

Ministère de l'enseignement supérieur et de la recherche scientifique
UNIVERSITE MOULOUD MAMMERI DE TIZI-OUZOU
FACULTE DE GENIE ELECTRIQUE ET D'INFORMATIQUE
Département d'automatique



Mémoire de fin d'études

**En vue de l'obtention du diplôme d'Ingénieur d'Etat
en Automatique**

THEME

Application de la transformée en ondelettes à la
reconnaissance de chiffres manuscrits.

Proposé et dirigé par :

M^{elle} AIT AIDER Malika

Présenté par :

M^r BOUAZZA Mourad

Promotion 2007 / 2008

Remerciements

Je tiens à remercier ma promotrice Melle AIT AIDER Malika, pour le temps précieux qu'elle m'a accordé, les conseils et le suivi tout au long du travail.

Je suis aussi reconnaissant envers tous les enseignants qui ont contribué dans toutes mes études. Ainsi que tout ceux qui m'ont aidé de près ou de loin tout au long de mon cursus.

Mes remerciements vont également aux membres du jury pour avoir accepté de juger ce travail.

Qu'ils trouvent ici toute l'expression de ma gratitude et de ma considération.

À ma chère famille

À tous mes amis

Et à tous ceux qui sont content pour moi.

MERCI.

Dédicaces

Je dédie ce modeste travail à :

Ma chère famille.

Mes amis (es) : BILAL, AMINA, SAID, YIDIR, MOULOUD, MASSI, KARIM'S.

Tous les gens avec qui j'ai passé de très bons moments pendant ces 5 ans (LES MEILLEURES ANNEES) et tout les résidents de la cité OUED AISSI.

Tous mes rancards et dieu sait combien j'en ai eu, beaucoup mais jamais assez Surtout SORAYA (qu'il nous pardonne à tous) ...

Toute la promotion 2008 d'Automatique surtout SAID qui m'a toujours donné les cours (j'ai trop séché) et aussi AMINA (rien à dire). Ne changez pas vous êtes trop bien, en plus je vous ai cité deux fois ! Sans oublier KARIM, FELLA, LAMIA (ke j'aime bien)...

A tous les autres à qui je pense (en écrivant ces dédicaces) mais que la page ne peut contenir.

Sommaire

Introduction générale	1
Chapitre I : Généralités sur la reconnaissance des formes.	
I – Introduction	3
II- Processus de reconnaissance	3
II.1. Le monde physique	4
II.2. Le codage	4
II.3. Le prétraitement	5
II.4. L'analyse	6
II.5. L'apprentissage	7
1- L'apprentissage supervisé	7
2- L'apprentissage non supervisé	7
II.6. La décision	7
III- Classification automatique	8
III.1 Approches mathématiques	8
1- Méthodes métriques	8
2- Méthodes statistiques	9
III.2 Approches structurelles	9
1- La squelettisation	9
2- Extraction de segments primitifs	10
III.3 Méthodes connexionnistes	11
IV- Domaines d'application.	11
Conclusion	13
 Chapitre II : Théorie des ondelettes	
I- Introduction	14
II- Analyse de Fourier	14
II.1. Transformée de Fourier	14
II.2. Transformée de Fourier à fenêtre glissante	15
III- Transformée en ondelettes	16
III.1. Exploitation de la corrélation	16
III.2. Définition	18
III.3. Transformée en ondelettes continue	20
III.4. Transformée en ondelettes discrète	20
VI- Analyse multirésolution	21
VI- 1. Fonction d'échelle	22
VI- 2. Fonction d'ondelettes et espace de détail	23
VI- 3. Fonction d'échelle et ondelettes à deux dimensions	24
VI- 4. Présentation de l'algorithme pyramidal	27
V- Exemples d'ondelettes pratiques	27
1- Chapeau mexicain	27
2- Ondelette de Haar	28

3- Ondelettes de Daubechies	29
VI- Conclusion	30

Chapitre III : Réseaux de neurones

I- Introduction	31
II- Le neurone biologique	31
III- Le neurone formel.	32
IV- Présentation des réseaux de neurones	34
IV-1. Structure d'interconnexion	35
1-1. Les réseaux bouclés	35
1-2. Les réseaux non bouclés	36
IV-2. L'apprentissage des réseaux de neurones	36
2-1. L'apprentissage supervisé	37
2- 2. L'apprentissage non supervisé	37
IV-3. Propriétés des réseaux de neurones	37
IV-4. Quelques modèles de réseaux de neurones	38
4-1. Le perceptron multicouche	38
4-1-1. Mise en œuvre du perceptron multicouche	39
4-1-2. La rétropropagation du gradient de l'erreur	39
4-2. Le modèle de Hopfield	42
4-3. Le modèle de Kohonen	43
V- Conclusion	44

Chapitre IV : Tests et résultats .

I- Introduction	45
II – Présentation de la base de données USPS	45
III- Choix du réseau de neurones	47
IV- Tests	48
IV-1- Données brutes	49
IV-2- Ondelettes de Haar	50
IV-3- Ondelettes de Daubechies	51
IV-4- Ondelettes Bior 3.7	52
V- Conclusion	53
Conclusion générale	54
Bibliographie	

La reconnaissance automatique des formes consiste à identifier ou à classer des formes ou des objets en se basant sur certaines de leurs caractéristiques. La notion d'apprentissage est au cœur de la plupart des techniques développées.

Il s'agit d'un domaine très vaste, qui trouve de plus en plus d'applications parmi lesquelles on trouve la reconnaissance de la parole, la vision artificielle, la reconnaissance de chiffres manuscrits...

Il est d'usage de distinguer deux activités en reconnaissance de chiffres manuscrits ou de caractères, de façon générale, la reconnaissance en ligne et la reconnaissance hors ligne. La reconnaissance en-ligne s'effectue à partir d'une acquisition spatio temporelle des caractères sur une tablette électronique c'est à dire que la reconnaissance se fait au fur et à mesure que le chiffre s'écrit. La reconnaissance hors-ligne, privée de l'information spatio-temporelle, est plus délicate. Elle concerne, à priori, tous les documents écrits puis numérisés (adresses postales, chèques bancaires, formulaires structurés ou manuscrits quelconques). Bien que ces deux domaines d'application soient fortement liés, nous ne nous intéresserons dans ce mémoire qu'à la reconnaissance hors-ligne.

La reconnaissance hors-ligne reste à ce jour un thème de recherche ouvert. En effet, bien que le nombre de classes naturelles soit réduit (chiffres 0 à 9), on trouve à l'intérieur de chacune d'entre elles, une très grande variabilité de l'écriture [1]. De plus, les conditions souvent précaires dans lesquelles sont écrits les chiffres (chèques écrits rapidement sur un coin de table) et la variabilité 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.

Lors de l'analyse d'une image, différents calculs seront nécessaires pour déterminer de manière optimale les paramètres qui serviront à caractériser chacune des formes. Il est important de considérer le fait que dans le cas d'une application temps réel, le temps de traitement est un facteur non négligeable dans l'étape de reconnaissance. Dans notre mémoire, nous nous sommes intéressés à la technique d'analyse fondée sur la transformation en ondelettes.

Notre mémoire s'articule autour de quatre chapitres. Le premier chapitre présentera la décomposition classique d'une chaîne de reconnaissance des formes, tout en détaillant chacune de ses étapes.

Dans le second chapitre, nous aborderons la théorie des ondelettes et l'analyse multirésolution sans faire appel à des notions approfondies.

Le troisième chapitre, sera consacré aux réseaux de neurones artificiels utilisés comme classifieurs. Nous nous intéresserons en particulier aux réseaux de neurones multicouches à apprentissage supervisé.

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

I. Introduction

La reconnaissance des formes est un domaine de l'informatique qui a pour but de percevoir et de reconnaître automatiquement des formes présentes dans des signaux et des images. Elle regroupe les tâches d'apprentissage et de décision qui jouent des rôles assez proches. 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 représentant la forme entrée. Le résultat de la décision est un avis sur l'appartenance ou non de la forme aux modèles de l'apprentissage.

II. Processus de reconnaissance [2]

Un système de reconnaissance des formes a pour objectif de prédire la classe d'appartenance d'une forme inconnue. Pour y parvenir, trois étapes sont nécessaires. La première étape dite de prétraitement consiste à isoler la forme à reconnaître. La seconde étape quant à elle consiste à effectuer un certain nombre de mesures qui vont permettre de caractériser la forme. Le vecteur de caractéristiques alors obtenu sera exploité pour prendre une décision lors de l'étape de classification. Par ailleurs, de manière à optimiser les différents paramètres de chacune de ces étapes, les approches classiques de reconnaissance des formes utilisent un ensemble de données étiquetées que l'on nomme base d'apprentissage.

La démarche classique suivie en reconnaissance des formes consiste à opérer selon le schéma général de la figure I.1 :

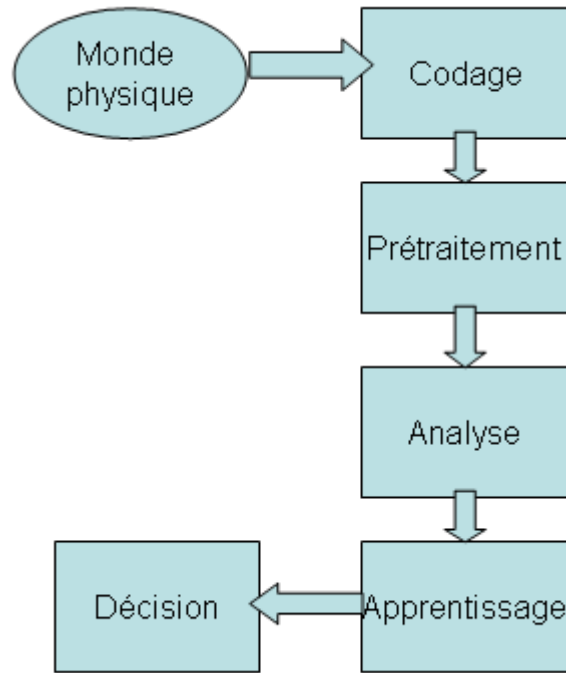


Figure I.1. Schéma général d'un système de reconnaissance des formes

II.1. Le monde physique

Le monde physique est un espace analogique de dimension infinie appelé espace des formes. Les objets, dans cet espace, peuvent être décrits de différentes façons avec une multitude de propriétés, dont il serait difficile de tenir compte pour effectuer une reconnaissance des formes. La loi de passage au monde discret nécessite forcément une sélection et par conséquent une certaine simplification.

II.2. Le codage

C'est une opération de conversion numérique du monde physique continu vers un monde numérique discret. Ce dernier, appelé aussi espace de représentation, a une dimension encore trop importante même si elle est finie. La dimension de cet espace est choisie volontairement grande de manière à pouvoir disposer d'un maximum d'informations sur la forme et à pouvoir y sélectionner des sous-ensembles relatifs à de

nombreuses caractéristiques. Le codage s'effectue à l'aide de capteurs branchés à une unité de conversion analogique-numérique.

II.3. Le prétraitement

L'étape de prétraitement consiste à sélectionner dans l'espace de représentation l'information nécessaire à l'application. Cette sélection passe souvent par l'élimination du bruit dû aux conditions d'acquisition, par la normalisation des données, ainsi que par la suppression de la redondance. Le nouvel espace de représentation a une dimension plus petite mais demeure tout de même un espace de grande dimension. Toutes les techniques élaborées dans ce sens se gardent de modifier les propriétés essentielles des formes. Parmi les techniques de prétraitement qui existent, nous citons :

1- L'égalisation d'histogramme

Un histogramme est un graphique statistique permettant de représenter la distribution des intensités des pixels d'une image (la densité de probabilité des niveaux de gris) , c'est-à-dire le nombre de pixels pour chaque intensité lumineuse. Par convention un histogramme représente le niveau d'intensité en abscisse en allant du plus foncé (à gauche) au plus clair (à droite). Ainsi, l'histogramme d'une image en 256 niveaux de gris sera représenté par un graphique possédant 256 valeurs en abscisses, et le nombre de pixels de l'image en ordonnées.

Son calcul se fait par quantification des niveaux de gris de l'image, puis de calculer, pour chaque niveaux de gris, le nombre de pixels correspondant à ce niveau. Ainsi, l'égalisation de l'histogramme permet un réglage automatique de l'intensité des pixels. On peut aussi avoir une description de l'image en paramètres et ce en utilisant le seuillage.

La figure ci-dessous illustre deux types d'histogrammes :

- L'histogramme b est concentré sur les bas niveaux de gris, c.-à-d. l'image est trop sombre ;
- L'histogramme h est concentré sur les hauts niveaux de gris, c.-à-d. l'image est trop claire ;

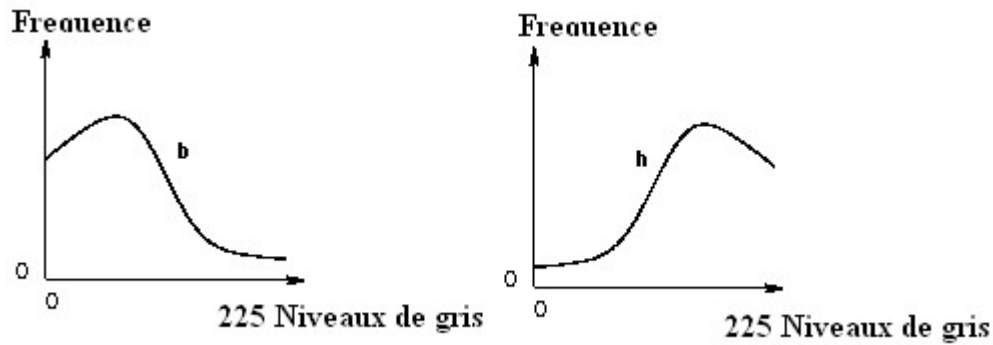


Figure I.2 : Représentation d'une image par histogramme

2- Le filtrage

Le filtrage permet de réduire les variations d'intensité au sein de chaque région de l'image. Elle se fait par application d'un filtre médian qui détecte les pixels blancs ou noirs isolés et leur affecte une valeur moyenne des pixels qui les entourent.

II.4. L'analyse

Une des difficultés associées à la reconnaissance des formes est l'espace de dimension trop élevé. A cet effet, l'analyse ou l'extraction de caractéristiques est une étape très importante puisque on cherche à extraire l'information utile à partir d'une très grande quantité d'informations. Un mauvais choix de caractéristiques va affecter négativement le système et conduire à une mauvaise décision.

Lors de cette étape, les techniques de reconnaissance des formes calculent un certain nombre de caractéristiques ou paramètres. Ces paramètres correspondent à des mesures de nature géométrique, topologique ou statistique et servent comme seules données représentant la forme. Ils sont généralement limités en nombre. Ainsi, l'espace obtenu est l'espace des paramètres de dimension (p) très petite par rapport à l'espace de représentation obtenu après prétraitement. Pour une bonne phase d'analyse, le concepteur doit suivre quelques règles :

- Extraire des informations pertinentes.
- Le vecteur de paramètres doit discriminer les classes entre elles.
- Le vecteur de paramètres doit être stable, en fonction du bruit.
- Le vecteur de paramètres doit atténuer l'effet du bruit.

II.5. L'apprentissage

L'apprentissage, ou classification, est une étape clé dans la chaîne de reconnaissance. Son rôle est d'éclairer la décision à l'aide de connaissances à priori sur les formes. A partir de critères spécifiques aux formes, l'apprentissage tente de définir des modèles de référence ou de caractériser des classes de décision. Cela exige de l'apprentissage de bien définir les classes de manière à distinguer les familles homogènes de formes et donc à identifier les nouvelles formes. On distingue deux types d'apprentissage :

II.5. 1. Apprentissage supervisé

Il est dit supervisé lorsque les différentes familles de formes sont connues à priori et si la tâche d'apprentissage est guidée par un professeur.

II.5.2. 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 le professeur. Ce mode d'apprentissage est aussi appelé suivant l'approche utilisée, classification automatique, inférence ou encore apprentissage sans professeur.

II.6. La décision

La décision, ou classement, est l'étape de reconnaissance proprement dite. Son rôle est d'identifier la forme de l'objet considéré à partir de l'apprentissage réalisé. La méthode de décision est souvent exhibée par l'apprentissage, ce qui veut dire que les critères utilisés pour la comparaison sont les mêmes que ceux utilisés pour l'apprentissage. En effet, il est évident que le choix de critères différents pour la décision, n'assurant par l'uniformité de la description, ne peut pas conduire avec certitude à un résultat cohérent.

III. Classification Automatique [3]

L'objectif de la classification automatique consiste à répartir les objets ou les formes ayant les mêmes caractéristiques en classes les plus distinctes possibles, de sorte que chaque individu ressemble beaucoup plus aux objets de sa classe qu'aux objets des autres classes. Parmi les approches les plus rencontrées, se trouvent :

III.1. Approche mathématique

On distingue les méthodes métriques et statistiques.

III.1.1. Méthodes métriques

Dans ce cas, la forme à reconnaître est décrite par un vecteur de caractéristiques, qui représente la valeur des paramètres calculés sur cet objet. Parmi les méthodes métriques, on trouve la méthode des k plus proches voisins [4].

Dans le cas de cette méthode, pour une forme inconnue x à classer (Fig.I.4), 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.

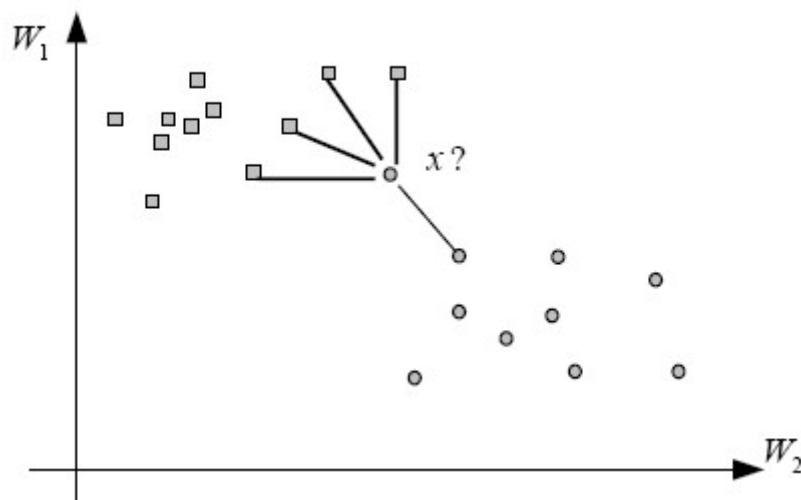


Figure I.4 : Méthode de décision des k plus proches voisins

Il existe plusieurs choix de distances possibles, la plus utilisée est la distance euclidienne.

III.1.2. Méthodes statistiques

Ces méthodes se basent sur la notion de fonction de densité de probabilité et la théorie de la décision de Bayes [5]. Dans le cas où l'on dispose de toute l'information en terme de probabilité à priori sur les objets, l'application du théorème de Bayes permet de transformer ces probabilités à priori en probabilités à posteriori, autrement dit, à trouver les classes les plus probables.

III.2. Approches structurelles

Par définition [6], une signature structurelle vise à représenter la structure de la forme. Souvent, cette signature repose sur une décomposition de la forme en primitives élémentaires et sur une description des relations entre ces primitives. On appelle primitive élémentaire ou primitive graphique tout regroupement de pixels de la forme : lignes, arcs de cercle, régions, vecteurs, . . . Techniquement, une signature structurelle peut contenir, soit un ensemble de primitives, soit un ensemble de relations topologiques (jonctions, parallélisme, . . .), ou bien encore les deux à la fois. Selon les éléments intégrés dans la signature structurelle, la description de l'objet graphique obtenue peut être plus ou moins élaborée. Outre les graphes, il existe aussi des approches syntaxiques basées sur l'utilisation de grammaires. Une grammaire est constituée d'un alphabet d'éléments terminaux (ensemble d'éléments de base), d'un alphabet d'éléments non terminaux (ensemble d'éléments plus complexes composés d'éléments de base), et d'un ensemble de règles de grammaires. Ces dernières modélisent la construction des éléments complexes de manière hiérarchique, à partir des éléments de base. Plusieurs types de grammaire sont utilisés en reconnaissance de formes : grammaire de chaînes, grammaire bidimensionnelle, grammaire stochastique.

III.2.1. La squelettisation (méthode d'amincissement)

Appliquée essentiellement en reconnaissance de chiffres manuscrits, l'algorithme de squelettisation de Marthon [7] a pour but d'obtenir un squelette d'épaisseur de 1 pixel. De façon itérative, il supprime à chaque fois les points n'appartenant pas au squelette (Fig I.11). Cette dernière ne permet pas en elle-même d'obtenir des primitives significatives

d'un symbole. Donc, il est nécessaire de compléter le processus par une seconde étape d'extraction ; il s'agit très souvent de construire un graphe du squelette.

Un point indispensable pour l'extraction de ce type de primitives structurelles est que les extrémités des branches ne sont pas des raccourcis, les courbures des formes sont bien représentées.

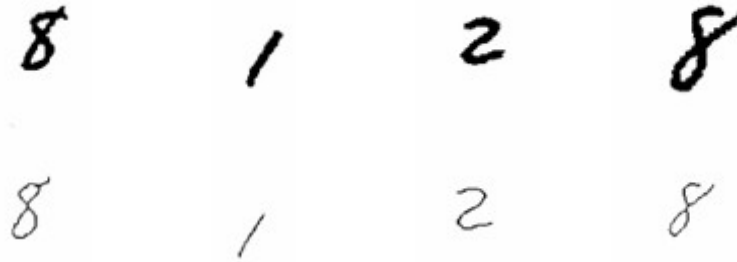


Figure I.11 : exemples de chiffres et de leurs squelettes

III.2.2 Extraction des segments primitifs

Une chaîne de reconnaissance est souvent amenée à évoluer par ajout ou modification d'extracteurs de primitives. L'objectif est d'obtenir, une représentation structurée de plus haut niveau.

Le premier plan informatif est constitué à partir de la décomposition de l'image en éléments de haut niveau facilement manipulables (trait, liaison inter-carrefour, boucle) . Ce plan donne accès à un deuxième plan informatif où l'on trouve la description de chaque élément de haut niveau sous la forme d'une liste des points le constituant ; longueur, nombre de pixels, etc.

Le deuxième plan permet l'accès au troisième et au dernier plan informatif, où l'on trouve la description des points (coordonnées, type du point, nombre de voisins sur un 8-voisinage). Ce plan permet l'extraction de primitives représentatives de points caractéristiques (points extrémités , points de jonction) ainsi que toutes les opérations de localisation d'un élément (élément dans la partie basse, haute du chiffre) . Il permet aussi une opération de projection qui a pour but la détection de traits horizontaux significatifs au niveau du chiffre .

Les méthodes de reconnaissance basées sur l'utilisation de signatures structurales intègrent donc une information haut niveau et peuvent ainsi être appliquées sur des images de symboles en contexte, c'est-à-dire sur des symboles non segmentés. Elles possèdent également l'avantage d'être invariantes à l'échelle et à la rotation. Cependant, elles sont assez sensibles au bruit et aux déformations et nécessitent souvent l'intégration de connaissances à priori sur les symboles traités.

Quand on estime que la phase d'apprentissage est terminée, le même algorithme ou une forme modifiée sera utilisée dans une phase de reconnaissance des formes proprement dite. La reconnaissance consiste soit à reconnaître un des exemples d'apprentissage plus ou moins déformé, soit à définir la classe d'appartenance la plus proche d'un échantillon inconnu.

III.3. Méthodes connexionnistes

La recherche sur les modèles connexionnistes a connu un développement important au cours de la dernière décennie et a déjà donné lieu à des applications intéressantes en reconnaissance des formes. Par exemple, en tri automatique du courrier. Le problème de reconnaissance est exécuté par un réseau de neurones à apprentissage supervisé qui exige un ensemble de formation d'exemples pour l'apprentissage du réseau.

Beaucoup d'applications doivent grouper des données multidimensionnelles, dont la relation mutuelle n'est pas clairement définissable. L'exploration ou la conceptualisation de données dans le procédé de découverte de la connaissance essaye d'identifier la structure des données et de diviser ces données en groupes distincts.

IV. Domaines d'application

La reconnaissance de formes possède de nombreux domaines d'application qui vont de la reconnaissance de caractères manuscrits, à la reconnaissance automatique de la parole ou encore à la reconnaissance de formes géométriques dans une image. L'intérêt de la reconnaissance des formes est illustré à travers quelques exemples d'application :

1. La vision artificielle

La vision par ordinateur est sans doute l'un des principaux champs d'application de la reconnaissance des formes. De nos jours, elle a trouvé une place prépondérante dans un domaine tel que l'imagerie biomédicale. Par exemple, dans la détection de cellules cancéreuses du cerveau, les images sont acquises en imagerie par résonance magnétique. On cherche ensuite à détecter d'éventuelles tumeurs à partir de caractéristiques extraites de régions de l'image.

Dans l'industrie, la technique de la vision artificielle est également intégrée dans un système de détection de défauts en vue de contrôler la qualité de fabrication de produits. Ainsi, on peut imaginer diverses pièces usinées passant sur un tapis roulant et filmées en continu par une caméra.

L'analyse d'images satellitaires ainsi que la biométrie (reconnaissance d'empreintes, de faces) qui ont de différentes débouchées dans les domaines sécuritaires, météorologiques.

2. La reconnaissance de l'écriture

La reconnaissance des caractères est une application classique de la reconnaissance de formes. Elle vise à transformer un texte manuscrit ou typographié, sur tout type de support, vers un texte numérique. Le document, sous forme papier par exemple, est préalablement numérisé par un scanner pour en obtenir une représentation sous la forme d'une image. On essaie ensuite d'extraire chacune des lettres présentes dans l'image et d'en déterminer un équivalent informatique.

Ce type d'application est couramment utilisé à des fins d'archivage, d'indexation, tri automatique du courrier, traitement de chèques .

3. La reconnaissance vocale

D'une manière générale, l'objectif de la reconnaissance de la parole est de faciliter la communication homme / machine en reconnaissant l'information parlée.

Conclusion

La reconnaissance hors-ligne de caractères manuscrits est un domaine de recherche exploré depuis déjà plusieurs décennies et débouche sur un nombre important d'applications industrielles. En outre, bon nombre de traitements liés aux particularités de l'écriture manuscrite non contrainte sont encore aujourd'hui des problèmes ouverts dans le domaine de la reconnaissance de l'écrit.

I- Introduction

La plupart des signaux du monde réel ne sont pas stationnaires, et 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.

Un aspect très important dans l'analyse des objets (que ce soit des images ou des signaux) est leurs représentations. On désire souvent représenter ces objets de telle façon à ce que les informations qu'on veut extraire soient facilement accessibles. Dans le traitement d'images, pour aboutir à une classification capable de bien discriminer les différentes classes, on désire une représentation qui permet une extraction d'indices capables de minimiser les confusions entre les différentes classes.

Or l'analyse de Fourier propose une approche globale du signal, donc toute notion de localisation temporelle (ou spatiales pour les images) disparaît dans l'espace de Fourier. Il faut donc trouver un compromis, une transformation qui renseigne sur le contenu fréquentiel tout en préservant la localisation afin d'obtenir une représentation temps/fréquence ou espace/échelle pour les images.

II- Analyse de Fourier

II.1- Transformée de Fourier

La transformée de Fourier 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.

La transformée de Fourier d'une fonction f dans L^2 est :

$$\hat{f}(\omega) = \int_{-\infty}^{+\infty} f(t) e^{-i\omega t} dt \dots\dots\dots \text{II.1}$$

La transformée de Fourier inverse représente la fonction $f(t)$ comme une sommation de sinusoïdes :

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{f}(\omega) e^{i\omega t} d\omega \dots\dots\dots \text{II.2}$$

La transformée de Fourier peut être vue comme une représentation à base de sinusoïdes. Ces sinusoïdes sont très bien localisées en fréquence, mais pas en temps, car leur support est infini. C'est une conséquence de leur périodicité. La transformée de Fourier est une représentation globale du signal. Elle ne permet pas d'analyser son comportement fréquentiel local, ni sa régularité locale.

II.2- Transformée de fourier à fenêtre glissante [8]

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.1). Elle a deux arguments : le temps, et la fréquence. Elle est définie par :

$$Sf(u, \omega) = \int_{-\infty}^{+\infty} f(t) g(t-u) e^{-i\omega t} dt \dots\dots\dots \text{II.3}$$

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

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

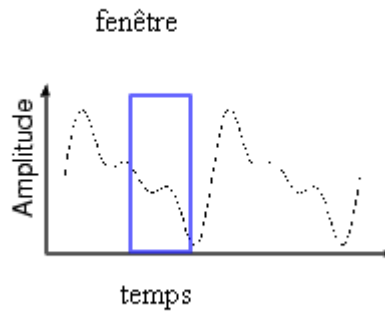


Figure II.1 : Transformée de Fourier à fenêtre

Une limitation importante de la transformée de Fourier à fenêtre glissante est que, du fait qu'une seule fenêtre est utilisée pour toutes les fréquences, la résolution de l'analyse est la même en tout lieu de l'espace temps-fréquence. Ceci constitue un handicap si on désire étudier des variations en basses fréquences durant une longue période et vis versa. Ce problème peut être résolu par la transformée en ondelettes.

III- La transformée en ondelettes

III.1. Exploitation de la corrélation

Etant donné un signal discret $f(x)$, caractérisé par des singularités que dans quelques points et représenté de façon constante par morceaux. Pour réduire l'espace mémoire de stockage, l'idée est de prendre une paire de points adjacents et de calculer leurs moyennes et leurs différences. Ainsi, la figure II.2, nous montre la décomposition du signal original en deux signaux : la moyenne et la différence. Le signal moyenne est une version grossière du signal original, quant au signal de différence code l'information perdue ou coefficients de différences lors du passage du signal original vers le signal moyenne. Il est possible de réitérer l'opération pour le signal moyenne, pour avoir d'autres signaux moyennes et différences. Cette transformation est connue sous le nom de transformation de Haar [9].

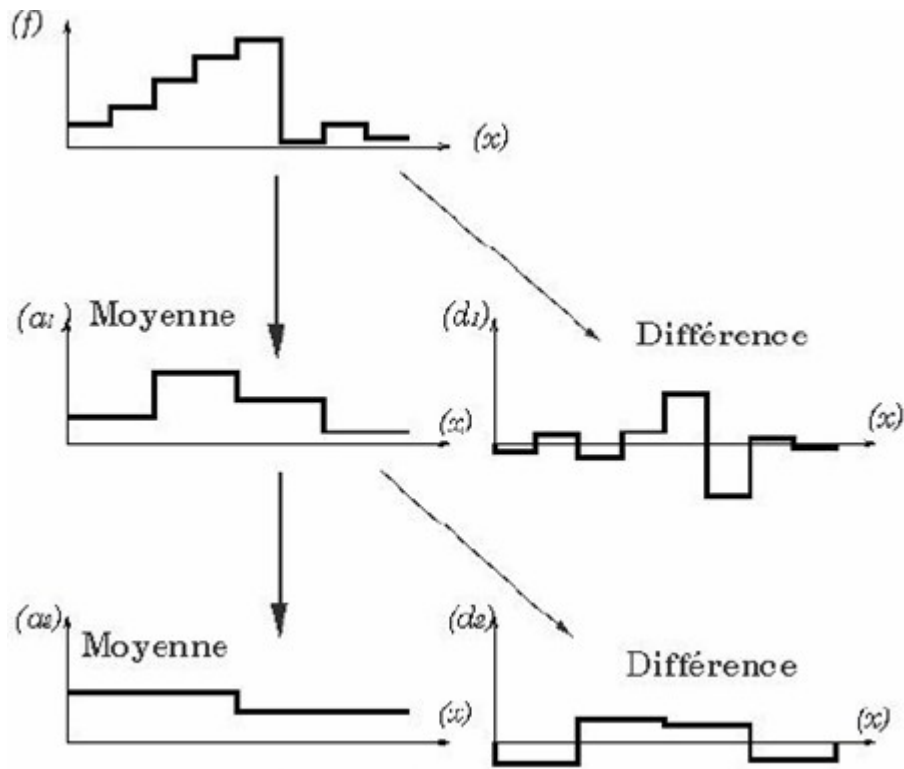


Figure II.2 : Décomposition d'un signal

En pratique, la transformation en ondelettes peut être considérée comme un processus de décomposition du signal original en approximations et en détails [10]. Le signal original $s(t)$, traverse deux filtres complémentaires, passe-haut et passe-bas, et émerge en temps que deux signaux : respectivement le signal d'approximations A et le signal de détails D (fig. II.3).

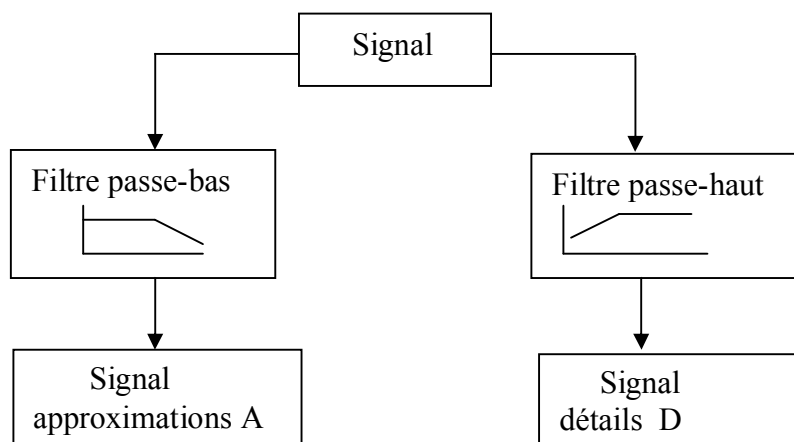


Figure II.3 : Décomposition du signal $s(t)$

III.2. Définition [11]

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 ψ . La forme générale de l'ondelette est donnée par :

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right); \quad u \in \mathbb{R}^+ \text{ et } s \in \mathbb{R}^{+*} \dots\dots\dots \text{II.4}$$

Deux opérations importantes appliquées à la fonction d'ondelette nous 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 , la deuxième opération est 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 $\psi(t)$.

La figure (II.4) montre l'effet de ces deux paramètres. L'ondelette utilisée est celle de Grossmann et Morlet [12], dont l'équation :

$$\psi(x) = \cos(5x) e^{-\frac{x^2}{2}} \dots\dots\dots \text{II.5}$$

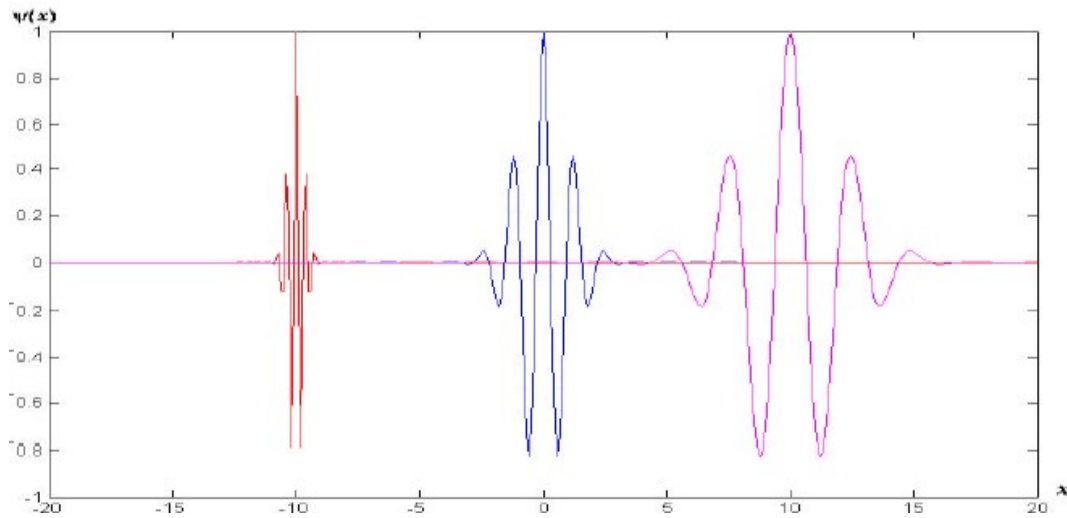


Figure II.4 : Effet de la dilatation et de la translation

L'ondelette au centre est la fonction de référence, elle correspond à une ondelette non dilatée et non traduite : $s = 1$ et $u = 0$, celle de gauche correspond à une ondelette comprimée avec $s = 0.3$ et traduite de $b = -10$, l'ondelette à droite, correspond à une dilatation de $s = 2$ et une translation de $b = 10$.

Rappelons qu'une ondelette est une fonction localisée aussi bien en temps qu'en fréquence. Ceci, se traduit dans le domaine fréquentiel par la condition d'admissibilité suivante :

$$C_\psi = \int_{-\infty}^{+\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty \quad \dots\dots\dots \text{II.6}$$

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

Dans la pratique, on utilise souvent une condition plus forte en imposant à l'ondelette un nombre p de moments nuls :

$$M_p = \int t^p \psi(t) dt = 0 \quad \dots\dots\dots \text{II.7}$$

III.3- Transformée en ondelettes continue [13]

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 :

$$WT_f(u, s) = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{s}} \bar{\psi}\left(\frac{t-u}{s}\right) dt \quad \dots\dots\dots \text{II.8}$$

$\bar{\psi}$: le complexe conjugué ψ

$WT_f(u, s)$: 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.

III.4 - Transformée en ondelettes discrète

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

$$\psi_{j,k}(t) = \frac{1}{\sqrt{s_0^j}} \psi\left(\frac{t - k u_0 s_0^j}{s_0^j}\right) \dots\dots\dots \text{II.9}$$

Où j et k sont entiers. Et on choisit généralement $s_0 = 2$ et $u_0 = 1$.

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

IV- Analyse multirésolution (AMR)

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 mulirésolution. 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 Meyer [14] ont défini la notion d'analyse multirésolution comme suit : Une suite $\{V_j\}_{j \in \mathbb{Z}}$ de sous-espaces vectoriels de $L^2(\mathbb{R})$ est une approximation multirésolution 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- $\lim_{j \rightarrow -\infty} V_j = \bigcup_{j \in \mathbb{Z}} V_j = L^2(\mathbb{R})$
- 6- $\exists \phi : \{\phi(\cdot - n)\}_{n \in \mathbb{Z}}$ forme une base orthonormée de V_0 .

Interprétation des propriétés

- 1- V_j est invariant par translation de 2^j : Il existe une grille temporelle sous-jacente par pas de 2^j .
- 2- Pour tout j , V_{j+1} est un sous-espace de V_j : Un signal basse résolution est aussi un signal à haute résolution.
- 3- V_{j+1} est l'image de V_j par une dilatation d'un facteur 2 : Il existe une grille fréquentielle sous-jacente en progression géométrique.
- 4- L'intersection des V_j est réduite à 0 dans L^2 : A résolution minimale, on perd toute l'image.
- 5- La réunion des V_j est dense dans L^2 : A résolution infinie, on reproduit parfaitement tous les signaux.
- 6- Chaque résolution est engendrée par une base de translatées.

IV-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 :

$$\Phi_{j,n}(x) = 2^{-\frac{j}{2}} \Phi(2^{-j}x - n); \quad (j,n) \in \mathbb{Z}^2 \dots\dots\dots \text{II.10}$$

La fonction Φ constitue une base orthogonale pour (V_j) .

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

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

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

$$\forall f \in L^2 : A_j f = \text{proj}_j(f) = \sum \langle f(x), \Phi_{j,n} \rangle \Phi_{j,n} \dots\dots\dots \text{II.11}$$

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 $a(j, n)$.

IV-2. Fonction d'ondelette et espace de détail

L'espace de détails 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 :

$$V_{j-1} = V_j \oplus W_j \quad \text{et} \quad V_j \perp W_j$$

$$\forall j, k \neq j \quad \text{alors} \quad W_j \perp W_k$$

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

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

Pour j fixé, 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.

IV.3. Fonction 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'échelles en dimension 1. La fonction d'échelle 2D obtenue est $\phi(x,y) = \phi(x)\phi(y)$ avec $\phi(x)$ la fonction d'échelle mono-dimensionnelle [15].

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

$$\psi^1(x, y) = \phi(x)\psi(y)$$

$$\psi^2(x, y) = \psi(x)\phi(y)$$

$$\psi^3(x, y) = \psi(x)\psi(y)$$

Avec l'analyse multirésolution séparable, une grande importance est donnée aux directions horizontale et verticale dans l'image. Pour de nombreux types d'analyse, cette sélectivité est utile. En pratique, la transformée est calculée en appliquant un banc de filtres à l'image (fig.II.5).

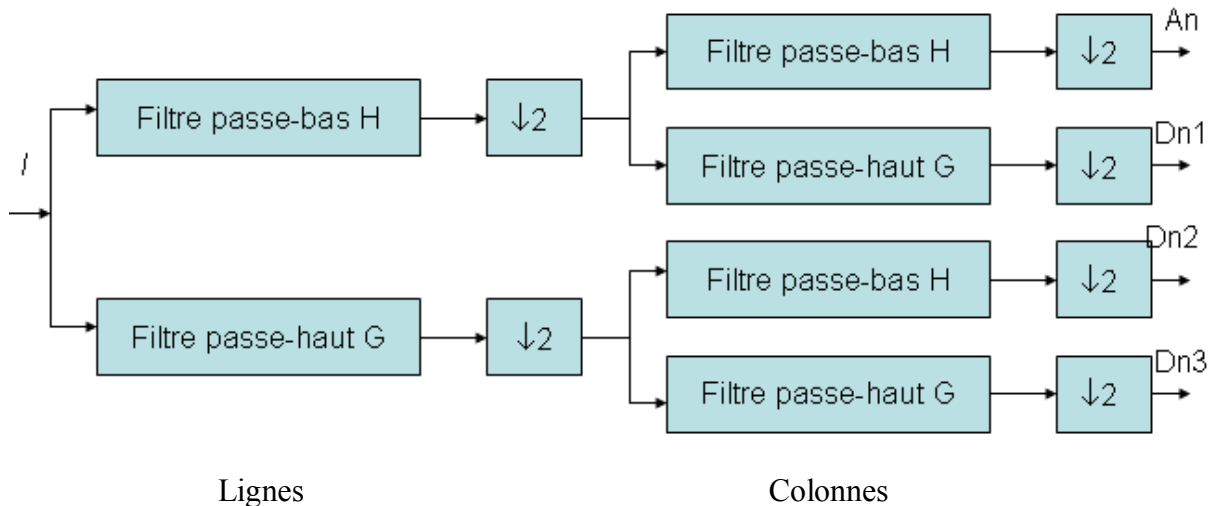


Figure II.5 : Banc de filtres à quatre sorties

$$A_n(b_i, b_j) = [H_x * [H_y * A_{n-1}] \downarrow_{2,1}] \downarrow_{1,2} (b_i, b_j)$$

$$D_{n1}(b_i, b_j) = [H_x * [G_y * A_{n-1}] \downarrow_{2,1}] \downarrow_{1,2} (b_i, b_j)$$

$$D_{n2}(b_i, b_j) = [G_x * [H_y * A_{n-1}] \downarrow_{2,1}] \downarrow_{1,2} (b_i, b_j)$$

$$D_{n3}(b_i, b_j) = [G_x * [G_y * A_{n-1}] \downarrow_{2,1}] \downarrow_{1,2} (b_i, b_j)$$

‘*’ 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 obtenue par filtrage passe-bas, et est donc l’image à basse résolution à l’échelle n .

Les coefficients D_{ni} sont obtenues 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 : $\{ A_n, D_{ni} \}$, $i=1,2,3$, 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 (voir figures II.6, II.7 et II.8).

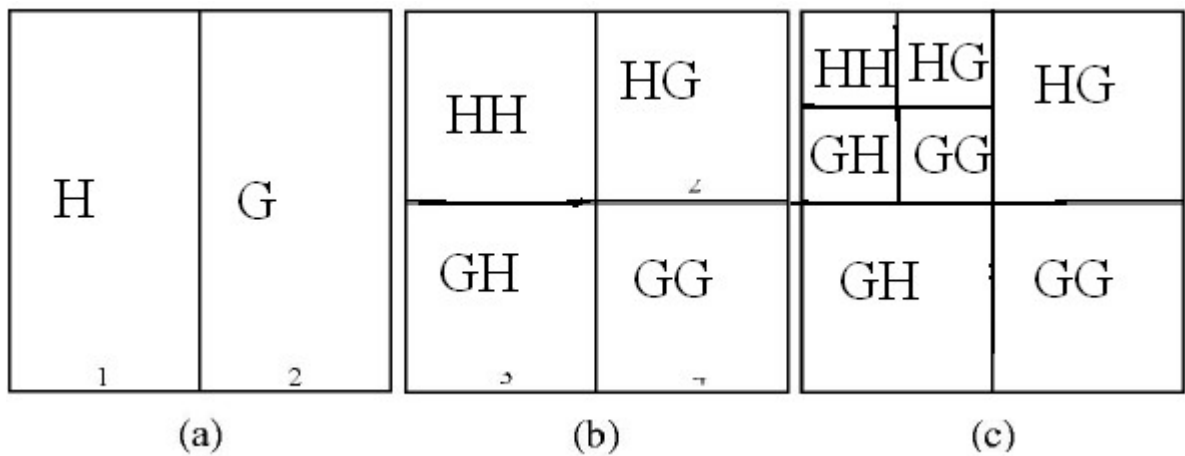


Figure II.6 : Organisation des sous images de détails et de l’image lissée à différentes résolutions

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 .

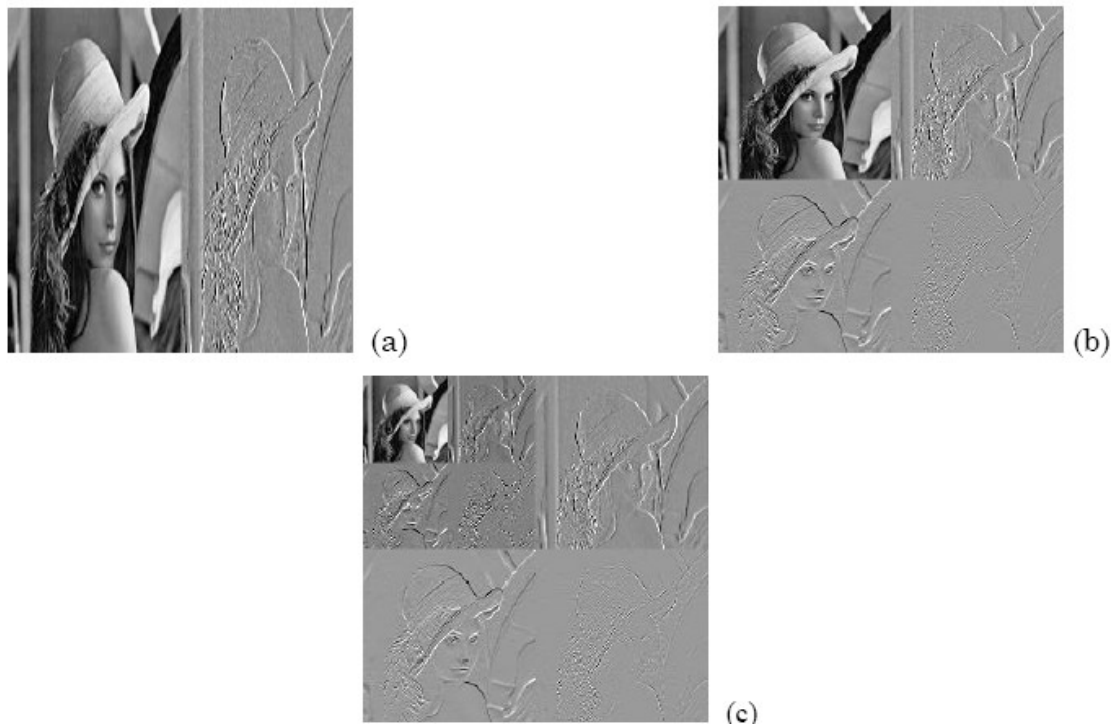


Figure II.7 : Exemple de décomposition d'une image.

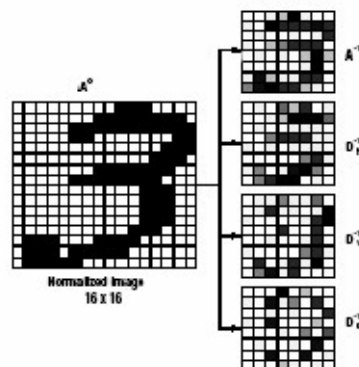


Figure II.8 : Exemple de décomposition d'un chiffre à la première résolution.

IV.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_1 sont obtenus par filtrage passe haut (filtre h_1) puis décimation par 2, les approximations sont obtenues de la même manière par filtrage passe bas (filtre g_1). Pour obtenir les coefficients de détails aux résolutions supérieures, il suffit de réitérer ces étapes sur les coefficients d'approximations.

V- Exemples d'ondelettes pratiques

1- Laplaciens itérés de Gaussienne [15]

Couramment appelé le chapeau mexicain, son expression est donnée par l'équation :

$$\Psi(x) = \left(\frac{2}{\sqrt{3}} \pi^{-\frac{1}{4}} \right) (1 - x^2) e^{-\frac{x^2}{2}} \dots\dots\dots \text{II}$$

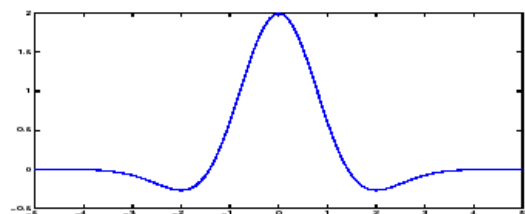
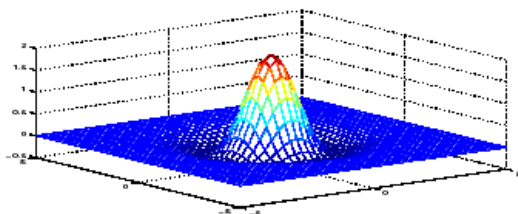


Figure II.9: Ondelette $h(x, y)$ (chapeau mexicain)

2-Ondelettes de Haar [16]

L'analyse de Haar est parmi les bases d'ondelettes orthonormées les plus populaires. Elle est appréciée en traitement d'images pour ses filtres linéaires en phase qui ont donc des propriétés de symétrie.

Ce qui correspond à la fonction :

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

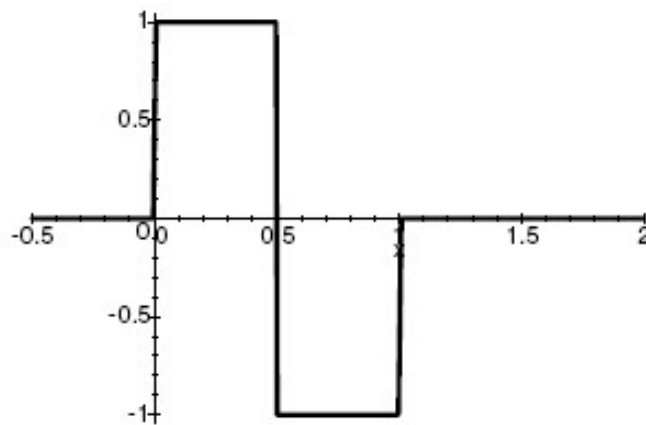


Figure II.10 : ondelette mère de l'analyse de Haar.

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

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 la compacité des fonctions de base et surtout des filtres associés est l'économie en coût 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'architecture.

La figure II.11 nous montre une représentation graphique d'une ondelette de Daubechies.

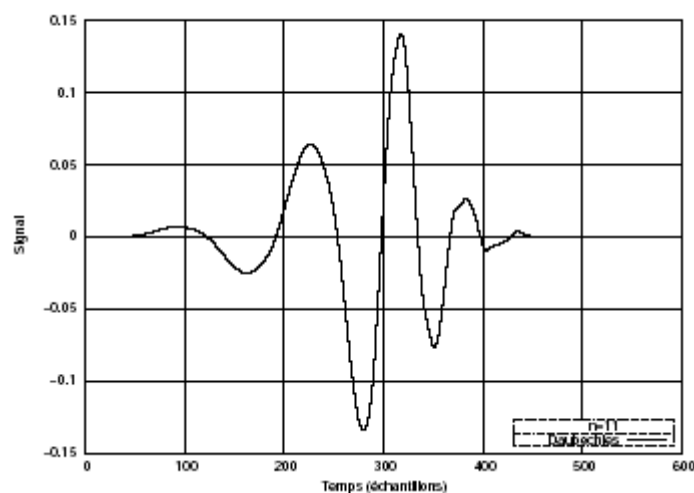


Figure II.11 : Ondelette de Daubechies

VI- Conclusion

A travers ce chapitre, nous avons fait une analyse plus au moins détaillée de la théorie des ondelettes, en commençant par la transformée en ondelettes continue, puis la transformée en ondelettes discrète, et enfin l'analyse multirésolution .

Nous avons pu conclure que l'utilisation des ondelettes permet de nombreuses applications au niveau de la reconnaissance des formes.

Dans le prochain chapitre, nous compléterons notre étude par le choix du classifieur neuronal le mieux adapté à notre schéma de reconnaissance.

I. Introduction

L'objectif de ce chapitre est multiple : il s'agit tout d'abord de rappeler les définitions fondamentales relatives aux réseaux de neurones ainsi que leurs propriétés mathématiques. Nous décrirons ensuite les principaux types de réseaux de neurones. Finalement, nous nous attacherons à détailler le type de réseaux de neurones utilisé dans notre mémoire, et plus particulièrement ses propriétés et sa mise en œuvre.

II. Le neurone biologique

Les réseaux de neurones artificiels ou formels sont à l'origine, une tentative de modélisation mathématique du cerveau humain. Les premiers travaux datent de 1943 et sont l'œuvre de MacCulloch et Pitts [21]. Ils présentent un modèle assez simple pour les neurones et explorent les possibilités de ce modèle. L'idée principale des réseaux de neurones artificiels est de donner une unité simple, un neurone, qui est capable de réaliser quelques calculs élémentaires.

Le modèle biologique illustré dans la figure III.1 décrit un modèle simple de neurone biologique qui a servi la mise en place des premiers neurones formels.

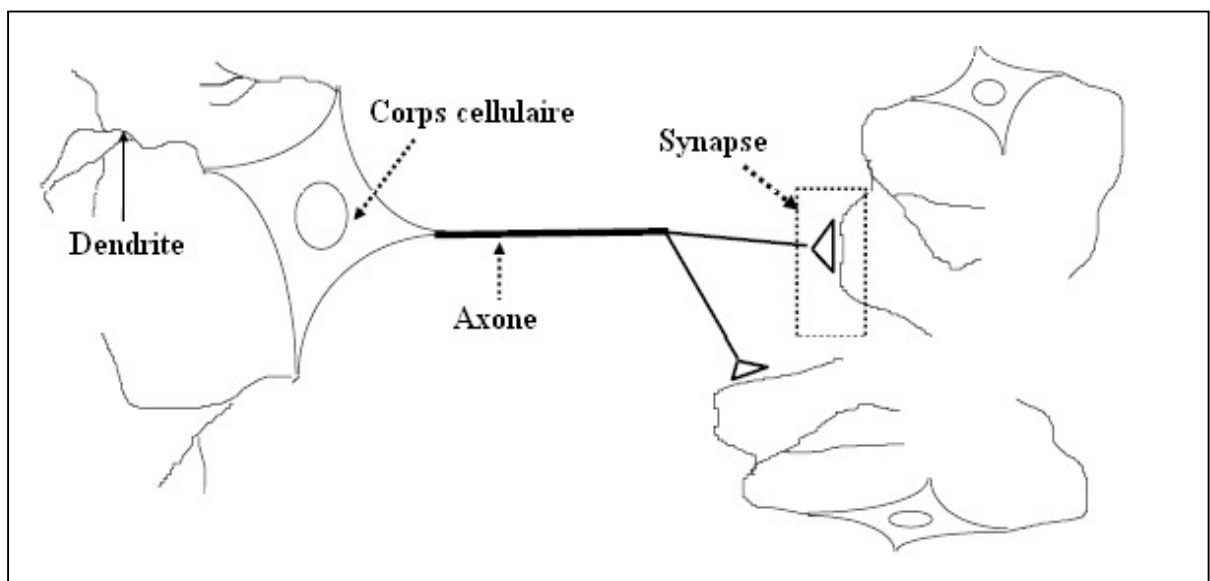


Figure III.1 : Neurones biologiques

Dans le cerveau, les neurones sont reliés entre eux par l'intermédiaire d'axones et de dendrites. En première approche, on peut considérer que ces filaments sont conducteurs d'électricité et peuvent véhiculer des messages depuis un neurone vers un autre. Les dendrites représentent les entrées du neurone et son axone sa sortie.

Un neurone émet un signal en fonction des signaux qui lui parviennent des autres neurones. On observe au niveau d'un neurone, une intégration des signaux reçus au cours du temps, c'est à dire une sorte de sommation des signaux. En général, quand la somme dépasse un certain seuil, le neurone émet à son tour un signal électrique. La notion de synapse explique la transmission des signaux entre un axone et une dendrite. Au niveau de la synapse, il existe un espace vide à travers lequel le signal électrique ne peut pas se propager. La transmission se fait alors par l'intermédiaire de substances chimiques.

III. Le neurone formel

Rappelons qu'un neurone formel est un processeur très simple, simulé sur ordinateur, imitant grossièrement la structure et le fonctionnement d'un neurone biologique. La figure III.2 montre la structure d'un neurone artificiel. Chaque neurone reçoit un nombre variable d'entrées en provenance des neurones amonts. A chacune de ces entrées est associée un poids représentatif de la force de la connexion.

Chaque neurone élémentaire est doté d'une fonction de transfert ou fonction d'activation qui donne une sortie unique Y_i et se ramifie ensuite pour alimenter un nombre variable de neurones avals.

Ainsi, les calculs de l'état interne et de la sortie s'expriment par les relations suivantes :

$$\text{Activité : } a_i = \sum_j w_{ij} x_j \dots\dots\dots \text{III.1}$$

$$\text{Sortie : } y_i = f(a_i) \dots\dots\dots \text{III.2}$$

Où : x_j est la sortie du neurone j connecté au neurone i

y_i , la sortie du neurone i

f est la fonction d'activation

a_i , l'activité du neurone i

w_{ij} désigne le poids synaptique de la connexion entrée j du neurone i .

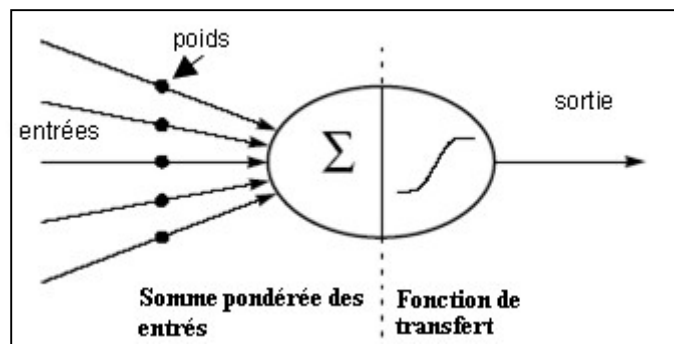


Figure III.2 : Modèle de neurone formel *i*

Il existe de nombreuses formes possibles pour la fonction de transfert [23]:

La figure III.3 présente des fonctions binaires à seuil qui correspondent à un mécanisme tout ou rien. Parmi celles-ci se trouve la fonction utilisée par McCulloch et Pitts.

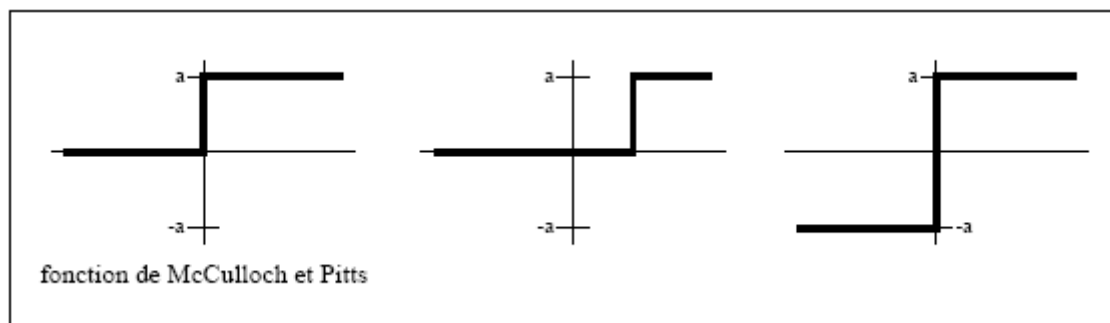


Figure III.3 : Exemples de fonctions binaires à seuil

La figure III.4 présente une généralisation de ces fonctions : les fonctions à saturation ou linéaires à seuil.

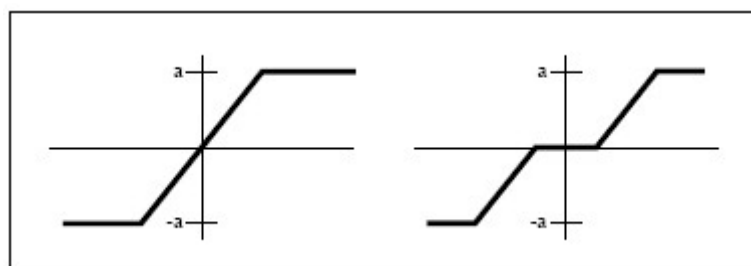


Figure III.4 : Exemples de fonctions à saturation

Les fonctions non linéaires dérivables représentées dans la figure III.5 sont actuellement les plus utilisées puisque la dérivabilité de ce type de fonctions est aujourd'hui une condition nécessaire au processus d'apprentissage.

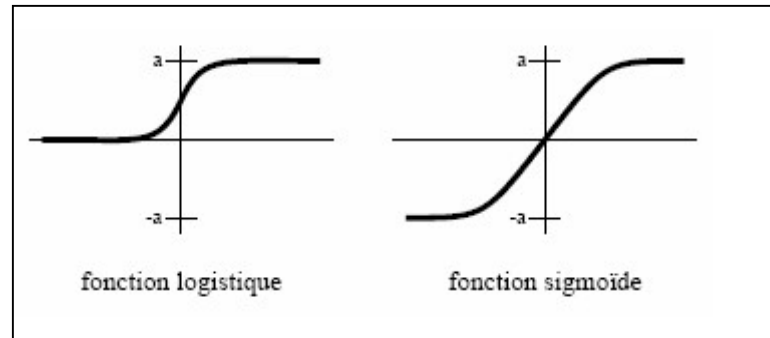


Figure III.5 : Fonctions non linéaires dérivables

La fonction logistique a pour expression : $f(x) = \frac{1}{1 + e^{-(k \times x)}}$ III.3

La fonction sigmoïde est définie par : $f(x) = \frac{e^{(k \times x)} - 1}{e^{(k \times x)} + 1}$ III.4

La plupart des réseaux de neurones ont une certaine capacité d'apprentissage. Cela signifie qu'ils apprennent à partir d'exemples. Le réseau peut ensuite dans une certaine mesure être capable de généraliser, c'est-à-dire de produire des résultats corrects sur des nouveaux cas qui ne lui avaient pas été présentés au cours de l'apprentissage.

IV. Présentation des réseaux de neurones :

Un neurone élémentaire est limité en ses applications. En effet, un neurone réalise une simple fonction non linéaire, paramétrée de ses variables d'entrée. L'intérêt des neurones réside dans la propriété qui résulte de leur association dans une structure, par une certaine logique d'interconnexion, formant un réseau de neurones. Le comportement collectif ainsi obtenu permet de réaliser des fonctions d'ordre supérieur par rapport à la fonction élémentaire réalisée par un neurone.

Dans un tel réseau, les entrées d'un neurone sont soit les entrées du réseau globale, soit les sorties d'autres neurones. Les valeurs des poids du réseau sont, en général, déterminées par une opération dite d'apprentissage.

IV.1. Structure d'interconnexion

Les connexions entre les neurones qui composent le réseau décrivent la topologie du modèle. Il existe deux familles de réseaux de neurones :

IV.1.1. Les réseaux bouclés [24]

Dans ce cas, Les sorties à l'instant t dépendent non seulement des entrées externes à l'instant t , mais aussi de toute la séquence des entrées externes précédentes (et de l'initialisation des variables d'état).

La structure d'un réseau de neurone bouclé est donnée par la figure ci-dessous :

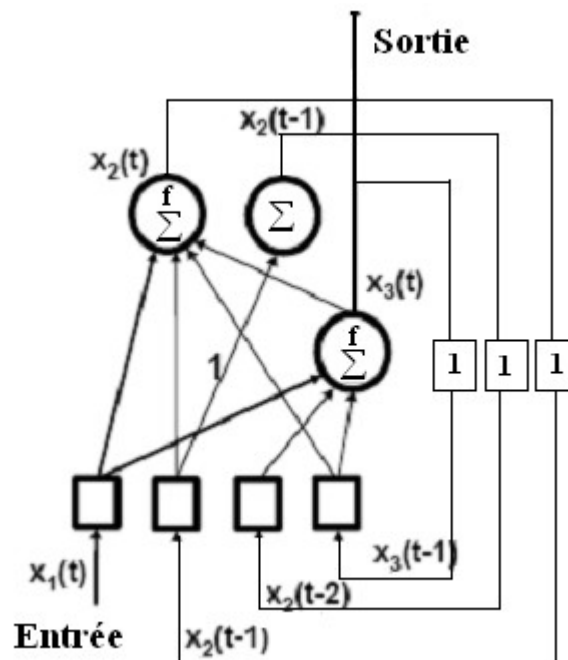


Figure III.6 : structure d'un réseau bouclé.

IV.1.2. Les réseaux non bouclés

Un réseau de neurones non bouclé réalise une ou plusieurs fonctions algébriques de ses entrées par composition des fonctions réalisées par chacun de ses neurones. Le réseau est représenté par un ensemble de neurones connectés entre eux. Dans un tel réseau, le flux de l'information circule des entrées vers les sorties sans retour en arrière.

Une autre façon d'assembler les neurones entre eux consiste à constituer des couches de neurones en interdisant toute connexion entre neurones de la même couche et aussi les connexions entre les neurones de deux couches non consécutives. Les réseaux de neurones à couches sont appelés perceptrons multicouches (MLP). Cette architecture est particulièrement utilisée dans notre travail car elle possède des propriétés intéressantes.

IV.2. L'apprentissage des réseaux de neurones

L'une des caractéristiques des réseaux de neurones est leur capacité à apprendre (par exemple à reconnaître une lettre, un son...). Mais cette connaissance n'est pas acquise dès le départ. La plupart des réseaux de neurones apprennent par l'exemple en suivant un algorithme d'apprentissage [25].

L'apprentissage est le but principal du développement de modèles à base de réseaux de neurones. Il est réalisé par la modification des poids de connexion du réseau, généralement par des algorithmes spécifiques utilisant des exemples d'apprentissage, afin d'obtenir des valeurs optimales appropriées à ces poids. A la fin de cette opération, on converge vers un fonctionnement de réseau, le plus possible adapté au problème à résoudre.

Une base d'apprentissage pour un réseau de neurones est constituée de N exemples. Chaque échantillon est composé d'un vecteur d'entrées et d'un vecteurs de sorties désirées correspondantes à l'entrée. Suivant la règle utilisée pour l'apprentissage, on distingue deux principaux types d'apprentissage : supervisé et non supervisé.

IV.2.1. L'apprentissage supervisé

L'apprentissage est dit supervisé si les différentes familles des formes sont connues à priori, et si la tâche d'apprentissage est guidée par un superviseur. Autrement dit, on fournit au réseau un couple (entrée, sortie désirée) et on modifie les poids en fonction de l'erreur entre la sortie désirée et la sortie obtenue par le réseau.

IV.2.2. L'apprentissage non supervisé

On l'appelle aussi classification automatique ou apprentissage sans professeur. Il s'agit de construire automatiquement les classes sans intervention de l'opérateur à partir d'échantillons de référence et de règles de regroupement. Dans ce type d'apprentissage, aucune connaissance à priori n'est fournie au réseau. A cet effet, le réseau s'auto organise pour détecter des points communs aux exemples présentés, par la modification des poids, afin de fournir la même sortie pour des entrées ayant des caractéristiques semblables.

IV.3. Propriétés des réseaux de neurones [26]

Diverses propriétés caractérisent les réseaux de neurones et font d'eux le meilleur choix pour résoudre beaucoup de problèmes, parmi elles, on cite :

- **La capacité de généralisation**

Le réseau de neurones sait résoudre plus qu'il n'a appris. En effet, il généralise à partir des exemples qui lui ont été fournis à condition qu'ils soient nombreux, et est en mesure de donner des réponses sensées à des cas de nouvelles données.

- **La rapidité**

L'utilisation des réseaux de neurones requiert une puissance informatique peu importante.

- **Une mémoire distribuée**

Une information n'est pas mémorisée dans un endroit précis (neurone ou connexion) d'un réseau de neurones. Comme dans le cerveau humain, la mémoire d'un fait est délocalisée en ce sens qu'elle est distribuée sur l'ensemble (ou tout au moins sur une partie) de la structure. Elle correspond à un état d'activation du réseau. La distribution de la

mémoire confère aux réseaux de neurones des propriétés importantes, comme la résistance au bruit ou à la dégradation.

- **La résistance au bruit**

Les réseaux de neurones permettent de traiter des données dégradées. En effet, des entrées entachées d'erreurs, imprécises ou incomplètes sont acceptables, à condition que la dégradation des données ne soit pas trop importante.

- **L'adaptabilité**

Un réseau de neurones évolue et s'adapte au fur et à mesure des changements constatés sur des problèmes évolutifs.

IV.4. Quelques modèles de réseaux de neurones

IV.4.1. Le perceptron multicouche

Parmi les types de réseaux de neurones les plus utilisés, on trouve le perceptron multicouches (MLP) avec son algorithme d'apprentissage supervisé, la rétropropagation du gradient.

Le perceptron multicouche, auquel on s'est intéressé, est un réseau de neurones artificiels organisé en couches, où l'information circule dans un seul sens, de la couche d'entrée vers la couche de sortie (figure III.7).

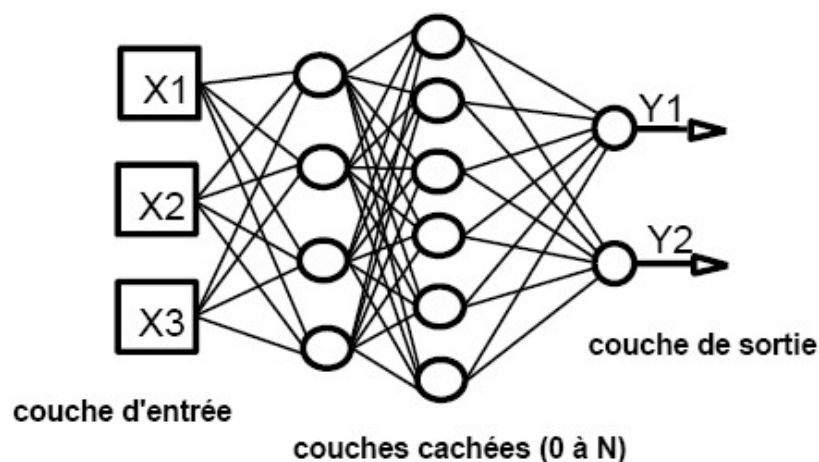


Figure III.7 : réseau multicouche.

IV.4.1.1. Mise en œuvre du perceptron multicouche

La mise en œuvre du réseau de neurones comporte une partie conception, dont l'objectif est de permettre de choisir la meilleure architecture possible, et une partie de calcul numérique pour réaliser l'apprentissage du réseau.

Dans le cas général, un perceptron multicouches peut posséder un nombre quelconque de couches et un nombre de neurones par couche également quelconque. En vue de perfectionner le fonctionnement du perceptron multicouches d'un côté et minimiser au maximum le temps de calcul d'autre part, on doit chercher, par test, une architecture optimale du point de vue nombre de couches et nombre de neurones par couche.

A partir d'une architecture de réseau de neurones donnée et des exemples disponibles, on détermine les poids optimaux, par l'algorithme de la rétropropagation du gradient de l'erreur, pour que la sortie du modèle s'approche le plus possible du fonctionnement désiré.

IV.4.1.2. la rétropropagation du gradient de l'erreur

L'apprentissage du perceptron multicouches se fait grâce à la méthode de la rétropropagation du gradient de l'erreur [27]. Cette méthode permet de rétropropager une erreur exacte sur plusieurs couches de connexions et d'adapter ces connexions par ajustement pour minimiser l'erreur. La figure (III.8), nous montre la structure d'un réseau multicouche caractérisé par :

La couche d'entrée à M cellules.

La couche cachée (h_1) à L_1 cellules.

La couche cachée (h_2) à L_2 cellules.

La couche de sortie (S) à N cellules.

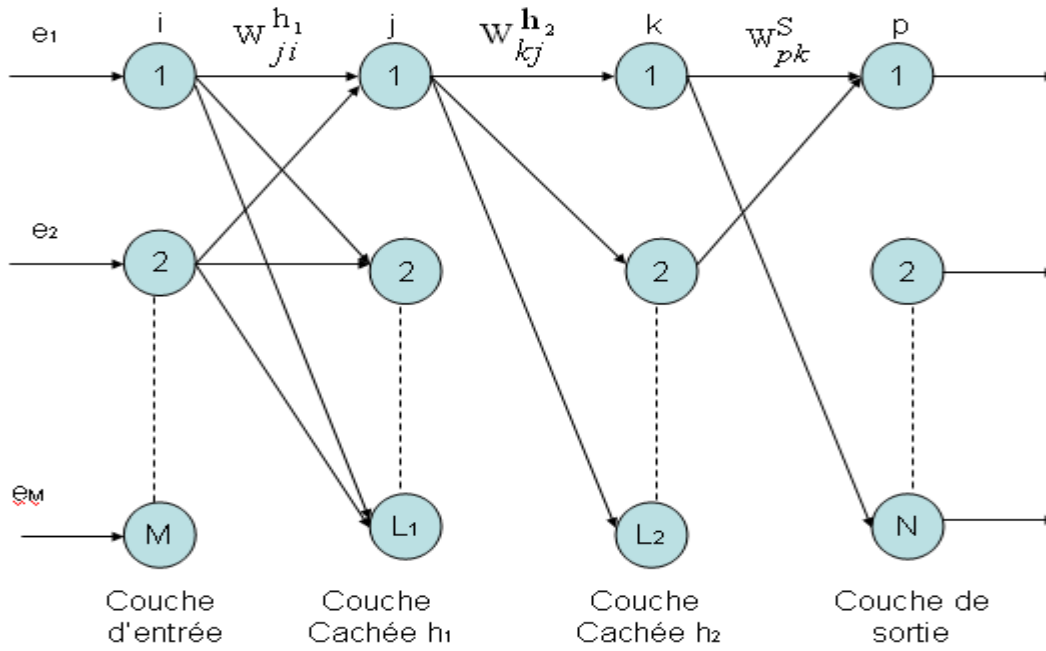


Figure III.8 : Structure d'un réseau multicouches

$w_{ji}^{h_1}$: Poids synaptique de la cellule i de la couche d'entrée avec la cellule j de la couche cachée (h_1).

$w_{kj}^{h_2}$: Poids synaptique de la cellule j de la couche cachée (h_1) avec la cellule k de la couche cachée (h_2).

w_{pk}^S : Poids synaptique de la cellule k de la couche cachée (h_2) avec la cellule p de la couche de sortie (S).

$$X_j^{h_1} = f \left(\sum_{i=1}^M w_{ji}^{h_1} e_i \right) : \text{Sortie de la cellule } j \text{ de la couche cachée } (h_1).$$

$$X_k^{h_2} = f \left(\sum_{j=1}^{L_1} w_{kj}^{h_2} X_j^{h_1} \right) : \text{Sortie de la cellule } k \text{ de la couche cachée } (h_2).$$

$$X_p^S = f \left(\sum_{k=1}^{L_2} w_{pk}^S X_k^{h_2} \right) : \text{Sortie de la cellule } p \text{ de la couche de sortie (S)}.$$

Algorithme

- 1) Initialiser les poids synaptiques par des valeurs aléatoires faibles.

$$W_{ji}, W_{kj}, W_{pk} \quad i: 1 \rightarrow M$$

$$j: 1 \rightarrow L_1$$

$$k: 1 \rightarrow L_2$$

$$p: 1 \rightarrow N$$

- 2) Appliquer un vecteur d'entrée $e = (e_1, e_2, \dots, e_M)$

- 3) Calculer : $X_j^{h_1}$, $X_k^{h_2}$, X_p^S

- 4) Calculer l'erreur quadratique de la couche de sortie :

$$E = \frac{1}{2} \sum_{p=1}^N \left(X_{dp}^S - X_p^S \right)^2$$

- 5) Calculer le gradient d'erreur de chaque cellule de la couche de sortie :

$$\Delta W_{pk}^S = \left(X_{dp}^S - X_p^S \right) f' \left(\sum_{k=1}^{L_2} W_{pk} X_k^{h_2} \right) \cdot X_k^{h_2}$$

- 6) Calculer le gradient d'erreur de chaque cellule des couches cachées :

$$\Delta W_{kj}^{h_2} = f' \left(\sum_{j=1}^{L_1} W_{kj}^{h_1} X_j^{h_1} \right) \cdot X_j^{h_1} \cdot \sum_{p=1}^N \Delta W_{pk}^S \cdot W_{pk}^S$$

$$\Delta W_{ji}^{h_1} = f' \left(\sum_{j=1}^M W_{ji}^{h_1} e_i \right) \cdot e_i \cdot \sum_{k=1}^{L_2} \Delta W_{kj}^{h_2} \cdot W_{kj}^{h_2}$$

7) Adaptation des poids de la couche de sortie :

$$W_{pk}^S(t+1) = W_{pk}^S(t) + \eta \Delta W_{pk}^S$$

8) Adaptation des poids des couches cachées :

$$W_{kj}^{h_2}(t+1) = W_{kj}^{h_2}(t) + \eta \Delta W_{kj}^{h_2}$$

$$W_{ji}^{h_1}(t+1) = W_{ji}^{h_1}(t) + \eta \Delta W_{ji}^{h_1}$$

9) Refaire les étapes 2 à 9 jusqu'à ce que $E \leq \varepsilon \leftarrow$ seuil.

Aspects pratiques de l'algorithme

La vitesse de convergence dépend essentiellement du pas d'apprentissage (η). Il varie entre 0 et 1. S'il est proche de 1, l'algorithme risque de diverger et s'il est très petit, l'apprentissage devient lent.

Pour accélérer la convergence, on utilise en pratique :

$$W_{pk}^S(t+1) = W_{pk}^S + \eta \Delta W_{pk}^S(t) + \alpha \Delta W_{pk}^S(t-1)$$

Avec : $0.9 \leq \alpha \leq 0.99$

$\alpha \Delta W_{pk}^S(t-1)$ représente le momentum.

IV.4.2. Le modèle de Hopfield

Il s'agit d'un réseau constitué de neurones à deux états (-1 et 1, ou 0 et 1), Chaque neurone du réseau est à la fois une entrée et une sortie, le réseau implantant ainsi une mémoire associative. Le processus d'apprentissage de ses connexions fait appel à la règle de Hebb [28] qui renforce la connexion entre deux neurones s'ils sont actifs simultanément (figure III.8).

Il est un des rares modèles connexionnistes à avoir des propriétés connues et démontrées. La plus intéressante de ces propriétés est la connaissance à priori du nombre de formes que le réseau pourra mémoriser. Un réseau de Hopfield de N neurones pourra ainsi mémoriser $0,14N$ formes distinctes.

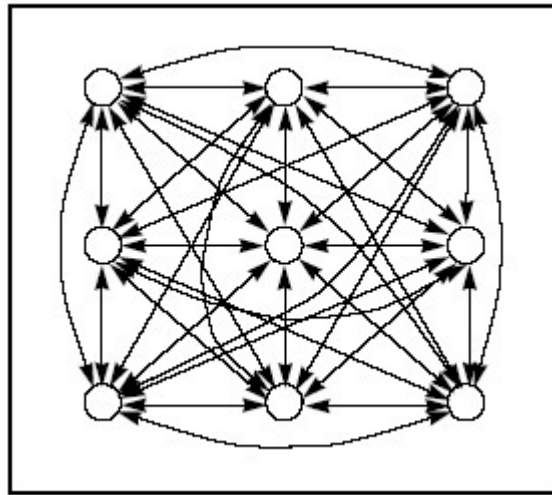


Figure III.9 : Architecture d'un réseau Hopfield

IV.4.3. Le modèle de Kohonen

Contrairement aux réseaux de Hopfield où les neurones sont modélisés de la façon la plus simple possible, on recherche ici un modèle de neurone plus proche de la réalité. En effet, le modèle de Kohonen trouve son inspiration dans une structure locale particulière existant dans certaines aires du cortex (figure III.9). Les neurones sont organisés en couches. A l'intérieur de chaque couche, chaque neurone émet vers ses voisins les plus proches des connexions excitatrices et vers les neurones plus éloignés des connexions inhibitrices. Tous les neurones reçoivent les mêmes entrées.

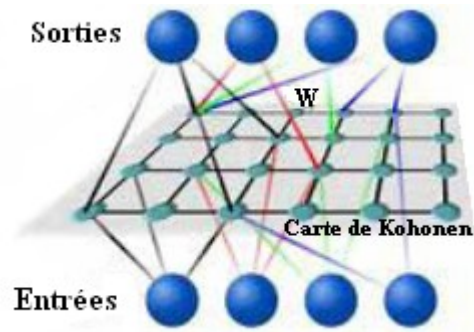


Figure III.10 : Réseaux de Kohonen

Kohonen a simulé le fonctionnement de ce type de réseau [29], en utilisant un modèle de neurones qui répond d'autant plus fortement que le vecteur d'entrée est proche de son vecteur poids.

Lorsqu'on présente une entrée, les sorties évoluent sous l'effet des connexions intra-couches, puis se stabilisent rapidement de telle sorte que le neurone qui répondait initialement le mieux à l'entrée, ainsi que ses voisins soient actifs, alors que les autres neurones sont inhibés.

V- Conclusion

A travers ce chapitre, nous avons exploré la capacité des réseaux de neurones artificiels dans le domaine de la modélisation physique. Du fait de leurs forte non linéarité et l'aspect de généralisation qu'ils présentent, les réseaux de neurones artificiels sont de bons candidats à la reconnaissance de chiffres manuscrits.

Pour notre application, nous avons choisi un réseau de neurones multicouches utilisant l'algorithme d'apprentissage la rétropropagation du gradient de l'erreur.

I- Introduction

L'objectif de ce chapitre constitue à présenter les résultats effectués sur un fichier de 2007 imageries de chiffres manuscrits composés de 10 classes (0 à 9).

II- Présentation de la base de données USPS (de l'anglais US Postal Service) :

Nous avons travaillé sur des chiffres stockés sur disque sous la forme d'imageries en niveau de gris de taille 16x16 pixels. Ces chiffres sont extraits de la base de données USPS. Celle-ci est construite à partir de chiffres manuscrits saisis à partir d'images d'enveloppes collectées au centre CEDAR à Buffalo (États Unis). La base de données contient 9298 images de chiffres dont 7291 images d'apprentissage servent à construire le classifieur et 2007 images de test servent à tester ses performances. Il est connu que l'ensemble de test de l'USPS est plutôt difficile puisque l'erreur humaine se situe autour de 2.5%. Dans ce mémoire, une fraction de 500 images a été extraite à partir de 7291 pour servir à l'apprentissage et pour le test, nous avons utilisé tous les prototypes (2007 images).

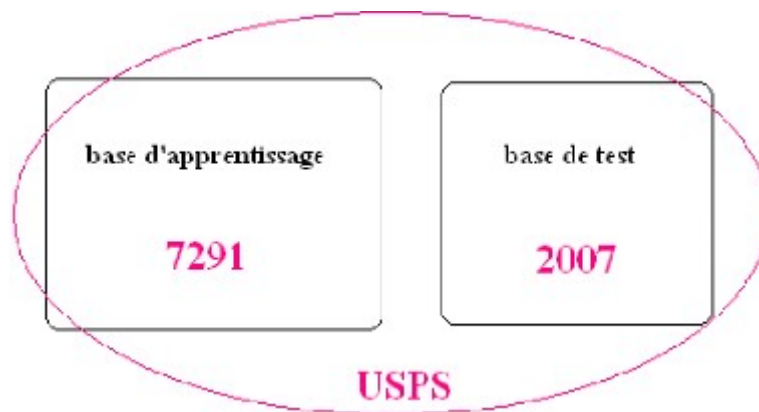


Figure IV.1 : Représentation de la base USPS.

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

III- Choix du réseaux de neurones

Notre choix du réseaux de neurones est porté sur le perceptron multicouches. En effet, la connaissance à priori de la réponse désirée ne peut pas être négligée. De ce fait, le perceptron multicouche est le meilleur choix puisqu'il nous permet d'exploiter cette connaissance et ce à l'aide de l'algorithme de rétropropagation.

Les paramètres du réseaux sont :

- Le nombre de couches cachées.
- Le nombre de neurones dans chaque couche.
- Le nombre d'itérations.
- Le pas d'apprentissage.
- L'erreur.
- Le momentum.

Il n'y a pas de règles générales pour le choix de ces critères. Tout se fait donc par test. Néanmoins, on connaît le nombre de neurones de la couche d'entrée qui sera le vecteur de caractéristiques obtenu de l'étape d'analyse, et le nombre de neurones de la couche de sortie qui sont au nombre de dix classes (0 à 9).

On a également fixé deux critères d'arrêt :

- Le nombre d'itérations : une fois le nombre d'itérations est atteint la simulation s'arrête. Ceci nous évitera toute divergence du système.
- L'erreur : si le système atteint la valeur de l'erreur, la simulation s'arrête évitant ainsi une perte de temps dans des itérations inutiles.

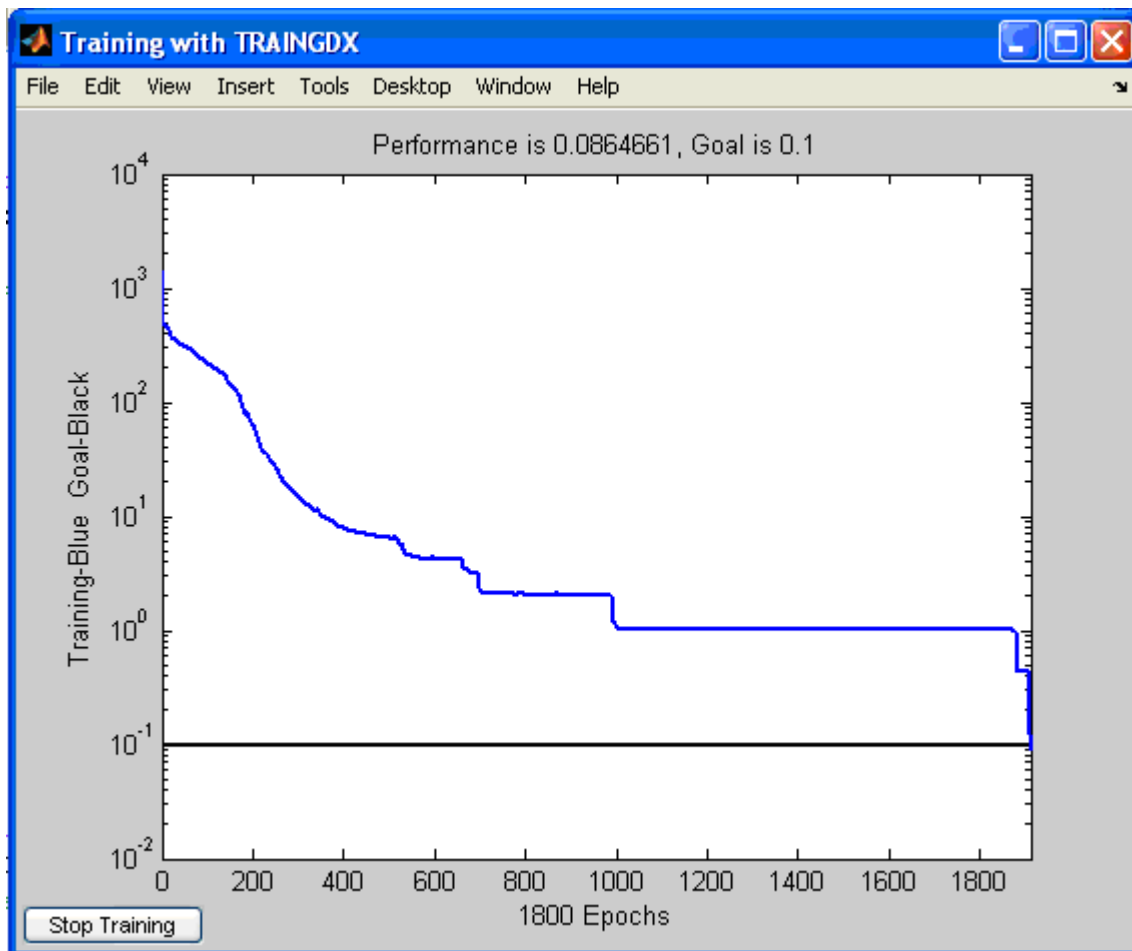


Figure IV.4 : L'apprentissage du réseau de neurone

IV- Tests

Nous avons effectué une étude comparative en implémentant 3 types d'ondelettes avec différentes approches pour chacune. Nous avons également fait un test en injectant des données brutes au réseau de neurones .

IV.1. Données brutes

Sur le tableau IV.1, nous présentons les taux de reconnaissance obtenus après avoir injecté les données brutes au réseau de neurones .

<i>Taux de reconnaissance</i>			
<i>Chiffre</i>	<i>Nombre total</i>	<i>Nombre de chiffres bien classés</i>	<i>Taux de reconnaissance</i>
0	359	291	81,07%
1	264	239	90,51%
2	198	116	58,58%
3	166	130	78,31%
4	200	152	76,00%
5	160	106	66,25%
6	170	146	85,88%
7	147	123	83,67%
8	166	118	71,08%
9	177	144	81,35%
<i>Totaux</i>	2007	1565	77,97%

Tableau IV.1

IV.2. Ondelettes de Haar

Cette approche exploite la première résolution de l'analyse de Haar. Le vecteur caractéristique est composé des coefficients de l'approximée. Les taux de reconnaissance sont indiqués pour chaque chiffre dans le tableau IV.2 :

<i>Taux de reconnaissance</i>			
<i>Chiffre</i>	<i>Nombre total</i>	<i>Nombre de chiffres bien classés</i>	<i>Taux de reconnaissance</i>
0	359	295	82,17%
1	264	240	90,90%
2	198	130	65,65%
3	166	130	78, 31%
4	200	130	65,00%
5	160	112	70,00%
6	170	134	78,82%
7	147	117	79,59%
8	166	102	61,00%
9	177	143	80,79%
<i>Totaux</i>	2007	1533	76,38%

Tableau IV.2

IV.3 Ondelettes de Daubechies

Dans cette approche, l'extraction du vecteur de caractéristiques se fait par application de l'ondelette DB-4 . les taux de reconnaissance obtenus par cette méthode sont notés dans le tableau IV.6

<i>Taux de reconnaissance</i>			
<i>Chiffre</i>	<i>Nombre total</i>	<i>Nombre de chiffres bien classés</i>	<i>Taux de reconnaissance</i>
0	359	288	80,22%
1	264	243	92,04%
2	198	142	71,71%
3	166	116	69,87%
4	200	144	72,00%
5	160	93	58,15%
6	170	147	86,47%
7	147	125	85,34%
8	166	111	66,87%
9	177	154	87,00%
<i>Totaux</i>	2007	1563	77,87%

Tableau IV.6

IV.4. Ondelettes Bior 3.7

Un autre type d'ondelettes est également implémenté pour le réseau de neurones. C'est l'ondelette Bior 3.7. Nous avons extrait un vecteur composé par la première résolution [ap1]. Les résultats sont dans le tableau IV.7

<i>Taux de reconnaissance</i>			
<i>Chiffre</i>	<i>Nombre total</i>	<i>Nombre de chiffres bien classés</i>	<i>Taux de reconnaissance</i>
0	359	306	85,23%
1	264	245	92,80%
2	198	151	76,26%
3	166	127	76,50%
4	200	152	76,00%
5	160	89	55,62%
6	170	145	85,29%
7	147	114	77,55%
8	166	121	72,89%
9	177	145	81,92%
<i>Totaux</i>	2007	1595	79,47%

Tableau IV.7

V- Conclusion

A travers les différents tests effectués, nous avons pu constater que les taux de reconnaissance obtenus sont satisfaisants comparativement à ceux obtenus par d'autres méthodes appliquées à la reconnaissance de chiffres manuscrits.

Ainsi, l'exploitation des ondelettes comme descripteurs dans le processus de reconnaissance s'est avérée très encourageante.

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

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

Dans la phase de prétraitement, le choix s'est porté sur une simple normalisation des données, les images de chiffres manuscrits sont ainsi plus lisibles.

Dans l'étape d'analyse, l'utilisation des ondellettes pour l'extraction des caractéristiques, s'est avérée une approche adéquate. Elle permet le choix de paramètres pertinents et évite ainsi une perte d'informations qui va se répercuter négativement sur le taux de reconnaissance.

Pour l'apprentissage, nous avons opté pour un réseau de neurones de type perceptrons multicouches, un choix judicieux et appuyé par les différentes théories sur le sujet. Nous avons également effectué des essais en vu de régler les paramètres du réseau.

Ainsi ,nous avons pu concevoir un système de reconnaissance automatique des chiffres manuscrits, avec un certain taux de reconnaissance assez acceptable, en comparant avec les différents résultats obtenus par d'autres méthodes, quelques fois plus performantes, car elles utilisent une combinaison de différents algorithmes pour garantir un taux optimal. Ce que nous n'avons pas pu concrétiser faute de temps limité.

Enfin, ce travail m'a permis d'approfondir mes connaissances théoriques et pratiques, dans le domaine de traitement d'images en général, et d'ajouter un plus à mon cursus de formation.

Bibliographie

- [1] G. R. CEHEUX (nom collectif), Stratégie pour l'interprétation de documents
- [2] Denis ARRIVault, Thèse de doctorat en traitement du signal et des images : Apport des Graphes dans la reconnaissance non contrainte de caractères manuscrits anciens, Université de Poitiers, Mars 2006.
- [3] International Journal on Document Analysis and Recognition, revue 2003.
- [4] Serge BEUCHER, Thèse de doctorat en morphologie mathématiques : Segmentation d'images et morphologie mathématique, Ecole Nationale des Mines de Paris, mars 1990.
- [5] Laurent Guigues, Thèse de doctorat en traitement d'images et du signal : Modèles multiéchelles pour la segmentation d'images, Université de Cergy-Pontoise, Décembre 2003.
- [6] Djamel GACEB, Master de recherche : Extraction de similarités de formes dans les images de traits, INSA de Lyon, Juin 2005.
- [7] Nicolas Ferlay, Thèse de doctorat en physique : Application de la méthode de Galerkin-Ondelette à l'Equation du Transfert Radiatif, Université Blaise Pascal, Clermont-ferrand, juillet 2003.
- [8] Valérie Perrier, Application de la théorie des ondelettes, Institut National Polytechnique de Grenoble, mars 2005.
- [9] Principes de Télécommunications : Les ondelettes, mai 2007.
- [10] Samuel Foucher, Application de la transformation en ondelette dans un espace géométrique, Centre de Recherche en Informatique de Montréal, Canada, 2003.

- [11] Yacine Oussar, Thèse de doctorat en robotique : Réseaux d'ondelettes et réseaux de neurones pour la modélisation statique et dynamique de processus, Université Pierre et Marie Curie, Juillet 1998.
- [12] Simon Bernard, Master Informatiques, Mathématiques et leurs Applications : Extraction de primitives structurelles pour la reconnaissance de symboles : Une approche robuste par Transformée de Hough, Université de la rochelle,
- [13] Frédéric Truchetet, Ondelettes pour le signal numérique, Editions Hermes, Paris, 1998.
- [14] S. Mallat, Multiresolution approximations and wavelet orthonormal bases of $L_2(\mathbb{R})$, 1989.
- [15] L'analyse multirésolution en vision informatique, 1992.
- [16] J.P.Gourret, Maillage multirésolution de surfaces pour la modélisation d'images, 2000.
- [17] François Mailo, Une construction d'ondelettes de Daubechies, Edition Hermes, 1998.
- [18] S. W. Lee, C.H. Kim, Y. Y. Tang : Multiresolution Recognition of Unconstrained Handwritten Numerals with Wavelet Transform and Multilayer Cluster Network, 1996.
- [19] U. Bhattacharya T. K. Das B. B. Chaudhuri : A Cascaded Scheme for Recognition of Handprinted Numerals , Kolkata, India, 2004.
- [20] Suzette. N, Correia, M. De Carvalho : recognition of unconstrained handwritten numerals using biorthogonal spline wavelets, Brazil, 2000..
- [21] G.Dreyfus, J.M.Martinez, L.Herold : Réseaux de neurones : méthodologie et applications, Edition Eyrolles, 2004.
- [22] P. Wunsch and A. F. Laine : Wavelet Descriptors for Multiresolution Recognition of Handprinted Characters, Pattern Recognition, 1995.

[23] T. Cour G. Giraud, A. Kodsi, T.A. Luong : Reconnaissance de formes par réseau de neurones, 2002

[25] Stéphanie GUILLAS : Thèse de doctorat en informatique, reconnaissance d'objects graphiques détériorés , la rochelle, 2007

[26] Laurent BUNIET : Thèse de doctorat en informatique, Traitement automatique de la parole en milieu bruité : étude de modèles connexionnistes statiques et dynamiques, Université de Nancy, 1997.

[27] Applications du perceptron multicouche, Cours, Université de Québec, 2006

[28] G. Y. Chen, T. D. Bui and A. Krzyzak: Contour-Based Handwritten Numeral Recognition using Multiwavelets and Neural Networks, 2003.

[29] Deux Approches Neuronales pour la Vérification Hors-Ligne de la Signature Manuscrite, ENI, Laboratoire de Systèmes et Traitement de Signal, Tunis, 2004.