

**REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE**  
**Ministère de l'Enseignement Supérieur et de la Recherche Scientifique**  
**Université Mouloud MAMMERI de Tizi-Ouzou**



**Faculté de Génie Electrique et d'Informatique**  
**Département d'Electronique**

## **MEMOIRE DE FIN D'ETUDES**

En vue de l'obtention du diplôme

***D'INGENIEUR D'ETAT EN ELECTRONIQUE***  
***OPTION CONTRÔLE***

*Thème :*

**Application de la méthode support  
vector machines (SVM) pour la  
reconnaissance des chiffres manuscrits.**

Proposé et dirigé par :

M<sup>elle</sup> AIT AIDER Malika

Réalisé par :

M<sup>elle</sup> HALICHE Noria

M<sup>elle</sup> ALLOUANE Aniya

*Promotion 2011*

# Remerciements

*NOUS REMERCIONS, AVANT TOUT, LE BON DIEU DE NOUS AVOIR DONNÉ LA SANTÉ, LE COURAGE ET LA VOLONTÉ POUR FINIR CE TRAVAIL.*

*NOUS TENONS À REMERCIER NOTRE PROMOTRICE, M<sup>ELLE</sup> AIT AIDER MALIKA, POUR SON AIDE, LE TEMPS QU'ELLE NOUS A CONSACRÉ, SES ORIENTATIONS ET POUR SA PATIENCE TOUT AU LONG DE CE TRAVAIL.*

*NOUS REMERCIONS, ÉGALEMENT, LES MEMBRES DE JURY QUI FERONT L'HONNEUR DE JUGER NOTRE TRAVAIL, D'APPORTER LEURS RÉFLEXIONS ET SUGGESTIONS SCIENTIFIQUES.*

*NOS REMERCIEMENTS LES PLUS CHALEUREUX S'ADRESSENT À NOS FAMILLES ET SURTOUT NOS PARENTS QUI SONT LA SOURCE DE CETTE RÉUSSITE ET QUI NOUS ONT SOUTENU ET ENCOURAGÉ POUR ALLER AU BOUT DE CE TRAVAIL.*

*NOS DERNIERS REMERCIEMENTS S'ADRESSENT, PRINCIPALEMENT, À M<sup>ELLE</sup> CHILALI OUARDIA ET À TOUS CEUX QUI ONT CONTRIBUÉ DE PRÈS OU DE LOIN À LA RÉALISATION DE CE TRAVAIL.*

# *Dédicace*

Je dédie ce modeste travail :

A mes *parents*.

A ma chère famille.

A tous mes amis (es).

Et à tous ceux qui sont content pour moi.

## ***SOMMAIRE***

<b>Introduction générale.....</b>	<b>01</b>
-----------------------------------	-----------

### **CHAPITRE I : Généralité sur la reconnaissance des formes.**

<b>I.1- Introduction.....</b>	<b>03</b>
<b>I.2- Méthodes de reconnaissance des formes.....</b>	<b>03</b>
<b>I.3- Domaines d'application.....</b>	<b>04</b>
<b>I.4- Reconnaissance des caractères manuscrits .....</b>	<b>06</b>
<b>I.5- Reconnaissance des chiffres manuscrits.....</b>	<b>08</b>
<b>I.5.1- Prétraitements.....</b>	<b>08</b>
<b>I.5.2- Segmentation de séquences des chiffres.....</b>	<b>15</b>
<b>I.5.3- Classification.....</b>	<b>18</b>
<b>I.5.4- Post-traitement.....</b>	<b>20</b>
<b>I.6- Conclusion.....</b>	<b>20</b>

### **CHAPITRE II : Théorie des ondelettes.**

<b>II.1- Introduction.....</b>	<b>21</b>
<b>II.2- Analyse de Fourier.....</b>	<b>21</b>
<b>II.2.1- Transformée de Fourier.....</b>	<b>21</b>
<b>II.2.2- Transformée de Fourier à fenêtre glissante.....</b>	<b>24</b>
<b>II.3- Transformée en ondelettes.....</b>	<b>25</b>
<b>II.3.1- Définition.....</b>	<b>25</b>
<b>II.3.2- Transformée en ondelettes continue (CWT).....</b>	<b>26</b>
<b>II.3.3- Transformée en ondelettes discrète (DWT).....</b>	<b>27</b>
<b>II.4- Analyse multi résolution (AMR).....</b>	<b>28</b>
<b>II.4.1- Fonction d'échelle.....</b>	<b>28</b>
<b>II.4.2- Fonction d'ondelette et espace de détail.....</b>	<b>29</b>
<b>II.4.3- Fonctions d'échelle et ondelettes à deux dimensions.....</b>	<b>30</b>

<b>II.4.4-</b> Présentation de l'algorithme pyramidal.....	33
<b>II.5-</b> Exemples d'ondelettes pratiques.....	34
<b>II.5.1-</b> Laplaciens itérés de Gaussienne.....	34
<b>II.5.2-</b> Ondelettes de Haar.....	34
<b>II.5.3-</b> Ondelettes de Daubechies.....	35
<b>II.5.4-</b> Les symmlets.....	36
<b>II.5.5-</b> Les coiflets.....	37
<b>II.6-</b> Conclusion.....	38

### **CHAPITRE III : Les Machines à Vecteurs de Supports (SVM).**

<b>III.1-</b> Introduction .....	39
<b>III.2-</b> Principe de la méthode des Machines à Vecteurs de Support (SVM).....	39
<b>III.3-</b> SVM linéaire .....	40
<b>III.4-</b> SVM non linéaire.....	44
<b>III.5-</b> Quelques noyaux classiques utilisés pour le SVM.....	45
<b>III.6-</b> SVM multi-classes .....	46
<b>III.6.1-</b> La stratégie un contre tous.....	46
<b>III.6.2-</b> Méthode un contre un .....	47
<b>III.7-</b> Les domaines d'applications.....	47
<b>III.8-</b> Conclusion.....	47

### **CHAPITRE IV : Tests et résultats.**

<b>IV.1-</b> Introduction.....	49
<b>IV.2-</b> Présentation des bases de données.....	49
<b>IV.3-</b> Les SVM.....	51
<b>IV.4-</b> Tests et Résultats.....	52
<b>IV.4.1-</b> Données artificielles (Base1).....	52
<b>IV.4.2-</b> Données réelles (base2).....	55

**IV.5-** Interprétation des résultats.....58

**IV.6-** Conclusion.....58

**Conclusion générale**.....60

**Bibliographie**

# **Introduction générale**

La reconnaissance de formes est une branche de l'intelligence artificielle qui a la capacité d'identifier ou de classer des formes ou des objets en se basant, essentiellement, sur l'étape de caractérisation que sur la technique de classification.

C'est un domaine très vaste, car on trouve plusieurs applications à savoir : la reconnaissance des empreintes digitales, identification de personnes, reconnaissance des défauts de pièces (contrôle de qualité), la reconnaissance de caractères manuscrits, etc. Dans cette dernière application on distingue la reconnaissance de chiffres manuscrits qui fera l'objet de notre travail.

On distingue deux volets dans la reconnaissance de chiffres manuscrits : reconnaissance en ligne et la reconnaissance hors ligne. Dans la première catégorie, la reconnaissance se fait en temps réel où le facteur temps de traitement est non négligeable. Par contre dans la deuxième, la reconnaissance est en temps différé. Dans ce travail, on s'est intéressé à la deuxième catégorie qui est une tâche plus complexe que la précédente.

La reconnaissance hors ligne reste à ce jour un thème de recherche ouvert. En effet, bien que le nombre de classes naturelles soit réduit à 10 (chiffres 0 à 9), on trouve à l'intérieur de chacune d'entre elles une très grande variabilité de l'écriture. De plus, les conditions, souvent précaires, du matériel utilisé (divers stylos, différentes qualités de papiers) tendent à rendre la reconnaissance plus complexe. Dans ces conditions, l'extraction des caractéristiques est une étape essentielle et délicate dans la construction d'un système de reconnaissance. Dans notre mémoire, l'extraction des caractéristiques est faite selon trois méthodes : la binarisation, la normalisation et la méthode d'ondelette. Les images ainsi traitées sont classifiées par le classifieur machine à vecteur de support (SVM). Les résultats de ces trois analyses sont comparés entre elles afin d'en déduire la plus performante.

Pour bien mener notre travail, nous avons organisé notre mémoire en quatre chapitres.



Le **premier chapitre** présentera la décomposition classique d'une chaîne de reconnaissance des formes, tout en se focalisant sur la reconnaissance des chiffres manuscrits.

Le **second chapitre** porte sur la théorie des ondelettes et l'analyse multi résolutions.

Le **troisième chapitre** sera consacré à la méthode des machines à vecteurs de support (SVM) utilisée comme classifieur. Nous nous intéresserons au SVM utilisant la technique un contre tous.

Quant au **quatrième chapitre**, nous présenterons les résultats obtenus, ainsi que leurs interprétations.

Nous terminerons notre travail par une conclusion générale.

**CHAPITRE I :**  
**Généralité sur la**  
**reconnaissance de**  
**formes**

**I.1- Introduction**

L'apparition de l'ordinateur a mis en sorte la réalisation d'un nombre important de projets dont la simulation des tâches assurées par les humains. Cela est même allé loin en ajoutant à ces tâches d'autres tâches dont l'exécution révèle des facultés mentales comme la perception et le raisonnement. Ce qui a entraîné un nouveau domaine de recherche qui l'intelligence artificielle.

La reconnaissance des formes, étant une branche de cette intelligence artificielle, a pour objectif d'associer une étiquette à une donnée qui peut se présenter sous forme d'une image ou d'un signal. Il s'agit, au fait, de repérer dans des signaux compliqués (signal de parole, images numérisées ou même des textes en langue naturelle ou séquence d'ADN, etc) la présence de certaines structures (formes, objets) et à les classer en un nombre restreint de classes pertinentes pour l'application envisagée. Elle fait largement appel aux techniques d'apprentissage automatique et aux statistiques [11][19].

La reconnaissance des formes représente un domaine de recherche actif et d'études variées [2]. Les formes à reconnaître peuvent être de nature très variée. Il peut s'agir de contenu visuel (code barre, visage, empreinte digitale, etc) ou sonore (reconnaissance de parole), d'images médicales (rayon X, EEG, IRM, etc) ou multi spectrales (images satellitaires) et bien d'autres [28].

Dans ce chapitre, nous allons essayer de contourner cette notion de reconnaissance des formes.

**I.2- Méthodes de reconnaissance des formes**

Des méthodes générales ont été développées, en reconnaissance des formes, pour extraire, automatiquement, des informations des données sensibles afin de caractériser les classes de formes (apprentissage) et d'assigner, automatiquement, des données à ces classes (reconnaissance) telles que :

- la mise en correspondance des graphes.
- la méthode Bayesienne.
- l'estimation Paramétrique.
- le classifieur linéaire.
- les réseaux de neurones.
- le local feature focus.
- SVM : Support Vector Machine.
- les polyèdres de contrainte.
- la méthode des hypercubes.

Divers algorithmes d'apprentissage automatique peuvent être effectués tels que:

- le réseau de neurones.
- l'analyse statistique.
- l'utilisation de modèles de Markov cachés.
- la recherche d'isomorphisme de graphes ou sous-graphes.

### **I.3- Domaines d'application**

La reconnaissance des formes s'est propagée à un nombre important de domaines. Dans le traitement du signal, les études ont été faites sur :

- la reconnaissance des paroles : appliquée dans les banques et commerce.
- l'électrocardiogramme, électro-encéphalogramme : dans le domaine médical.

Dans le traitement d'images, les domaines d'études sont distincts. On trouve :

- la reconnaissance des caractères machines ou manuscrits : employées dans la saisie de textes, bureautique, tri-postal, chèque et compression télécopie.

- reconnaissance des signatures : appliquées dans les banques, postes et commerce.
- reconnaissance des empreintes digitales, des visages : mis en œuvre dans les banques, commerces et police.
- analyse de radiographie, échographie, reconnaissance chromosomes et le comptage globules : appliqués dans les contrôles systématiques de santé.
- détection de défauts, circuits intégrés, pièces métalliques : manufacturées dans le contrôle de qualité industrielle.
- identification d'objets : appliquée au tri d'objets industriels et surveillance militaire.
- localisation d'objets : mise en œuvre dans le guidage de robots industriels et le guidage de missiles.
- analyse d'image de satellite : mise en œuvre dans la météorologie, agriculture, ressources terrestres et surveillance militaire.
- analyse de photos aériennes : en agriculture ou en surveillance militaire.
- analyse d'échos radar : dans la poursuite de cibles et pilotage missiles.

Parmi les domaines d'étude de formes, notre intérêt s'est porté sur la reconnaissance des caractères manuscrits. C'est un traitement informatique ayant pour but de traduire un texte écrit en un texte codé numériquement. Cette application a pris le succès après l'apparition du premier ordinateur dans les années soixante avec l'apparition des premiers systèmes de lecture automatique de textes imprimés [5]. Dans la même période [23], l'un des premiers systèmes de reconnaissance de l'écriture manuscrite a été introduit. Ce grand pas, dans la technologie de reconnaissance, a contribué à révolutionner et simplifier la simulation de l'être humain que ce soit dans le tri automatique de courrier, le traitement automatique de dossiers administratifs, des formulaires d'enquêtes ou encore l'enregistrement des chèques bancaires [26]. Reste à savoir que c'est un domaine de recherche source de nombreuses publications [4] [10] [21] [23] [24][27][29][30].

**I.4- Reconnaissance des caractères manuscrits**

Les systèmes de reconnaissance de l'écriture manuscrite sont créés pour faciliter l'interaction homme-machine. Un système de reconnaissance de l'écriture, doit, idéalement, localiser, reconnaître et interpréter n'importe quel texte ou nombre écrit sur un support de qualité, arbitrairement, variable tel que des cartes, des formulaires, des agendas, des vieux manuscrits, etc. Ce type de systèmes aide à convertir l'information, contenue dans les vieux manuscrits, pour être enregistrée dans des bases de données que tout le monde peut interroger. Par la suite, les chercheurs se sont vite aperçus de la difficulté de réaliser des systèmes capables de reconnaître tout type de texte. Ainsi, leurs recherches se sont tournées vers des systèmes dédiés, où le style d'écriture, que le système doit traiter, est connu à priori. Par exemple, il existe des systèmes pour la lecture automatique des chèques bancaires, la reconnaissance du code et de l'adresse postale, la lecture des formulaires d'impôts, etc [3]. La figure I.1 montre un exemple de système de reconnaissance de l'écriture manuscrite du chiffre neuf.



Figure I.1: reconnaissance de l'écriture manuscrite pour le chiffre neuf.

Deux types d'applications se présentent :

- Reconnaissance d'écriture manuscrite en ligne.
- Reconnaissance d'écriture manuscrite hors-ligne.

Dans le cadre de la reconnaissance en ligne, le système prend en compte l'information chronologique des mouvements du bras du scripteur, les posés et les levés du stylo et éventuellement l'inclinaison et la vitesse (il détermine l'ordre de traits constituant un caractère).

Les échantillons sont beaucoup plus informatifs. Seulement, cela revient coûteux et impose de fortes contraintes au scripteur puisque la capture de l'encre doit se faire au moment de la saisie (capture synchrone).

Le matériel adapté à cette application est spécifique comme le cas des stylos numériques, stylets sur agendas électroniques, les tablettes PC, le DPA et le Pocket PC [7][32].

Les systèmes actuels ont l'aptitude d'effectuer une comparaison de l'échantillon à reconnaître avec ceux contenus dans une base de données. Cette base de données peut être créée de toutes pièces ou être l'objet d'une phase d'apprentissage [32].

En revanche, dans la reconnaissance hors ligne, l'écriture de l'utilisateur est acquise par un scanner. L'entrée de ces systèmes est une image. Dans ce type de reconnaissance, il est impossible de savoir comment ont été tracés les caractères. Contrairement à la reconnaissance en ligne, des informations supplémentaires ne sont pas en possession. D'autant plus, ces systèmes doivent être capables de traiter de larges bases de données avec un taux de reconnaissance élevé. C'est le cas des traitements asynchrones tels que la classification automatique des lettres, la lecture des chèques bancaires et tri postal [7].



Figure I.2 : reconnaissance en ligne et hors ligne.

**I.5- Reconnaissance des chiffres manuscrits**

Les grandes étapes qui composent une chaîne de reconnaissance des chiffres manuscrits sont : prétraitements, normalisations et segmentation.

**I.5.1- Prétraitements**

Cette étape prépare l'image entrée pour l'étape de reconnaissance. Les systèmes de reconnaissance de l'écriture prennent les images numérisées, qui peuvent être binarisées ou en niveaux de gris. Lorsque les images sont en niveaux de gris, les systèmes incluent, généralement, une étape de binarisation de l'image, qui consiste à séparer l'information utile (l'écriture) du fond.

- ***Binarisation***

L'image d'entrée se représente en niveau de gris et les algorithmes de reconnaissance courants travaillent, souvent, sur des images binaires. Donc, il faut appliquer un seuillage [23]. On distingue en général deux approches : le seuillage global et le seuillage adaptatif.

- **Seuillage global**

Il consiste à prendre un seuil ajustable, mais identique pour toute l'image. Chaque pixel de l'image est comparé à ce seuil et prend la couleur blanche ou noire selon la valeur supérieure ou inférieure. Cette classification ne dépend alors que du niveau de gris du pixel considéré [23].

Cette méthode convient pour les documents simples et de bonne qualité. Néanmoins, elle n'est plus applicable lorsque la qualité d'impression du texte n'est pas constante dans toute la page, des caractères peuvent être partiellement perdus. Des



problèmes surviendront également si le fond est bruité ou non homogène, dans ce cas des taches parasites peuvent apparaître.

- **Seuillage adaptatif**

Dans les documents pour lesquels l'intensité du fond et l'intensité de la forme peuvent varier au sein du document, un seuillage global est inadapté. Il devient nécessaire de choisir le seuil de binarisation de manière locale. On calcule un seuil de binarisation pour chaque pixel de l'image, en fonction de son voisinage.

J. Bernsen [23], propose le calcul qui détermine le seuil de binarisation de chaque pixel. Cette méthode est suffisante pour les images dont le fond varie lentement (sans transitions brusques dans le fond) de façon non uniforme.

Les travaux récents de B. Gatos et al., appliqués à la binarisation de documents dégradés (tâches, ombres), incluent une phase de filtrage, une première estimation grossière du découpage fond/forme et la phase de calcul des seuils locaux (une phase de post-traitements spécifiques) [23].

Les travaux de Y. Xi et al., combinent deux méthodes anciennes (Niblack et Palumbo ). Le seuillage de Niblack conserve bien les caractères, mais a tendance à laisser beaucoup de bruits de fond dans l'image. Le seuillage de Palumbo filtre une grande partie des bruits de fond, mais a parfois tendance à supprimer, également, des parties de caractères. En combinant les deux méthodes, le système offre des résultats intéressants [23].

• ***Squelettisation***

Ce traitement aboutit à la représentation de l'information indépendamment de l'épaisseur initiale de l'écriture. Elle permet le passage d'une image à sa représentation en squelette.

Le squelette a un pixel d'épaisseur et permet d'extraire les caractéristiques essentielles telles que les intersections, le nombre de tracés et leurs positions relatives.

Il est possible, aussi, de renormaliser l'épaisseur de l'écriture à partir du squelette. Ce dernier doit remplir les trois conditions suivantes :

- finesse (typiquement, 1 pixel d'épaisseur) ;
- respect de la connexité ;
- centré dans la forme qu'il représente.

Le squelette ne doit comporter qu'une composante connexe (d'un pixel d'épaisseur) de la forme initiale (figure I.3).

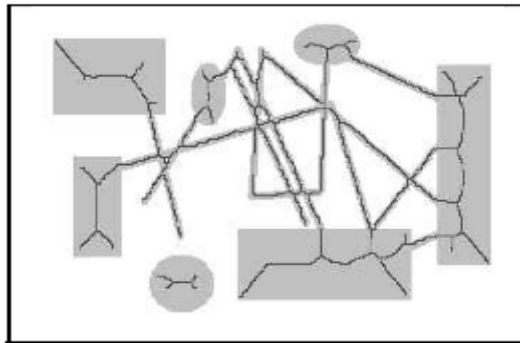


Figure I.3 : Squelettisation d'une forme quelconque.

Il existe de nombreuses méthodes de squelettisation. L'une des manières d'évaluer le squelette est de calculer l'axe médian de la forme, comme l'ensemble des centres des boules maximales de cette forme (figure I.4).

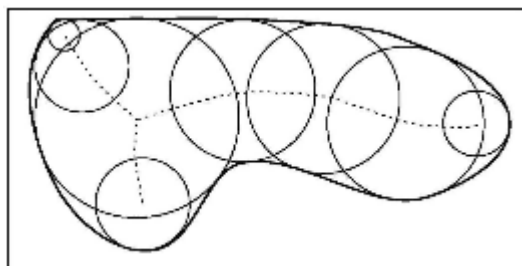


Figure I.4 : Axe médian d'une forme continue (trait pointillé) dont seul le contour est représenté.

Un autre moyen d'obtenir le squelette est de procéder par affinages successifs : la forme est épluchée (peeling) de manière itérative, en maintenant valide le critère de connexité. L'algorithme de Hilditch [23] exploite ce principe.

On définit le voisinage du pixel P1 comme :

<b>P9</b>	<b>P2</b>	<b>P3</b>
<b>P8</b>	<b>P1</b>	<b>P4</b>
<b>P7</b>	<b>P6</b>	<b>P5</b>

On définit également :

$A(P_1)$  : nombre de transitions de 0 vers 1 dans la séquence P2, P3, P4, P5, P6, P7, P8, P9.

$B(P_1)$  : nombre de pixels noirs dans le voisinage de P1 (P1 étant exclu).

$$B(P_1) = \sum_{i=2}^{i=9} P_i \quad \text{avec} \quad \begin{cases} P_i = 1 & \text{si le pixel est noir} \\ P_i = 0 & \text{si le pixel est blanc} \end{cases}$$

L'image est parcourue plusieurs fois et un pixel noir est marqué comme effaçable s'il respecte les quatre conditions suivantes [23]:

– Condition1 :  $2 \leq B(P_1) \leq 6$ , la première inégalité s'assure que P1 n'est ni un point isolé, ni une extrémité. La deuxième inégalité s'assure que P1 est bien un point du contour de la forme.

– Condition2 :  $A(P_1) = 1$ , maintient le critère de connectivité.

– Condition3 :  $P_2 \cdot P_4 \cdot P_8 = 0$  ou  $A(P_2) \neq 1$ , cette règle s'assure que les lignes verticales de 2 pixels de large ne sont pas totalement érodées.

– Condition4 :  $P2.P4.P6 = 0$  ou  $A(P4) \neq 1$ , cette règle s'assure que les lignes horizontales de 2 pixels de large ne sont pas totalement érodées.

Tous les points marqués comme effaçables sont ensuite effacés, et on réitère l'opération sur la nouvelle image, jusqu'à ce que plus aucun point ne soit effaçable.

Après la squelettisation, un grand nombre de petits segments du squelette, appelés barbules, peuvent s'avérer non pertinents, comme le montre la figure I.5.

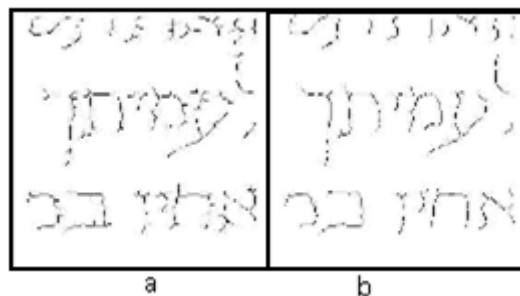


Figure I.5 : Barbules. (a) : squelette brut.  
(b): squelette nettoyé de ses barbules.

Ces barbules peuvent être nettoyées pour rendre le squelette plus lisible. Plusieurs critères ont été proposés pour supprimer ces traits superflus. Par exemple, on peut utiliser un simple critère de taille : tous les arcs de squelette dont la longueur est inférieure à un seuil sont considérés comme un bruit et sont supprimés. B.K.

Les intersections entre deux segments de droites se retrouvent parfois scindées, comme le montre la figure I.6-b. En parcourant l'image selon plusieurs directions, D.X.Zhong et H.Yang [31] proposent une méthode pour déterminer les zones d'embranchements. Ils font, ainsi, converger, en un point unique, tous les arcs de squelette qui sont connectés dans cette zone (figure I.6-c).

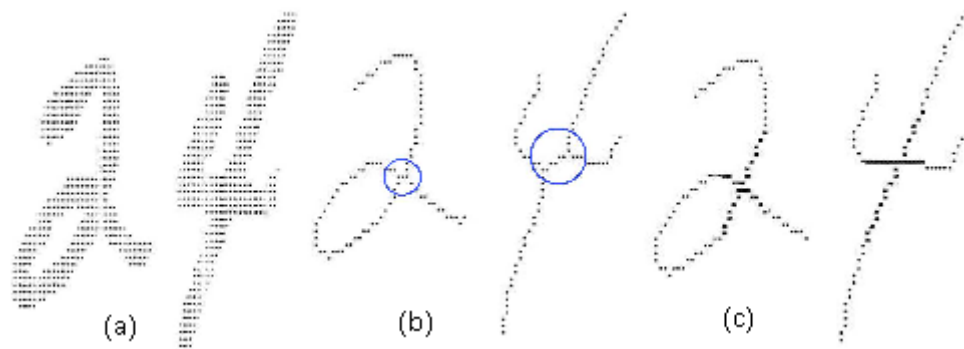


Figure I.6 : Correction du squelette afin de mieux représenter les intersections.

(a) image initiale. (b) squelette initial. (c) squelette corrigé.

- **Normalisation**

La normalisation a pour objectif de rendre l'écriture la plus indépendante possible du scripteur.

Trois techniques de normalisation sont présentées :

- la correction de l'inclinaison des lignes : skew.
- la correction de l'inclinaison des lettres : slant.
- la normalisation en hauteur.

Cette phase de normalisation a pour objectif de faire en sorte que l'extraction de caractéristiques soit la moins perturbée possible par la variabilité des styles d'écriture.

- **Correction de l'inclinaison des lignes**

Cette application a pour objet de redresser horizontalement les lignes d'écriture oblique.

Différents procédés existent, les plus appliqués sont :

- La transformée de HOUGH (TH) : c'est une technique, qui permet de détecter dans des images la présence de courbes ayant une forme paramétrique (droite, cercle, etc.). Elle est appliquée sur le contour de l'image binaire d'un mot manuscrit (appliquée sur le centre de gravité des composantes connexes), afin de détecter sa

ligne de base. Cette transformation permet de transformer les points caractéristiques de l'image vers l'espace de Hough [9][23].

➤ Les histogrammes de projection ; la méthode des histogrammes consiste à parcourir l'image selon des directions  $d$  proches de l'horizontale, et à compter le nombre de pixels noirs selon ces directions pour chaque ligne. La qualité de l'histogramme, ou sa pertinence, est estimée par son entropie. La direction la plus probable est celle qui maximise l'entropie. L'histogramme d'entropie maximale est celui dont les extremas sont les plus marqués. L'angle du document (ou de la ligne)  $\theta$  est celui qui correspond à l'histogramme d'entropie maximale. Pour corriger cette inclinaison, il suffit d'appliquer une rotation de l'image d'angle  $\theta$  [23] :

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (1)$$

- **Correction de l'inclinaison des lettres (chiffres)**

Il est parfois nécessaire de redresser les caractères lorsque l'écriture est penchée, leur direction générale forme un angle avec la verticale. Il convient de corriger cette inclinaison de l'écriture pour la rendre la plus indépendante possible des spécificités d'écriture du scripteur. Plusieurs approches ont été proposées [23] :

➤ dans l'utilisation des chaînes de Markov, le principe consiste à estimer l'orientation du texte en ne considérant que les ascendants ou descendants suffisamment grands compris entre deux lignes verticales reliées aux bords supérieur et inférieur de l'image (notion de fenêtre qui se déplace de gauche à droite dans le temps).

➤ R. M. Bozinovic et S. N. Srihari utilisent des portions d'écriture proches de la direction verticale pour évaluer l'inclinaison.

➤ d'autres travaux utilisent des histogrammes de projection, de la même manière que pour la correction du skew, mais cette fois dans des directions proches de la

verticale. Là encore, l'histogramme d'entropie maximale est celui dont les extremas sont les plus marqués.

- **Lissage du contour**

Le lissage du contour est appliqué sur les images bruitées. Ces bruits qui se traduisent par la présence d'irrégularités le long des contours des lettres peuvent engendrer des troubles dans la reconnaissance.

Des travaux simples et rapides ont été faits pour réduire l'influence de ces bruits.

=	=	=	X	=	=	=	=	X	X	X
=	P	=	X	P	=	=	P	X	=	=
X	X	X	X	=	=	=	=	=	=	=

Tableau I.1 : Exemple de masques de filtrage permettant de lisser une image.

Le principe qu'a proposé M. Chariet et al. [23], se base sur quatre masques simples, le pixel 'P' prend la valeur des pixels '='. Les règles sont appliquées sur les pixels de forme ou les pixels de fond.

### **I.5.2- Segmentation de séquences des chiffres**

Le but de la segmentation est de proposer, soit le découpage correct en chiffres, soit une liste d'options de segmentation dont on espère trouver la segmentation correcte. Les chiffres peuvent se présenter isolés, déformés ou liés par un ou plusieurs points de contact (Figure I.7).

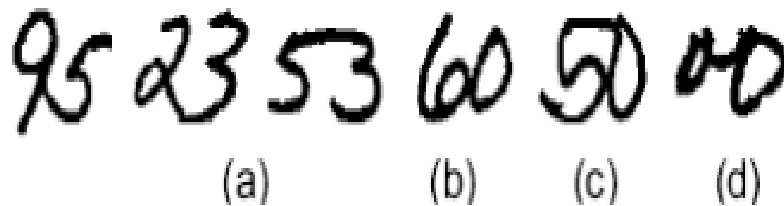


Figure I.7 : Types de contacts: (a) Un seul point de contact. (b) Une portion de trait en commun. (c) Doublement connectés. (d) Une ligature.

De nombreux algorithmes ont été proposés pour estimer la segmentation de couples de chiffres liés. Certains travaux sont basés sur les pixels de la forme, d'autres s'appuient sur les pixels du fond. D'autres encore combinent les deux approches, en exploitant à la fois les squelettes du fond et de la forme, ce qui leur permet d'obtenir de meilleures performances. Ces algorithmes permettent dans certains cas de supprimer les ligatures superflues, qui pourraient perturber la reconnaissance (figure I.8) [23].

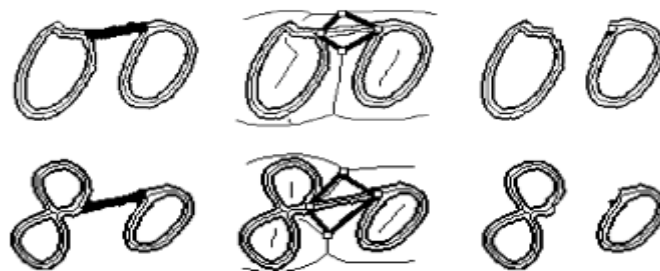


Figure I.8 : Segmentation, à l'aide des squelettes, de la forme et du fond.

D'autres travaux s'appuient sur le nombre, la taille, le centre de gravité et la position des réservoirs d'eau pour déduire des points de coupure potentiels, en incluant un grand nombre d'heuristiques [23].

C. Renaudin et al., couplent leur système avec un reconnaisseur. Et comme le montre la figure I.10, cette proposition aide à résoudre certains problèmes complexes



comme l'imbrication partielle des caractères. Dans cette approche les données sont générées de façon pseudo-aléatoire à partir d'images de chiffres isolés de la base IRON/OFF. Ces exemples ne prennent donc pas en compte les ligatures (traits supplémentaires provoqués par l'absence de levé de stylo) [23].

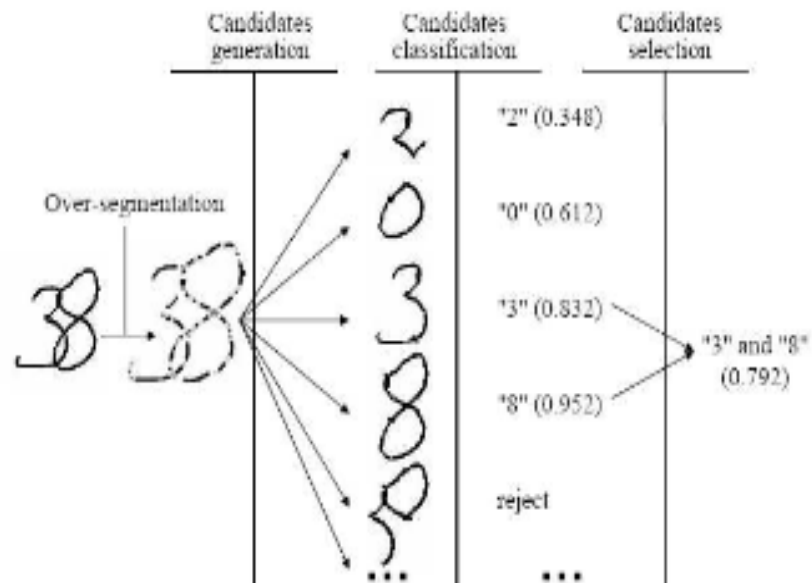


Figure I.9 : La meilleure segmentation est celle pour laquelle le produit des probabilités des symboles obtenus est maximal.

D'autres approches ont, également, été proposées. Elles sont basées sur la segmentation implicite par fenêtre glissante (figure I.10) [23].

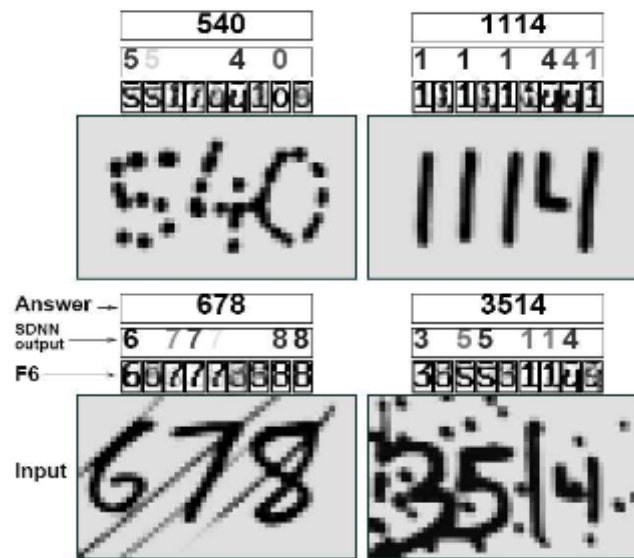


Figure I.10 : Système de reconnaissance à fenêtre glissante de type SDNN  
(Space Displacement Neural. Network).

### I.5.3- Classification

La phase de classification consiste à interpréter les données obtenues et à décider à quelle classe appartient le chiffre analysé. Cette phase regroupe les deux tâches d'apprentissage et de décision. En effet, à partir de la même description de la forme en paramètres, elles tentent, toutes les deux d'attribuer cette forme à un modèle de référence. Le résultat de l'apprentissage est soit la réorganisation ou le renforcement des modèles existants en tenant compte de l'apport de la nouvelle forme, soit la création d'un nouveau modèle de l'apprentissage. Le résultat de la décision est un avis sur l'appartenance ou non de la forme aux modèles de l'apprentissage [14][25].

Généralement, un système de reconnaissance est construit à partir des connaissances à priori sur le problème. Le fait de classer un objet correspond donc à prendre une décision sur base d'une ou plusieurs règles. Dès lors, une des premières approches pour automatiser le traitement, fut d'extraire la connaissance experte de spécialistes sous forme de règles. Ainsi, pour chaque catégorie, on disposait d'un ensemble de règles permettant de déterminer l'appartenance d'un objet à une classe.

Bien que cette façon ait été largement utilisée, on en conçoit aisément les limites. En effet, quand le nombre d'exemples et de classes est très important, il devient difficile de rassembler toutes les connaissances expertes nécessaires aux prises de décision. Ainsi, d'autres travaux se sont tournés vers l'approche Machine Learning (ML) ou apprentissage machine (automatique) qui est devenu très populaire.

L'apprentissage est une étape clé dans la machine de reconnaissance. Son rôle est d'éclairer la décision à l'aide des connaissances à priori sur les formes. A partir de critères spécifiques aux formes, l'apprentissage tente de définir des modèles de référence ou de caractère des classes de décision.

Cela exige de bien définir les classes de manière à distinguer les familles homogènes de formes et donc à identifier ces dernières. On distingue deux types d'apprentissage :

- ***L'apprentissage supervisé***

Il est dit supervisé lorsque les différentes familles de formes sont connues à priori et la tâche d'apprentissage est guidée par un professeur. Parmi les nombreuses méthodes supervisées qui ont été développées dans le ML distinguons les réseaux de neurones et les machines à vecteur de support (SVM) que nous avons adopté.

- ***L'apprentissage non supervisé***

Dans ce cas, les données ne sont pas classées à priori. C'est à partir d'échantillons de référence et de règle de regroupement qu'on construit, automatiquement, les classes ou les modèles de référence, sans l'intervention de l'opérateur ou de professeur. Parmi les nombreuses méthodes développées citons les Nuées dynamiques (ou K-means).

**I.5.4- Post-traitement**

Cette étape aide à réduire considérablement des erreurs. Elle permet la validation des décisions de l'analyse sur la base de connaissances. Cependant, ce n'est pas une étape complètement séparée des autres étapes [7][25].

**I.6- Conclusion**

Nous avons évoqué dans ce chapitre les généralités concernant la reconnaissance des formes. Nous nous sommes arrêtées sur la reconnaissance de l'écriture manuscrite et précisément, les chiffres manuscrits.

Dans ce qui suit, nous allons développer les étapes qui ont mené à l'analyse en ondelettes.

# **CHAPITRE II :**

## **Théorie des ondelettes**

## **II.1- Introduction**

La plupart des signaux du monde ne sont pas stationnaires. C'est justement dans l'évolution de leurs caractéristiques (statistiques, fréquentielles, temporelles, spatiales) que réside l'essentiel de l'information qu'ils contiennent. Les signaux vocaux et les images sont, à ce titre, exemplaires.

Dans l'analyse des objets (images ou signaux), on désire souvent extraire des informations de façon à ce qu'elles soient facilement accessibles. Dans le traitement d'image, pour aboutir à une classification capable de bien discriminer les différentes classes, on désire une représentation qui permet une extraction d'indices capable de minimiser les confusions entre les différentes classes.

Cela dit, l'analyse de Fourier propose une approche globale du signal, donc toute notion de localisation temporelle (ou spatiale) disparaît dans l'espace de Fourier. Il a donc fallu développer de nouveaux outils mathématiques qui permettent de traiter les signaux et d'en extraire facilement l'information utile. C'est là que la transformée en ondelettes entre en scène. La démarche suivie ci-dessous permet de mieux comprendre comment cet outil a été élaboré

## **II.2- Analyse de Fourier**

### **II.2.1- Transformée de Fourier**

La transformée de Fourier (TF) analyse le contenu fréquentiel d'un signal. Ses nombreuses propriétés la rendent adaptée à l'étude des opérateurs linéaires stationnaires. C'est une représentation globale du signal dans le domaine fréquentiel[6].

La transformée de Fourier d'une fonction  $f \in L^2(\mathbb{R})$  est:

$$F(\omega) = \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} dt \quad (1)$$

$\omega = 2\pi\nu$   $\omega$  et  $\nu$  étant respectivement la pulsation et la fréquence du signal  $f(t)$ .

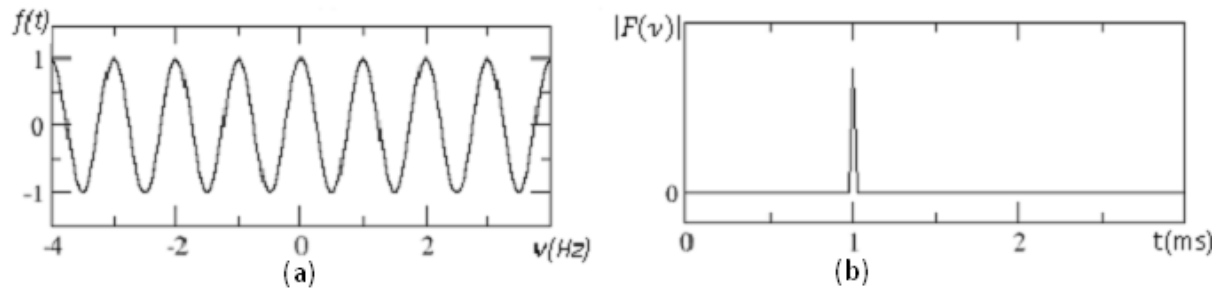


Figure II.1 : (a) Signal périodique. (b) Représentation de la TF de (a).

Les figures II.1, II.2 représentent un signal stationnaire, un signal non-stationnaire et leurs TF respectivement. Il s'est ainsi avéré que la TF soit peu efficace sauf pour les signaux stationnaires : signaux dont le contenu de fréquence ne change pas au cours du temps.

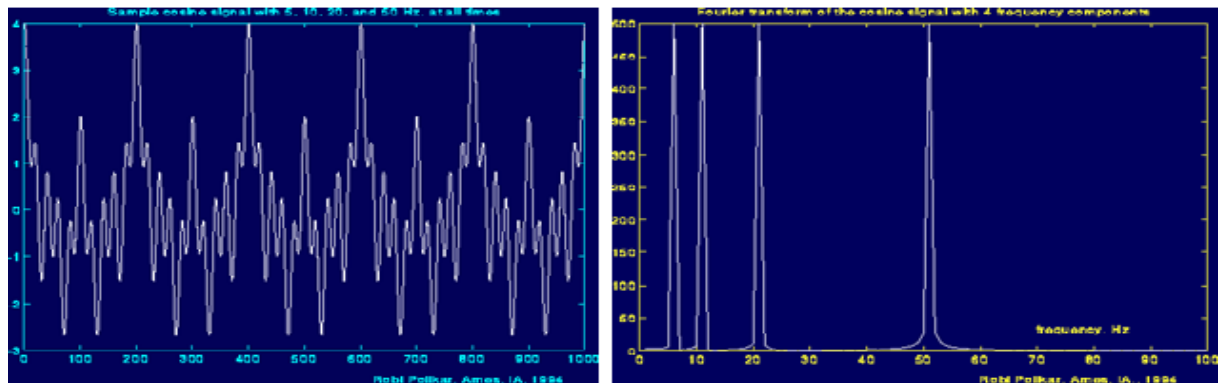


Figure II.2 : Signal stationnaire et sa transformée de Fourier.

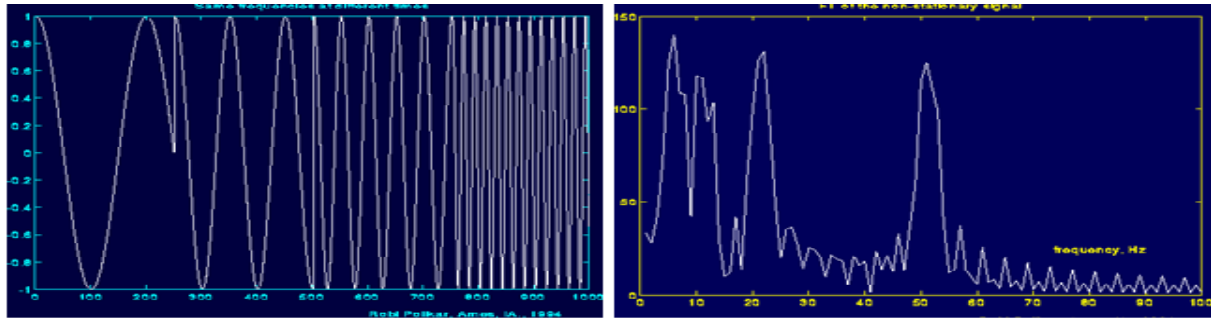


Figure II.3 : Signal non stationnaire et sa transformée de Fourier.

L'intégrale s'effectue sur toutes les valeurs du temps  $]-\infty, +\infty[$  peu importe où la composante (fréquence) apparaît, le résultat n'en sera pas affecté. Les fréquences mesurées ne peuvent être situées dans le temps. La correspondance temps-fréquence n'existe pas, c'est-à-dire deux signaux différents peuvent avoir la même transformée de Fourier (figure II.4) [13].

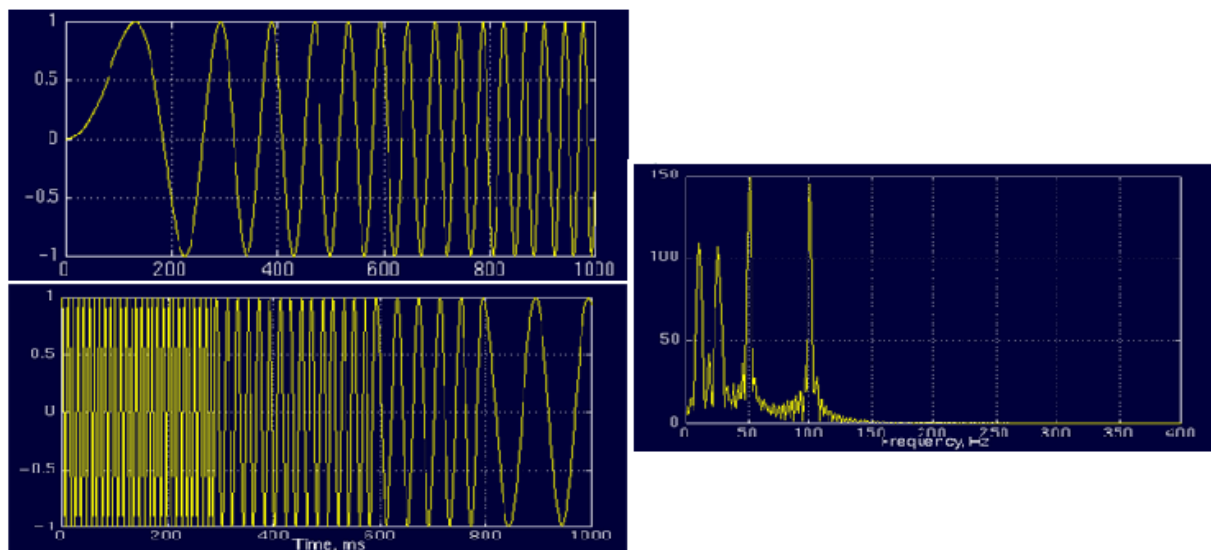


Figure II.4 : Deux signaux peuvent avoir la même transformée de Fourier.

L'idée de l'amélioration de la transformée de Fourier était de modifier la base de fonctions analysantes de manière à être capable de localiser les éventuelles non-stationnarités du signal. Ceci en modifiant la transformée de Fourier pour lui donner ce pouvoir de localisation, c'est l'idée de la transformation de Fourier à fenêtre [6].



### II.2.2- Transformée de Fourier à fenêtre glissante

La transformée de Fourier fenêtrée remplace la sinusoïde de la transformée de Fourier par le produit d'une sinusoïde et d'une fenêtre localisée en temps (figure II.5). Elle a deux arguments : le temps et la fréquence. Elle est définie par :

$$F(u, \omega) = \int_{-\infty}^{+\infty} f(t)g(t-u)e^{-j\omega t}dt \quad (2)$$

Où  $g(t-u)$  est une fenêtre centrée sur  $u$ . Alors que la transformation de Fourier utilise les fonctions élémentaires  $e^{-j\omega t}$  pour décomposer le signal.

La transformée de Fourier à fenêtre glissante utilise les fonctions  $g(t-u)e^{-j\omega t}$  pour décomposer le signal dans un plan à deux dimensions : le plan temps-fréquence.

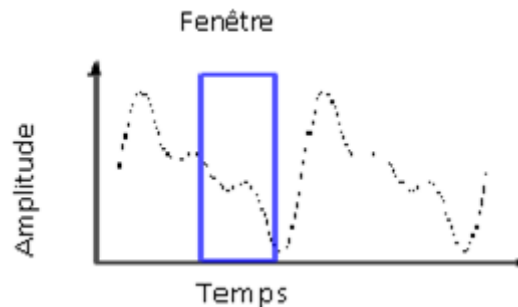


Figure II.5 : Transformée de Fourier à fenêtre.

L'inconvénient de la transformation de Fourier à fenêtre glissante est lié au fait que la taille de la fenêtre est fixe et ne dépend donc pas des fréquences analysées. Par conséquent, la résolution de l'analyse est la même en tout lieu de l'espace temps-fréquence.

L'idée dans la théorie des ondelettes est d'adapter la résolution temporelle aux fréquences analysées. Ainsi, au lieu de chercher à analyser le signal utile avec une fenêtre de taille fixée et contenant un nombre d'oscillations qui varie avec la fréquence,

la fenêtre utilisée est de taille dépendante de la fréquence analysée mais avec un nombre d'oscillations fixe. Ces fonctions analysantes sont précisément les ondelettes qui permettent à la fois :

- d'observer les hautes fréquences avec une haute résolution temporelle et donc de fournir une des informations précises sur la localisation des phénomènes brefs.
- et d'observer les basses fréquences sur une durée suffisante d'analyse pour rendre compte des phénomènes lents [18].

## II.3- Transformée en ondelettes

### II.3.1- Définition

Pour avoir une représentation temps-fréquence adapté à la fois pour les grandes et basses variations, on utilise les ondelettes. Une ondelette est une fonction continue qui disparaît au voisinage de plus ou moins l'infini. On dit que les ondelettes sont des ondes localisées. Cette propriété de localisation lui permet d'agir localement. On la désigne souvent par la lettre  $\Psi$ . La forme générale de l'ondelette est donnée par :

$$\Psi_{u,s}(t) = \frac{1}{\sqrt{s}} \Psi\left(\frac{t-u}{s}\right); u \in \mathbb{R}^+ \text{ et } s \in \mathbb{R}^{+*} \quad (3)$$

Deux opérations importantes appliquées à la fonction d'ondelette permettent de générer une infinité de variations de celle-ci : la transformation qui correspond au déplacement de l'ondelette le long de l'axe des x et la dilatation. Le paramètre  $u$  est relatif à la translation, tandis que  $s$  est relatif à la dilatation. Ainsi, chaque élément de la base  $\{\Psi_{u,s}\}$  sera obtenu par translation et dilatation d'une ondelette mère  $\Psi(t)$  [6]. La figure II.6 représente une ondelette mère de Morlet après dilatation ou compression.

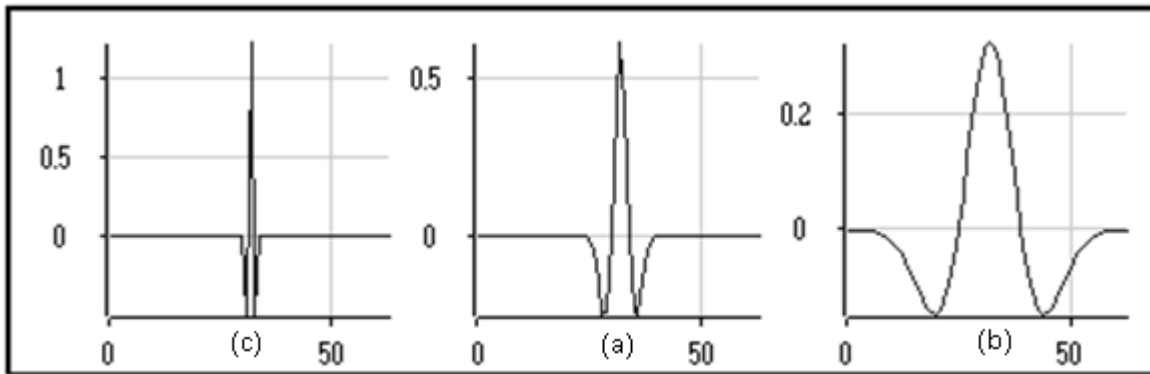


Figure II.6 : a : Ondelette mère (de Morlet), b : après dilatation, c: après compression.

L'ondelette est une fonction localisée aussi bien en temps qu'en fréquence. Sa condition d'admissibilité dans le domaine fréquentiel est la suivante :

$$C_{\Psi} = \int_{-\infty}^{+\infty} \frac{|\hat{\Psi}(\omega)|}{|\omega|} d\omega < \infty \quad (4)$$

Où  $\hat{\Psi}(\omega)$  est la transformée de Fourier de  $\Psi(t)$ .

En pratique, on utilise souvent une condition plus forte en imposant à l'ondelette un nombre  $p$  de moment nul :

$$M(p) = \int t^p \Psi(t) dt = 0 \quad (5)$$

### II.3.2- Transformée en ondelettes continue (CWT)

Une transformée en ondelettes d'une fonction  $f \in L^2(\mathbb{R})$  est dite continue lorsque les paramètres structurels des fonctions utilisées (les translations et dilatations) peuvent prendre n'importe quelle valeur réelle (les dilatations peuvent néanmoins être positives). Lorsqu'on calcule la transformée en ondelettes continue, on fait glisser de manière continue, une fonction de base le long de l'axe réel et on calcule sa convolution avec la fonction  $f(t)$ . Les coefficients ainsi obtenus sont alors très corrélés

entre eux et donc redondants. L'expression de la transformée en ondelettes est donnée par :

$$WT_f(u, s) = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{s}} \overline{\Psi}\left(\frac{t-u}{s}\right) dt \quad (6)$$

$\overline{\Psi}$  : le complexe conjugué  $\Psi$

$WT_f(u, s)$  : est un ensemble d'éléments représentant les contributions locales du signal à des instants et échelles divers qui dépendent du choix des paramètres  $u$  et  $s$ .

Vue la continuité des paramètres de translation et de dilatation, l'information contenue dans la transformée en ondelettes est extrêmement redondante, ce qui engendre une transformée en ondelettes inverse qui n'est pas toujours unique [6].

### II.3.3- Transformée en ondelettes discrète (DWT)

L'utilisation de la transformée en ondelettes discrète permet de réduire fortement la redondance. Les ondelettes discrètes sont obtenues en ne permettant que des translations avec des pas discrets :

$$\Psi_{j,k}(t) = \frac{1}{\sqrt{s_0^j}} \Psi\left(\frac{t - k u_0 s_0^j}{s_0^j}\right) \quad (7)$$

Où  $j$  et  $k$  sont entiers. On en choisit généralement  $s_0 = 2$  et  $u_0 = 1$ .

Dans la pratique, la transformée en ondelettes discrète est calculée par application de filtre passe-bas et passe-haut à l'image, suivis du sous-échantillonnage par 2, c'est-à-dire en omettant une valeur sur 2. Ceci est effectué sur les lignes et les colonnes. L'opération de sous échantillonnage est également appelée la décimation [6].

## II.4- Analyse multi résolution (AMR)[6]

Elle consiste à représenter un signal comme une limite de ses approximations successives, où chaque approximation est une version lissée de la précédente. Les approximations successives sont présentées à différentes résolutions d'où le nom multi résolution. Il suffit alors de passer d'un espace à un autre pour changer de résolution, et modifier ainsi la taille de l'objet manipulé.

Les ondelettes permettent de représenter les détails gagnés lors du passage d'une résolution à la résolution suivante plus fine.

Mallat et Mayer ont défini la notion d'analyse multi résolution comme suit :

« Une suite  $\{V_j\}_{j \in \mathbb{Z}}$  de sous-espaces vectoriels de  $L^2(\mathbb{R})$  est une approximation multi résolution si elle vérifie les propriétés suivantes :

$$1-\forall (j,k) \in \mathbb{Z}^2, f \in V_j \Leftrightarrow f(\cdot - 2^j k) \in V_j$$

$$2-\forall j \in \mathbb{Z}, V_{j+1} \subset V_j$$

$$3-\forall j \in \mathbb{Z}, f \in V_j \Leftrightarrow f\left(\frac{\cdot}{2}\right) \in V_{j+1}$$

$$4-\lim_{j \rightarrow -\infty} V_j = \bigcap_{j \in \mathbb{Z}} V_j = 0$$

$$5-\lim_{j \rightarrow \infty} V_j = \overline{\bigcup_{j \in \mathbb{Z}} V_j} = L^2(\mathbb{R})$$

$$6-\exists \emptyset: \{\emptyset(\cdot - n)\}_{n \in \mathbb{Z}} \text{ forme une base orthonormée de } V_0. \gg.$$

### II.4.1- Fonction d'échelle

On construit une analyse multi résolution à l'aide de sous-espaces  $V_j$  emboîtés les uns dans les autres, tel que le passage de l'un à l'autre soit le résultat d'un changement d'échelle.

Pour réaliser les approximations d'une fonction  $f \in L^2$ , on fait appel à une fonction d'échelle ou père d'ondelette  $\Phi(x) \in L^2(\mathbb{R})$  telle que :

$$\Phi_{j,n}(x) = 2^{-\frac{j}{2}} \Phi(2^{-j}x - n); (j, n) \in \mathbb{Z}^2 \quad (8)$$

La fonction  $\Phi$  constitue une base orthogonale pour  $V_j$ .

Rappelons que le produit scalaire de deux fonctions réelles ou complexes est défini par :

$$\langle f, g \rangle = \int_{-\infty}^{+\infty} f(x)g(x)dx \quad (9)$$

L'approximation de  $f$  avec une résolution de  $j$  correspond à sa projection dans l'espace  $V_j$  :

$$\forall f \in L^2: A_j f = \text{proj}_{V_j}(f) = \sum \langle f(x), \Phi_{j,n} \rangle \Phi_{j,n} \quad (10)$$

On pose :  $a(j, n) = \langle f, \Phi_{j,n} \rangle$

Ainsi, l'approximation à la résolution  $j$  de la fonction  $f$  sera définie par la suite discrète des nombres réels ou complexes  $a(j, n)$ .

#### II.4.2- Fonction d'ondelette et espace de détail

L'espace de détail vient compléter l'analyse. En effet, on peut définir pour chaque espace  $V_j$  son complément orthogonal  $W_j$  dans  $V_{j-1}$  tel que :

- $V_{j-1} = \overline{V_j \oplus W_j}$  et  $V_j \perp W_j$
- $\forall k \neq j$  alors  $W_j \perp W_k$

$$L^2(\mathbb{R}) = \bigoplus_{j \in \mathbb{Z}} W_j$$

Les sous espaces  $W_j$  ne forment pas des sous espaces emboîtés.

Pour  $j$  fixe, les coefficients  $(D_{j,n}(f) = \langle f, \Psi_{j,n} \rangle_{n \in \mathbb{Z}})$  représentent la différence entre deux approximations de  $f$  à des résolutions différentes.

### II.4.3- Fonctions d'échelle et ondelettes à deux dimensions

La méthode la plus simple pour engendrer des bases d'ondelettes, en dimension 2, consiste à faire des produits tensoriels sur les ondelettes et des fonctions d'échelle en dimension 1. La fonction d'échelle 2D obtenue est  $\Phi(x, y) = \Phi(x)\Phi(y)$  avec  $\Phi(x)$  la fonction d'échelle monodimensionnelle [15].

Soit  $\Psi(x)$  l'ondelette associée à  $\Phi(x)$ , les trois ondelettes sont données par :

$$\begin{aligned}\Psi^1(x, y) &= \Phi(x)\Psi(y) \\ \Psi^2(x, y) &= \Psi(x)\Phi(y) \\ \Psi^3(x, y) &= \Psi(x)\Psi(y)\end{aligned}\tag{11}$$

Avec l'analyse multi résolution séparable, une grande importance est donnée aux directions horizontale et verticale dans l'image. Pour de nombreux types d'analyses, cette sélectivité est utile. En pratique, la transformée est calculée en appliquant un banc de filtres à l'image (Figure II.7).

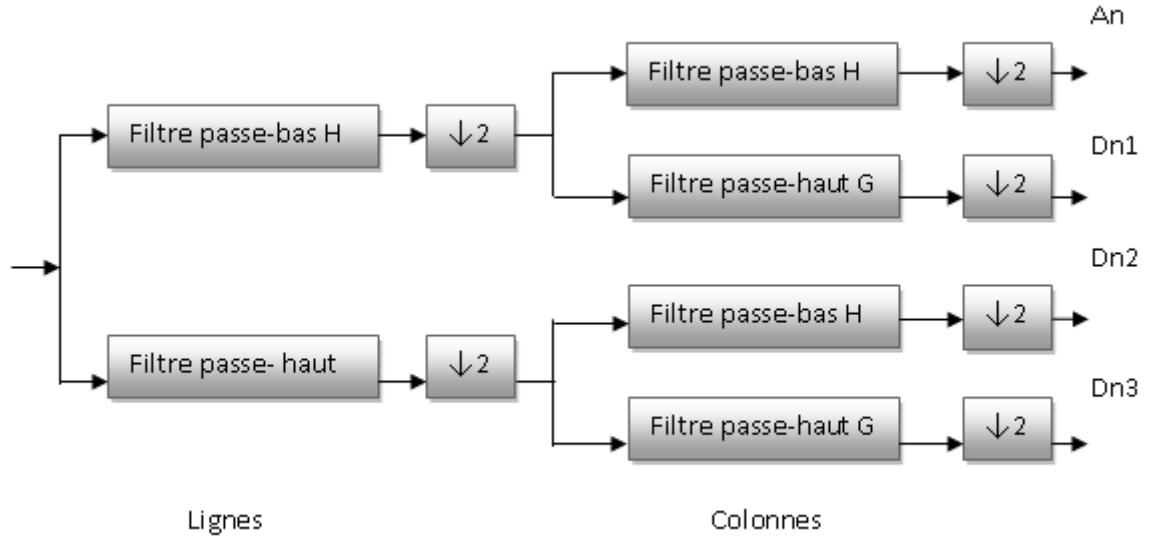


Figure II.7 : Banc de filtres à quatre sorties.

$$\begin{aligned}
 A_n(b_i, b_j) &= [H_x * [H_y * A_{n-1}] \downarrow_{2,1}] \downarrow_{1,2} (b_i, b_j) \\
 D_{n1}(b_i, b_j) &= [H_x * [G_y, A_{n-1}] \downarrow_{2,1}] \downarrow_{1,2} (b_i, b_j) \\
 D_{n2}(b_i, b_j) &= [G_x * [H_y * A_{n-1}] \downarrow_{2,1}] \downarrow_{1,2} (b_i, b_j) \\
 D_{n3}(b_i, b_j) &= [G_x * [G_y * A_{n-1}] \downarrow_{2,1}] \downarrow_{1,2} (b_i, b_j)
 \end{aligned} \tag{12}$$

Avec : ‘\*’ est l’opérateur de convolution,  $\downarrow_{2,1} (\downarrow_{1,2})$  représente le sous échantillonnage le long des lignes (des colonnes),  $A_0 = I(x, y)$  est l’image d’origine,  $b_i$  est la fréquence spatiale.

- Les filtres H et G sont des filtres passe-bas et passe-haut, respectivement.

$A_n$  est obtenu par filtrage passe bas, et est donc l’image à basse résolution à l’échelle  $n$ .



Les coefficients  $D_{ni}$  sont obtenus par un filtrage passe-bande dans une direction spécifique ; par conséquent, ils contiennent les informations de détails.

L'image d'origine  $I$  est donc représentée par un ensemble de sous-images à différentes échelles :  $\{A_n, D_{ni}\}, i = 1, 2, 3, \text{ et } n = 1, \dots, d$ . Ceci est la représentation multi échelles de profondeur  $d$  de l'image  $I$ .

La notion de multi résolution nous permet de considérer l'image sur plusieurs échelles. Les échelles consécutives sont séparées par un facteur de résolution 2 (Figure II.8, II.9 et II.10).

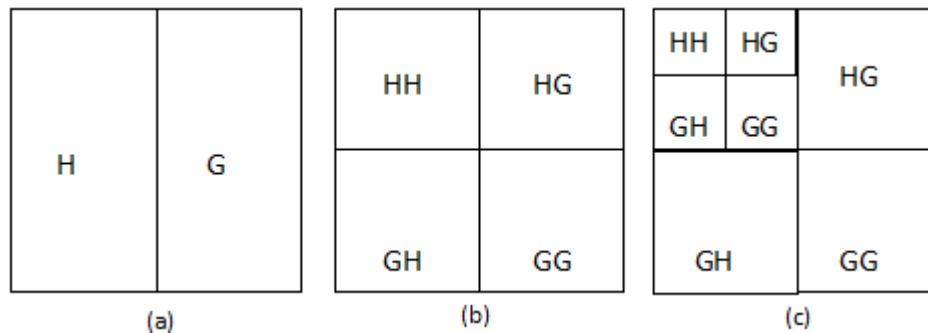


Figure II.8 : Organisation des sous-images de détails et de l'image lissée à différentes résolutions.

Dans le cadre de l'analyse des images, la transformée en ondelettes sera, essentiellement, utilisée par le biais de filtres séparables [16]. Effectuer une transformée en ondelettes d'une image revient, donc, à la convoluer avec un filtre obtenu par discrétisation d'une ondelette. La séparabilité du filtre permet de traiter les lignes puis les colonnes de l'image sans en perdre la vision globale.

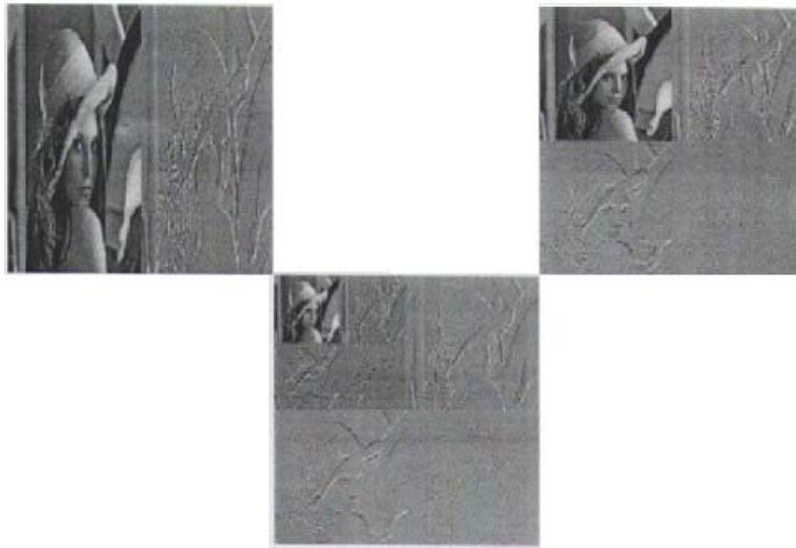


Figure II.9 : Exemple de décomposition d'une image.

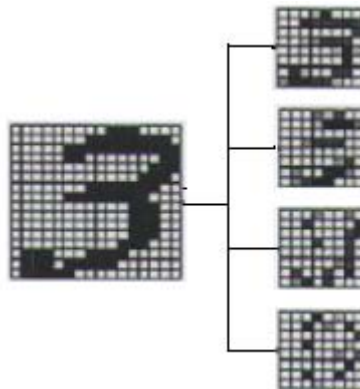


Figure II.10 : Exemple de décomposition d'un chiffre  
à la première résolution.

#### II.4.4- Présentation de l'algorithme pyramidal

Nous avons vu que l'analyse multi résolution d'un signal revient à le décomposer à différentes échelles, en approximations et en détails.

S. Mallat propose un algorithme rapide permettant de calculer les coefficients de détails et d'approximations en utilisant des filtrages et décimations successifs.

Les coefficients de détails correspondant à l'espace  $W_l$  sont obtenus par filtrage passe haut (filtrage  $h_1$ ) puis décimation par 2. Les approximations sont obtenues de la même manière par filtrage passe bas (filtrage  $g_1$ ). Pour obtenir les coefficients de détails aux résolutions supérieures, il suffit de retirer ces étapes sur les coefficients d'approximations.

## II.5- Exemples d'ondelettes pratiques

### II.5.1-Laplaciens itérés de Gaussienne [15]

Couramment appelé le chapeau mexicain (figure II.11), son expression est donnée par l'équation :

$$\psi(x) = \left( \frac{2}{\sqrt{3}} \pi^{-\frac{1}{4}} \right) (1 - x^2) e^{-x^2/2} \quad (13)$$

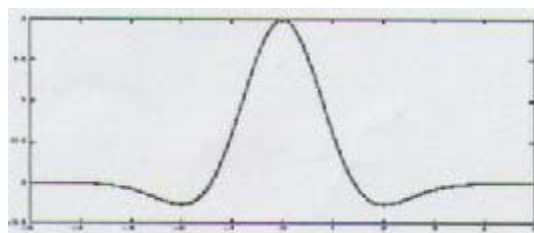


Figure II.11 : Ondelette  $h(x,y)$  (chapeau mexicain).

### II.5.2-Ondelettes de Haar [16]

L'analyse de Haar est parmi les bases d'ondelettes orthonormées les plus populaires (figure II.12). Elle est appréciée en traitement d'image pour ses filtres

linéaires en phase qui ont donc des propriétés de symétrie. Ce qui correspond à la fonction :

$$\Psi(\pi t) = \begin{cases} 1 & 0 \leq t \leq \frac{1}{2} \\ -1 & \frac{1}{2} \leq t \leq 1 \\ 0 & \text{sinon} \end{cases} \quad (14)$$

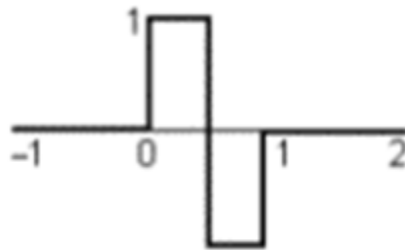


Figure II.12 : Ondelette mère de l'analyse de Haar.

L'inconvénient majeur de cette analyse est sa très mauvaise résolution fréquentielle. Son intérêt majeur est la grande simplicité de l'algorithme de calcul.

### II.5.3- Ondelettes de Daubechies [17]

La mathématicienne Ingrid Daubechies a cherché dans ses travaux à concilier deux contraintes restrictives : l'orthogonalité de la base d'ondelettes et la compacité du support de l'ondelette-mère (ce qui implique que toute ondelette de la base est à support compact). De plus, elle a imposé à ses ondelettes une troisième condition : avoir n moments nuls.

Inconvénients : l'abandon de la linéarité en phase ou plus simplement de la symétrie droite-gauche très recherchée par les traiteurs d'images. Dans un certain

nombre de cas, l'abandon de cette symétrie n'a pas des conséquences dramatiques et ces bases ont des applications tout à fait intéressantes en analyse des images.

L'avantage de compacité des fonctions de base et surtout des filtres associés est l'économie en cout de calcul qu'elle procure. Le filtrage peut être effectué, directement, par l'application de masques de convolution. Des applications temps réel sont facilement envisageables même si dans certains cas la précision requise sur les coefficients réels pèse malgré tout sur l'architecteur. La figure II.13 nous montre une représentation graphique d'une ondelette de Daubechies.

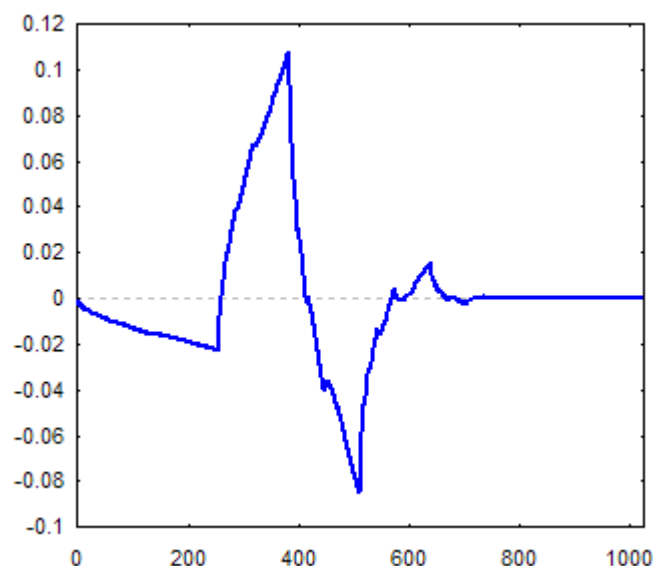


Figure II.13 : Ondelette de Daubechies.

#### II.5.4- Les symmlets

Daubechies a construit des ondelettes à support compact les plus symétriques possibles. En effet, il n'existe pas d'ondelettes à support compact dans une AMR orthogonale qui soient symétriques, exceptées l'ondelette de Haar qui est antisymétrique. La figure II.13 montre la représentation graphique de l'ondelette symmlet 4.

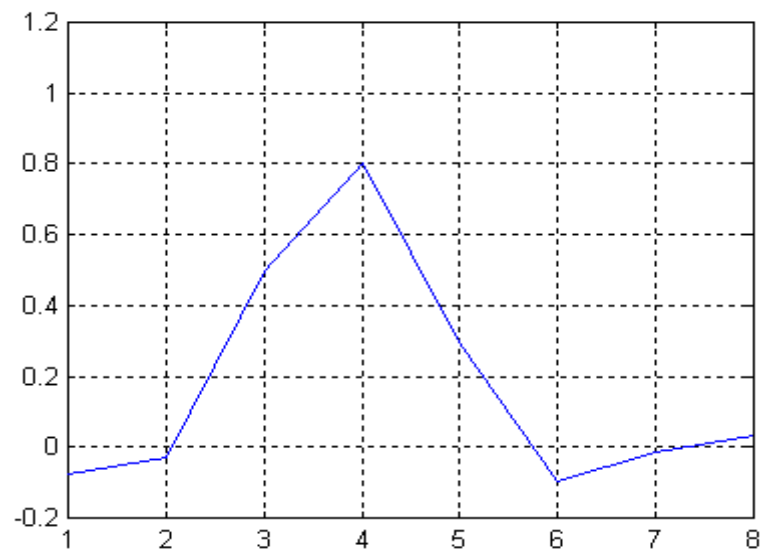


Figure II.14 : Ondelette symmlet 4.

### II.5.5- Les coiflets

Coifman suggéra à Daubechies de construire une famille telle que la fonction d'échelle ait, elle aussi, des moments nuls. La figure II.14 nous montre une représentation graphique de l'ondelette Coiflet 2.

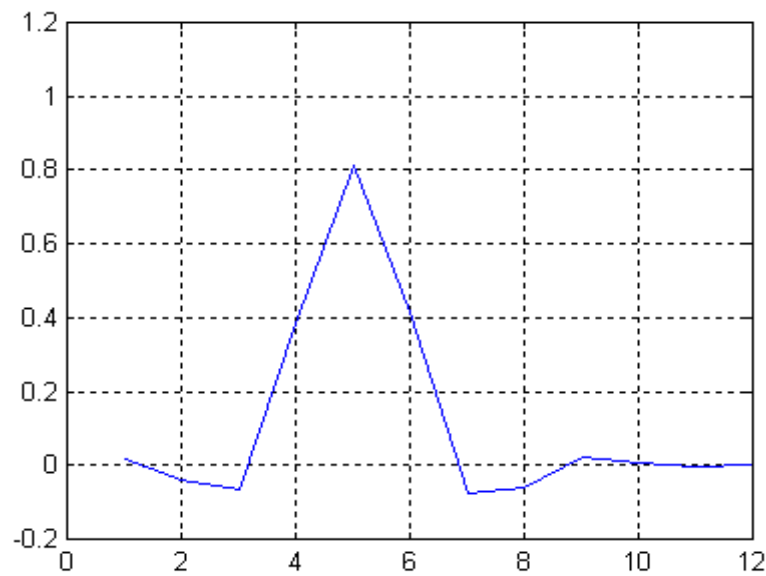


Figure II.15 : Ondelette Coiflet 2.

## II.6- Conclusion

Dans ce chapitre, nous sommes brièvement passées par les différentes étapes qui ont mené à la création des ondelettes et sur ces dernières, nous avons suivi une analyse plus ou moins détaillée sur la théorie des ondelettes, en commençant par la transformée en ondelettes continue puis la transformée en ondelettes discrète et enfin l'analyse multi résolution.

Les ondelettes ont montré leur robustesse dans différentes applications. Cette constatation, nous a incité à les introduire dans notre système de reconnaissance de chiffres manuscrits SVM. Le prochain chapitre abordera les machines à vecteur de support (SVM).

**CHAPITRE III :**  
**Les Machines à**  
**Vecteurs de Supports**  
**(SVM)**



### **III.1- Introduction**

L'origine des machines à vecteurs de support (SVM) remonte à 1975 lorsque Vapnik et Chervonenkis proposèrent le principe du risque structurel et la dimension VC pour caractériser la capacité d'une machine d'apprentissage. A cette époque, ce principe n'a pas trouvé place et il n'existait pas encore un modèle de classification solidement appréhendé pour être utilisable. Il a fallu attendre jusqu'à l'an 1982 pour que Vapnik propose un premier classificateur basé sur la minimisation du risque structurel baptisé SVM . Ce modèle était toutefois linéaire et l'on ne connaissait pas encore le moyen d'induire des frontières de décision non linéaires. En 1992, Boser et al. proposent d'introduire des noyaux non-linéaires pour étendre le SVM au cas non-linéaire . En 1995, Cortes et al. proposent une version régularisée du SVM qui tolère les erreurs d'apprentissage tout en les pénalisant [3].

### **III.2- Principe de la méthode des Machines à Vecteurs de Support (SVM)**

A l'origine, la méthode des machines à vecteurs de supports était destinée à la classification binaire. Elle a été présentée dans plusieurs travaux [22],[20]. D'une manière simple, soient  $x_1, x_2, \dots, x_n$  un ensemble d'observations décrites par des attributs prédéfinis. Soient  $C1$  et  $C2$  deux classes possibles. Étant donné un étiquetage connu des  $n$  observations en termes de classes, la méthode SVM cherche à estimer les paramètres d'un classifieur pour étiqueter une nouvelle observation. En effet, cela revient à chercher l'hyperplan optimal, celui dont la distance aux exemples d'apprentissage est maximale. Cette distance est appelée marge entre l'hyperplan et les exemples. Les points les plus proches qui suffisent à déterminer cet hyperplan sont appelés vecteurs de support. Comme on cherche à maximiser cette marge, on parlera de méthode des séparateurs à vaste marge [5] (voir figure III.1).

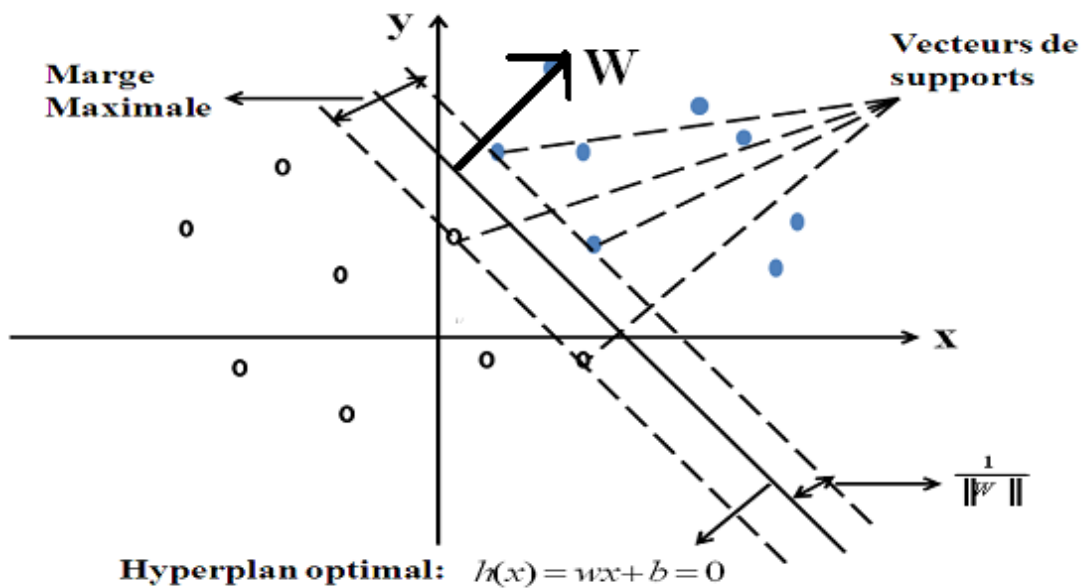


Figure III.1 : Hyperplan optimal séparant les points de deux classes.

On distingue deux types de données : données linéairement séparables et données non linéairement séparables.

### III.3- SVM linéaire

Dans le cas de données linéairement séparables qui représente un cas simple, s'il existe généralement une infinité d'hyperplans qui permettent de séparer les données d'apprentissage, tous ne permettent pas de bien généraliser, c'est-à-dire de bien classer les données à venir. L'idée consiste alors à rechercher la frontière de décision qui permettrait non seulement de bien classer l'ensemble des données d'apprentissage, mais aussi toutes les données de test. Cet hyperplan peut être considéré comme optimal dans le sens où il maximise la marge de séparation.

En pratique, un hyperplan est constitué de l'ensemble des points  $x$  appartiennent à  $R^d$ , où  $d$  est la dimension de l'observation  $x$  vérifiant :

$$w \cdot x + b = 0 \quad (1)$$

Qui est donc caractérisé par un vecteur de poids  $w \in \mathbb{R}^d$  et une valeur de biais  $b$ .

La fonction de décision associée à un hyperplan est donc :

$$f(x) = w.x + b \quad (2)$$

La règle de décision consiste alors à classer la donnée  $x$  dans la première classe, caractérisée par le label  $y = 1$ , si  $f(x) \geq 0$ , ou dans la seconde classe caractérisée par le label  $y = -1$ , si  $f(x) < 0$ .

Ainsi, étant donné un ensemble de données d'apprentissage  $x_i \in \mathbb{R}^d : i = 1, \dots, n$  et de labels  $y_i \in \{-1, 1\}, i = 1, \dots, n$ . Si les données des deux classes sont linéairement séparables, il existe au moins un hyperplan qui vérifie :

$$f(x_i)y_i > 0, i = 1, \dots, n \quad (3)$$

La recherche de l'hyperplan optimal, tel que défini précédemment, consiste donc à déterminer les valeurs de  $w$  et  $b$  qui vérifient les contraintes (3) et maximisent la marge de séparation qui correspond à la distance minimale entre les données d'apprentissage et leurs projections sur cet hyperplan.

Ce qui signifie que pour l'ensemble des vecteurs de support, la valeur de la fonction de décision  $f(x)$  est égal soit à 1, soit à -1, selon la classe du vecteur de support. En outre, la marge de séparation est alors égale à  $\frac{1}{\|w\|}$ .

Déterminer l'hyperplan optimal revient donc à résoudre un problème d'optimisation qui dans sa formulation primale consiste à minimiser :

$$\frac{1}{2} \|w\|^2 \quad (4)$$

sous les contraintes :

$$y_i (w.x_i + b) \geq 1, \quad i = 1, \dots, n \quad (5)$$

Une manière de résoudre ce type de problème d'optimisation sous contraintes est d'introduire les multiplicateurs de Lagrange :

$$\alpha_i \geq 0, \quad i = 1, \dots, n \quad (6)$$

et le Lagrangien :

$$L(w, b, \alpha) = \frac{\|w\|^2}{2} - \sum_{i=1}^n \alpha_i [y_i (w.x_i + b) - 1] \quad (7)$$

Le Lagrangien  $L$  doit alors être minimisé par rapport à  $w$  et  $b$  et maximisé par rapport aux  $\alpha_i$ . Par conséquent, au point d'équilibre, nous devons obtenir :

$$\frac{\partial L(w, b, \alpha)}{\partial w} = 0 \quad \text{et} \quad \frac{\partial L(w, b, \alpha)}{\partial b} = 0 \quad (8)$$

ce qui entraîne :

$$\sum_{i=1}^n \alpha_i y_i = 0 \quad (9)$$

$$w = \sum_{i=1}^n \alpha_i y_i x_i \quad (10)$$

La formulation duale du problème d'optimisation revient ainsi à maximiser :

$$\sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \quad (11)$$

sous les contraintes (6) et (9).

Notons par ailleurs que seuls les vecteurs de support ont des  $\alpha_i \neq 0$  et que l'hyperplan optimal est donc caractérisé par ce sous-ensemble des données d'apprentissage, ainsi que par les multiplicateurs de Lagrange associés à chacune des données de ce sous ensemble.

Même si les données d'apprentissage sont linéairement séparables, en présence de données aberrantes, la maximisation de la marge de séparation peut conduire à l'obtention d'un hyperplan qui n'est plus du tout optimal.

Pour contourner ce problème et tolérer ainsi que certaines données d'apprentissage violent les contraintes de séparation (5), des variables d'écart  $\zeta_i$ , ont été introduites dans le classifieur SVM.

$$\zeta_i \geq 0, \quad i = 1, \dots, n \quad (12)$$

et de nouvelles contraintes :

$$y_i (w \cdot x_i + b) \geq 1 - \zeta_i, \quad i = 1, \dots, n \quad (13)$$

La formulation primale du problème d'optimisation revient alors à minimiser sous les contraintes (12) et (13) la fonction objective suivante :

$$\frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \zeta_i \quad (14)$$

Où  $C$  est une constante positive, dite de régularisation. Ainsi, plus la valeur de  $C$  est

grande, moins l'algorithme tolère que les données violent les contraintes de séparation (5).

Finalement, dans le cas de la formulation duale, on maximise toujours l'expression (11) sous les contraintes (6) et (9), aux quelles on ajoute les contraintes :

$$\alpha_i \leq C, \quad i = 1, \dots, n \quad (15)$$

### **III.4- SVM non linéaire**

Dans le cas de problèmes de classification où les données sont non linéairement séparables, la recherche de l'hyperplan optimal devient difficile par méthode de SVM linéaire. Ce problème est résolu par la recherche d'une transformation non-linéaire  $\phi$  définie dans l'espace de Hilbert pour redécrire le problème dans un espace de plus grande dimension, dans lequel elles pourront être séparées par un hyperplan. Ainsi, l'idée de base consiste alors à trouver une fonction telle que :

$$\phi(x_i) \cdot \phi(x_j) = K(x_i, x_j) \quad (16)$$

Il est possible de déterminer directement la fonction de décision correspondant à la frontière de séparation non linéaire, sans jamais passer par l'espace de redescription.

En définitif, l'apprentissage d'un SVM consiste donc à minimiser la fonction objective:

$$\sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (17)$$

sous les contraintes :

$$\sum_{i=1}^n \alpha_i y_i = 0 \quad \text{et} \quad 0 \leq \alpha_i \leq C, \quad i = 1, \dots, n \quad (18)$$

Après l'apprentissage, la fonction de décision du SVM est égale à :

$$f(x) = \sum_{i=1}^n y_i \alpha_i K(x, x_i) + b \quad (19)$$

Et la règle de décision consiste, comme dans le cas de l'hyperplan, à classer la donnée  $x$  dans la classe caractérisée par le label  $y = 1$  si  $f(x) > 0$ , ou dans la classe caractérisée par le label  $y = -1$  si  $f(x) < 0$ .

En d'autres termes, le label d'une donnée inconnue  $x$  est égal à :

$$y = \text{signe}(f(x)) \quad (20)$$

Ainsi, le SVM est caractérisé par les données d'apprentissage pour lesquels les coefficients  $\alpha_i \neq 0$  que l'on nomme vecteurs de support et par un biais  $b$ .

### III.5- Quelques noyaux classiques utilisés pour le SVM

Les premiers à avoir élucidé le concept de manipulation de produits scalaires via un noyau non-linéaire  $K(x, y)$  ont été M.A. Aizerman, F.M. Braverman et L.I. Rozner [3]. Ici, nous introduisons quelques formules des noyaux utilisés dans les différentes méthodes de SVM.

Noyau	Formule
Linéaire	$K(x, y) = x \cdot y$
Sigmoïde	$K(x, y) = \tan(\text{h}(ax \cdot y + b))$
Polynomial	$K(x, y) = (ax \cdot y + b)^d$
RBF	$K(x, y) = \exp(-\ x - y\ ^2 / \sigma^2)$
Laplace	$K(x - y) = \exp(- y  \ x - y\ )$

### III.6- SVM multi-classes

A l'origine la machine à vecteurs de support est conçue pour la classification de données à deux classes. Bien entendu, de nombreuses solutions ont été proposées pour étendre l'utilisation des SVM aux problèmes multi-classes. Cependant, aucune des approches existantes ne s'est réellement imposée. Ainsi, la résolution de problèmes multi-classes à l'aide de SVM reste encore aujourd'hui un sujet de recherche actif. Dans notre cas, nous nous intéressons aux SVM multi-classes pour la classification de dix classes à savoir les dix chiffres manuscrits. Parmi les méthodes qui traitent le problème de classification de multi-classes : « un contre un » et « un contre tous ».

#### III.6.1- La stratégie un contre tous

La solution la plus simple pour résoudre un problème multi-classes à l'aide de SVM consiste à le décomposer en un ensemble de sous-problèmes binaires et à construire indépendamment un SVM pour chacun d'entre eux. Cette stratégie consiste à construire autant de SVM qu'il y a de classes. Chaque SVM est alors entraîné à séparer les données d'une classe qui seront étiquetées +1, de celles de toutes les autres



classes qui seront étiquetées  $-1$ . Chaque SVM est ainsi associé à une classe et sa sortie avant seuillage peut être considérée comme une mesure d'appartenance à la classe. La règle de décision généralement utilisée consiste donc à attribuer la donnée inconnue à la classe correspondant au SVM ayant la plus grande valeur de sortie.

### **III.6.2- Méthode un contre un**

Une autre stratégie classique consiste à construire un SVM pour chaque paire de classes, soit  $K(K-1)/2$  SVM pour un problème à  $K$  classes. Chaque classifieur est donc entraîné à séparer les données d'une classe de celles d'une autre classe. Nous parlerons alors de stratégie « un contre un ». La règle de décision généralement utilisée est le vote majoritaire.

### **III.7- Les domaines d'applications**

Les SVM est une méthode de classification qui montre de bonnes performances dans la résolution de problèmes variés. Cette méthode a montré son efficacité dans de nombreux domaines d'applications tels que le traitement d'images, la catégorisation de textes ou le diagnostics médicales et ce même sur des ensembles de données de très grandes dimensions [15]. Plusieurs communautés de chercheurs ont utilisé le SVM dans leurs recherches. Blanz et al. ont utilisé le SVM pour la classification d'objets 2D de vues différentes. Schmidt et al. ont exploré la tâche de reconnaissance d'orateur. Osuna et al. ont développé un système de reconnaissance de visages [3]. Cortes et al., Scholkopf et al., Burges et al. ont appliqué le SVM à la reconnaissance de chiffres manuscrits isolés[3].

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

En premier lieu, nous avons évoqué l'évolution que SVM à ramené aux domaines de son utilisation. Suivi d'un bref historique dont on a introduit l'arbre de classification des méthodes d'apprentissage à base de noyaux. Nous somme passés

après à la mise au point de la minimisation du risque structurel par VAPNIK ; puis à la propagation du SVM au domaine non linéaire, procédé des classes de la machines à vecteur de support ; dans la mono-classe, on n'a pas manqué de préciser sa formulation mathématique. Le chapitre s'est achevé par une brève vue sur quelques domaines d'application de cette méthode.

Dans notre travail, nous avons utilisé la méthode un contre tous du SVM multiclasse. Dans l'étape de la classification des chiffres manuscrits isolés.

# **CHAPITRE IV :**

## **Tests et résultats**

## **IV.1- Introduction**

L'objectif de ce chapitre consiste à présenter les résultats effectués sur deux bases de données des imagerie de chiffres manuscrits, composés de dix classes.

## **IV.2- Présentation des bases de données**

Pour évaluer notre méthode, nous avons utilisé deux bases de données de chiffres: une base artificielle et une base réelle.

La première base contient des chiffres allant de 0 à 9 qui générés de manière artificielle.

Cette base contient 150 images d'apprentissage et 110 images de test. Chaque image est de taille 16x10 pixels.

Les chiffres de la deuxième base sont extraits de la base de données USPS (de l'anglais US Postal Service). Les chiffres sont choisis d'une manière aléatoire. Celle-ci est construite à partir de chiffres manuscrits saisis à partir d'images d'enveloppes collectées au centre CEDAR à Buffalo (Etats Unis). La base de données contient 500 images de chiffres dont 350 images d'apprentissage (servent à construire le classifieur) et 150 images de test (servent à tester ses performances). chaque image est de taille 16x16 pixels. Il est connu que l'ensemble de test de l'USPS est plutôt difficile puisque l'erreur humaine se situe autour de 2.5%.

Il faut noter que les chiffres peuvent présenter les défauts suivants :

- présence de chiffres fragmentés ;
- formes provenant des cas de chevauchement de chiffres ;
- présence de formes de chiffres inconnus ;
- cas de confusion entre le chiffre 1 et le 7, le chiffre 2 et le 9 et le chiffre 3 et 5.

Les deux figures IV.1 et IV.2 montrent certaines imperfections qui se trouvent dans les deux bases de données

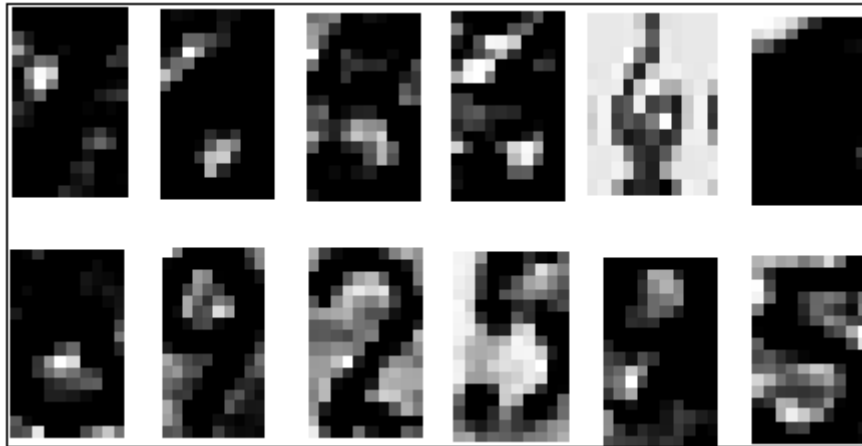


Figure IV.1 : Les défauts des chiffres dans la base1.

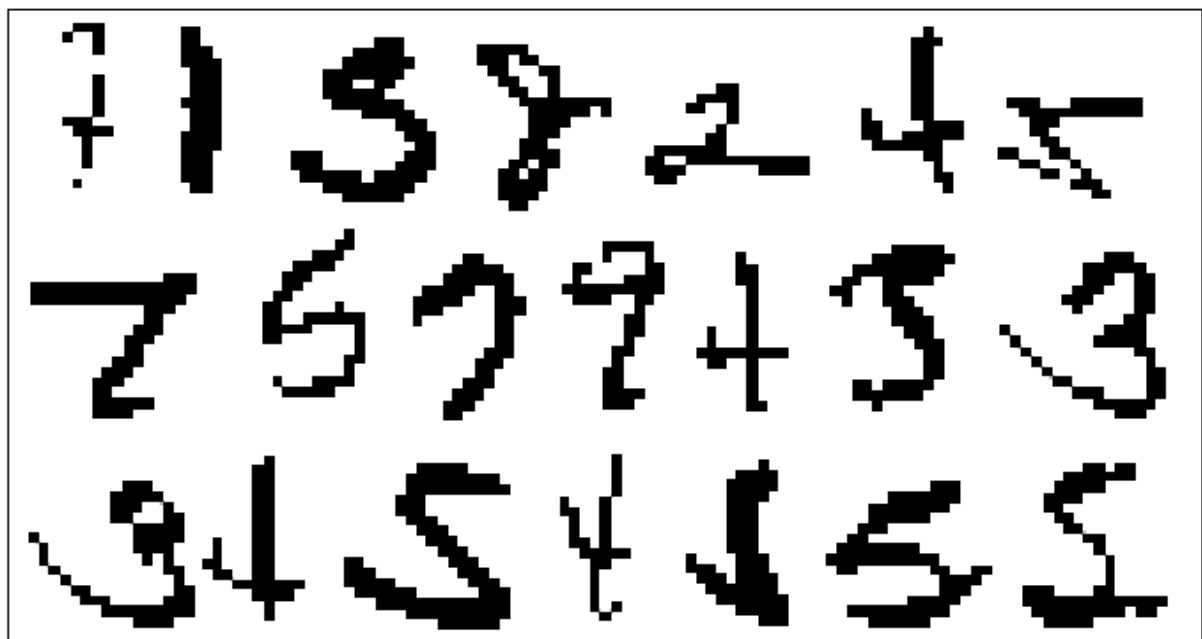


Figure IV.2 : Les défauts des chiffres dans la base2.

Le fichier de la base1 sous forme matricielle est présenté dans la figure suivante :

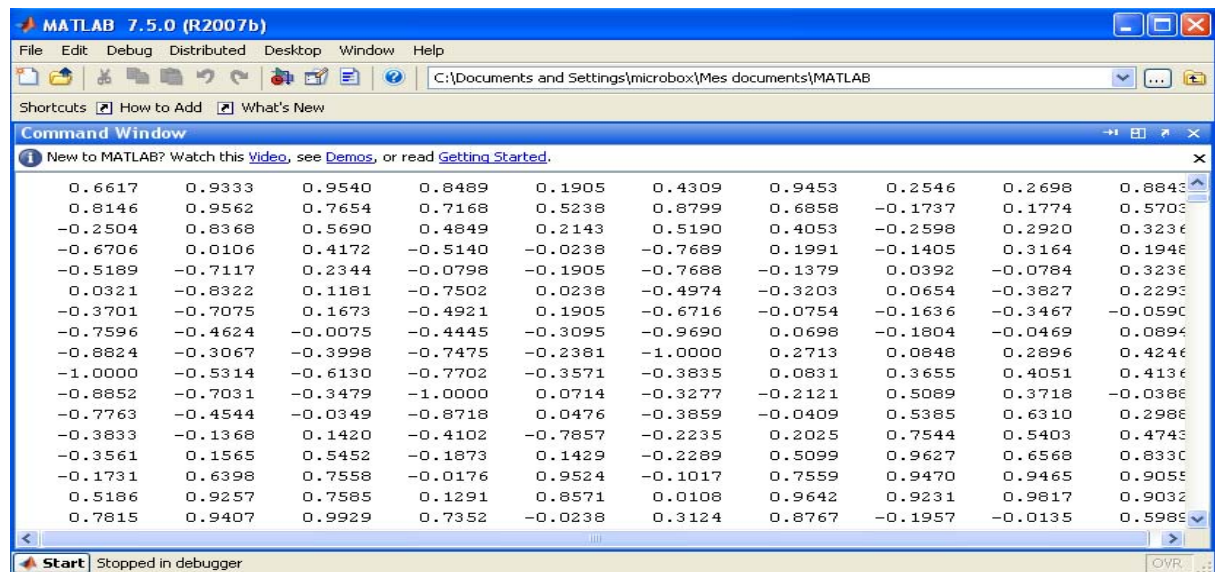


Figure IV.3 : Présentation de base1 sous forme matricielle.

Dans la figure IV.4 nous présentons le fichier d'apprentissage de la base USPS.

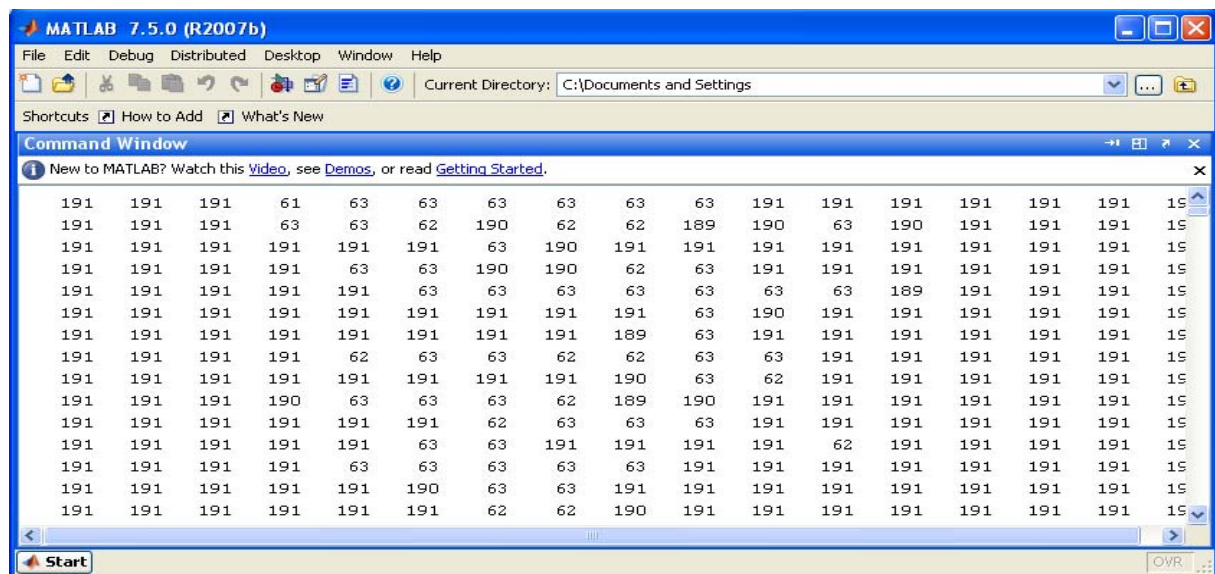


Figure IV.4 : Présentation du la base2 sous forme matricielle.

### IV.3- Les SVM

Pour résoudre les problèmes de la classification de ces chiffres, nous avons utilisé les SVM multi classes un contre tous. La technique des SVM est une méthode de classification qui montre de bonnes performances dans la résolution des problèmes variés. Nous avons effectué une étude comparative en employant quatre types d'ondelettes qui sont : Haar, Symlet, Coiflet et Daubechies. Nous avons aussi fait un

test en injectant à l'entrée du classifieur les données brutes, après avoir effectué un ensemble de prétraitement de binarisation et de normalisation.

Notant que les différents paramètres du classifieur SVM sont fixés expérimentalement (telle que C qui est la fonction de coût de mauvaise classification et sigma qui est le paramètre de noyau utilisé). Nous avons opté pour le noyau gaussien pour ces bons résultats et son temps d'exécution rapide.

#### IV.4- Tests et Résultats

##### IV.4.1- Données artificielles (Base1)

- *Premier test :*

Dans ce test nous avons injecté les valeurs de l'image initiale après avoir procédé à la binarisation et à la normalisation.

Sur le tableau IV.1, nous présentons le taux de reconnaissance ((Taux de reconnaissance = nombre des images bien classées / nombre total) x 100) obtenus après avoir binarisé et normalisé la base de données artificielle.

Taux de reconnaissance	
La binarisation	La normalisation
92.72%	93.63%

Tableau IV.1. Résultats obtenus sur la base1 avec la binarisation (C=2, sigma=4) et la normalisation (C=1, sigma=3).

- *Deuxième test :*

La décomposition de l'image initiale par application de la transformée en ondelettes a donné en sortie 4 cartes de fréquence (1 carte de basse fréquence et 3 autres cartes de haute fréquence.). Dans ce travail, on s'est intéressé seulement à la

carte de basse fréquence. Ainsi le vecteur de caractéristiques est composé des coefficients de l'approximée. Nous présentons les résultats de l'application de quatre ondelettes sur la base1.

La valeur de l'hyper paramètre  $C$  est fixée d'une manière expérimentale, et nous avons opté pour le noyau RBF dans tous nos tests effectués. Quant à la valeur de l'écart type sigma ( $\sigma$ ) celle-ci est choisi également d'une manier expérimentale.



Les paramètres du SVM	L'ondelette	Le taux de reconnaissance
C=2, Sigma=1	'haar'	95.45%
C=1, Sigma=2	'db4'	95.45%
C=2, Sigma=3	'db3'	94.45%
C=4, Sigma=2	'db2'	94.45%
C=1, Sigma=1	'db1'	95.45%
C=2, Sigma=1	'Sym1'	95.45%
C=2, Sigma=2	'Sym2'	95.54%
C=1, Sigma=2	'Sym3'	94.54%
C=1, Sigma=2	'Sym4'	93.63%
C=2, Sigma=2	'Coif1'	95.45%
C=3, Sigma=3	'Coif2'	94.54%
C=3, Sigma=4	'Coif3'	93.63%
C=1, Sigma=3	'Coif4'	93.60%

Tableau IV.2. Résultats obtenus avec les ondelettes.

## IV.4.2- Données réelles (base2)

- *Premier test :*

Dans ce test nous avons injecté les valeurs de l'image initiale après avoir procédé à la binarisation et à la normalisation.

Nous présentons le taux de reconnaissance de la binarisation et de la normalisation de la base USPS dans le tableau IV.3.

Taux de reconnaissance	
La binarisation	La normalisation
78.66%	78%

Tableau IV.3. Résultats obtenus sur la base2 avec la binarisation ( $C=4$ ,  $\sigma=5$ ) et la normalisation ( $C=5$ ,  $\sigma=5$ ).

- *Deuxième test :*

La décomposition de l'image initiale par application de la transformée en ondelettes a donné en sortie 4 cartes de fréquence (1 carte de basse fréquence et 3 autres cartes de haute fréquence.). Dans ce travail, on s'est intéressé seulement à la carte de basse fréquence. Ainsi le vecteur de caractéristiques est composé des coefficients de l'approximée. Nous présentons les résultats de l'application de quatre ondelettes sur la base2.

La valeur de l'hyper paramètre  $C$  est fixée d'une manière expérimentale, et nous avons opté pour le noyau RBF dans tous nos tests effectués. Quant à la valeur de l'écart type sigma ( $\sigma$ ) celle-ci est choisie également d'une manière expérimentale.

Les résultats de la classification de la base2 (la base USPS) avec les ondelettes, sont donnés dans le tableau IV.4.

Les paramètres	L'ondelette	Le taux
C=2, sigma=1	'haar'	82%
C=1, Sigma=2	'db4'	84%
C=3, Sigma=2	'db3'	84%
C=2, Sigma=2	'db2'	83.33%
C=2, Sigma=2	'db1'	81.33%
C=100, Sgma=2	'Sym1'	82%
C=100, Sgma=2	'Sym2'	83.33%
C=100, Sigma=2	'Sym3'	82.66%
C=100, Sigma=2	'Sym4'	81.33%
C=100, Sigma=3	'Coif1'	82%
C=100, Sigma=2	'Coif2'	84%
C=100, Sigma=3	'Coif3'	84%
C=100, Sigma=4	'Coif4'	82%

Tableau IV.4.Résultats obtenus sur la base2 avec les ondelettes.

### **IV.5- Interprétation des résultats**

Différents tests sont réalisés, dans le but de chercher les descripteurs les plus discriminants.

D'après les résultats obtenus dans la binarisation et la normalisation, nous constatons que le taux de reconnaissance varie selon la base utilisée et le type de données injectées au classifieur SVM.

En vue d'avoir des résultats plus performants, nous avons injecté au classifieur SVM les paramètres, des images, extraits avec les ondelettes. En effet, dans la première base, un taux de reconnaissance de 95.45% est obtenu en utilisant les ondelettes de Haar, db4, db1, sym1 et sym2.

Par contre, dans le cas de la base réelle, on a obtenu un taux de reconnaissance très satisfaisant qui est de 84% (db4, db3, coif2, coif3).

La différence entre les deux résultats s'explique par le fait que, dans base réelle, il y'a présence d'une variété de styles d'écritures (âge, sexe, niveau intellectuel, etc.). Par contre, dans la base artificielle, les données sont choisies soigneusement pour avoir un taux de reconnaissance assez acceptable.

A partir de ces résultats, on constate que le choix de l'ondelette dépend étroitement de type de l'information qu'on cherche à extraire.

### **IV.6- Conclusion**

Dans ce chapitre, nous avons présenté l'approche adoptée ainsi que les résultats obtenus sur deux bases de données réelle et artificielle. Plusieurs tests ont été réalisés d'une part, pour chercher le type d'ondelette qui convient mieux à notre application et d'autre part, pour régler des paramètres du classifieur SVM qui permettent d'améliorer davantage le taux de reconnaissance.

Pour améliorer le taux de reconnaissance d'une manière efficace, d'une part, on procédera à l'élargissement de la base de données en rajoutant de nouvelles données, afin de couvrir tous cas possibles de styles d'écriture et d'autre part, il est impératif de traiter soigneusement, au niveau de la phase de décision, les cas de rejet qui se présentes.

# **Conclusion**

## **générale**

Les problèmes posés par la vision par ordinateur sont multiples et encore loin d'être résolus car ils sont souvent très complexes. Lors de la reconnaissance d'objets, plusieurs problèmes sont posés comme l'adaptation à une grande variété de situations ( le changement de fond, les variations de luminance, les occultations ou les changements de points de vue, etc.).

L'objectif de notre mémoire était de concevoir un système de reconnaissance automatique de chiffres manuscrits. A cet effet, nous avons consacré un chapitre sur la reconnaissance de forme en général, tout en présentant les étapes constituant le système.

Pour augmenter le taux de reconnaissance et rendre le système plus performant, nous avons agi sur les différentes étapes de processus de reconnaissance.

Pour l'apprentissage, nous avons opté pour la méthode SVM un contre tous, nous avons aussi effectué des essais en vue de régler les paramètres de SVM.

Ainsi, nous avons pu concevoir un système de reconnaissance des chiffres manuscrits qui donne des résultats très intéressants par la méthode SVM avec les paramètres des ondelettes.

Enfin, ce travail nous a permis d'approfondir nos connaissances théoriques et pratiques, dans le domaine de traitement d'image en général, et d'ajouter un plus à notre cursus de formation.



## ***Bibliographie***

- [1] R.André-Obrecht, Reconnaissance des Formes et Ingénierie des Technologies Vocales.
- [2] F. Appriou, président du club 29 de la SEE. La reconnaissance des formes. Quelle méthode pour quelle application ? . 2006.
- [3] N.E. Ayat, mémoire doctorat, Sélection de modèles automatiques des machines à vecteur de support : application à la reconnaissance d'image de chiffres manuscrits ; Ecole de technologie supérieur, Université du Québec. 2004.
- [4] H. S. Baird, Recognition technology frontiers; AT&T Bell Laboratoires, USA. CNED'94. 1993.
- [5] G.J. Balm, mémoire doctorat, An introduction to optical character reader considerations. Université Helsinki de Technologie. 1997.
- [6] M. Bouazza, mémoire d'ingénieur d'état, Application de la transformée en ondelettes à la reconnaissance de chiffres manuscrits, département d'automatique, UMMTO, 2008.
- [7] A.Boucher, N.A.Tuan, Reconnaissance d'écriture manuscrite. 2005.
- [8] P. Bourgeat. Segmentation d'images de semi-conducteurs appliquée à la détection de défauts. Université de Bourgogne, le 9 juillet 2004
- [9] K.Bouriel, F.Bouafif Samoud, H. El Abed, S.Snoussi Maddouri. Stratégies d'évaluation et de comparaison de méthodes d'extraction de la ligne de base de mots manuscrits arabes. Ecole nationale d'ingénieurs, Laboratoire des systèmes et traitement du signal, Tunis.

- [10] D. Brucq, K. Romeo, M. Lachkar, N. Feray, et F. Bagui. Une méthode d'évaluation du codage en reconnaissance de chiffres manuscrits ; UFR des Sciences et des Techniques, Mont-Saint-Aignan, ; CNED'94, 1982.
- [11] O.Catoni. Mathématiques et reconnaissance des formes.
- [12] P.HSUEN CHEN, C-J. LIN, B. SCHÖLKOPF. A tutorial on v-support vector machines.
- [13] DJENDER Samy, GHEDDACHE Krim. Mémoire d'ingénieur d'état. Segmentation d'image par la méthode JSEG et les ondelettes. Département Automatique UMMTO, 2010.
- [14] N. feray, D. De Bruq. Système de Reconnaissance de Chiffres Manuscrits Hors Lignes. PSI - La31 – Itepea, UFR des Sciences et des Techniques BP 118, 76821 Mont Saint Aignan Cedex.
- [15] M.HASAN & F.BORIS. SVM : Machines a Vecteurs de Support ou Séparateurs à Vastes Marges. Versailles St Quentin, France, 16 janvier 2006
- [16] L.Heutte. Reconnaissance de formes définitions et applications .
- [17] E.JAKOBOWICZ. La discrimination a plus de deux classes comparaison de plusieurs approches issues des supports vector machines (SVM). École polytechnique fédérale de Lausanne.
- [18] S.Lasaulce. de la transformée de Fourier à la transformée en ondelettes. Module Ondelettes du DEA TIS, 2010.
- [19] Luminey. Brève introduction à la reconnaissance des formes. Ecole Jeunes Chercheurs CNRS 'Interaction HOMME-MACHINE'. 1997.

- [20] H. Majdoulayne. Thèse doctorat. Extraction de caractéristiques de texture pour la classification d'images satellites. Université Toulouse III – Paul Sabatier, 2009.
- [21] J. Mantas. Mémoire doctorat. An overview of character recognition methodologies. National Centre of Medical Documentation, Greece ; CNED'94. 1985.
- [22] Ali. Mcheih. Mémoire doctorat. Modélisation statistique du Speckle en OCT : application à la segmentation d'images de la peau. INP TOULOUSE. 2010.
- [23] F. Menasri. Mémoire doctorat. Contribution à la reconnaissance d'écriture arabe manuscrite. Université de Paris, UFR de Mathématiques et Informatique. 2008.
- [24] S. Mori, K. Yamamoto Member IEEE, and M. Yasuda Member IEEE. Research on machine recognition of hand printed characters. 1992.
- [25] H. Nemmour, R. Diboune, Y. Chibani, H. Fellah. Mémoire d'ingénieur d'état. Utilisation des machines à vecteur de support (SVM) pour la reconnaissance des chiffres manuscrits. Université des Sciences et de la Technologie Houari Boumediene. 2007.
- [26] S. Nemouchi, N. Farah. Reconnaissance de l'écriture Arabe par Systèmes Flous. Laboratoire de Gestion Electronique du Document (LABGED), Annaba.
- [27] C. Olivier, O. Colot et P. Courtellemont. Choix de paramètres pertinents pour la reconnaissance de chiffres manuscrits. La3i-LACI UFR des Sciences, Université de Rouen, France ; CNED'94, 1995.
- [28] D. Poisson, S. Mahjoub. Mémoire doctorat. Comment améliorer les techniques

de reconnaissance de formes 3D ?. Ecole supérieur de génie informatique, Paris. 2010.

[29] P. Telmosse, and R. Palmondon. Caractérisation et classification de nombres manuscrits par méthodes statistique. Ecole polytechnique de Montréal, Laboratoire SCRIBENS. IPAR '86. 1998.

[30] M. H.Thien, M. Zimmermann, and H. Bunke. Off-line handwritten numeral string recognition by combining segmentation-base and segmentation-free methods. Université de Berne, Institut des sciences informatiques et mathématiques, Neubruckstr, Suisse; CNED'94. 1997.

[31] [http://www.clopinet.com/isabelle/](http://www.clopinet.com/isabelle/Projects/SVM/applist.html) Projects/SVM/applist.html.

[32]. Reconnaissance de l'écriture manuscrite WIKIPEDIA.