méthodes pour le latin cursif pourraient être appliquées à l'arabe, en général, elles ne donnent pas de bons résultats [66]. La segmentation et la reconnaissance de l'écriture arabe dépendent fortement du contexte de l'information et non seulement des paramètres topologiques des caractères.

L'analyse de l'écriture cursive demande une description structurelle par la reconnaissance des lettres arabes, la segmentation des caractères dans le mot et la détection des primitives de chaque lettre. Ce problème n'est pas posé seulement pour l'ordinateur mais aussi pour l'être humain qui fait des erreurs de l'ordre de 4% dans la lecture sans contexte []. Les erreurs dépendent principalement des variabilités des formes des caractères d'une personne à une autre [86].

L'être humain peut facilement segmenter le mot arabe en lettres bien que l'ordinateur ne peut pas le faire. Segmenter en lettres parfaites n'est pas toujours une opération possible surtout dans le cas des ligatures (chevauchements de deux lettres) dans la position verticale le deuxième caractère du mot peut commencer avant la fin du premier ou même avant son début Figure.5.1.



Figure 5.1: Exemple de ligature (chevauchement).

Pour cela, la plus part des recherches considèrent les ligatures comme une seule forme à reconnaître, ce qui va étendre le nombre de classes .certaines méthodes segmentent le mot et les lettres en petits morceaux (graphèmes) puis utilisent un algorithme de fusion de ces graphème pour former les lettres.

5.2 Les méthodes classiques de segmentation

Les travaux de Parhami et Traghi [17] en 1981 sont les premiers essais pour segmenter les caractères arabes. Depuis ce temps plusieurs méthodes on été proposées, la projection verticale à été employée pour cette raison. Ensuite la détection du squelette du mot et le trace systématique du mot en cherchant les points de segmentation possible [22]. Cette méthode a été suivie par un essai de segmentation du mot en traçant le contour et la fonction de distance supérieure. Les réseaux de neurones, le graphe d'adjacence des lignes et les opérations morphologiques sont aussi utilisés[87].

5.2.1. Méthode de segmentation basée sur la projection verticale

Le but de la méthode de la projection verticale et de simplifier un système de reconnaissance des caractères, en réduisant une information bidimensionnelle en une information monodimensionnelle. Cette méthode est apparue au début de l'OCR [20].elle fonctionne mieux avec l'écriture imprimée, spécialement avec les sources qui ne forme pas de ligatures comme « arabic transparent » et « simplified arabic » ainsi que pour les sources comme « traditionnel arabic » qui contiennent plusieurs ligatures. Nous ne notons pas beaucoup de pour l'écriture manuscrite. Cette méthode est basée sur le fait que les ligatures de connexion sont toujours moins larges que les autres parties de mot.

5.2.2. Méthode de segmentation basée sur la fonction de distance supérieure

La fonction de distance supérieure est l'ensemble des points les plus hauts de chaque colonne. Pour chaque fonction de distance supérieure, KURDY et JOUKHADAR [21] ont détermine la ligne de base et chaque sous mot est mesuré. Finalement, une des trois mentions (supérieur, inférieur, moyenne) est donnée à chaque point, les mentions sont reliées à la distance verticale entre la ligne de base et le max de chaque colonne de chaque point et la distance verticale du point précédent. En utilisant une grammaire, il passe la séquence des mentions d'un sous mot pour trouver les point de connexion. Dans cette méthode les points

adjacents ayant la même mentions construisent un passage. si un passage satisfait certaines conditions, telle que La longueur du passage est plus large que le $\frac{1}{3}$ que la taille du stylo, le dernier point du passage est marqué comme un point de segmentation potentiel. Les chercheurs ont prouvé que l'avantage de cette méthode réside dans le fait que le caractère peut être complètement obtenu en une seule pièce, donc le nombre des formes différentes est minimal ,ce qui facilite la reconnaissance.

5.2.3. Méthode de segmentation basée sur le squelette

Dans la reconnaissance de caractère, l'information essentielle de la forme est préservée dans son squelette [23]. Dans la méthode de Ahmed.Z [25], la ligne de base du squelette est trouvée en premier, ensuite seulement les colonne qui n'ont pas de pixels au dessus de la ligne de base sont considérés dans la recherche. Les points de segmentation seront au milieu du segment de connexion.

5.2.4. Méthode de segmentation basée sur la reconnaissance

Elle est différente des méthodes décrites qui sont considérées comme des méthodes de segmentation explicites. Alors que cette méthode est implicite [26]. Dans la segmentation explicite, les mots sont explicitement ou extérieurement segmentés en caractères ou pseudo caractères, qui sont ensuite reconnue individuellement, Un niveau élevé de connaissance contextuelles et ensuite utilisé pour la reconnaissance des mots. Cette approche est très conteuse.

5.3. Méthode de segmentation proposée pour un mot

La méthode de segmentation proposée consiste à utiliser et exploiter la ligne de base qui présente une particularité importante dans l'écriture arabe manuscrite, car la majorité des lettres sont connectées sur cette ligne et présentent la partie basse de chaque lettre, pour cela on peut définir pour notre système de segmentation les éléments de sortie qui sont soit des ligatures qui peuvent contenir plusieurs lettres qui se chevauchent verticalement, ou une partie haute de la lettre (Figure.5.2), Certains types d'écritures ont omis quelques parties des caractères, pour cela on utilise des méthodes mathématiques pour les marquer, basées sur l'étude des fonctions de localisation de la position de ces éléments. Aant d'utiliser ces méthodes, nous présentons ci-dessous quelques fondements mathématiques.

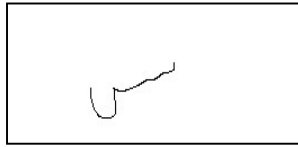


Figure 5.2: la lettre sin.

5.3.1. Fondements mathématiques

5.3.1.1. Définition du maximum

On dit que la fonction $f(x)$ admet un maximum au point x_1 si la valeur de la fonction $f(x)$ est en ce point plus grande qu'en tout autre point d'un certain intervalle contenant le point x_1 . En d'autres termes, la fonction $f(x)$ admet un maximum au point $x=x_1$ si

$$f(x_1 + \Delta x) < f(x_1) \dots \dots \dots (5.1)$$

Pour les Δx (positifs ou négatifs) suffisamment petits en valeur absolus. Par exemple, la fonction $y=f(x)$, dans le graphique représenté sur la Figure (5.3), admet un maximum pour $x=x_1$.

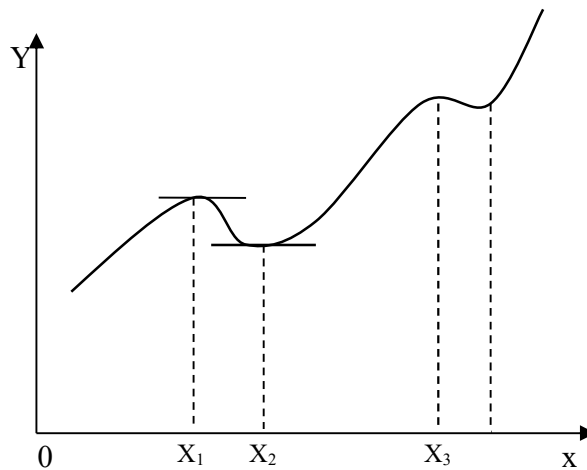


Figure.5.3: La courbe d'une Fonction $y=f(x)$ (5.2)

5.3.1.2. Définition du minimum

On dit que la fonction $f(x)$ admet un minimum pour $x=x_2$ si

$$f(x_2 + \Delta x) > f(x_2) \dots \dots \dots (5-3)$$

Pour tout les Δx (positifs ou négatifs) suffisamment petit en valeur absolu (figure 5.4).

Par exemple, la fonction $y = x^n$, admet un minimum pour $x=0$, puisque $y=0$ pour $x=0$, et $y>0$ pour toutes autres valeurs de x (Figure.5.4).

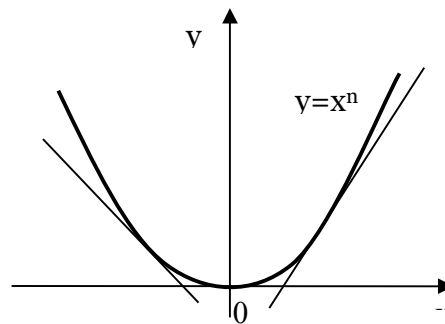


Figure.5.4: La courbe d'une Fonction $y=x^n$

5.3.1.3. Théorème

Soit $y=f(x)$ une fonction dont la dérivée s'annule au point $x=x_1$, c'est à- dire $f'(x_1)=0$
 Supposons, en outre, que la dérivée seconde $f''(x)$ existe et soit continue dans un voisinage du point x_1 .

Soit $f'(x_1)=0$; alors la fonction a un maximum au point $x=x_1$ si $f''(x_1) < 0$ et un minimum si $f''(x_1) > 0$

5.3.2. Extraction de la ligne de base

Après l'extraction du squelette on élimine les éléments secondaires (les points) des lettres en utilisant le même algorithme de filtrage mais en changeant le seuil de façon à éliminer ces points, puis on procède à la recherche des points de position plus bas qui présentent la plus grande partie de la ligne de base sous forme d'une fonction bijective, puis La séparation de la ligne de base est fait par l'algorithme suivant :

Début

1-lecture de l'image

2-création d'un vecteur ligne Y dont la taille est égale à la largeur de l'image et l'initialiser à zéros

3-elimination de la ligne de base : par balayage de l'image du bas vers le haut, de la droite vers la gauche Colonne par colonne

si on trouve un point en initialise notre fonction et en marque l'image imb de la ligne de base par un point spécifié par les cordonnées du point trouvé et on passe à la colonne suivante.

La fonction trouvé est une fonction discontinue qui peut se présenter comme une image.

la soustraction de cette image à l'image initiale donne une image segmenté.

Fin.



Figure.5.5 Image originale« سطيف »



Figure.5.6 : Image sans les éléments secondaires (les points).



Figure.5.7 : la ligne de base.



Figure.5.8: mot segmenté.

5.3.3. Fonction de la ligne de base

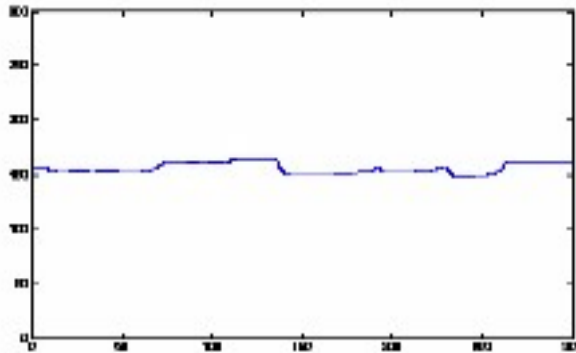


Figure.5.9:La forme de la courbe $Y = f(x)$

5.3.4. Calcul de la première et la deuxième dérivé

On calcule la première et la deuxième dérivée de la fonction pour trouver les valeurs des (Minimum//maximum) de la fonction: $Y = f(x)$

On détermine La forme de la courbe $(x, f(x))$ avec la plus petite longueur correspond une approche spécifique.

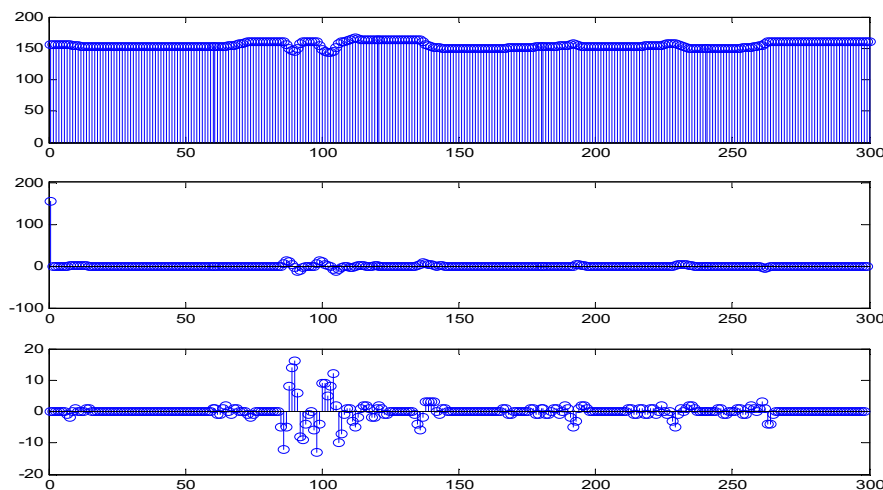


Figure.5.10 : la dérivée de la fonction

5.4 Segmentation du caractère :

Après le filtrage de l'image nous calculons les surfaces des différents composantes connexes de l'image ; la composante de la plus grande surface est affectée à la partie primaire du caractère le reste sont les points (partie secondaire) la figure 5.11 montre un exemple du caractère (Dha) ظ

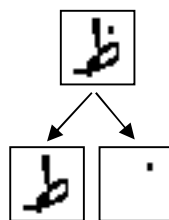


Figure5.11: Segmentation d'un caractère arabe

5.4.1 La position des points

La majorité des caractères arabe se compose de deux parties : une partie primaire et une partie secondaire (les points) qui se situent en haut, en bas, ou bien au milieu.

Pour détecter la position de ces points il faut appliquer l'algorithme suivant :

5.4.2 l'algorithme de position

Le principe de l'algorithme est de faire un balayage à partir des points :

*Si par le balayage vers le **haut** en rencontre la forme*

*alors Le point est en **bas***

*Si par le balayage vers le **bas** en rencontre la forme*

*alors Le point est en **n haut***

*Si par le balayage vers le **haut** en rencontre la forme &*

*Si par le balayage vers le **bas** en rencontre la forme*

*alors le point est en **milieu** .*

5.5 Conclusion

La segmentation est l'étape la plus délicate dans un système de reconnaissance des caractères arabes, ce domaine est très vaste, aucun système qui utilise la segmentation n'est opérationnel à nos jours. Dans ce chapitre nous avons proposé une nouvelle méthode qui prend en considération les caractéristiques de l'écriture arabe et exploite la ligne de base sur laquelle toutes les lettres sont connectées dans le mot, cette méthode fait appel à des outils mathématiques et est considérée comme un nouvel essai qui peut être amélioré par la suite. Pour la segmentation de caractères isolés la séparation de la partie primaire de la partie secondaire; nous avons proposé un algorithme robuste qui utilise l'extraction des composante connexes dans les images ainsi que la détection de la position des points par rapport à la partie primaire et leur nombre, car certains caractères ne se diffèrent que par la position et le nombre des points.

Chapitre 6

Paramètres de classification

6.1 Introduction

L'objet de l'extraction des paramètres de l'objet dans le domaine de la reconnaissance consiste à exprimer ces paramètres numériques ou symboliques appelées *vecteur des primitives*, permettant la discrimination des classes différentes. Un bon choix du vecteur d'attributs se présente comme un vecteur de nombre minimal de paramètres pour éliminer la redondance et l'ambiguïté, mais qui contient le maximum d'informations utiles pour la classification. Parmi les techniques utilisées pour extraire les paramètres de discrimination, on trouve celles basées sur l'image (utilisation des moments, techniques d'histogrammes, techniques de distribution des pixels) ou même à niveau de gris ou bien celles qui sont basées sur des informations extraites à partir de l'image du contour [59].

Parmi les techniques utilisées pour extraire les paramètres de discrimination, on trouve :

6.2 Méthode des moments invariants

Les descripteurs obtenus par moments géométriques sont utilisés pour caractériser et reconnaître des images indépendamment de certaines transformations géométriques, cette indépendance est exploitée dans la reconnaissance, afin de caractériser chaque forme par des paramètres propres à celle-ci, qui permettent de distinguer de toutes les formes celles qui lui sont peu semblables [59]. Les moments invariants dépendent de trois critères qui sont :

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

On répondant à ces trois critères, les moments invariants ont prouvé leur efficacité dans la reconnaissance des formes.

Les moments des fonctions bidimensionnelles $f(x, y)$ d'ordre $(p+q)$ sont définis par :

$$m_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^p y^q f(x, y) dx dy \quad (6.1)$$

Pour $p, q=0, 1, 2, \dots$,

Si $f(x, y)$ est continue par tranches des valeurs non nulles dans une partie finie du plan $x - y$, alors tous les ordres des moments (m_{pq}) existent et déterminent exclusivement $f(x, y)$.

6.3 Méthode des moments centrés

Les moments centrés (u_{pq}) peuvent être exprimés par :

$$u_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x-\bar{x})^p (y-\bar{y})^q f(x, y) dx dy \dots \dots \dots (6.2)$$

Tels que : $\bar{x} = \frac{m_{10}}{m_{00}}, \quad \bar{y} = \frac{m_{01}}{m_{00}}.$ (6.3)

Pour une image discrète l'équation (2) et (3) devient :

$$u_{pq} = \sum_x \sum_y (x-\bar{x})^p (y-\bar{y})^q f(x, y) \quad (6.4)$$

Les moments centrés d'ordre 3 sont donnés par :

$$u_{10} = m_{10} - \frac{m_{10}}{m_{00}}(m_{00}) = 0 \quad (6.5)$$

$$u_{11} = m_{11} - \frac{m_{10}}{m_{00}}(m_{01}) \quad (6.6)$$

$$u_{02} = m_{02} - \frac{m_{10}^2}{m_{00}} \quad (6.7)$$

$$u_{20} = m_{20} - \frac{m_{10}^2}{m_{00}} \quad (6.8)$$

$$u_{30} = m_{30} - 3 \frac{m_{10}}{m_{00}}(m_{20}) + 2 \frac{m_{10}^2}{m_{00}^2} \quad (6.9)$$

$$u_{12} = m_{12} \frac{2m_{01}(m_{11})}{m_{00}} \frac{m_{10}}{m_{00}} (m_{02}) + 2 \frac{m_{01}^2}{m_{00}^2} (m_{10}) \quad (6.10)$$

Les moment centrées et normalisés sont définis par :

$$\eta_{mn} = \frac{\mu_{pq}}{\mu_{00}^\gamma} \quad (6.11)$$

$$\text{Avec : } \gamma = \frac{1}{2}(m+n) + 1 \quad (6.12)$$

A partir des moment du second et du troisième ordre, un ensemble de sept moments invariants peuvent être obtenu [24] :

$$h_1 = \eta_{20} + \eta_{02} \quad (6.13)$$

$$h_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (6.14)$$

$$h_3 = (\eta_{30} - 3\eta_{12})^2 + (\eta_{03} - 3\eta_{21})^2 \quad (6.15)$$

$$h_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{03} - \eta_{21})^2 \quad (6.16)$$

$$h_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) \left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{03} + \eta_{21})^2 \right] + (3\eta_{21} - \eta_{03})(\eta_{03} + \eta_{21}) \times \left[3(\eta_{30} + \eta_{12}) - (\eta_{03} + \eta_{21})^2 \right] \quad (6.17)$$

$$h_6 = (\eta_{20} - \eta_{02}) \left[(\eta_{30} + \eta_{12})^2 - (\eta_{03} + \eta_{21})^2 \right] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{30} + \eta_{21}) \quad (6.18)$$

$$h_7 = (3\eta_{12} - \eta_{30})(\eta_{03} + \eta_{21}) \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{03} + \eta_{21})^2 \right] + (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12}) \left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{03} + \eta_{21})^2 \right] \quad (6.19)$$

6.4 détection des cavités

Ici nous allons nous intéresser à la détection automatique de la forme d'un caractère. On définit cinq types de cavités : Ouest, Nord, Sud, Est et Centrale. Cela suppose la définition de 4 directions cardinales dans l'image : Ouest vers la gauche, Nord vers le haut, Est vers la droite et le sud vers le bas [11].

Par exemple, un point de l'image appartient à une cavité Ouest si et seulement si les

conditions suivantes sont simultanément vérifiées :

1- Ce point n'appartient pas au tracé.

2- À partir de ce point, en se déplaçant en ligne droite vers l'Ouest, on ne croise pas le tracé.

3- À partir de ce point, en se déplaçant en ligne droite vers le nord, le sud, ou l'Est, on croise le tracé.

Par analogie, on peut facilement en déduire les définitions des cavités Nord, Sud et Est. Pour une cavité Centrale, la définition est la suivante. Un point de l'image appartient à une cavité Centrale si et seulement si les condition suivantes sont simultanément vérifiées :

1- Ce point n'appartient pas au tracé.

2- A partir de ce point, en se déplaçant en ligne droite vers l'Ouest, le nord, le sud, ou l'Est, on croise le tracé.

Les cavités peuvent être détectées en faisant des intersections de dilatations par des demi-droites. Appelons E la demi-droite qui s'étend vers l'Est, W celle qui s'étend vers l'Ouest, N vers le Nord et S vers le Sud.

Si on dilate l'image d'un caractère avec la demi-droite E, on obtient une image sur laquelle les points blancs sont :

- Soit des points du tracé ;

- Soit des points à partir desquels, en se déplaçant vers l'Est, on rencontre le tracé. Il s'ensuit que l'image CW des cavités Ouest est obtenue ainsi :

$$CW = (I \oplus E) \text{and} (\overline{I \oplus W}) \text{and} (I \oplus N) \text{and} (I \oplus S) \text{and} \overline{I} \quad (6.20)$$

Où *and* désigne le ET logique, pixel, entre deux images binaires, $I \oplus E$ désigne la dilatation de l'image I par la demi-droite E , et la barre supérieure désigne la négation, c'est-à-dire l'inversion des 0 et des 1. Cette équation ne fait que traduire la définition des cavités Ouest.

$$CE = (\overline{I \oplus E}) \text{and} (I \oplus W) \text{and} (I \oplus N) \text{and} (I \oplus S) \text{and} \overline{I} \quad (6.21)$$

$$CN = (I \oplus E) \text{and} (I \oplus W) \text{and} (\overline{I \oplus N}) \text{and} (I \oplus S) \text{and} \overline{I} \quad (6.22)$$

$$CS = (I \oplus E) \text{and} (I \oplus W) \text{and} (I \oplus N) \text{and} (\overline{I \oplus S}) \text{and} \overline{I} \quad (6.23)$$

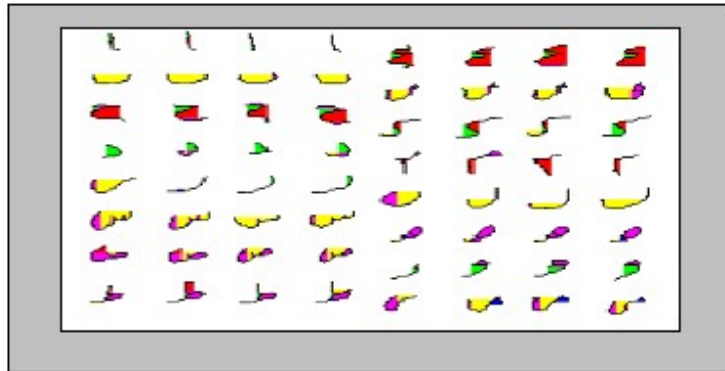


Figure 6.1 : exemple des cinq cavités

6.5 Descripteur de Fourier

6.5.1 Descripteur par tangente :

Dans cette approche on considère le contour comme étant une courbe continue qui peut être décrite par son abscisse curviligne s à partir d'une origine. On paramètre la courbe par l'angle fait par le vecteur tangent en chaque point et celui au point origine : $\Phi(s)$ et on crée la variable réduite t qui prend ses valeurs entre 0 et 2π : $t = 2\pi s/L$, où L est la longueur complète du contour. On construit alors la fonction $\Phi(t)$:

$$\Phi(t) = \Phi[2\pi.s/L] - 2\pi s/L \quad (6.24)$$

Le terme correctif prenant en compte l'enroulement de 2π de la tangente pour un tour de contour [57]. La fonction $\Phi(t)$ est une fonction périodique sur $[0, 2\pi]$ qui admet donc une série de Fourier :

$$\phi(t) = \sum_{k=0}^{\infty} a_k \exp(-ikt) \quad (6.25)$$

On appelle descripteurs de Fourier l'ensemble des modules qui bénéficient des propriétés suivantes :

- ils sont invariants par translation de la forme,
- ils sont invariants par changement d'échelle (puisque t est normalisé),
- ils sont invariants par rotation, puisque l'on a choisi la différence d'angle entre 2 tangentes,

- ils sont invariants par changement d'origine, car passer d'une origine A à un origine A' revient à :
 - retrancher $\Phi(t_{A'})$ à toutes les valeurs $\Phi(t)$,
 - changer t en $t - t_{A'}$

Donc : $\phi_{A'}(t) = \phi_A(t) * \delta(t - t_{A'}) - \phi(t_{A'})$ et : $a_k \rightarrow a_k \exp(-kt_{A'})$ si $k \neq 0$

Pour comparer des formes on compare leurs descripteurs par ordre croissant si de plus on veut simplifier le contour, il suffit de supprimer les ordres k élevés dans le développement. Malheureusement, dans cette représentation, si un contour est fermé, le contour obtenu en filtrant les hautes fréquences ne l'est généralement plus.

6.5.2 Représentation complexe

Dans cette représentation, on décrit la forme par un ensemble $\{M_j\}$ de points de contours, et on représente la forme dans le plan complexe. On attache donc à chaque M_j un nombre complexe $z_j = x_j + iy_j$. On appelle alors descripteurs de Fourier, les coefficients de la TF Z de z :

$$Z_k = \sum_{j=1}^N z_j \exp(-2\pi ijk) \quad (6.26)$$

Les coefficients Z_k , pour $k \in [-N/2 + 1, N/2]$, jouissent d'intéressantes propriétés

- pour $k = 0$, Z_0 est le centre de gravité de la forme. Si l'on l'omet, la description est invariante par translation.
- Si tous les Z_k sont nuls sauf pour $k = 1$, la forme est un cercle de rayon Z_1 (ou un polygone régulier à N côtés), donc Z_1 joue le rôle de facteur d'échelle. La normalisation par Z_1 rend la forme invariante par homothétie.
- Les coefficients Z_k et Z_{1-k} (pour $k \neq 1$ et $k \neq 0$) jouent des rôles symétriques (mais opposés) de la façon suivantes :
- l'ordre k indique le nombre d'action sur le cercle unité (entre 0 et 2π) : 1 action pour $k = 2$ et $k = -1$, 2 actions pour $k = 3$ et $k = -2$, Et 3 actions pour $k = 4$ et $k = -3$, etc. Ces actions sont réparties régulièrement autour du cercle unité,

- Les valeurs de $k > 0$ indiquent des actions de traction sur la courbe, pour la déformer vers l'extérieur du cercle unité, les valeurs de $k < 0$ indiquent des actions de pression sur la courbe, pour creuser la courbe vers son centre,
- La phase du nombre complexe $Z_k : \Phi_k$ exprime le lieu sur le cercle unité.
- Plus les coefficients sont nombreux, plus la forme est complexe. Et plus les coefficients sont élevés plus les détails sont fins la courbe.

6.5.3 Les descripteurs de Fourier pour contours fermés

Les descripteurs de Fourier sont obtenus à partir de la décomposition en série de Fourier d'une grandeur unidimensionnelle, appelée signature, extraite du contour.

6.5.3.1 Echantillonnage :

Plusieurs signatures du contour d'une forme peuvent être extraites. La plus courante est la distance au centre de gravité de la courbe, cependant les positions complexes ou les variations angulaires sont également de bonnes candidates.

Toutes ces techniques font appel à un échantillonnage de la courbe soit angulaire, soit linéique. Cet échantillonnage doit être réalisé à pas constants pour rester dans le cas unidimensionnel. De plus, un échantillonnage angulaire est impossible sur les courbes non convexes.

Pour les contours 4-connexes l'échantillonnage à longueur constante est simple à réaliser puisqu'il suffit de considérer tous les points de la courbe qui sont par définition espacés deux à deux par une longueur constante de 1.

L'erreur introduite par l'imprécision de l'échantillonnage est à peu près du même ordre que celle induite par la prise en compte de tous les points de contour.

6.5.3.2 Signature angulaire :

On utilise la différence angulaire segment à segment d'un contour décrit dans le sens anti-trigonométrique. Les descripteurs suivants peuvent être calculés :

$$a_n = \frac{-1}{n\pi} \sum_{k=1}^m \Delta\Phi_k \sin\left(\frac{2\pi nk}{m}\right). \quad (6.27)$$

$$b_n = \frac{1}{n\pi} \sum_{k=1}^m \Delta\Phi_k \cos\left(\frac{2\pi mk}{m}\right) \quad (6.28)$$

Où m est le nombre de segments décrivant le contour et $\Delta\Phi_k$ correspond à la différence angulaire des segments $[Z(k+1), Z(k)]$ et $[Z(k), Z(k-1)]$. Dans le cas d'un contour 8-connexe, peut également être remplacé par la courbure $k(t) = \frac{\Delta\Phi_k}{\Delta(t)}$. Les coefficients a_n et b_n sont invariants par changement d'échelle et translation. L'invariance par rotation s'obtient, quant à elle, en utilisant les amplitudes des moments complexes.

$$A_n = \sqrt{a_n^2 + b_n^2} \dots\dots\dots (6.29)$$

$$\alpha_n = \tan\left(\frac{a_n}{b_n}\right) \dots\dots\dots (6.30)$$

6.5.3.3 Descripteur de Fourier de distance centroides.

La distance des points de la courbe au centre de gravité (x_c , y_c) est établie en chaque point par l'équation :

$$R(t) = \sqrt{[x(t) - x_c]^2 + [y(t) - y_c]^2} \dots\dots\dots (6.31)$$

De la DFT seul le module est conservé au vu de son invariance en rotation. Les fréquences négatives ne sont pas non plus conservées puisque $R(t)$ est à valeurs réelles, les coefficients fréquentiels sont symétriques par rapport à la composante continue.

$$F_i = \frac{1}{m} \sum_{t=0}^{m-1} R(t) \exp\left(\frac{-j2\pi it}{m}\right) \quad i=0,1,\dots, m-1 \dots\dots\dots (6.32)$$

Le vecteur d'attributs ainsi défini possède des coefficients normalisés par $|F_0|$ pour s'affranchir des Changements d'échelle :

$$f_R = \left\{ \frac{|F_1|}{|F_0|}, \frac{|F_2|}{|F_0|}, \dots, \frac{|F_{\frac{m}{2}}|}{|F_0|} \right\} \dots\dots\dots (6.33)$$

Au final les f_R sont invariants par translation (distance au centre de gravité), rotation (seul le

module est conservé) et changement d'échelle (normalisation par la composante continue).

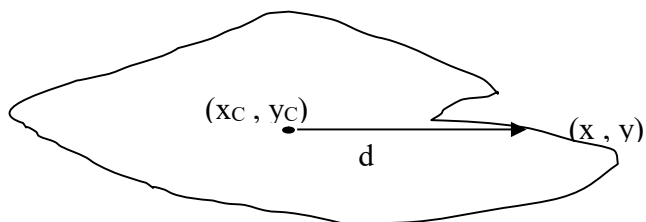


Figure 6.2 : Application de la distance

6.6 Conclusion

Le processus d'identification de la meilleure méthode d'extraction de caractéristiques n'est pas évident. Il est nécessaire d'effectuer pour chaque problème de reconnaissance une évaluation expérimentale de quelques méthodes d'extraction de paramètres les plus prometteuses. L'utilisation d'une seule méthode d'extraction de caractéristique n'est pas suffisante pour obtenir de bonne performance du système de reconnaissance. La solution évidente est de combiner plusieurs méthodes d'extraction afin de donner une meilleure description de la forme.

Chapitre 7

classification

7.1 Introduction

La classification est l'élaboration d'une règle de décision qui transforme les attributs caractérisant les formes en appartenance à une classe (passage de l'espace de codage vers l'espace de décision).

7.2 Notion de Classificateur

Soit la représentation d'un objet quelconque au moyen d'un vecteur de caractéristiques $X = [x_1 x_2 \dots x_d]^T$.

Tous les vecteurs qui représentent l'ensemble des objets peuvent être positionnés dans l'espace Euclidien R_d , où ils correspondent chacun à un point. Ceux-ci peuvent alors être regroupés en amas, chacun de ces amas étant associé à une classe particulière.

Le rôle d'un classificateur est de déterminer, parmi un ensemble fini de classes, à laquelle appartient un objet donné [17].

Un classificateur doit être capable de modéliser au mieux les frontières qui séparent les classes les unes des autres. Cette modélisation fait appel à la notion de fonction discriminante, qui permet d'exprimer le critère de classification de la manière suivante:

« Assigner la classe ω_i à l'objet représenté par le vecteur X si, et seulement si, la valeur de la fonction discriminante de la classe ω_i est supérieure à celle de la fonction discriminante de n'importe quelle autre classe ω_j ».

Ou encore, sous forme mathématique:

$$X \in \omega_i \Leftrightarrow \Phi_i(X) \geq \Phi_j(X) \quad \forall j = 1, 2, \dots, C, j \neq i.$$

Où $\Phi_i(X)$ est appelé fonction discriminante de la classe ω_i , et C est le nombre total de classes.

Avant qu'un modèle de décision ne soit intégré dans un système de reconnaissance de l'écriture, il faut avoir procédé auparavant à deux étapes : l'étape d'apprentissage et l'étape de test.

7.3 Apprentissage

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

Il y a plusieurs cas possibles:

1) - Nombre de classes connues

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

2) - Nombre de classes connu ou non

- On ne connaît pas la classe des échantillons (ce qu'on sait c'est seulement que l'ensemble des échantillons représente ce que doit connaître). On parle alors d'apprentissage non supervisé (ou sans « professeur ») converge vers une solution vérifiant $a^t \times y_i$ si les deux classes sont linéairement séparables.

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

7.4 Test

L'étape de test permet d'évaluer la performance du classifieur pour un apprentissage donné. C'est une étape importante car elle peut mettre en cause le choix des primitives ou le choix de la méthode d'apprentissage. En effet, il est difficile de trouver a priori les primitives pertinentes et la méthode d'apprentissage la plus adaptée au problème posé d'où l'utilité de procéder par itérations successives. Ces itérations consistent à extraire des primitives jugées utiles au problème de reconnaissance à résoudre et de tester la performances du système. Si les performances souhaitées ne sont pas atteintes alors il suffit de trouver à nouveau une nouvelle famille de primitives ou de combiner les primitives extraites avec de nouvelles primitives.

7.5 Les méthodes de classification

Pour classifier des formes qui sont représentés par les vecteurs de caractéristiques, on peut utiliser plusieurs méthodes de classification :

7.5.1 Les méthodes paramétriques

Les méthodes paramétriques supposent la connaissance des lois de probabilité des observations et des classes. En effet, tout vecteur x de IR^d suit, dans une classe donnée w_i , une loi de probabilité $f(x/w_i)$. par ailleurs $w_j (j = 1, \dots, M)$ sont caractérisées par leurs probabilités a priori $p_r(w_j)$. ces probabilités vérifient la condition d'orthogonalité suivante :

$$\sum_{j=1}^M p_r(w_j) = 1 \quad (8.1)$$

Dans ce contexte, la règle de décision la plus couramment utilisée est la règle de Bayes [Dubuisson, 90].

a. Règle de Bayes et décision Bayésienne

La règle de Bayes permet de calculer ce qu'on appelle les probabilités à posteriori des classes, c'est à dire $p(w_1/x)$ et $p(w_2/x)$ à partir des probabilités à priori des classes $P(w_1)$ et $P(w_2)$ et des fonctions de densité des probabilités $p(x/w_1)$ et $p(x/w_2)$

$$p(w_i/x) = \frac{p(x/w_i) \times p(w_i)}{p(x)} \quad \forall i = 1, 2 \quad (8.2)$$

Avec

$$p(x) = \sum_{j=1}^2 p(x/w_j) \times p(w_j) \quad (8.3)$$

L'interprétation de $p(w_1/x)$ est la suivante: ayant obtenu une forme (décrite par une valeur x de luminance) $p(w_1/x)$ donne la probabilité d'avoir alors la classe w_1 .

Dans notre exemple, avec les lois de densité de probabilité obtenues et les probabilités à priori données précédemment, on obtient approximativement les densités de probabilité à posteriori suivantes

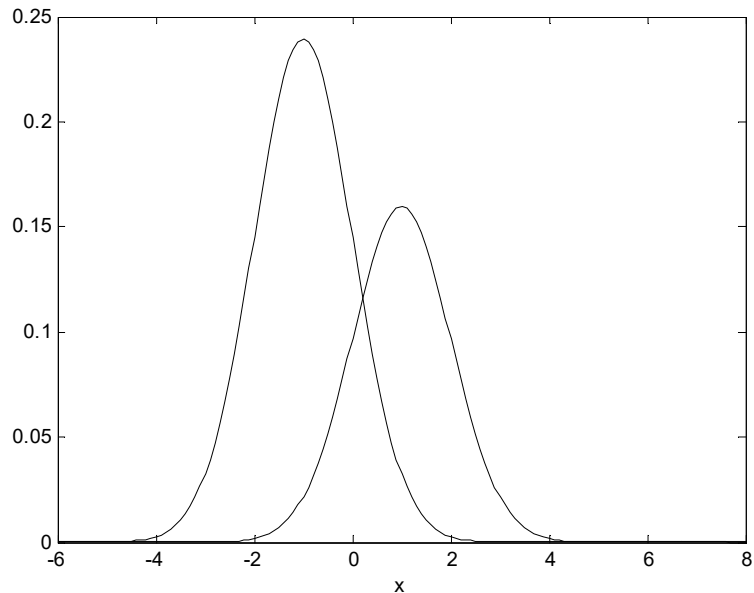


Figure 7.1 : Espace d'observation R avec deux classes « gaussiennes »

La règle de décision Bayésienne est alors pratiquement immédiate:

Ayant remarqué une forme inconnue x , on dira que la forme inconnue x est de la classe w_1 si

$P(w_1/x) > P(w_2/x)$ sinon elle est de la classe w_2 Soit sous une forme simplifiée:

REGLE de décision Bayésienne

X observé non décidé

W_1 si $p(w_1/x) \geq p(w_2/x)$

W_2 sinon

Ou encore

W_1 si $p(x/w_1) \times p(w_1) \geq p(x/w_2) \times p(w_2)$

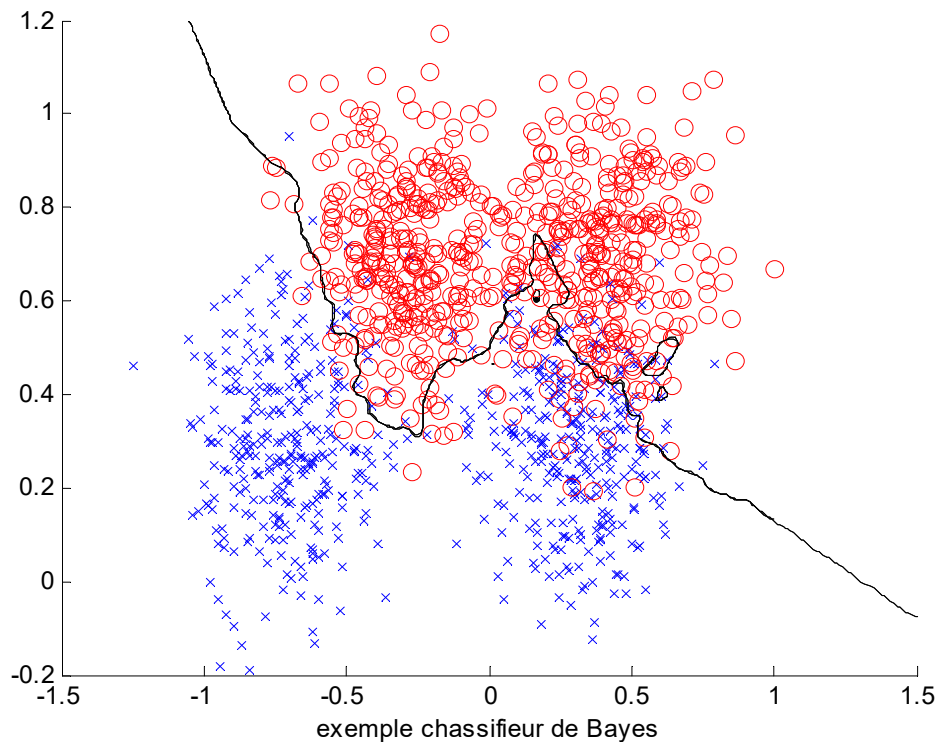


Figure 7.2: exemple d'un classifieur de Bayes

b. Probabilité d'erreur et fonctions discriminantes bayésiennes

Voyons d'abord le cas de deux classes et du classifieur bayésien:

On a donc deux régions R_1 et R_2

Le classifieur se trompe dans les cas suivant:

- x tombe dans R_1 mais en fait x appartient à w_2

- x tombe dans R_2 mais en fait x appartient à w_1

On a donc

$$\begin{aligned}
 P(\text{erreur}/x) &= P(x \in R_2, w_1) + P(x \in R_1, w_2) \\
 &= P(x \in R_2 / w_1) \times P(w_1) + P(x \in R_1 / w_2) \times P(w_2) \quad (8.4)
 \end{aligned}$$

Soit en calculant la probabilité d'erreur sur toutes les valeurs de x

$$P(\text{erreur}) = \int_{R_1} P(x / w_1) \times P(w_1) \times dx + \int_{R_2} P(x / w_2) \times P(w_2) \times dx \quad (8.5)$$

On peut aussi voir le problème par une maximisation de probabilité de classification correcte.

Ainsi si on a s classes possibles on a

$$p(\text{correct} / x) = \sum_{i=1}^s p(x \in R_i, w_i)$$

$$= \sum_{i=1}^s p(x \in R_i / w_i) \times p(w_i) \quad (8.6)$$

Ce qui donne

$$p(\text{correct}) = \sum_{i=1}^s \int_{R_i} p(x / w_i) \times p(w_i) \times dx \quad (8.7)$$

Le classifieur de Bayes maximise cette probabilité

Dans cette règle on voit de façon particulièrement claire qu'il est nécessaire et suffisant de connaître pour toutes les classes cherchées la probabilité à priori et la loi de densité de probabilité de cette classe.

7.5.2 Méthodes non paramétriques

Il a été vu précédemment que les méthodes paramétriques étaient basées sur l'existence des lois de probabilités régissant les observations et les classes. Toutefois si cette connaissance est incomplète, il est préférable d'orienter la procédure de décision vers l'estimation de la loi de probabilité ou de celle des probabilités a posteriori.

Parmi les estimateurs usuels, on distingue l'estimateur de Parzen et l'estimateur des k plus proches voisins (k _ppv)

a. Principe des méthodes non paramétriques :

Considérons l'ensemble d'apprentissage constitué des N vecteurs indépendants x_1, x_2, \dots, x_N . Ces vecteurs obéissent à la même loi $f(x / w_i)$.

Pour estimer cette densité de probabilité au point x à partir de l'ensemble d'apprentissage, comparant le nombre k d'échantillons appartenant à D par rapport à l'ensemble des N échantillons, on démontre que l'estimation de $f(x / w_i)$ prend la forme :

$$\hat{f}(x / w_i) = \frac{k}{N.V} \quad (8.8)$$

Toutefois les conditions suivantes doivent être vérifiées pour que la convergence de $\hat{f}(x/w_i)$ soit assurée :

$$\lim_{N \rightarrow \infty} V = 0$$

$$\lim_{N \rightarrow \infty} k = \infty \quad (8.9)$$

$$\lim_{N \rightarrow \infty} \frac{k}{N} = 0$$

Le choix du couple (k,v) permettra de réaliser ces trois conditions . Deux approches sont possibles :

Les différentes zones sont définies, chacune de même volume V fonction de N (par exemple $V = \frac{1}{\sqrt{N}}$ de manière à remplir la première condition). Il suffit alors de s'assurer que la valeur de k permet de vérifier les deux autres conditions. c'est la méthode utilisée par l'estimateur de Parzen.

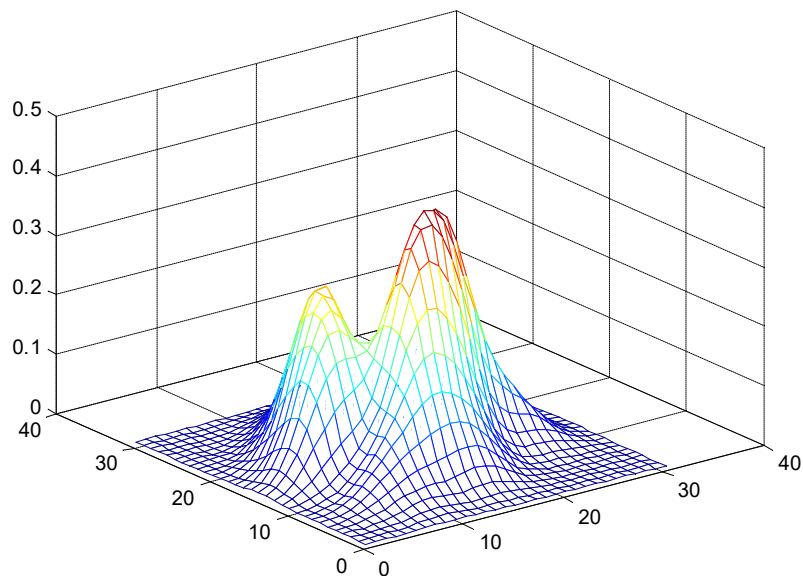


Figure 7.3 densité de probabilité à posteriori des classes

Le nombre d'échantillons k est choisi en fonction de N (par exemple $K = \sqrt{N}$).

Le volume V est ajusté, autour du point x , pour que le domaine D contienne exactement ces k échantillons. C'est la méthode des k _plus proche voisins (k _ppv)

b. Règle de décision du plus proche voisin

Cette méthode suppose disposer d'une distance dans l'espace des formes, et on est donc capable de calculer une distance entre une forme x quelconque à un échantillon.

Le principe de décision consiste tout simplement à calculer la distance de la forme inconnue x à tous les échantillons fournis. La forme est alors classée dans la classe de l'échantillon le plus proche (par exemple ci dessous x serait classé dans la classe w_2):

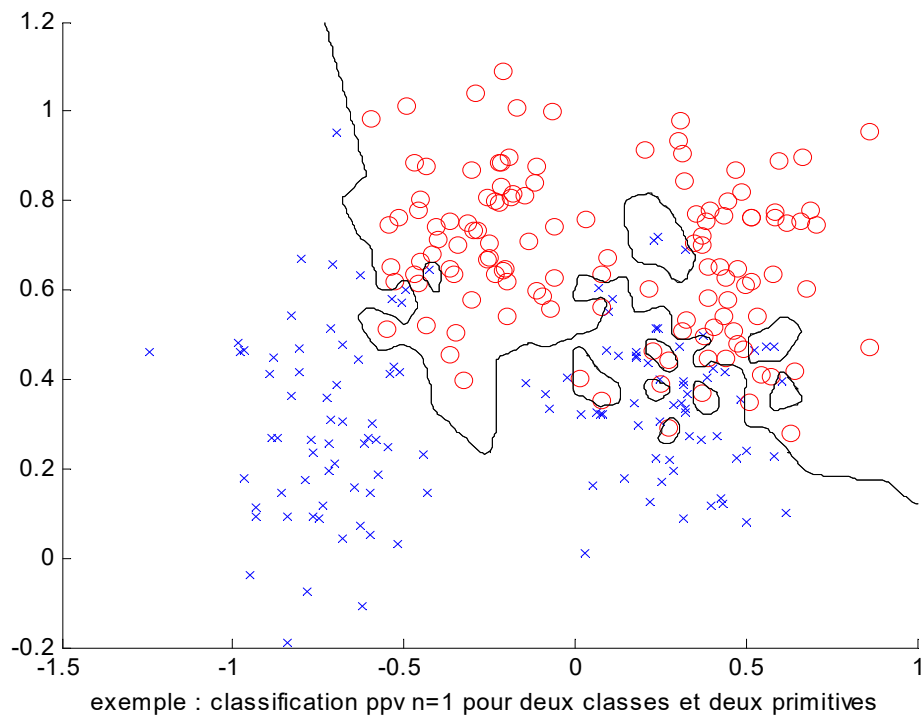


Figure 7.4 $k_{ppv} n=1$ pour deux classes

b. Règle de décision des k plus proches voisins

Il s'agit d'une extension de la méthode précédente (plus proche voisin) : pour une forme inconnue x à classer, on va examiner la distance de x à tous les échantillons (qui définissent

toutes les classes), puis on sélectionne les k plus proches échantillons et on affecte x à la classe majoritaire parmi ces k échantillons.

Dans l'exemple suivant et pour 5 voisins on classerait x dans la classe w_1 :

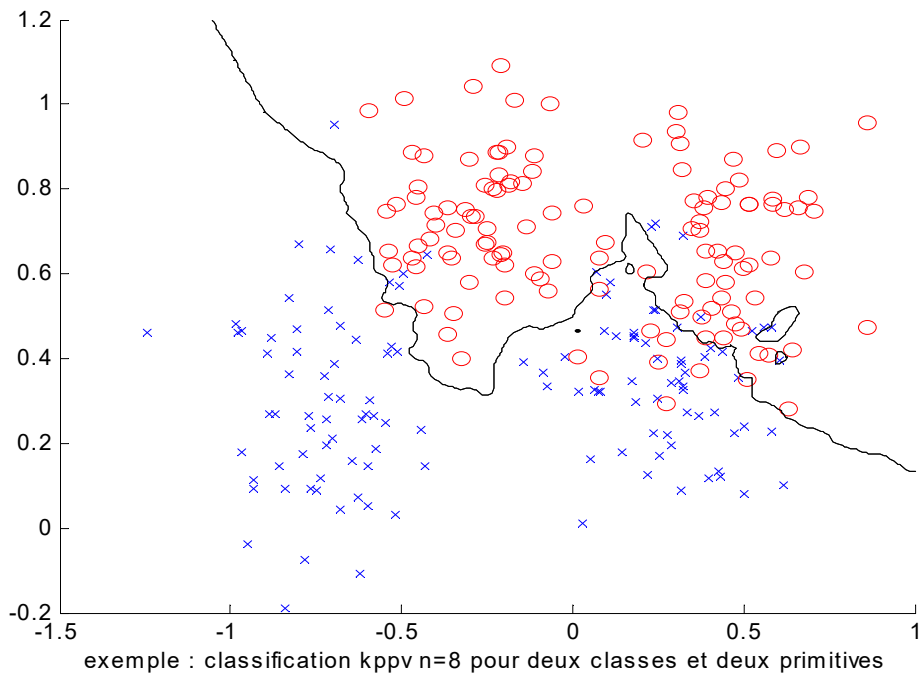


Figure 7.5 kppv $n=8$ pour deux classes

Choix possible pour les distances:

- distance Euclidienne
- distance "city block" (somme des valeurs absolues)
- distance de Tchebycheff (le sup.)
- distance de Mahalanobis

c. Les réseaux connexionnistes

représentent en fait une famille de méthodes de discriminations linéaires dont les fonctions de discrimination ne sont pas connues au départ mais élaborées par apprentissage. L'apprentissage correct des réseaux reste un problème important (Stéphane Boucheron va plus

loin dans sa Théorie de l'Apprentissage) , en étudiant le Modèle de Valant sous les aspects statistiques et sous les aspects calculatoires et la sémantique des fonctions de discrimination ne peut pas être précisée.

7.5.3 Les méthodes structurelles

Reposent sur des connaissances plus riches que les précédentes méthodes, elles permettent de prendre en compte l'information sur la structure et le contexte de la forme. La description peut alors rendre compte des composants, de leurs propriétés, de leurs interrelations (qui peuvent être spatiales et temporelles...).

Dans le domaine de la reconnaissance de forme, deux structures sont citées les plus souvent:

- Les structures de graphe pour décrire et/ou mettre en correspondance des objets formes. Les méthodes classiques appliquées aux graphes sont utilisées pour diminuer la complexité de recherche d'isomorphisme (le plus souvent incomplets) . Un exemple connu est l'étiquetage de scènes par Walt dans ce qu'il appelle le monde des blocs

Les structures syntaxiques qui sont inspirées de la théorie des langages formels. Dans ce dernier cas, l'apprentissage de la grammaire est plus difficile à réaliser automatiquement mais possible. Pour éviter une trop grosse sensibilité aux erreurs dues à l'incomplétude de la grammaire, une recherche de similarité entre " phrase" décrivant un objet et la grammaire peut être menée en évaluant ce qu'il faut modifier dans la phrase pour qu'elle devienne "correcte" au sens de la grammaire mise en comparaison.

7.6 Conclusion

La décision ou classement est l'étape de reconnaissance proprement dite. Son rôle est de classer la forme ciblée à partir de l'apprentissage réalisé. Les critères utilisés pour la décision sont habituellement les mêmes que ceux utilisés pour l'apprentissage. Parmi les techniques utilisées, certaines sont fondées sur la notion de proximité et nécessitent de calculer une distance ou une probabilité de ressemblance avec les modèles définis. Dans les deux cas, la réponse peut être accompagnée d'un taux ou score de confiance.

Le classifieur est l'élément adaptatif du système de reconnaissance pour fonctionner, il doit au préalable, avoir été réglé par apprentissage. L'apprentissage consiste à donner au classifieur des exemples de caractères. On utilise les méthodes (PPV, kPPV,...). Dans les systèmes modernes de reconnaissance, le classifieur le plus récent utilise une méthode de SVM (support vector machine) machine à vecteurs support, C'est un dispositif de calcul simple à élaborer et donne des résultats satisfaisants qui sera décrit dans le chapitre qui suit.

Chapitre 8

Machines à Vecteurs Support (SVM)

8.1 Introduction

SVM (*support vector machines*) ou machines à vecteurs support a été introduit par Vapnik en 1995 [90]. Le succès de cette méthode est justifié par les solides bases théoriques qui la soutiennent.

Il existe en effet un lien direct entre la théorie de l'apprentissage statistique et l'algorithme d'apprentissage de SVM. La plupart des techniques du ML (Machine Learning) possèdent un (trop) grand nombre de paramètres d'apprentissage à fixer par l'utilisateur (structure d'un réseau de neurones, coefficient de mise à jour du gradient,...).

De plus, avec ces méthodes, le nombre de paramètres à calculer par l'algorithme d'apprentissage est en relation linéaire, voire exponentielle, avec la dimension de l'espace d'entrée. La formulation élégante de SVM laisse très peu de place aux paramètres utilisateurs et le nombre de paramètres est linéaire en la taille de l'ensemble d'apprentissage. SVM est donc une méthode de classification particulièrement bien adaptée pour traiter des données de très haute dimension telles que les textes et les images. Dans ce chapitre nous présentons les aspects théoriques de la méthode SVM.

Du point de vue conceptuel, la notion de risque structurel introduite par Vladimir Vapnik dans les années 1990 donne une borne de l'erreur de test en fonction de l'erreur d'apprentissage et de la complexité du modèle. En pratique, les machines à vecteurs de support (*Support Vector Machines* ou *SVM*), présentées en 1992, offrent un moyen opérationnel pour minimiser le risque structurel, ce qui explique l'engouement suscité dans la communauté scientifique par cette méthode

8.2 Principe

Le principe théorique des SVM comporte deux points fondamentaux :

- la transformation non linéaire Φ des exemples de l'espace d'entrée vers un espace dit de redescription de grande dimension muni d'un produit scalaire (espace de Hilbert),

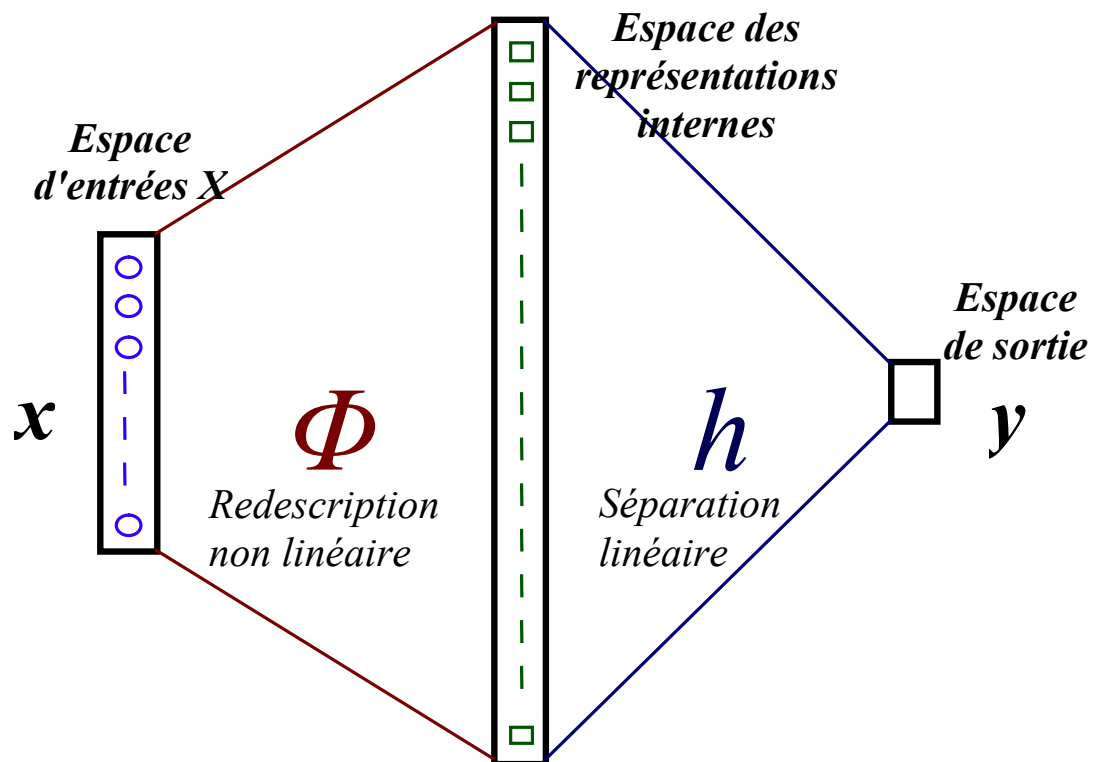


Figure 8.1 : Transformation d'espace

La transformation non linéaire des données peut permettre une séparation linéaire des exemples dans un nouvel espace.

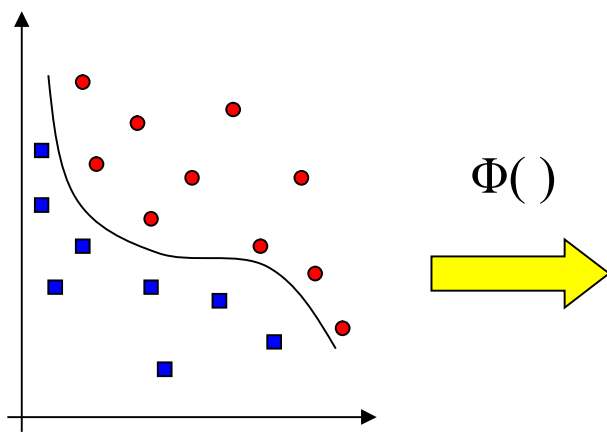


Figure 8.2 : Espace d'entrée

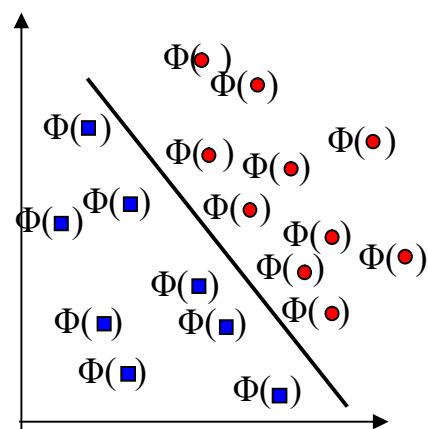
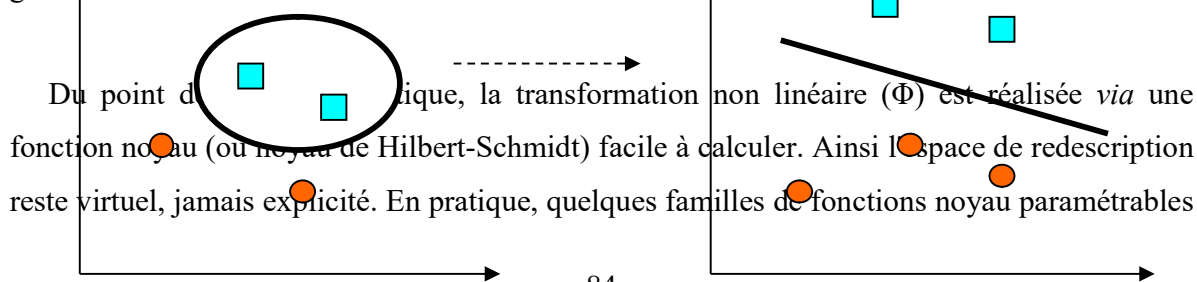


Figure 8.3 : Espace de redescription

La détermination d'un hyperplan permettant une séparation linéaire optimale dans cet espace de grande dimension.

L'intérêt est que dans l'espace de redescription la reconnaissance de formes peut s'avérer tâche aisée : en effet, intuitivement, plus la dimension de l'espace de redescription est grande, plus la probabilité de pouvoir trouver un hyperplan séparateur entre les exemples est élevée.

L'idée des SVM : transformer un problème de séparation non linéaire dans l'espace de représentation en un problème de séparation linéaire dans un espace de re-description de grande dimension.



Du point de vue mathématique, la transformation non linéaire (Φ) est réalisée via une fonction noyau (ou noyau de Hilbert-Schmidt) facile à calculer. Ainsi l'espace de redescription reste virtuel, jamais explicité. En pratique, quelques familles de fonctions noyau paramétrables

Figure 8.4 : Transformation d'un espace non linéairement séparable en un espace linéairement séparable

sont connues et il revient à l'utilisateur de SVM d'effectuer des test pour déterminer celle qui conviennent le mieux pour son application (il s'agit de traduire le maximum de connaissances préalables dont on dispose sur le problème étudié et sur les données).

- **Exemple: le cas du XOR:**

Le cas du XOR n'est pas linéairement séparable, si on place les points dans un plan à deux dimensions, on obtient la Figure8.5 suivantes:

Index i	X	Y
1	0	0
2	0	1
3	1	0
4	1	1

Tableau 8.1 : Coordonnées des points en 2 dimensions

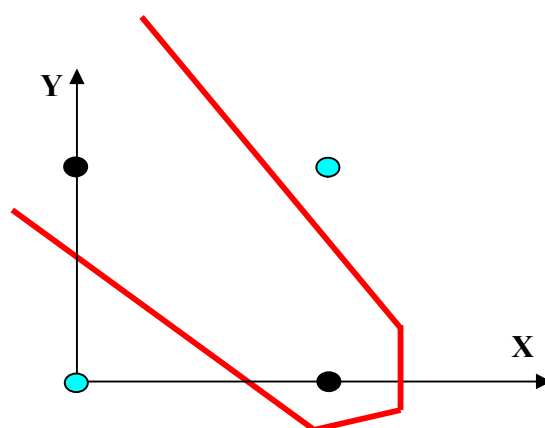


Figure 8.5: espace description de deux dimensions

Si on prend une fonction polynomiale $(x,y) \rightarrow (x,y,x.y)$ qui fait passer d'un espace de dimension 2 à un espace à dimension 3, on obtient un problème à trois dimensions

linéairement séparable Figure-8.6:

Index i	x	Y	X.Y
1	0	0	0
2	0	1	0
3	1	0	0
4	1	1	1

Tableau 8.2: Coordonnées des points 3 dimensions :

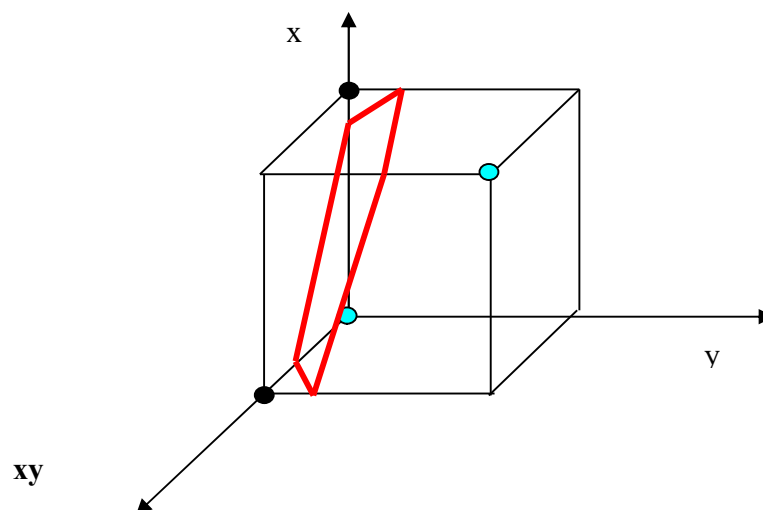


Figure 8.6: Espace de description à 3 dimensions

8.3 classifieur

8.3.1 Définition:

Un classificateur est dit linéaire lorsqu'il est possible d'exprimer sa fonction de décision par une fonction linéaire en x . Dans la suite, nous supposons que les exemples nous sont fournis dans le format vectoriel. Notre espace d'entrée X correspond donc à \mathbb{R}^n ou n est le nombre de

composantes des vecteurs contenant les données. On peut, en toute généralité, exprimer une telle fonction comme ceci :

$$f(x)=(w,x)+b \quad (8.1)$$

$$f(x) = \sum_{i=1}^n w_i x_i + b \quad (8.2)$$

ou: $w \in \mathfrak{R}^n$ et $b \in \mathfrak{R}$ sont des paramètres, et $x \in \mathfrak{R}^n$ est une variable.

Pour décider à quelle catégorie un exemple \bar{x} appartient, il suffit de prendre le signe de la fonction de décision :

$$y = \text{sign} (f(\bar{x})) \quad (8.3)$$

Géométriquement, cela revient à considérer un hyperplan qui est le lieu des points x satisfaisant

$$(w,x)+b=0. \quad (8.4)$$

En orientant l'hyperplan, la règle de décision correspond à observer de quel côté de l'hyperplan se trouve l'exemple \bar{x} .

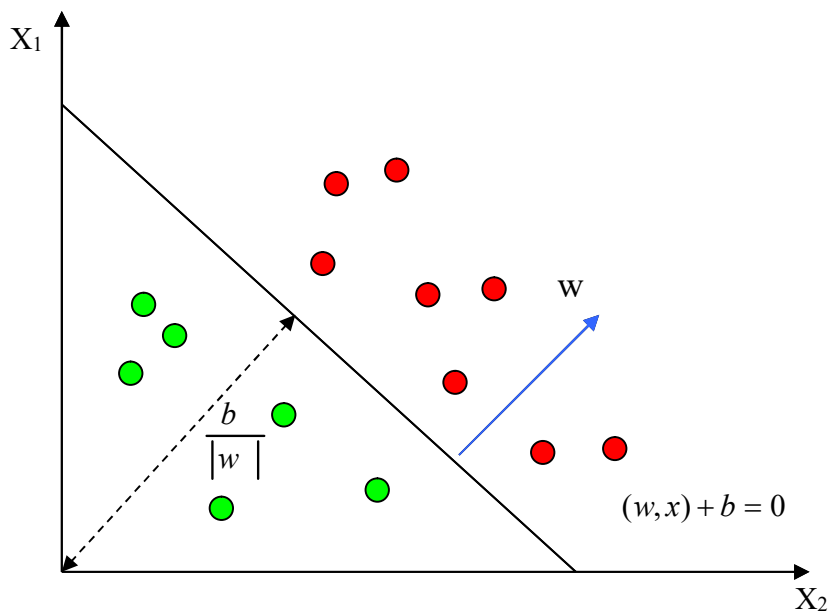


Figure 8.7 :représentation dans \mathfrak{R}^2 de l'hyperplan correspond à la fonction de la décision d'un classifieur linéaire

On voit que le vecteur w définit la pente de l'hyperplan : w est perpendiculaire à l'hyperplan. Le terme b quant à lui permet de translater l'hyperplan parallèlement à lui-même

8.3.2 Marge de l'hyperplan:

Avant de définir la marge, il est utile d'introduire la notion de training set linéairement séparable.

Définition: si un training set $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ est linéairement séparable:

$$\exists w \in \mathfrak{R}^n, b \in \mathfrak{R} : y_i((w, x_i) + b) \geq 0 \quad \forall i=1 \dots n \quad (8.5)$$

La définition consiste à dire qu'il doit exister un hyperplan laissant d'un côté toutes les données positives et de l'autre, toutes les données négatives. Dans le cas de données

linéairement séparables, il existe plusieurs méthodes pour trouver un tel hyperplan. La première d'entre elles et la plus connue est l'algorithme du perceptron de Rosenblatt [75].

La notion de marge peut être relative à un exemple particulier ou à l'ensemble du training set. De plus, on considère deux types de marges : fonctionnelle et géométrique.

a. marge fonctionnelle d'un exemple:

La marge fonctionnelle d'un exemple x_i , par rapport à l'hyperplan caractérisé par w et b est la quantité :

$$y_i((w, x_i) + b) \quad (8.6)$$

La marge géométrique quant à elle représente la distance euclidienne prise perpendiculairement entre l'hyperplan et l'exemple x_i . En prenant un point quelconque x_p se trouvant sur l'hyperplan, la marge géométrique peut s'exprimer par $w/\|w\| \cdot (x_i - x_p)$

La figure 8.8 illustre la situation. En utilisant la marge fonctionnelle et le fait que x_p est sur l'hyperplan, on peut calculer la valeur de la marge géométrique:

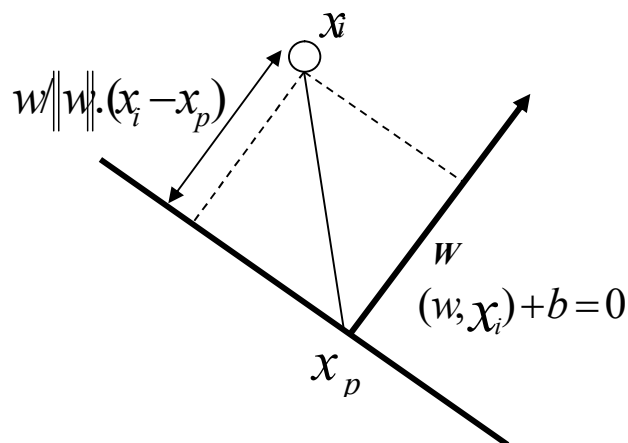


Figure 8.8 : marge d'un exemple

$$(w, x_i) + b = f(x_i) \quad (8.7)$$

$$(w, x_i) + b = 0 \quad (8.8)$$

$$w \cdot (x_i - x_p) = f(x_i) \quad (8.9)$$

$$\Rightarrow w / \|w\| \cdot (x_i - x_p) = f(x_i) / \|w\|. \quad (8.10)$$

b -marge géométrique d'un exemple:

La marge géométrique d'un exemple x_i , par rapport à l'hyperplan caractérisé par w et b est la quantité:

$$\psi_{w,b}(x_i, y_i) = y_i (w / \|w\| \cdot x_i + b / \|w\|) \quad (8.11)$$

Remarquons que la marge d'un exemple mal classé correspond à la distance négative qui le sépare de l'hyperplan.

c- marge géométrique de l'ensemble d'apprentissage

La marge du training set par rapport à l'hyperplan caractérisé par w et b est définie comme suit:

$$\psi_{w,b} = \min \psi_{w,b}(x_i, y_i) \quad (8.12)$$

Par la suite nous utiliserons uniquement la marge géométrique, aussi emploierons nous le terme “**marge**” pour désigner cette dernière. Les classificateurs ayant pour critère de maximiser la marge du training set sont appelés classificateurs à marge maximale. SVM, en particulier, en est un.

8.4 Hyperplans canoniques:

Dans le cadre de classificateurs à marge maximale, l'hyperplan séparateur correspond à la médiatrice du plus petit segment de droite reliant les enveloppes

$$(w, x) + b = 1 \quad (8.13)$$

et : $(w,x)+b=-1$ (8.14)

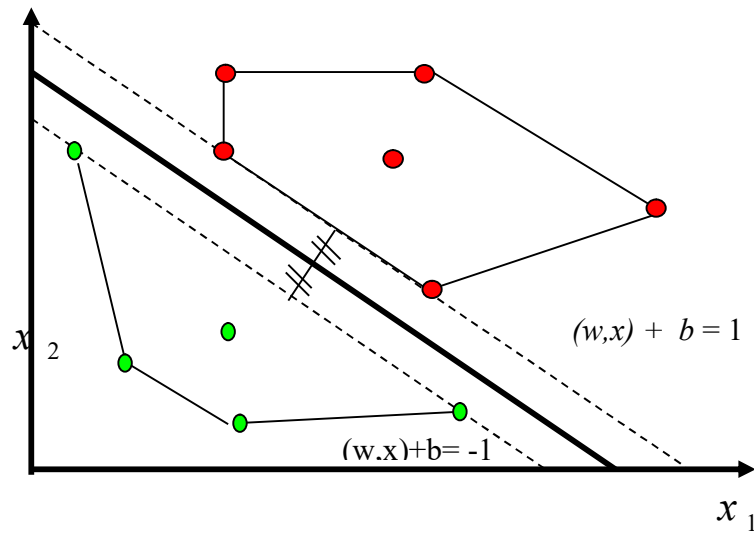


Figure 8.9 : Hyperplans canoniques

Ces deux hyperplans sont appelés hyperplans canoniques. Notons que la marque des hyperplans canoniques est: $1 / \|w\|$.

Le vecteur w possède à présente une signification géométrique très claire.

8.5 Relation avec l'apprentissage statistique:

Intuitivement, le fait d'avoir une marge plus large procure plus de "sécurité" lorsque l'on classe un exemple inconnu [73]. La partie gauche de la figure 8.10 nous montre qu'avec l'hyperplan optimal, un nouvel exemple reste bien classé alors qu'il tombe dans la marge. On constate sur la partie droite qu'avec une plus petite marge, l'exemple se voit mal classé.

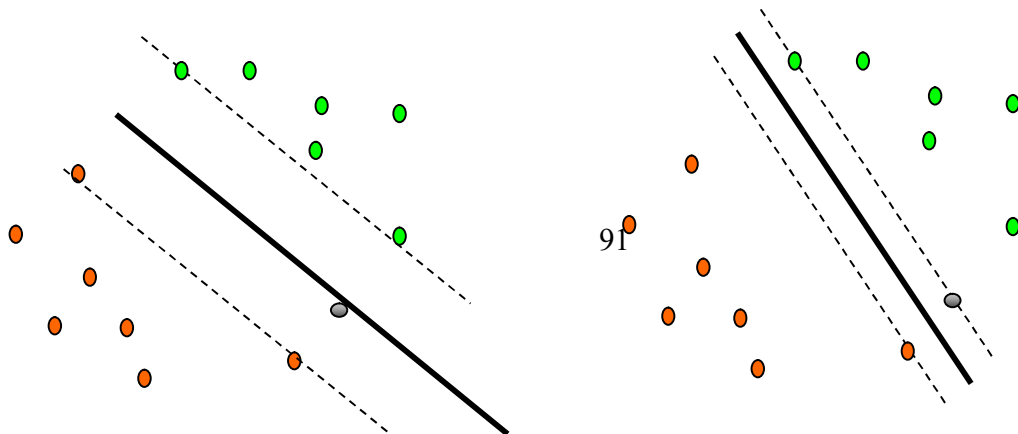


Figure 8.10 : intérêt de la marge maximale

Dans la Figure 8.1 : Le même jeu de données classé par un hyperplan à marge maximale (à gauche) et un hyperplan quelconque ne commettant pas d'erreur sur l'ensemble d'apprentissage (à droite). La boule grisée est un exemple de test.

8.6 Classificateur à marge maximale

Maintenant que nous avons défini les notions de marges et d'hyperplans canoniques, nous pouvons formuler un problème d'optimisation mathématique tel que sa solution nous fournisse l'hyperplan optimal (maximisant la marge) :

$$\text{Minimiser} \quad W(w,b) = 1/2 \|w\|^2 \quad (8.15)$$

$$\text{tel que:} \quad y_i ((w, x_i) + b) \geq 1 \quad (8.16)$$

Il s'agit d'un problème quadratique dont la fonction objective est à minimiser. Cette fonction objective est le carré de l'inverse de la double marge. L'unique contrainte stipule que les exemples doivent être bien classés et qu'ils ne dépassent pas les hyperplans canoniques. Il existe d'autres formulations du problème ; celle que nous avons présentée est la plus répandue.

Dans cette formulation, les variables à fixer sont les composantes w_i et b . D'un point de vue ML (Machine Learning) ou machine d'apprentissage, ces variables correspondent aux α_i de la machine d'apprentissage.

Le vecteur w possède un nombre de composantes égal à la dimension de l'espace d'entrée. En gardant cette formulation telle quelle, nous souffrons des mêmes problèmes que les méthodes classiques du ML.

Pour éviter cela, il est nécessaire d'introduire une formulation dite duale du problème. On dira qu'un problème dual est un problème fournissant la même solution que le primal mais dont la formulation est différentielle. On appellera variables primales, les variables du problème primal, et variables duales, les variables du problème dual qui n'interviennent pas dans le primal.

Pour dualiser (8.15) (8.16), nous devons former ce que l'on appelle le Lagrangien. Il s'agit de faire rentrer les contraintes dans la fonction objective et de pondérer chacune d'entre elles par une variable duale :

$$L(w, b, \alpha) = 1/2 \|w\|^2 - \sum_{i=1}^n \alpha_i [y_i ((w, x_i) + b) - 1] \quad (8.17)$$

Les variables duales α_i intervenant dans le Lagrangien sont appelées multiplicateurs de Lagrange. Selon le type de contraintes qu'ils représentent, les multiplicateurs doivent respecter les règles de signes suivants :

- Contrainte du type $C_i(x) \geq 0$: $\alpha_i \geq 0$
- Contrainte du type $C_i(x) = 0$: α_i sans restriction de signe.
- Contrainte du type $C_i(x) \leq 0$: $\alpha_i \leq 0$

En pratique, pour le dernier cas on considère des multiplicateurs positifs mais ces derniers interviennent avec un signe “-” dans le Lagrangien. On peut donner une signification physique à ces multiplicateurs : la variable α_i représente la “force” avec laquelle la solution appuie sur la contrainte i . Ainsi, un hyperplan qui violerait la contrainte pour x_i (il classe cet exemple du mauvais côté) rendrait α_i très grand ce qui ferait fortement augmenter la fonction objective (L). Cette solution ne pourrait donc pas être retenue comme solution optimale.

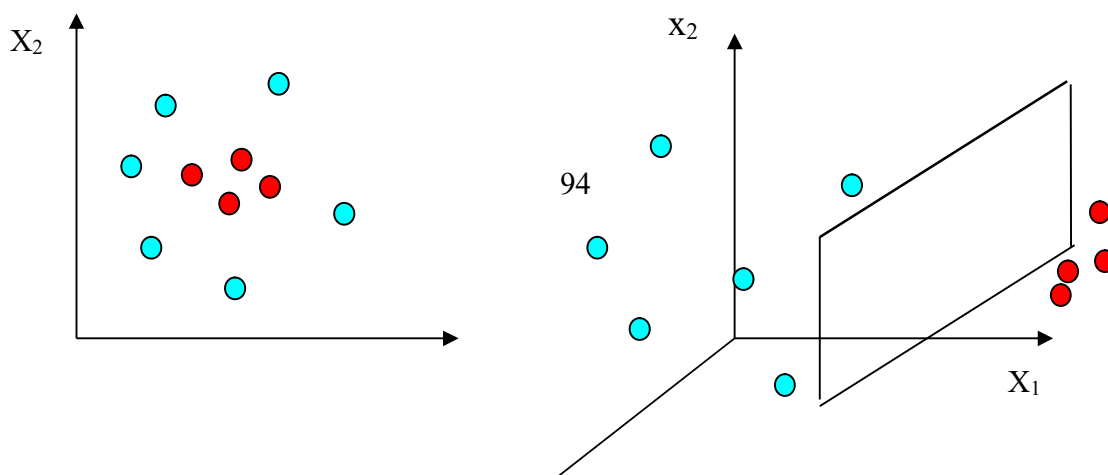
Notons que L doit être minimisé par rapport aux variables primales et maximisé par rapport aux variables duales.

8.7-Les fonctions Kernels:

8.7.1 Généralités:

Le classificateur à marge maximale que nous venons de présenter, permet d'obtenir de très bons résultats lorsque les données sont linéairement séparables. L'intérêt principal d'un classificateur de ce type réside dans le fait que l'on en contrôle facilement la capacité et donc le pouvoir de généralisation. Naturellement, un grand nombre de jeux de données sont non-linéairement séparables. Pour classer ce genre de données on pourrait utiliser une fonction de décision non-linéaire. Géométriquement, cela reviendrait à avoir une (hyper)courbe qui marquerait la frontière entre les exemples positifs et négatifs. Le danger de ce genre de méthode vient du fait que le nombre de paramètres à calculer est très élevé et provoque souvent de l'overfitting.

L'idée retenue dans SVM va dans un autre sens : on va tenter de trouver un mapping de l'espace d'entrée vers un autre espace, l'espace de représentation, dans lequel les données sont linéairement séparables. La figure 8.11 donne une représentation imagée de ce genre de mapping. Nous pouvons alors appliquer une méthode à marge maximale dans l'espace de représentation et garder le contrôle (qui est indépendant de la dimension de l'espace dans lequel nous travaillons) que nous avons sur la capacité. La dimension de l'espace de représentation est généralement très élevée. Cela ne pose pas de problème pour notre classificateur à marge maximale vue que sa formulation duale fixe le nombre de variables à déterminer en fonction de la taille du training set . Nous noterons l'espace de représentation F , et le mapping vers cet espace :



$$\phi: x \Rightarrow F$$

Figure 8.11 : Un mapping ϕ rendant les exemples linéairement séparables

Comme on peut le remarquer à la figure 8.11, l'espace de représentation est de dimension supérieure à celle de l'espace de départ. Les nouveaux axes contiennent une surgénération d'information par rapport aux précédents, ce qui permet idéalement d'effectuer une discrimination linéaire là ou auparavant ce n'était pas possible.

8.7.2 Exemple de kernels:

Le kernel polynomial:

La forme générique de ce kernel est:

$$k(x,z)=[a^*(x,z)+b]^d \quad (8.18)$$

Prenons une instance simple de cette fonction

$$: k(x,z)=(x,z)^2 \quad (8.19)$$

Tentons de trouver un candidat ϕ tel que :

$$(x,z)^2 = [\phi(x),\phi(z)] \quad (8.20)$$

En supposant que l'espace de représentation est de dimension 2 on peut trouver les mapping suivants:

$$\phi: \mathfrak{R}^2 \Rightarrow \mathfrak{R}^3 \quad (8.21)$$

$$\phi_2 : \mathfrak{R}^2 \Rightarrow \mathfrak{R}^4 \quad (8.22)$$

$$\phi(x) = \begin{pmatrix} x_1^2 \\ \sqrt{2}x_1x_2 \\ x_2 \end{pmatrix} \quad (8.23)$$

$$\phi_2(x) = \begin{pmatrix} x_1^2 \\ x_1x_2 \\ x_1x_2 \\ x_2 \end{pmatrix} \quad (8.24)$$

On peut déduire de ceci que le mapping et l'espace de représentation que le kernel induit ne sont pas uniques. Sans rentrer dans les détails, l'espace de représentation \mathbf{F} est un espace de Hilbert que l'on peut voir comme une généralisation de l'espace Euclidien.

Regardons maintenant une propriété intéressante du kernel polynomial. Supposons que nous ayons un ensemble d'exemples, $x_1, x_2, \dots, x_n \in \mathfrak{R}^N$ et que l'information discriminante se trouve dans les monômes d'ordre d , construits à partir des composantes de \mathbf{x} .

$$x_{i1}, x_{i2}, \dots, x_{id}$$

On peut extraire cette information en utilisant un mapping ϕ qui va amener les données dans un espace dans lequel chaque monôme de degré d est un axe. Si on **nomme** D_x , la dimension de l'espace d'entrée, la dimension de l'espace de représentation pour d fixé est :

$$D_F = \frac{(D_x + d - 1)!}{d!(D_x - 1)!} \quad (8.25)$$

Le calcul d'un kernel polynomial de la forme:

$$k(x,z) = (x,z)^d \quad (8.26)$$

Où: $x, z \in \mathfrak{R}^N$ revient à calculer le produit scalaire de ces deux vecteurs mappés dans L'espace de tous les produits (de composantes) d'ordre d .

Cette propriété est en pratique très importante. En effet, le coût d'une telle opération revient à

celui d'effectuer une élévation à la puissance d du produit scalaire (dans l'espace d'entrée) entre deux vecteurs. Nous terminons cette section en mentionnant qu'utiliser un kernel de la forme:

$$k(x, z) = (x, z) + 1 \quad (8.27)$$

Le kernel RBF (Radial Basis Function):

La forme générique de ce kernel est :

$$k(x, z) = \exp\left(-\frac{\|x - z\|^2}{2\sigma^2}\right). \quad (8.28)$$

Le mapping ϕ induit par ce type de kernel est un peu spécial. En fait, un exemple va être mappé sur une fonction gaussienne représentant la similarité de l'exemple avec tous les points de X .

L'espace d'arrivée F (l'espace de représentation) de la fonction ϕ est de dimension infinie étant donné que ϕ fait correspondre une fonction continue à chaque exemple, Le kernel RBF permet donc de calculer des similarités dans un espace de dimension infinie.

Avec le kernel RBF, tous les exemples sont placés sur une sphère de rayon 1 :

$$k(x, x) = \exp\left(-\frac{\|x - x\|^2}{2\sigma^2}\right) = 1 \quad (8.29)$$

Le paramètre σ permet de régler la largeur de la gaussienne. En prenant un σ grand, la similarité d'un exemple par rapport à ceux qui l'entourent sera assez élevée, alors qu'en prenant un σ tendant vers 0, l'exemple ne sera similaire à aucun autre. En resserrant fortement la gaussienne, un classificateur (faisant usage de ce kernel) peut arriver à apprendre n'importe quel training set sans commettre d'erreur. On sent tout de suite que le danger de l'apprentissage par cœur n'est pas loin. En fait, σ est un autre paramètre permettant de contrôler la capacité d'un classifieur.

8.7.3 Composition de kernels:

Il est possible de composer des nouveaux kernels en utilisant des kernels existants. En prenant $k_1(.,.)$ et $k_2(.,.)$ des fonctions satisfaisant à la condition de **Mercer**, $\mathbf{a} \in \mathbf{R}^+$ et \mathbf{B} une matrice définie positive.

Alors les fonctions suivantes des kernels:

$$1 - k(x, z) = k_1(x, z) + k_2(x, z)$$

$$2 - k(x, z) = ak_1(x, z)$$

$$3 - k(x, z) = k_1(x, z)k_2(x, z)$$

$$4 - k(x, z) = x^T Bz$$

8.7.4 Kernels spécifiques:

Si les données que l'on traite ne se trouvent pas dans \mathfrak{R}^n où que l'on dispose d'une connaissance a priori sur le sujet, il peut être utile de développer un kernel spécifique. Pour créer un nouveau kernel, on démarre souvent à partir des caractéristiques des objets pour définir un produit scalaire. Dans ce cas, la condition de **Mercer** est d'emblée vérifiée.

8.8 Formulation de SVM:

8.8.1 Cas linéairement séparables:

La méthode SVM consiste en un classificateur à marge maximale dans lequel le produit scalaire a été remplacé par le kernel. Effectuons dès lors cette substitution dans le problème suivant:

$$\text{Maximiser} \quad W(\alpha) = \sum_{i=0}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n y_i y_j \alpha_i \alpha_j k(x_i, x_j) \quad (8.30)$$

$$\text{Tel que} \quad \begin{cases} \sum_{i=1}^n \alpha_i y_i = 0 \\ \alpha_i \geq 0 \forall i = 1 \dots n \end{cases} \quad (8.31)$$

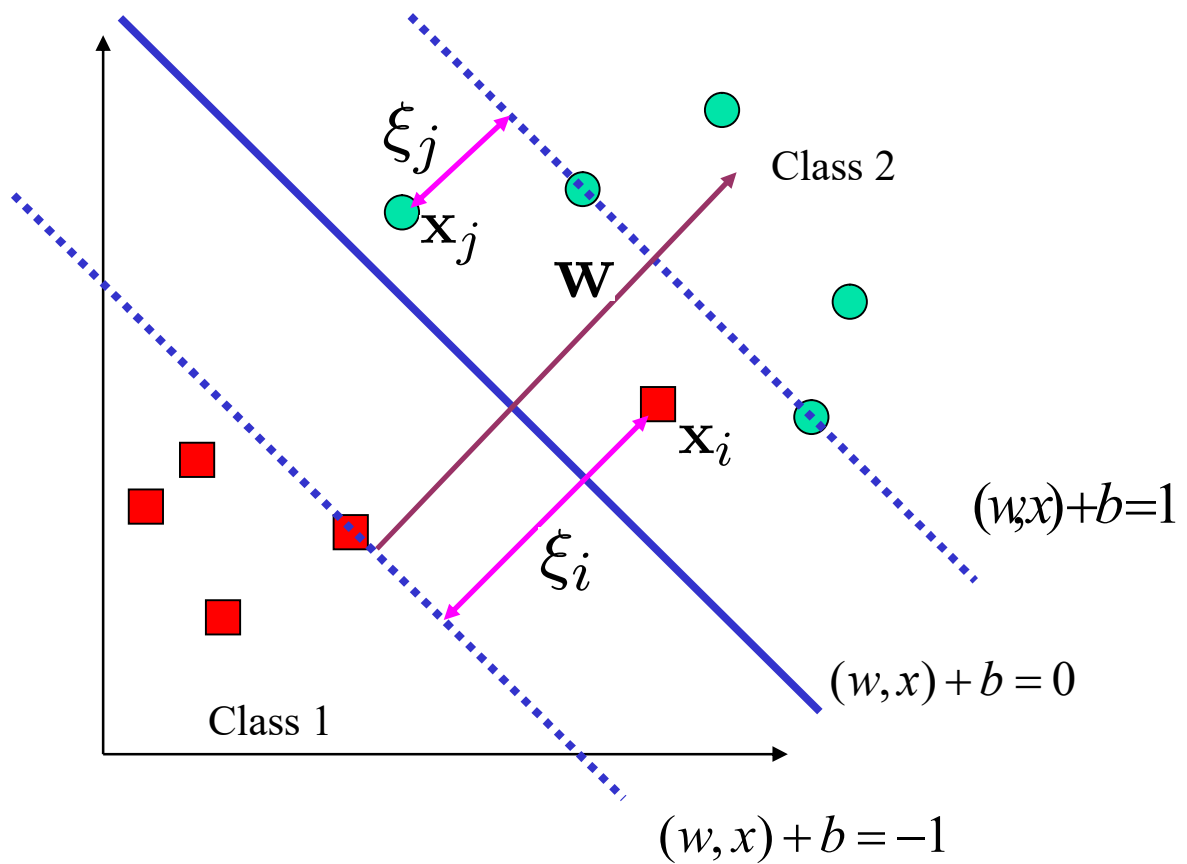
8.8.2 Formulation Soft Margin:

Il reste cependant un problème : nous avons toujours conservé l'hypothèse que les données sont linéairement séparables. Notons que dans ce cadre, la notion de linéairement séparable s'applique dans l'espace de paramètre. La plupart du temps, les données d'entraînement comportent du bruit (erreur d'acquisition, erreur sur la catégorie ...). Par conséquent même s'il existe une relation linéaire (dans F) entre les données et leur catégorie, un classificateur linéaire pourrait commettre des erreurs.

On pourrait trouver un l'espace de représentation induit par un kernel RBF suffisamment resserré qui rendrait les données linéairement séparables. Cependant, cela reviendrait à apprendre le bruit des exemples et donc à perdre une grande partie du pouvoir de généralisation (overfitting). Au lieu de cela, il paraît plus raisonnable d'admettre que certains exemples (supposés bruités) soient mal classés par notre classificateur. On appelle souvent ces exemples des outliers[62] .

Du point de vue de notre problème primal, cela revient à relaxer la contrainte imposant que tous les exemples soient bien classés. Pour ce faire, on peut introduire une «erreur tolérable» ξ_i dans la classification (figure 8.12).

$$\begin{aligned} y((w, x_i + b) &\geq 1 - \xi_i & \forall i = 1..n & \quad (8.32) \\ \xi &\geq 0 & \forall i = 1..n & \end{aligned}$$



Pour assurer que le nombre d'outliers reste raisonnable, nous allons intégrer les variables d'écart dans la fonction objective :

$$\text{Minimiser } \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \quad (8.33)$$

ou: C est une constante.

Géométriquement, la variable d'écart ξ_i divisée par $\|w\|$ correspond à la distance euclidienne prise perpendiculairement entre l'hyperplan canonique du côté de la catégorie de l'exemple et cet exemple. Notons que pour les exemples correctement classés, ξ_i est nul.

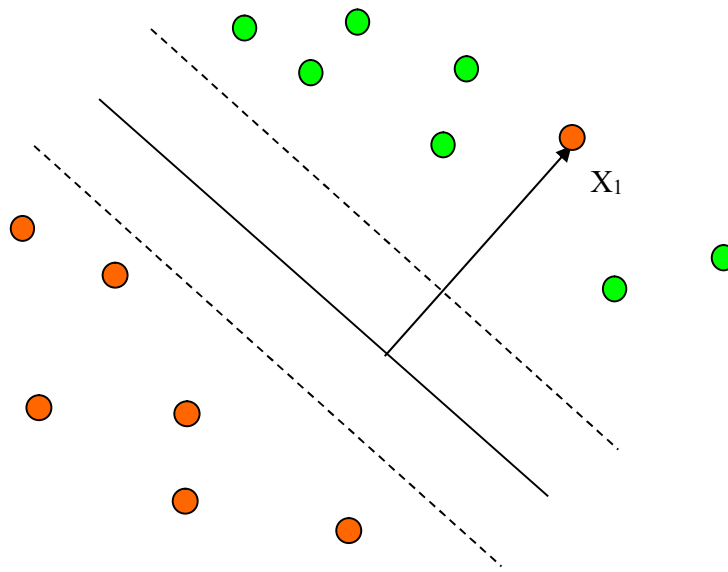


Figure 8.13 : la distance séparant un outlier et l'hyperplan canonique:

La distance est : $\xi_i / \|w\|$

8.8.2 Cas multi-classe

Plusieurs méthodes ont été proposées pour étendre le schéma ci-dessus au cas où plus de deux classes sont à séparer. Ces schémas sont applicables à tout classifieur binaire, et ne sont donc pas spécifiques aux SVM. Les deux plus connues sont appelées *one versus all* et *one versus one*. Formellement, les échantillons d'apprentissage et de test peuvent ici être classés dans M

classes $\{C_1, C_2, \dots, C_M\}$.

La méthode un contre tous (**one-versus-all**) appelée parfois **one-versus-the-rest**) consiste à construire M classifieurs binaires en attribuant le label 1 aux échantillons de l'une des classes et le label -1 à toutes les autres. En phase de test, le classifieur donnant la valeur de confiance (e.g la marge) la plus élevée remporte le vote.

La méthode un contre un (**one-versus-one**) consiste à construire $M(M - 1) / 2$ classifieurs binaires en confrontant chacune des M classes. En phase de test, l'échantillon à classer est analysé par chaque classifieur et un vote majoritaire permet de déterminer sa classe. Si l'on note x_t l'échantillon à classer et $h_{ij}(\cdot)$ le classifieur SVM séparant la classe C_i et la classe C_j et renvoyant le label attribué à l'échantillon à classer, alors le label attribué à x_t peut formellement se noter .

Card ($\{ h_{i,j}(x_t) \} \cap \{k\} ; i, j, k \in [1, M], i < j$)

C'est la classe qui sera le plus souvent attribuée à x_t quand il aura été analysé par tous les h_{ij} .

Ces méthodes souffrent toutes deux de défauts. Dans la version one-versus-all, rien n'indique que les valeurs du résultat de classification des M classifieurs soient comparables (pas de normalisation, donc possibles problèmes d'échelle)[36]. De plus le problème n'est plus équilibré, par exemple avec $M=10$, on utilise seulement 10% d'exemples positifs pour 90% d'exemples négatifs.

Dans le cas du one-versus-one, il peut exister une ambiguïté dans le résultat du comptage, s'il n'existe pas de vote majoritaire.

8.9 Conclusion:

Dans ce chapitre, nous avons tenté de présenter de manière simple et complète le concept de système d'apprentissage introduit par V.Vapnik, les «support vector machine» Nous avons donné une version générale des SVM. Cette méthode de classification est basée sur la recherche d'un hyperplan qui permet de séparer au mieux des ensembles de données.

Nous avons exposé les cas linéairement séparable et les cas non-linéairement séparable qui nécessitent l'utilisation de fonction noyau (Kernel) pour changer l'espace.

Il existe des extensions que nous n'avons pas présenté, parmi lesquelles l'utilisation des SVM pour des tâches de régression, c'est-à-dire de prédiction d'une variable continue en fonction d'autres variables, comme c'est le cas par exemple dans la prédiction de consommation électrique en fonction d'une période de l'année, de la température,...etc. Le champ d'application des SVM est donc large et représente une méthode de classification intéressante.

Chapitre 9

Analyse en composante principale (ACP)

9.1 Introduction

Le traitement de l'image s'appuie sur des données représentées par des mesures que l'on a accumulées sur des objets. Au sein du traitement de l'image, il existe une discipline, la statistique, qui a pour objet de décrire et d'analyser ces données. Dans la pratique, on peut être confronté à un problème de dimensionnalités élevées

L'analyse en composantes principales (ACP), permet d'étudier les données dans un espace de dimension réduite [91]. L'ACP revient à remplacer les attributs qui sont corrélés, par nouvelles variables : **les composantes principales**.

9.2 But :

Le but de l'ACP est d'obtenir une représentation la plus fidèle possible du nuage N en le projetant sur un espace de faible dimension. Pour cela, on cherche à minimiser les « écarts » entre les points de N et leurs projections. Les espaces de représentation choisis sont des espaces affines (droite, plan,.....).

9.3 Réduction de l'espace de représentation :

La réduction de l'espace de représentation permettra de diminuer le temps de classification pour toute nouvelle observation et d'éviter la dégradation des performances de la règle de décision. Les paramètres sélectionnés vont jouer un rôle important dans la phase d'apprentissage. En effet, ceux-ci devront caractériser au mieux les différents modes de fonctionnement. En d'autres termes, les paramètres jugés comme étant les plus pertinents seront ceux qui permettront de distinguer les classes dans l'espace de représentation. La réduction de l'espace de représentation peut être réalisée soit par des méthodes d'extraction de paramètres soit par des méthodes de sélection de paramètres.

L'extraction de paramètres consiste en la recherche de paramètres ($d' < d$), en fonction

des paramètres initiaux. Les points sont alors projetés dans un sous-espace \mathbb{R}^d . Le nombre de paramètres à calculer pour caractériser chaque observation, sera toujours égal à d . Seulement, l'extraction de paramètres fournit des combinaisons linéaires de ces paramètres initiaux. Parmi les différentes méthodes d'extraction de paramètres, on distingue l'analyse en composantes principales (ACP) ou transformée de Karhunen-Loève.

La sélection de paramètres consiste à en rechercher les plus pertinents et à conserver pour définir l'espace de représentation. Ces paramètres sélectionnés devront rendre optimum un critère de « séparabilité » entre classes et de compacité des classes comme par exemple le critère de Fisher [32].

9.4 Description de l'Analyse en composantes principales :

L'analyse en composantes principale (ACP) est une technique qui permet de caractériser un système complexe multidimensionnel à une autre dimension inférieure.

Dans le cadre de notre étude, le but est de définir un sous-espace de dimension d' , inférieure à d , pour mieux représenter l'ensemble d'apprentissage

Soit $x = \{x_1, x_2, \dots, x_N\}$ l'ensemble des N vecteurs (matrice des données) de dimension d . Si les composantes de ces vecteurs ne sont pas homogènes, alors les résultats de l'ACP et de la classification risquent d'être biaisés [41].

Afin de parer à cette éventualité, les données sont normalisées de la manière suivante :

Le centre de gravité de l'ensemble d'apprentissage est le vecteur défini par :

$$m = 1/N \sum_{k=1}^N X_k \quad (9.1)$$

Et les composantes de la variance sont :

$$\lambda^2_j = 1/N \sum_{k=1}^N (X_{kj} - m_j)^2 \quad (9.2)$$

Avec $j=1, \dots, d$

Ainsi, à partir des deux relations précédentes on en déduit les vecteurs formes normalisés (c'est-à-dire de moyenne nulle et d'écart type unité) :

$$x_{k\text{norm}} = \Delta^{-1/2} (x_k - m) \quad (9.3)$$

Avec Δ : la matrice diagonale des variances Γ_j^2 :

$$\begin{bmatrix} \lambda_1^2 & & 0 \\ & \lambda_2^2 & \\ 0 & & \lambda_d^2 \end{bmatrix} \quad (9.4)$$

Pour la suite de l'étude, les données seront considérées comme étant normalisées. La base des sous espace est définie par les vecteurs propres de la matrice de variance- covariance.

$$S = 1 / N \sum_{k=1}^N x_k \cdot x_k^t \quad (9.5)$$

S une matrice carrée (d×d) réelle et symétrique, c'est aussi la matrice d'inertie du nuage de points considéré. Ses éléments diagonaux sont les variances de l'ensemble d'apprentissage. Elles valent 1 car les données sont centrées et réduites. Et les éléments non diagonaux sont les covariances.

Les vecteurs propres (6) de la matrice de variance-covariance S sont les vecteurs V_j tels que $S.V_i = \lambda_j \cdot V_j$. Les scalaires λ_j , repris dans la matrice diagonale (7), sont les valeurs propres,

$$V = [V_1 \quad V_2 \quad \dots \quad V_d \quad \dots \quad V_d] \quad (9.6)$$

$$\begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \lambda_d \end{bmatrix} \quad (9.7)$$

Les premiers d' vecteurs propre, rangés par ordre décroissant des valeurs propres vont constituer la base du sous espace de dimension d'. Le vecteur x_j^p dont les N composantes sont les projections des N vecteurs x_k sur le vecteur v_j est appelé j^{ème} composante principale[15].

Chaque composante x_{kj}^p du vecteur x_j^p aura pour expression :

$$x_{kj}^p = x_k^t \cdot v_j \quad \text{Avec } k=1, \dots, N \quad (9.8)$$

La nouvelle matrice des données s'écrit : $x^p = [x_1^p \ x_2^p \ \dots \ x_d^p \ \dots \ x_d^p]$ (9.10)

Les axes définis par ces vecteurs x_j^p sont dits factoriels, et sont les vecteurs propres.

Ces axes sont les combinaisons linéaires des d axes initiaux.

On aboutit ainsi à une représentation optimale de la matrice des données x dans un espace de dimension inférieure.

L'ACP garantit que la déformation du nuage de points est minimale au sens d'un critère J_p [Dubuisson, 90] basé sur la moyenne des carrés des distances entre les vecteurs x_k et leurs projections notées x_k^p (telles que $x_k^p = [x_{k1}^p, x_{k2}^p, \dots, x_{kd}^p]^t$:

$$J_p = 1/N \sum_{k=1}^N d^2(x_k, x_k^p) \quad (9.11)$$

Les résultats de L'ACP peuvent être visualisés en considérant la projection du nuage de points sur les deux ou trois premiers axes factoriels. Autrement dit, en visualisant la matrice des données dans l'espace défini par les deux ou trois premiers vecteurs propres.

Enfin la qualité de la représentation peut être quantifiée par le calcul du pourcentage d'inertie expliqué sur chaque axe factoriel ou sur le sous espace de dimension d' obtenu. Ce

- Pourcentage d'inertie sur l'axe i :

$$\Gamma_i = \lambda_i / \sum_{j=1}^d \lambda_j \quad (9.12)$$

- Qualité de la représentation dans le sous espace de dimension d' :

$$\Gamma = \sum_{i=1}^{d'} \lambda_i / \sum_{j=1}^d \lambda_j \quad (9.13)$$

Ce rapport doit tendre vers 1 pour une représentation la plus proche possible de la réalité. Mais en pratique l'utilisateur fixe un seuil de qualité autour de 80%, 90%, puis sélectionne les vecteurs propres correspondant aux valeurs propres ayant permis d'atteindre ce seuil, afin de définir la base du sous espace. Une fois le sous espace $R^{d'}$ défini, il est possible d'évaluer la qualité de la représentativité de la projection x_k^p du vecteur x_k , dans ce sous espace, en calculant le cosinus de l'angle a_k que

forme ce dernier avec R^d :

$$\text{Cos}^2(a_k) = \frac{\sum_{i=1}^{d'} (x_{ki}^p)^2}{\sum_{j=1}^d (x_{kj})^2} \quad (9.14)$$

Ainsi plus $\text{cos}^2(a_k)$ est proche de l'unité plus la projection de x_k en x_k^p est représentative.

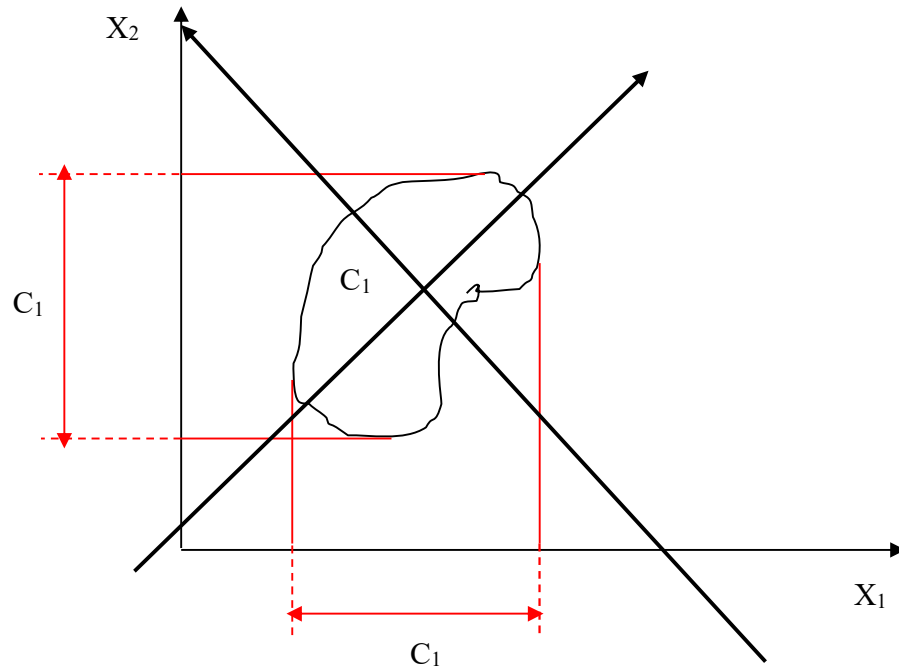


Figure 9.1 : Exemple de l'ACP

9.5 Sélection de paramètres :

L'ACP ou plus généralement les méthodes d'extraction de paramètres permettent de construire un sous espace de dimension inférieure à partir de l'espace initial pour entre autre décrire et visualiser l'ensemble d'apprentissage. Seulement cette procédure ne permet pas de réduire le nombre de paramètres à calculer pour chaque observation (étant donnée que les composantes des vecteurs projetés sont des combinaisons linéaires des composantes des vecteurs initiaux). On réduire le nombre de paramètres offre plusieurs

avantages :

- le temps de calcul est réduit non seulement pour l'extraction de paramètres, mais aussi pour la décision concernant toute nouvelle observation.
- Les performances de la classification sont améliorées.

Par la suite l'utilisation de l'analyse en composantes principales sera essentiellement réservée à la visualisation des données de l'ensemble d'apprentissage pour réduire le nombre de paramètres et, par la même occasion, la dimension de l'espace de représentation. On utilise des méthodes de sélection de paramètres [13].

L'objectif est de rechercher, parmi les paramètres de départ, un ensemble de paramètres donné pour être utilisé de la manière la plus efficace possible.

En d'autres termes, l'utilisateur doit connaître le nombre de classes existantes ainsi que les observations correspondantes. En outre les sous ensemble de paramètres devront maximiser un critère rendant compte à la fois :

-d'une faible dispersion des points dans chaque classe, ce qui suggère une variance intra-classe minimale.

-d'un éloignement maximal entre les différentes classes, donc d'une variance interclasses maximale.

Les matrices de variance intra classes et interclasses traduisent ces deux conditions. En effet l'expression de la matrice de variance intra classe fait intervenir les échantillons de chaque classe ainsi que les centres de gravité :

$$s_w = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^{N_i} (x_{ij} - m_i)(x_{ij} - m_i)^t \quad (9.15)$$

alors que la matrice de dispersion inter-classes d s'exprime en fonction des centres des classes et du centre de gravités de l'ensemble d'apprentissage :

$$S_B = \frac{1}{N} \sum_{i=1}^x \sum_{i=1}^M (m_i - m)(m_i - m)^t \quad (9.16)$$

avec M nombre de classes

N : nombre de vecteurs de l'ensemble d'apprentissage

N_i : nombre de vecteurs dans la classe w_i

m_i : centre de gravité de la classe w_i

m : centre de gravité

x_{ij} : j^{ième} vecteur de la classe w_i

Par ailleurs, la matrice variance covariance qui traduit la dispersion totale de l'ensemble vaut :

$$S = S_W + S_B \quad (9.17)$$

Le critère de Fisher appliqué à chacun des paramètres permet de quantifier le pouvoir de discrimination de ceux-ci. Dans le cas de deux classes et pour un paramètre α donné, le critère s'exprime de la manière suivante :

$$J(\alpha) = \frac{m_1(\alpha) - m_2(\alpha)}{N_1 \Gamma_1^2(\alpha) - N_2 \Gamma_2^2(\alpha)} \quad (9.18)$$

Avec :

- $m_i(\alpha)$ ($i = 1, 2$), centre de gravité de la classe w_i en ne considérant que le paramètre α
- $\Gamma_i^2(\alpha)$: variance de la composante α des vecteurs de classe w_i .

$$m_i(\alpha) = \frac{1}{N_i} \sum_{k=1}^{N_i} x_{ki}(\alpha) \quad (9.20)$$

$$\Gamma_i^2(\alpha) = \frac{1}{N_i} \sum_{j=1}^{N_i} [x_{ij}(\alpha) - m_i(\alpha)]^2 \quad (9.21)$$

D'une manière générale, pour M classes, $J(\alpha)$ s'écrit :

$$J(\alpha) = \sum_{i=1}^M \sum_{j=1}^{N_i} \frac{m_i(\alpha) - m_j(\alpha)}{N_i \Gamma_i^2(\alpha) - N_j \Gamma_j^2(\alpha)} \quad (9.21)$$

Le numérateur de l'expression (20) traduit la séparation des classes (variance inter-classes) alors que le dénominateur reflète la compacité des classes (par analogie avec la variance intra-classe). Plus le critère est grand, plus les classes sont séparées.

Ainsi une première solution pour la sélection des paramètres consiste à prendre en compte les d' premiers paramètres vérifiant :

$$J(\alpha_1) \geq J(\alpha_2) \geq \dots \geq J(\alpha_{d'}) \geq J(\alpha_{d'}) \quad (9.22)$$

Pour définir la dimension finale d' il suffit par exemple de visualiser la valeur du critère pour chacun des paramètres.

9.6 Conclusion

Dans notre travail nous avons utilisé l'ACP qui se présente comme une méthode statistique de fouille des données (Data Mining), pour l'extraction des paramètres pertinents et la réduction de vecteurs de paramètres avant la classification.

Chapitre 10

Application sur les Caractères Arabes Manuscrits Isolés

10.1 Introduction

Dans ce chapitre nous exposons quelques résultats des différentes méthodes développées appliquées aux caractères arabes manuscrits isolés : les prétraitements et l'extraction des paramètres ainsi que pour la classification

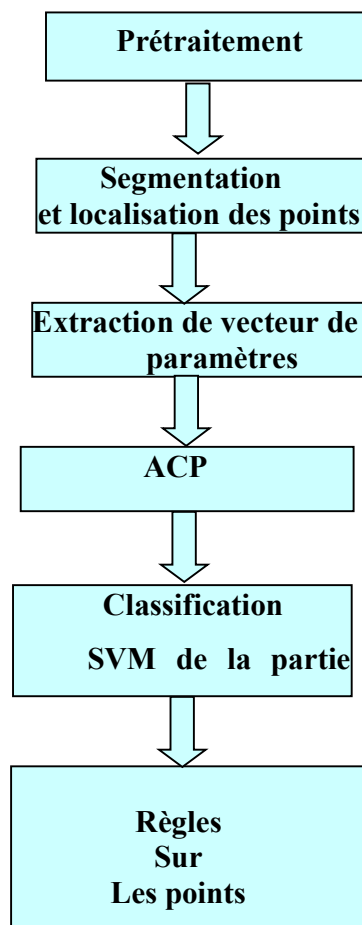


Figure 10.1 : différentes étapes d'analyse d'un caractère

10.2 Base de donnée

La base de donnée utilisée est constitué 1000 caractères manuscrits préservant uniquement la forme principale c'est-à-dire on ne tient pas compte aux point. Dans notre travail la base de donnée est composée de 16 classes :

- 800 caractères de cette base sont utilisés pour l'apprentissage.
- 200 pour le test

Le type d'écriture de cette base est mono-scripteur c'est-à-dire écrite par une seule personne

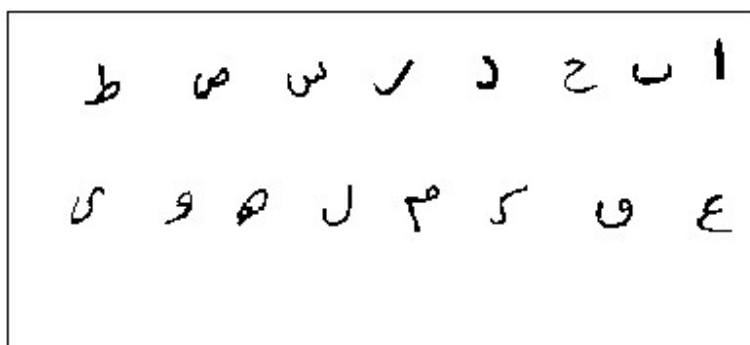


Figure-10.2: exemple de la base de donnée

les caractères sans point peuvent être classés en 16 classes nous considérons que les caractères ba ب , ta ت , tha ث , na ن appartiennent à la même classe

les différentes classes :

classe1 : أ

classe2 : ن ث ب

classe3 : خ ح ج

classe4 : ذ د

classe5 : ز ر

classe6 : ش س

classe7 : ض ص

classe8 : ط ظ

classe9 : غ ع

classe10 : ق ف

classe11 : ك

classe12 : ل

classe13 : م

classe14 : ه

classe15 : و

classe 16 : ي

10.3-Prétraitement

Afin d'améliorer les résultats de notre système, l'image du caractère subit les opérations suivante:

10.3.1 La squelettisation

Au lieu d'utiliser directement l'image binaire, on utilise son squelette qui préserve la totalité de l'information avec un minimum de données.

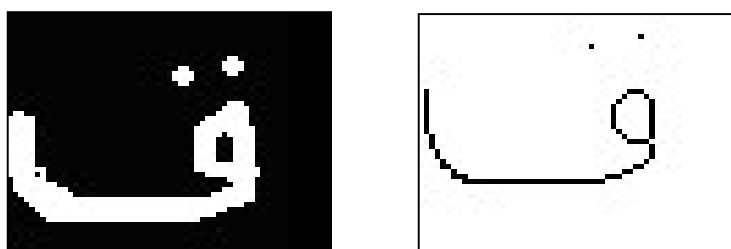


Figure 10.3: La squelettisation d'images

10.3.2 Filtrage

Cette opération consiste à éliminer les bruits dans l'image binaire, il suffit d'appliquer la méthode décrite dans le **chapitre 4**.

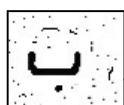


Image originale

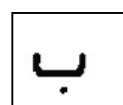


Image filtrée

Figure 10.4 : exemple de filtrage

On remarque bien que ce type de filtrage permet d'éliminer le bruit sans toucher aux points, et peut aussi éliminer les bruits sous forme de ligne car l'estimation de l'épaisseur d'écriture de ces points est moindre que l'épaisseur global.

10.4 Segmentation :

Cette étape consiste à diviser le caractère en deux parties : partie primaire et partie secondaire .

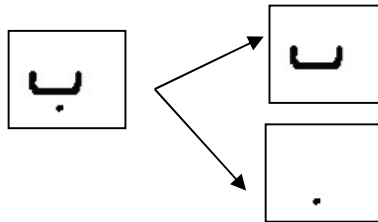


Figure-10.5: segmentation d'un caractère

10.5 La position des points

Il existe Trois positions possibles des points dans les caractères arabe (haut, bas, milieu)

point haut

les caractères arabes ou le où les points sont en haut : (ز.ن.ق.ف.غ.ظ.ض.ش.ذ.خ.ث. ت)

point bas

On a deux caractère arabe à point bas : (ب , ي)

Point milieu

On a un seul caractère à point milieu : caractère (ج)

10.6 Extraction de caractéristique

10.6.1-Les cavités

On définit les cavités en cinq types (Ouest, Nord, Sud, Est, centrale). Cela suppose la définition de quatre directions cardinales dans l'image (gauche, droite, haut, bas). Pour la visualisation des résultats, on peut créer une table de couleurs à sept lignes (cinq couleurs pour les cavités, plus une pour le fond et une pour le tracé). Où L'est en rouge, le West en vert, le Sud en bleu, le Nord en jaune et le centre en violet

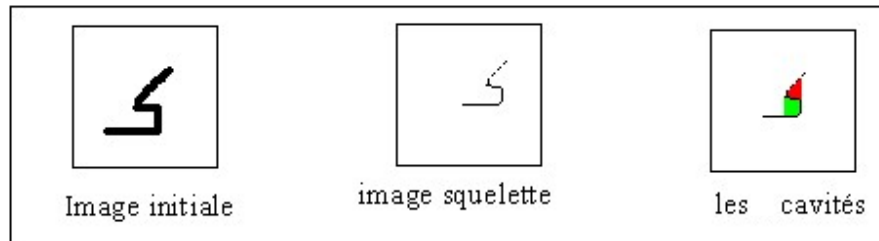


Figure 10.6: Exemple de cavités du squelette du caractères (kaf)

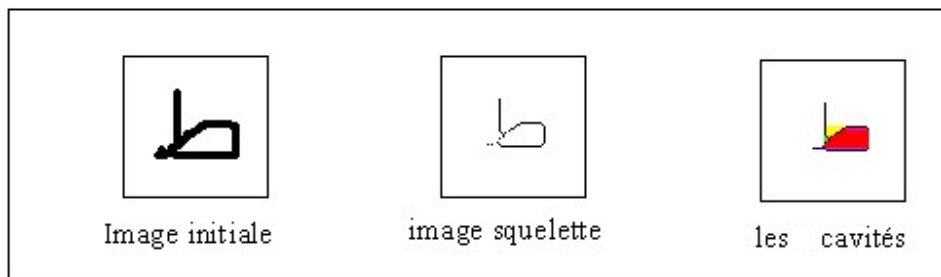


Figure-10.7 :Exemple de cavités du squelette du caractères (taa)

10.6.2-Les profils

On distingue quatre types de projection :

- **Profil bas**

Le profil bas est la distance entre l'axe horizontal bas est le premier pixel détecté verticalement dans chaque colonne.

- **Profil haut**

Le profil haut est la distance entre l'axe horizontal haut est le premier pixel détecté verticalement dans chaque colonne.

- **Profil gauche**

Le profil gauche est la distance entre l'axe vertical gauche est le premier pixel détecté horizontalement dans chaque ligne .

- **Profil droite**

Le profil droite est la distance entre l'axe vertical droit est le premier pixel détecté horizontalement dans chaque colonne

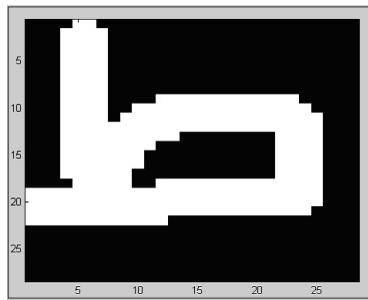


Figure 10.8; caractère taa

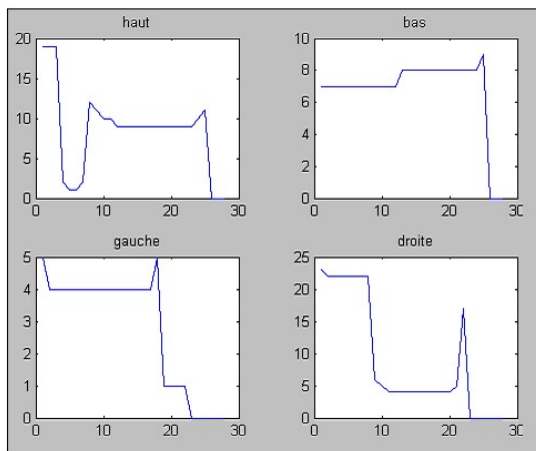


Figure 10.9: les quatre profils du caractère taa

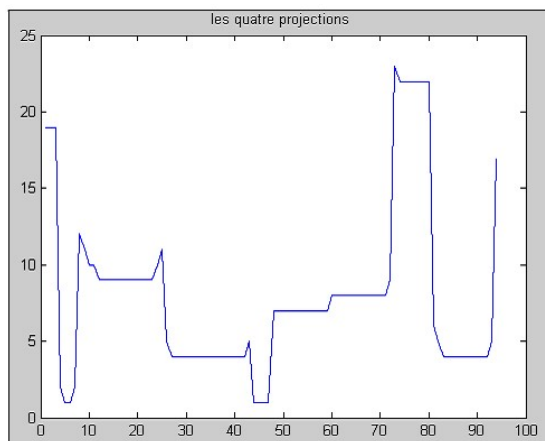


Figure 10.10 : les quatre profils du caractère taa dans un même vecteur

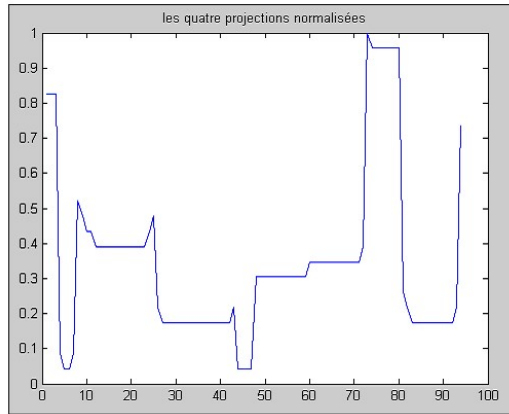


Figure 10.11 : profils normalisées du caractère taa

10.6.3 Exemples:

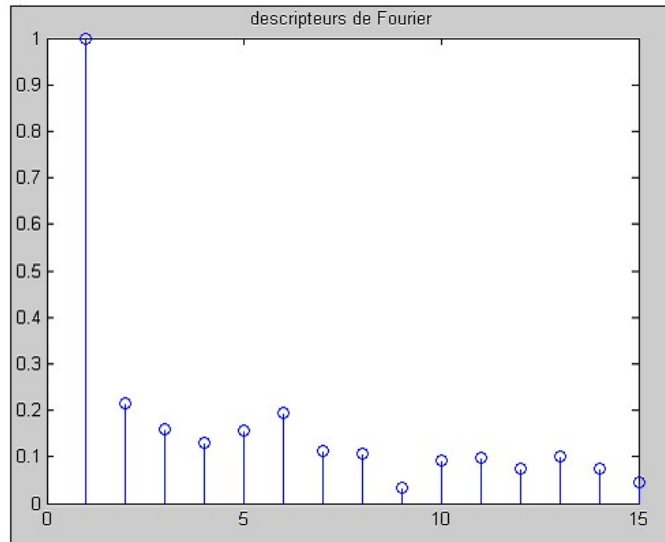


Figure 10.12: Descripteurs de Fourier des profils du caractère taa

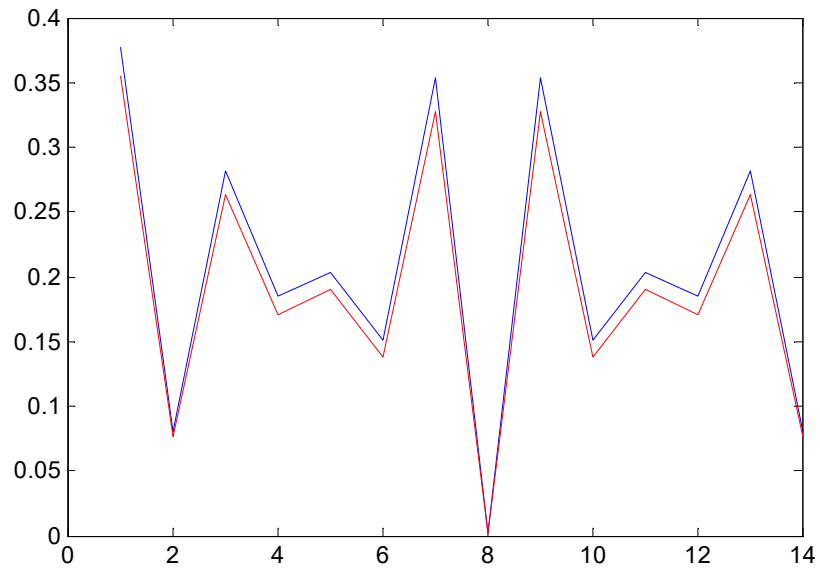


Figure 10.13 : Descripteurs de Fourier des 7 moment invariant du caractère taa en rouge et ba en bleu

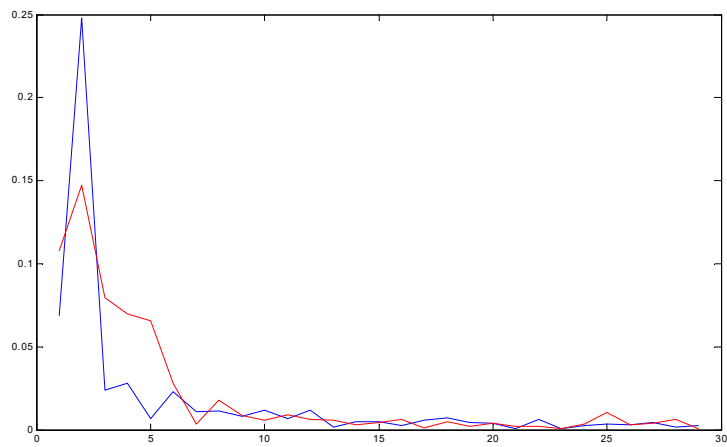


Figure 10.14 Descripteur de Fourier des distances centroides de les lettres (baa) ب and (jim) ج

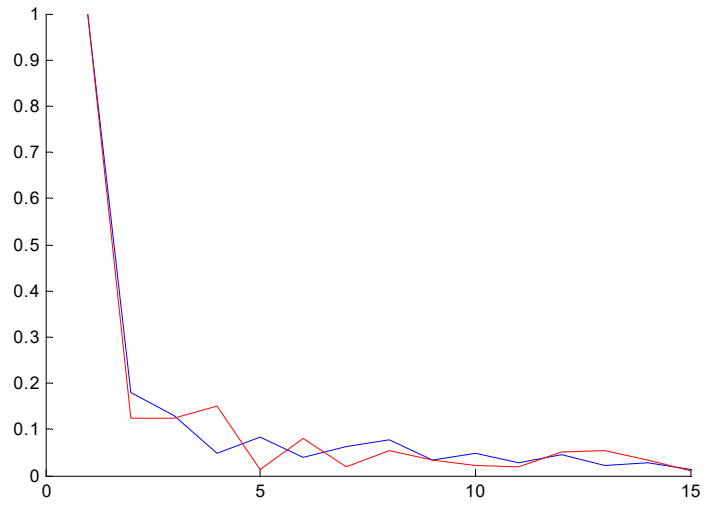


Figure 10.15 Descripteurs de Fourier des profils du caractère baa en bleu et HAA en rouge

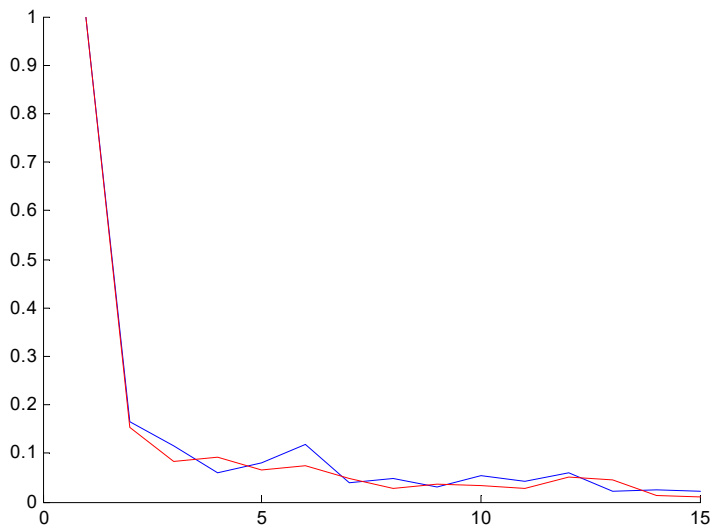


Figure 10.16 Descripteurs de Fourier des profils du caractère sine en bleu et saad en rouge

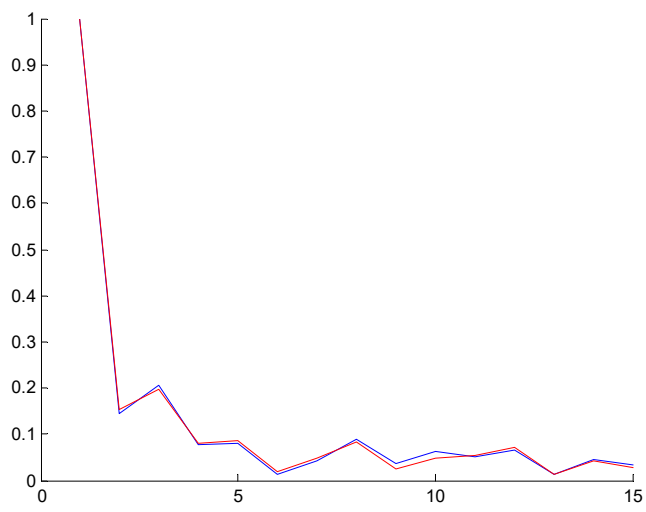


Figure 10.17 Descripteurs de Fourier des profils de deux écritures différentes du caractère ain

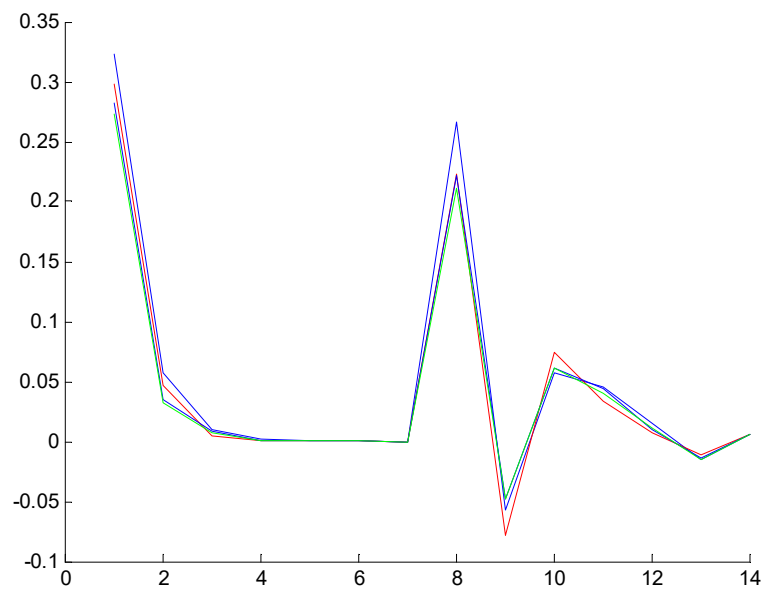


Figure 10.18: moment invariants de 4 caractères mim

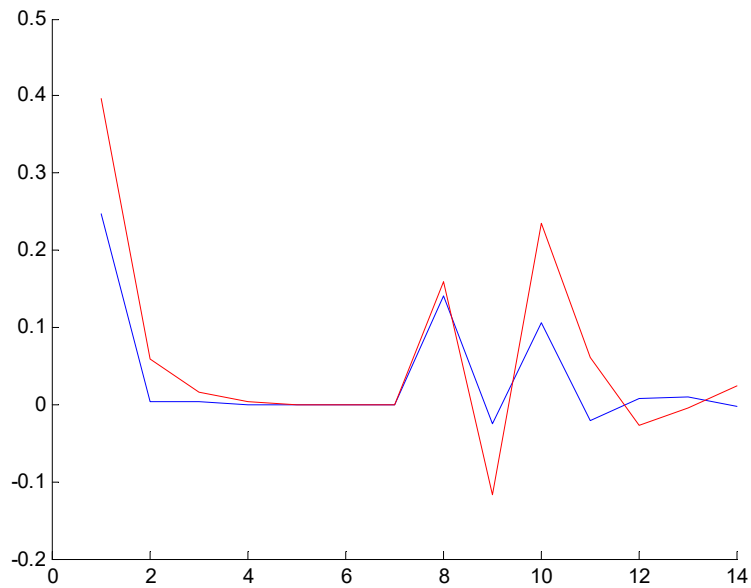


Figure 10.19 : Moments invariants du caractère Mim en bleu et Noun en rouge.

10.7 Classification

Pour la classification nous avons utilisé l'ACP avant la classification par SVM avec différents paramètres (cavité, moment invariant, Descripteurs de Fourier et profils),

En utilise toute la base de donnée de 16 classes (1000 caractères, 800 pour l'apprentissage, 200 pour le test.

10.8 Resultats obtenus:

METHODE	TEMPS	PARAMETRES	TAUX DE RECONNAISSANCE
SVM sans PCA avec RBF	rapide	Moment, cavités et Descripteurs de Fourier profiles	94%
SVM avec PCA avec RBF	rapide	Moment et Descripteurs de Fourier profils	96%

Tableau 10.1 Résultats obtenus

Les résultats obtenus montre que l'utilisation d'un classifieur SVM avec la fonction noyau RBF et après un traitement des données en utilisant l'ACP donne le meilleur taux de reconnaissance.

10.8 Conclusion :

Après ce test nous pouvons conclure que le SVM utilisant les fonction noyau RBF après l'application de l'ACP sur un vecteur de 58 paramètres et avec la méthode un contre un est le meilleur classifieur avec un taux de reconnaissance 96% L'avantage essentiel des méthodes basées sur SVM c'est qu'elle possèdent le pouvoir de généralisation à l'inverse des systèmes connexionnistes, qui peuvent donner un taux très élevé pour la base d'apprentissage mais ce n'est pas le cas pour la base de test.

Le seul problème lié aux SVM c'est que le choix des fonctions noyau les plus adaptées n'est pas évident, et reste à identifier par l'expérience.

Conclusion Générale et Perspectives

Dans ce travail de thèse, nous nous sommes principalement intéressés à l'analyse de l'écriture arabe manuscrite, pour cela nous avons développé des algorithmes pour les mots arabes cursifs et d'autre pour les caractères isolés. Dans notre étude nous avons utilisé deux bases de données; une pour les mots cursifs et l'autre pour les caractères isolés.

Plusieurs travaux ont appliqué des algorithmes développés pour le latin, mais cela a négligé les particularités de l'écriture arabe. Vu sa nature cursive et sa variabilité, des prétraitements spécifiques ont été élaborés pour les mots et d'autre pour les caractères isolés.

Dans la partie prétraitement nous avons développé un algorithme original qui permet d'éliminer les points de bruit sans toucher aux points qui font partie des caractères ; cet algorithme est basé sur l'estimation de l'épaisseur d'écriture qui dépend du stylo utilisé. Ce type de filtrage est appliqué pour les mots et les caractères isolés.

La ligne de base dans l'écriture arabe présente un avantage très important qui peut être utilisé par la suite dans les différentes étapes d'analyse. En effet, d'analyse car cette ligne peut être approximée par une ligne horizontale. La méthode développée permettant la correction de cette ligne de base est basée sur la transformé de Hough, qui elle permet de détecter les lignes dans une image.

Les travaux de recherches les plus récents présentent des tentatives pour la segmentation de l'écriture arabe manuscrite, qui sont encore au stade de l'expérimentation et on ne trouve pas de logiciel commercialisé utilisant la segmentation, à cet effet nous avons exposé une nouvelle tentative basée sur des théories mathématiques appliquées sur la ligne de base. Cette méthode peut être améliorée par une combinaison avec d'autres algorithmes.

Pour les caractères isolés nous avons développé un nouvel algorithme qui sert à séparer la forme principale des points, nous avons utilisé dans cette méthode l'extraction des composantes connexes dans l'image, la composante dont la surface est la plus grande est attribuée à la partie primaire du caractère le reste pour la partie secondaire (les points).

Les moments invariants et les descripteurs de Fourier des projections ont été utilisés comme paramètre de description (vecteur de primitive) de la partie primaire.

Nous avons utilisé une méthode de traitement des données en utilisant l'ACP l'analyse en

composantes principales qui est fondé sur la théorie de fouillé des données (data mining) afin d'extraire des paramètres pertinents ce qui a permis de réduire la taille du vecteur avec le même pouvoir de description. Le vecteur issu de cette étape est utilisé comme vecteur d'entrée pour un classifieur SVM machine à vecteurs support qui est considéré comme le meilleur classifieur développé à nos jours ce classifieur a été utilisé pour classifier la partie primaire. La position ainsi que le nombre des points ont été utilisés par la suite pour l'identification du caractère. Le taux de reconnaissance est de 96%.

Les résultats peuvent être amélioré en utilisant des combinaisons de méthodes et l'ajout d'autres paramètres descriptifs.

Références Bibliographiques

- [1] Adnan Amin. offline arabic character recognition a survey. In ICDAR '97 : Proceedings of the 4th International Conference on Document Analysis and Recognition, pages 596–599. IEEE Computer Society, 1997.
- [2] F. Bouchareb, M. Bedda Preprocessing Methods for Hand-written Arabic Word Recognition ICTTA06 IEEE Syrie Avril 2006
- [3] F. Bouchareb, R. Hamdi, M. Bedda Handwritten Arabic character recognition based on SVM Classifier “ Conference internationale en Syrie ICTTA 08 IEEE Avril 2008”.
- [4] F. Bouchareb, R. Hamdi, New Methods of Image Analysis Applied to Hand-Written Arabic Character Recognition “International Review on Computers and Software vol juin 2008
- [5] : Najoua Ben Amara, Abdel Belaïd et Noureddine Ellouze : Mémoire de recherche : Utilisation des modèles markoviens en reconnaissance de l'écriture arabe : Etat de l'art. 2000.
- [6] : Fouad Soufiane Douzidia : Université de Montréal, Mémoire de recherche : Résumé automatique de texte arabe.P.7. 2004
- [7] Sofien Touj, Najoua Essoukri Ben Amara, and Hamid Amiri. Reconnaissance hors ligne de caractères arabes isolées manuscrits. In CIFED, 2002.
- [8] B.Isabelle: Le traitement des images, “Morphologie mathématique. *Ecole Nationale Supérieure des Télécommunications - CNRS UMR 5141 LTCI Paris – France.* 1995
- [9] : Corinne Vachier ' Morphologie Mathématique et Analyse d'Image' 2002.
- [10] : M.coster –J.L.Cherment Pricis : Analyse d'image.2000
- [11] : Gilles Burel : ‘Introduction au Traitement d’images sous Matlab’.
- [12]: DUDA R.O, Hart, PE. “Use of the Hough transformation to detect lines and curves in pictures”. Comm. of AMC, 15(1).1972. P.11-15.
- [13]: Hough P.V.C., “Method and means for recognizing complex patterns”. United States patent 3069654.1962.
- [14]: R. Muniz, L. Junco, and A. Otero, A robust software barcode reader using the Hough transform, International Conference on Information Intelligence and Systems. (1999).
- [15] : S. El mejdani, R. Egli, F. Dubeau : Champ de hauteurs de la transformée de Hough Standard.

- [16]: Markus Ulrich, Carsten Steger, Albert Baumgartner; and Heinrich Ebner, "Real-Time Object recognition using a modified generalised Hough transform". *In Eckhardt sefert, editor, photogrammetrie – Fernerkundung – geoinformation: Technische Jahrestagung der DGPE*, Berlin, , pp571-578 .2001.
- [17]: B. Parhami and M. Taraghi, Automatic Recognition of Printed Farsi Texts, *Pattern Recognition*, 14(1-6), pp. 395-403, 1981.
- [18]: Adnan. A and Masini G., Machine Recognition of Arabic Cursive Words, SPIE 26th International Symposium on Instrument Display, Application of Digital Image Processing IV, Vol. 359, an Diego, pp. 286-292, Aug 1982.
- [19]: A. M. Obaid and T. P. Dobrowiecki, Heuristic Approach to the Recognition of Printed Arabic Script, IEEE International Conference on Intelligent Engineering Systems (INES'97), Budapest, Hungary, 15-17 Sep 1997.
- [20]: M. S. Khorsheed and W. F. Clocksin, Structural Features of Cursive Arabic Script, 10th British Machine Vision Conference BMVC'99, Univ of Nottingham, UK, Vol. 2, pp. 422-431, Sep 1999.
- [21]: B. M. Kurdy and A. Joukhadar, Multifont Recognition System for Arabic Characters, 3rd International Conference and Exhibition on Multi-lingual Computing (Arabic and Roman Script), University of Durham, UK, pp. 7.3.1-7.3.9, Dec 1992.
- [22]: A. Amin, Arabic Character Recognition, In *Handbook of Character Recognition and Document Image Analysis*, (Chapter 15), Edited by H. Bunke and P. S.P. Wang, World Scientific, Singapore, pp. 397-420, May 1997.
- [23]: I. Abuhaiba, S. Mahmoud and R. Green, Recognition of Handwritten Cursive Arabic Characters, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(6), pp. 664-672, Jun 1994.
- [24] Wai-Hong Wong, Wan-Chi Siu, and Kin-Man Lam. Generation of moment invariants and their uses for character recognition. *Pattern Recogn. Lett.*, 16(2) :115– 123, 1995.
- [25]: Ahmed M. Zeki: The Segmentation Problem in Arabic Character Recognition The State Of The Art. Department of Computer Science International Islamic University Malaysia Jln. Gombak 53100, Kuala Lumpur, Malaysia amzeki (1iiu.edu.my
- [26]: G. Abo Samra, K. Jambi, H. Al-barhamtoshy, R. Amer, and I. Al-Bidewi, A Comprehensive Algorithm for Segmenting Handwritten Arabic Scripts in Off-Line Systems, 17th National Computer Conference, KFUPM, 1-13 Nov 1997.
- [27] I. Al-Sughaiyer and I. Al-Kharashi. Arabic morphological analysis techniques : a

comprehensive survey. *J.Am.Soc.Inf.Sci.Technol.*,55(3):189–213, 2004.

- [28] A. Amin. Recognition of printed and handwritten arabic characters. In *BSDIA '97 : Proceedings of the First Brazilian Symposium on Advances in Document Image Analysis*, pages 40–59, London, UK, 1997. Springer-Verlag.
- [29] Abdel Belaïd and Christophe Choisy. Human reading based strategies for offline arabic word recognition. *Summit on Arabic and Chinese Handwriting Recognition 2006 - SACH'06*, 2006.
- [30] R. M. Bozinovic and S. N. Srihari. offline cursive script word recognition. *IEEE Trans. Pattern Anal. Mach. Intell.*, 11(1) :68–83, 1989.
- [31] C.Brechbuhler and S.Ho. Reconstruction of a closed curve from its elliptic descriptor.
- [32] Christopher J. C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2) :121–167, 1998.
- [33] P. Burrow. *Arabic Handwriting Recognition*. PhD thesis, Master of Science, School of Informatics University of Edinburgh, 2004.
- [34] T. Caesar, J.M. Gloger, and E. Mandler. Preprocessing and feature extraction for a handwriting recognition system. *Document Analysis and Recognition, 1993.*, Proceedings of the Second International Conference on, pages 408–411, 20-22 Oct 1993.
- [35] R.G. Casey and E. Lecolinet. A survey of methods and strategies in character segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 18(7) :690–706, Jul 1996.
- [36] Chih C. Chang and Chih J. Lin. *LIBSVM : a library for support vector machines*, 2001.
- [37] Yi-Kai Chen and Jhing-Fa Wang. Segmentation of handwritten connected numeral string using background and foreground analysis. *icpr*, 02 :2598, 2000.
- [38] M. Cheriet. Strategies for visual arabic handwriting recognition : issues and case study. In *ISSPA 2007, International Symposium on Signal Processing and its Applications*, 12 -15 February 2007, Sharjah, United Arab Emirates, Feb 2007.
- [39] Mohammed Cheriet, Nawwaf Kharma, Cheng lin Liu, and Ching Suen. *Character Recognition Systems : A Guide for Students and Practitioners*. Wiley-Interscience, 2007.
- [40] Nello Cristianini and John Shawe-Taylor. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press, March 2000.

- [41] Janez Demsar, Blaz Zupan, Gregor Leban, and Tomaz Curk. Orange : from experimental machine learning to interactive data mining. In PKDD '04 : Proceedings of the 8th European Conference on Principles and Practice of Knowledge Discovery in Databases, p537–539, New York, NY, USA, Springer-Verlag New York, Inc 2004.
- [42] X. Dupre. Contributions à la reconnaissance de l'écriture cursive à l'aide de modèles de Markov cachés. PhD thesis, Univ René Descartes-Paris V, 2003.
- [43] R. El-Hajj, L. Likforman-Sulem, and C. Mokbel. Arabic handwriting recognition using baseline dependant features and hidden markov modelling. In ICDAR, pages 893–897, 2005.
- [44] R. El-Hajj, C. Mokbel, and L. Likforman-Sulem. Reconnaissance de l'écriture arabe cursive : combinaison de classifieurs mms à fenêtres orientées. In CIFED, 2006.
- [45] M. Shridhar F. Kimura, S. Tsuruoka and Z. Chen. Context directed handwritten word recognition for postal service applications. Proc. Fifth US Postal Service Technology Conf., 1992.
- [46] A. M. Gouda and M. A. Rashwan. Segmentation of connected Arabic characters using hidden markov models. In Computational Intelligence for Measurement Systems and Applications, 2004. CIMSA. 2004 IEEE International Conference on, pages 115–119, 2004.
- [47] Latifa Hamami and Daoud Berkani. Recognition system for printed multi-font and multi-size arabic characters. In The Arabian Journal for Science and Engineering, volume 27 1B. IEEE Computer Society, avril 2002.
- [48] A. Harifi and A. Aghagolzadeh. A new pattern for handwritten persian/arabic digit recognition. Proceedings Of World Academy Of Science, Engineering And Technology, 3, 2005.
- [49] Nabil Hassan. Recognition of Arabic cursive handwriting. In GMAI, pages 135–140, 2006.
- [50] L. Heutte, J.V. Moreau, T. Paquet, Y. Lecourtier, and C. Olivier. Combining structural and statistical features for the recognition of handwritten characters. Pattern Recognition, Proceedings of the 13th International Conference on, 2 :210–214 vol.2, 25-29 Aug 1996
- [51] M.K. Hu. Pattern recognition by moment invariants. In Proc. IRE, page 1428, Sept. 1961.
- [52] M.K. Hu. Visual pattern recognition by moment invariants. In IRE Trans Inform, pages 179–187, 1962.
- [53] Lei Huang, Genxun Wan, and Changping Liu. An improved parallel thinning algorithm.

icdar, 02 :780, 2003.

- [54] B.K. Jang and R.T. Chin. One-pass parallel thinning : analysis, properties, and quantitative evaluation. *Pattern Analysis and Machine Intelligence*, IEEE Transactions on, 14(11) :1129–1140, Nov 1992.
- [55] Wady Kammoun and Abdel Ennaji. Reconnaissance de textes arabes à vocabulaire ouvert. In 8^{eme} colloque international francophone sur l'écrit et le document (CIFED'2004), June 2004.
- [56] M. S. Khorsheed. Recognising handwritten Arabic manuscripts using a single hidden markov model. *Pattern Recogn. Lett.*, 24(14) :2235–2242, 2003.
- [57] M. S. Khorsheed and W. F. Clocksin. Spectral features for Arabic word recognition. In ICASSP '00 : Proceedings of the Acoustics, Speech, and Signal Processing, 2000. on IEEE International Conference, pages 3574–3577. IEEE Computer Society, 2000.
- [57] F. P. Kuhl and C. R. Giardina. Elliptic fourier feature of a closed contour. In *Computer Vision, Graphics and Image Processing*, volume 18, pages 236–258, 1982.
- [58] Yann LeCun, Marc'Aurelio Ranzato, YLan Boureau, FuJie Huang, and Sumit Chopra. Learning a deep hierarchy of sparse and invariant features, 2006.
- [59] Y. Li. Reforming the theory of invariant moments for pattern recognition. *PR*, 25(7) :723–730, July 1992.
- [60] Liana M. Lorigo and Venu Govindaraju. offline Arabic handwriting recognition : A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(5) :712–724, 2006.
- [61] Volker.M and Haikal ElAbed. Arabic handwriting recognition competition. In ICDAR, pages 1274–1278. IEEE Computer Society, 2007.
- [62] David Meyer, Friedrich Leisch, and Kurt Hornik. The support vector machine under test. *Neurocomputing*, 55(1-2) :169–186, 2003.
- [63] H. Miled. Reconnaissance de l'écriture semi-cursive : Application aux mots manuscrits arabes. PhD thesis, PSI-La3i Univ Rouen, LIVIA ETS Montreal, 1998.
- [64] H.Miled, Mohamed Cheriet, and C.Olivier. Multi-level Arabic handwritten words recognition. In SSPR '98/SPR '98 : Proceedings of the Joint IAPR International Workshops on Advances in Pattern Recognition, pages 944–951, London, UK, 1998. Springer-Verlag.

- [65] J.J.Oliveira, J.deCarvalho, C.Freitas, and R.Sabourin. Feature sets evaluation for handwritten word recognition, 2002.
- [66] G. Olivier, H. Miled, K. Romeo, and Y. Lecourtier. Segmentation and coding of arabic handwritten words. In ICPR96, pages III : 264–268, 1996.
- [67] N.Otsu. A threshold selection method from gray-level histograms. In IEEE Trans. System, Man Cybernetics9, pages 62–66, 1979.
- [68] P.Swaminathan P.W.Palumbo and S.N.Srihari. Document image binarization: Evaluation of algorithms. Proc. SPIE, 697 :278–285, 1986.
- [69] M. Pechwitz and V. M'argner. Baseline estimation for arabic handwritten words. In IWFHR '02 : Proceedings of the Eighth International Workshop on Frontiers in Handwriting Recognition(IWFHR'02), page 479. IEEE Computer Society, 2002.
- [70] Pavithra Babu Sargur Srihari, Harish Srinivasan and Chetan Bhole. Spotting words in handwritten arabic documents. In In Procs. of SPIE, San Jose, CA, USA, Jan 2006.
- [71] Toufik Sari, Labiba Souici, and Mokhtar Sellami. offline handwritten arabic character segmentation algorithm :Acsa.In IWFHR '02
- [72] K. Sayre. Machine recognition of handwritten words : A project report. Pattern Recognition, 5 :213–228, 1973.
- [73] Bernhard Scholkopf and Alexander J. Smola. Learning with kernels : Support vector machines, regularization, optimization and beyond. MIT Press, 2002.
- [74] M. Sezgin and B. Sankur. Survey over image thresholding techniques and quantitative performance evaluation, 2004.
- [75] Patrice Simard, David Steinkraus, and John C. Platt. Best practices for convolutional neural networks applied to visual document analysis. In ICDAR, pages 958–962. IEEE Computer Society, 2003.
- [76] Ching Y. Suen. N-gram statistics for natural language understanding and text processing. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1(2) :164–172, apr 1979.
- [77] M. Suwa. Segmentation of connected handwritten numerals by graph representation. Document Analysis and Recognition, 2005. Proceedings. Eighth International Conference on, pages 750–754 Vol. 2, 29 Aug.-1 Sept. 2005.
- [78] C.C. Tappert, C.Y. Suen, and T. Wakahara. Online handwriting recognition-a survey. Pattern Recognition, 1988., 9th International Conference on, pages 1123– 1132 vol.2, 14-17 Nov 1988.

- [79] Y. Tay, P. Lallican, M. Khalid, C. Viard-Gaudin, and S. Knerr. An offline cursive handwritten word recognition system, 2001.
- [80] C. Aloulou Y. Bahou, L. Hadrich Belguith and A. Ben Hamadou. Adaptation et implémentation des grammaires hpsg pour l'analyse de textes arabes non voyellés. In Actes du 15e Congrès de Reconnaissance des Formes et Intelligence Artificielle (RFIA'2006), Tours, January 2006.
- [81] Sofien Touj, Najoua Essoukri Ben Amara, and Hamid Amiri. Generalized hough transform for arabic optical character recognition. In ICDAR '03 :
- [82] Oivind. and Anil K. Jain. Goal-directed evaluation of binarization methods. IEEE Transactions on Pattern Analysis and Machine Intelligence, 17(12) :1191–1201, 1995.
- [83] Oivind Due Trier and Torfinn Taxt. Evaluation of binarization methods for document images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 17(3) :312–315, 1995.
- [84] A.Vinciarelli and J.Luettin. A new normalization technique for cursive handwritten words. Pattern Recognition Letters, 22(9) :1043–1050, 2001.
- [85] Jiren Wang, Maylor K. H. Leung, and Siu Cheung Hui. Cursive word reference line detection. Pattern Recognition, 30(3) :503–511, 1997.
- [86] B. Yanikoglu and P. Sandon. Segmentation of off-line cursive handwriting using linear programming, 1998.
- [87] A.Zahour, L.Likforman-Sulem, W.Boussellaa, and B.Taconet. Text line segmentation of historical arabic documents. In ICDAR '07 : Proceedings of the Ninth International Conference on Document Analysis and Recognition(ICDAR2007) Vol 1, pages 138–142. IEEE Computer Society, 2007.
- [88] T.Y.Zhang and C.Y.Suen. Afast parallel algorithm for thinning digital patterns. Communications of the ACM, 27(3) :236–239, 1984.
- [89] David X. Zhong and Hong Yan. Pattern skeletonization using run-length-wise processing for intersection distortion problem. Pattern Recogn. Lett., 20(8) :833– 846, 1999.
- [90] V . Vapnik. The Nature of statistical Learning Theory. Springer Verlag New York 1995
- [91] B.pinkowsky Principal component analysis of speech spectrogram images patter recognition vol 30 n 5 pp777-787 1997

