

UNIVERSITE SAAD DAHLEB DE BLIDA

Faculté des Sciences de l'Ingénieur

Département d'Electronique

MEMOIRE DE MAGISTER

Spécialité : Images & Paroles

L'ANALYSE DES DOCUMENTS PAR LA METHODE DES ONDELETTES

Par

Faiza ABDAT

Devant le jury composé de :

H.Salhi	Maître de Conférences, U. de Blida	Président
A.Guessoum	Professeur, U. de Blida	Examineur
K.Achour	Directeur de recherche (CDTA)	Examineur
Z.Benslama	Chargé de cours, U. de Blida	Examineur
N.Khorissi	Chargé de cours, U. de Blida	Rapporteur
A.Sehad	Chargé de recherche (CDTA)	Co-Rapporteur

Blida, Septembre 2005

RESUME

Les documents électroniques sont souples en matière de stockage et de recherche d'information, pour cela il est nécessaire de passer du format papier vers un format électronique, cette conversion est souvent réalisée par un système d'analyse de documents. Ce dernier est composé de plusieurs parties : la numérisation, le prétraitement (élimination du bruit, la détection et la correction de l'inclinaison,...), l'analyse de la structure physique (segmentation et classification) et l'analyse de la structure logique.

Notre travail consiste à détecter l'inclinaison en utilisant la transformée de Hough, et à analyser la structure physique en utilisant la transformée en ondelettes qui se base sur la texture particulière du texte, ainsi que sur l'utilisation de l'analyse multi-échelle, pour rendre compte de l'information statique. Nous affinons notre travail avec un post traitement afin d'améliorer le résultat de la segmentation.

Mots clés : analyse et reconnaissance des documents, détection et correction de l'inclinaison, analyse de la structure physique, segmentation, classification et analyse de la structure logique, Transformée de Hough, Transformée en ondelettes.

الملخص :

الوثائق الإلكترونية سهلة الإستعمال من حيث الحفظ و البحث لهذا يجب المرور من الوثائق المطبوعة إلى تلك الإلكترونية, هذا المرور يتم غالبا بواسطة نظام تحليل الوثائق, هذا الأخير يتكون من عدة أجزاء : تصوير الوثائق, المعالجة الأولية (إزالة المناطق الدخيلة, تصحيح ميلان الوثيقة) تحليل البنية الفيزيائية (تقسيم و تصنيف), تحليل البنية المنطقية.

تقتصر دراستنا على تصحيح ميلان الوثيقة بإستعمال محولة هوق و على تقسيم الورقة بإستعمال محولة المويجات التي تعتمد على النسج الخاص للنص. بغية تحسين نوعية التقسيم أجرينا معالجة نهائية.

مفاتيح :التحليل و التعرف على الوثائق, تصحيح ميلان الوثيقة, تحليل البنية الفيزيائية, تقسيم , تصنيف, تحليل البنية المنطقية, محولة هوق, محولة المويجات.

ABSTRACT

Electronic documents are very flexible in storage and information research. For that purpose, it is necessary to go from the printed format to the electronic format; this conversion is often realized by a document analysis system, which is composed of: digitalisation, pre-processing (noise reduction, binarisation, skew detection and skew correction), physical structure analysis (segmentation and classification) and logical structure.

Our purpose concerns in the estimation of the skew angle with the Hough transform and the analysis of the physical structure by using the wavelet transform, which is based on the particular texture of the text, as well as the use of the multi-scale analysis to account for static information. This work is refined by a post treatment in order to improve the results of segmentation.

Key words: document analysis and recognition, skew estimation, physical structure analysis, segmentation, classification, logical structure analysis, Hough Transform, Wavelet Transform.

REMERCIEMENT

Je tiens en premier lieu, à remercier Dieu pour sa miséricorde, sa bonté, son obligeance et la faveur qui m'a donné afin de m'engager dans la voie de la recherche.

Je remercie mon directeur de thèse N.Khorissi C.C à l'université de Blida, mon codirecteur A.Sehad C.R au CDTA pour tous leurs conseils et directives tout au long de mon parcours de magister.

Je remercie A.Guessoum; H.Salhi; Z.Benslama; K.Achour pour avoir accepté d'examiner ce travail.

Je remercie H.Bessalah, directeur du CDTA, pour tous les efforts qu'il fait pour la recherche scientifique.

Je remercie M.Chriet pour sa collaboration.

Je remercie M.Bencherif pour toute son aide.

Je remercie l'ensemble des chercheurs et travailleurs du CDTA, particulièrement Mme Alim et Melle Bouzid pour le cadre agréable de travail qu'elles m'ont procurées.

Je remercie le corps enseignant et dirigeant de l'institut d'électronique.

DEDICACES

Je dédie ce travail à mes parents qui m'ont tant donné sans se laisser.

Je dédie ce travail à ma petite sœur Leila et à Kamel, Samir et à Ahcène et sa famille.

Et enfin à toutes mes amies de la cité universitaire de blida.

Faiza ...

TABLE DES MATIERES

RESUME.....	
REMERCIEMENTS.....	
TABLE DES MATIERES.....	
LISTE DES ILLUSTRATIONS GRAPHIQUE ET TABLEAUX.....	
INTRODUCTION	6
1. GENERALITES SUR L'ANALYSE DE DOCUMENTS	
1.1. Introduction.....	9
1.2. Généralités sur l'analyse des documents	10
1.3. Conclusion.....	24
1.2 LES METHODES D'ANALYSE DE LA STRUCTURE PHYSIQUE	
2.1. Introduction.....	25
2.2. Les méthodes de segmentation des documents	25
2.3. Les méthodes de classification régions	35
2.4. Représentation de la structure physique des documents	36
2.5. Conclusion.....	36
3. INTRODUCTION A LA THEORIE DES ONDELETTES	
3.1. Introduction.....	38
3.2. L'analyse temps fréquence :.....	38
3.3. la transformée en ondelettes.....	40
3.4. Intérêt en traitement d'image.....	54
3.5. Conclusion.....	55

4. CONCEPTION ET MISE EN ŒUVRE

4.1. Introduction	56
4.2. Système de segmentation de document	57
4.3. Conclusion.....	71

5. TESTS ET RESULTATS

5.1. Introduction.....	72
5.2. Environnement d'évaluation.....	72
5.3. Démarche d'expérimentation.....	74
5.4. Évaluation	84
5.5. Discussion critique	84
5.6. Conclusion	85

CONCLUSION.....	89
-----------------	----

Appendice A	91
-------------------	----

Appendice B	92
-------------------	----

Appendice C	93
-------------------	----

REFERENCES	94
------------------	----

LISTE DES ILLUSTRATIONS, GRAPHIQUES ET TABLEAUX

Figure 1.1	Structure d'un système d'analyse des documents	11
Figure 1.2	Texte incliné	12
Figure 1.3	Détection de l'angle d'inclinaison par la méthode de projection.	13
Figure 1.4	Méthode de groupement des plus proches voisins	15
Figure 1.5	Représentation d'une droite dans l'espace de Hough	16
Figure 1.6	Méthode de codage des séries de pixels noirs.	18
Figure 1.7	Méthode des composantes connexes	18
Figure 1.8	Méthode d'extraction des liaisons.	19
Figure 1.9	Sélection d'une région à partir d'un document avec deux colonnes	20
Figure 1.10	Histogrammes du document	20
Figure 1.11	Décomposition en ondelettes d'un document image	22
Figure 2.1	Exemple d'application de la méthode de maculage	26
Figure 2.2	Exemple d'application de l'algorithme Recursive X-Y cut.	27
Figure 2.3	Les composantes connexes.	29
Figure 2.4	Application de la projection des rectangles englobants sur un document	30
Figure 2.5	Application d'un filtre morphologique sur un document.	31
Figure 2.6	Application de la transformée de Haar sur un document.	33
Figure 2.7	Segmentation et structure physique.	36
Figure 3.1	La transformée de Fourier.	39
Figure 3.2	La transformée de Fourier à fenêtre	40
Figure 3.3	La transformée en ondelettes.	41
Figure 3.4	L'analyse temps fréquence	41
Figure 3.5	La transformée en ondelettes comme « <i>zoom</i> mathématique »	41
Figure 3.6	Modèle graphique du calcul du coefficient d'ondelette $W(a, b)$.	43
Figure 3.7	Comportement du coefficient d'ondelette.	44
Figure 3.8	Une fonction d'échelle B-spline	47
Figure 3.9	Une ondelette B-spline	47
Figure 3.10	Pyramide multi-résolution	49

Figure 3.11	Schéma de L'algorithme de Décomposition de S.Mallat	50
Figure 3.12	Schéma de la décimation	51
Figure 3.13	Processus de reconstruction du signal	51
Figure 3.14	Processus d'interpolation du signal	51
Figure 3.15	Schéma de décomposition/reconstruction multi-résolution (cas bidimensionnel)	52
Figure 3.16	Matrices d'une transformation de niveau 3	53
Figure 3.17	Une image et ses coefficients d'ondelettes.	54
Figure 4.1	Structure d'un système d'analyse des documents.	57
Figure 4.2	Les étapes de segmentation	61
Figure 4.3	Segmentation en blocs de différentes tailles	62
Figure 4.4	Décomposition en ondelettes d'un document image	63
Figure 4.5	Comparaison entre Haar et db2	65
Figure 4.6	Histogramme des coefficients de Haar	65
Figure 4.7	Exemple de dilatation	71
Figure 5.1	Exemple de quelques documents de la base 1.	73
Figure 5.2	Exemple de quelques documents de la base de données de Oulu base 2	73
Figure 5.3	Exemple d'un document incliné utilisé pour le test	74
Figure 5.4	Les différentes étapes de la détection/correction d'inclinaison	75
Figure 5.5	Document multi colonnes avec des images	78
Figure 5.6	Document avec tableau	78
Figure 5.7	Document avec la partie image dominante	78
Figure 5.8	Document latin	79
Figure 5.9	Exemple qui présente une fausse détection	79
Figure 5.10	Les différentes étapes de segmentation d'un exemple de la première base	80
Figure 5.11	Les différentes étapes de segmentation d'un exemple de la deuxième base	81
Figure 5.12	Histogramme de la variance des blocs pour un document complet	82
Figure 5.13	Histogramme de la variance des blocs pour une partie texte du document	82
Figure 5.14	Histogramme de la variance des blocs pour une partie photo du document	82
Figure 5.15	Exemple d'un document qui présente un cas défavorable.	83
Figure 5.16	Exemple d'un document avec fond non uniforme	85
Figure 5.17	Exemple d'un document en niveau de gris	86
Figure 5.18	Exemple d'un document structuré.	86

Figure 5.19	Exemple d'un document avec différents fonds et fontes	86
Figure 5.20	Exemple d'un document avec différentes orientations	87
Figure 5.21	Exemple d'un document avec texte de différentes couleurs	87
Figure 5.22	Exemple d'un document avec fond de différentes couleurs	87
Figure 5.23	Exemple d'un document dégradé	88
Figure 5.24	Exemple d'un document avec texte inclus dans les graphiques	88

Tableau 3.1	Comparaison entre la transformée en ondelettes, la transformée de Gabor (Fourier à fenêtre), et la transformée de Fourier	51
Tableau 4.1	Variation des histogrammes avec la taille du bloc homogène.	67
Tableau 4.2	Variation des histogrammes avec la taille du bloc hétérogène	68
Tableau 5.1	Exemple de quelques inclinés	76
Tableau 5.2	Tableau comparatif du nombre de points utilisés	77
Tableau 5.3	Taux de reconnaissance de l'angle d'inclinaison	77
Tableau 5.4	Temps de calcul et le rapport entre le temps de calcul total et le temps de calculs des différentes étapes de la segmentation	84

INTRODUCTION

Malgré les prédictions d'un monde sans papier qui devait accompagner le développement de l'ordinateur individuel, de la messagerie électronique et autres moyens facilitant le stockage, la conception et la diffusion des documents sous forme électronique, nous constatons que l'effet de substitution du papier ne s'est pas produit, mais qu'il y a plutôt redoublement des supports : communication par télécopie, brouillons manuscrits, impression des messages électroniques et des versions intermédiaires d'un document qui seront corrigées manuellement... Ceci s'explique par la nature même du support papier qui présente des qualités irremplaçables, surtout que lorsqu'il s'agit de lire ou d'annoter quelques feuilles, en plus de la qualité, la facilité de transport, la légèreté, l'autonomie ainsi que l'analogie matérielle avec le document final. A cela il faut ajouter que le papier a servi à mémoriser des connaissances depuis des siècles et qu'il n'est pas question de les oublier [1].

Pour profiter pleinement des avantages proposés par l'informatique, qui concernent le stockage, la transmission, la conception ou la gestion des documents, il y a lieu d'établir un pont de transfert entre supports papier et électronique. Cette tâche complexe a pour but ultime de permettre la saisie d'un document quelconque et la construction d'un objet machine structuré permettant des opérations simplifiées de conception et de manipulation. Cet objectif n'est pas atteint à ce jour, mais il existe des solutions partielles qui, dans des champs d'applications spécifiques, débouchent sur des réalisations fonctionnant en situations réelles.

Transformer le document papier en document électronique n'a pas pour but de mettre en mémoire de l'ordinateur son image brute ou compressée suivant des normes plus ou moins indépendantes de la nature de cette image [1]. Mais cette forme de mémorisation présente un double désavantage, un encombrement (la place mémoire occupée par l'image d'une feuille de texte A4 se compte en Mo) et l'absence d'une vision globale, non

structurée, du document excluant toute possibilité d'action sur le contenu ou sur la présentation. C'est à un transfert intelligent, c'est-à-dire à la construction d'un objet électronique décrit en terme de structures signifiantes du point de vue de la communication écrite, que le traitement du document papier est confronté.

Quant à la reconnaissance de l'écriture, qu'il s'agisse ou non de fabriquer un document électronique, elle aura d'autant plus de portée sur le plan des applications qu'elle sera intégrée à un système permettant la sélection des objets à reconnaître. Les montants des chèques et les adresses postales illustrent parfaitement cette nécessité de traiter un document préalablement à une reconnaissance de l'écriture.

Avec le développement des nouvelles technologies, ces dernières années ont vu se multiplier les grandes bandes d'images, texte avec différentes orientations, fond non uniforme et des structures complexes....

Les nombreuses recherches effectuées dans le domaine de l'extraction de l'information des images se sont concentrées sur la segmentation de l'image en zones ayant différentes propriétés, comme les zones images, les tableaux, le texte,...

Chaque zone d'intérêt présente des problèmes intrinsèques qui tendent à l'assimiler avec les autres zones.

Les recherches concernant la détection du texte se basent sur les principes essentiels :

- les textes sont monochromes ;
- les textes sont fortement contrastés par rapport au fond ;
- les textes ne sont composés que d'une seule texture ;
- les textes sont horizontaux.

Toutefois, ces principes ne prennent pas en compte, en même temps, l'inclinaison du texte, la multi-fonte, le fond non uniforme ainsi que la structure complexe qui pourrait contenir une image et texte mélangé.

Afin de solutionner ces problèmes capitaux en même temps, nous avons utilisé dans ce travail la segmentation par la transformée en ondelette, qui est basée sur la détection de la texture particulière du texte ; en vue d'améliorer les performances des ondelettes, nous

avons adjoint l'analyse multi-échelle en vue de prendre en compte l'information statique des blocs de travail qui apparaît dans le document multi-fontes.

L'envoi du texte vers l'OCR est ensuite sujet à la détection de son angle d'inclinaison par la transformée de Hough.

Plusieurs méthodes d'évaluation de l'angle d'inclinaison ont été proposées, nous pouvons les classer suivant l'approche adoptée :

- Les techniques basées sur la projection.
- Les techniques basées sur la transformée de Hough.
- Les techniques basées sur les k plus proches voisins.
- Les techniques basées sur la corrélation des lignes.

Afin de solutionner le problème de détection de l'angle d'inclinaison en minimisant les coûts de calculs, en considérant seulement les zones pouvant contenir l'angle dominant et en augmentant la précision de détection, nous avons utilisé la transformée en ondelette pour l'extraction des points qui vont être traités par la transformée de Hough.

Notre mémoire est organisée de la façon suivante :

Le premier chapitre a pour but de présenter des généralités sur le domaine de l'analyse des documents ainsi les méthodes de détection de l'inclinaison.

Ensuite les méthodes utilisées pour l'analyse de la structure physique seront décrites dans le deuxième chapitre.

Nous présentons dans le troisième chapitre une introduction sur la théorie des ondelettes.

La conception et les résultats expérimentaux seront discutés dans le quatrième et le cinquième chapitre.

Nous concluons ce mémoire, en résumant notre contribution et nous proposons des quelques perspectives à ces travaux.

CHAPITRE 1

GENERALITES SUR L'ANALYSE DE DOCUMENTS

1.1 Introduction:

L'analyse d'images de documents concerne tout le processus de conversion d'un document sous format papier vers un format électronique compréhensible et réutilisable. Ce processus est relatif à toutes les questions autour du langage écrit et sa transformation numérique : reconnaissance de caractères, formatage de texte, structuration du contenu et accès à l'information pour des applications d'indexation.

Plusieurs prétraitements sont nécessaires à la reconnaissance des documents, parmi lesquels, nous pouvons noter la détection d'inclinaison qui est un problème fréquent dans un système d'analyse des documents.

Définition d'un document:

Plusieurs définitions peuvent être attribuées à ce mot, à savoir:

- Ensemble de pages à usage professionnel, de loisir, ou social ;
- Moyen de communication de la pensée, de la connaissance, de l'information et du savoir.

Format d'un document:

- Se présente sous la forme de textes, tableaux, dessins, photos, ...
- La forme et le fond d'un document sont souvent très codifiés : règle de mise en page, typographie,...

Le document est donc un objet très riche à la fois dans sa forme, sa structure, dans son fond et son contenu [2].

Définition d'un document électronique :

Un ensemble cohérent d'objets numériques (textes, graphiques, photos, images animées et sons) stockés sur des machines informatiques interconnectées, ou stockés sur des supports informatiques amovibles et transportables[3].

En devenant numérique ou virtuel, les documents incorporent de nouvelles fonctionnalités : expressions sonores ou visuelles, mais également de nouvelles façons d'être manipulé ou d'être lu.

1.2 Généralités sur l'analyse des documents

L'analyse des documents consiste à localiser le texte, les graphiques et les images et extraire les informations logiques de la structure du document. Elle permet aussi de stocker le texte en une suite de code ASCII ou Unicode et de stocker les images sous un format BMP, TIFF ou autre format d'image.

En général, un système d'analyse des documents est représenté comme illustré dans la figure1.1 [4].

1.2.1 Numérisation des documents

L'étape de numérisation consiste à transformer le document papier en document électronique. Le coût de cette numérisation est fonction d'une part du matériel de base de numérisation disponible et d'autre part du niveau qualitatif du résultat souhaité [5].

1.2.2 Prétraitement

A l'issue de la phase de numérisation, vient celle du *pré-traitement*. Elle consiste à préparer les données reçues vers la phase d'analyse suivante. Ce passage n'est possible, que si les données sont dénuées de bruit et d'erreurs dues à la numérisation. Le prétraitement consiste à binariser et à filtrer l'image du document.

- ✓ *La binarisation* est utilisée pour convertir une image à différents niveaux de gris en une image à deux niveaux de gris (le noir et le blanc). [6]
- ✓ *Le filtrage* est utilisé pour réduire le bruit qui se présente dans les documents scannés, il est dû à différentes raisons ; comme :
 - la non propreté du scanner ou du document imprimé.
 - les défauts de papier (s'il est plié ou photocopié).
 - le processus de numérisation peut aussi introduire l'effet « sel et poivre », c'est à dire la présence des pixels blancs dans des régions noires ou la présence de pixels noirs dans les régions blanches [6].

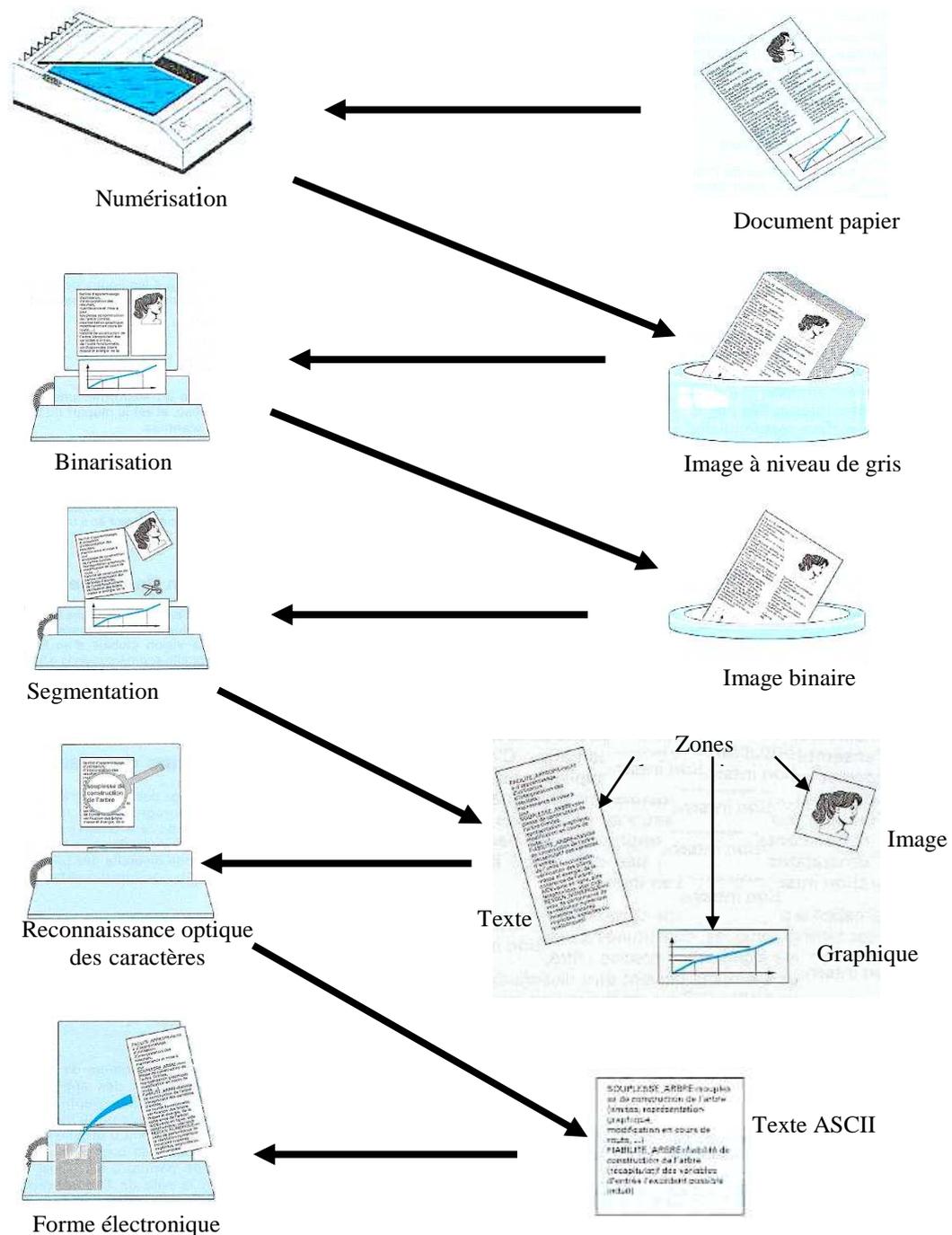


Figure 1.1 : Structure d'un système d'analyse des documents.

1.2.3 La détection et la correction de l'inclinaison

La détermination de l'inclinaison du document est un problème fréquent dans les systèmes d'analyse des documents. En général, ce problème se produit lorsque :

- le document est mal placé sur le scanner
- le document est mal photocopié.
- le document contient des zones textuelles inclinées.

Les techniques d'évaluation de l'angle d'inclinaison ont été développées pour répondre essentiellement à deux besoins [7]:

1. Corriger le défaut de positionnement du papier dans le scanner.
2. Corriger l'inclinaison interne de certaines parties du document qui peuvent apparaître.

En effet la plupart des algorithmes de reconnaissance ne fonctionnent bien que si le document est bien redressé. D'où l'intérêt du redressement [7].

L'angle d'inclinaison θ_s est l'angle entre la ligne horizontale du document et l'orientation des lignes du texte [6].

La détection de l'angle d'inclinaison consiste à déterminer :

- l'angle d'inclinaison θ_s du document, (voir figure1.2)
- et à faire une rotation du texte dans la direction opposée avec le même angle θ_s .

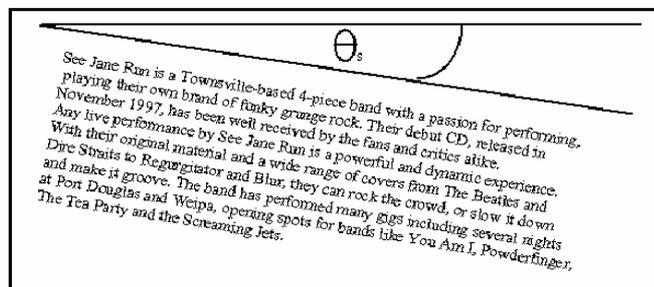


Figure 1.2 : Texte incliné.

Les techniques de détection d'inclinaison peuvent être classées [8]:

- soit suivant l'approche de base qu'ils adoptent:

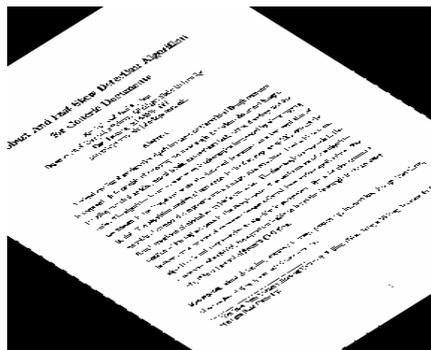
1. L'analyse de la projection.
 2. La transformée de Hough.
 3. Groupement des k plus proches voisins.
 4. La corrélation entre les lignes.
- Ou bien, en se basant sur l'analyse :
1. du gradient,
 2. du spectre de Fourier,
 3. morphologique.

Notons que, toutes les méthodes citées travaillent sur des images binaires [8], [9].

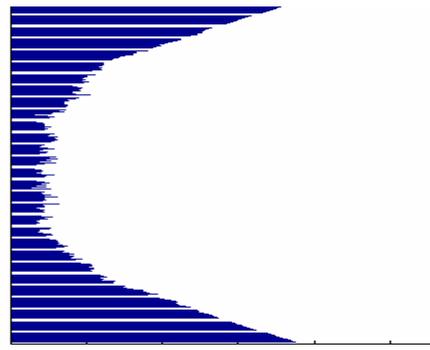
1.2.3.1 La méthode de projection

Cette méthode permet de calculer pour chaque angle dans l'intervalle de détection l'histogramme horizontal du document. Ceci consiste à projeter les pixels de l'image horizontalement afin de tracer l'histogramme représentatif du nombre de pixels noirs tout au long d'une ligne (voir figure 1.3). Pour un document, dont les lignes s'étendent horizontalement, la projection horizontale du document contiendra des pics avec des largeurs égales à la hauteur du caractère.

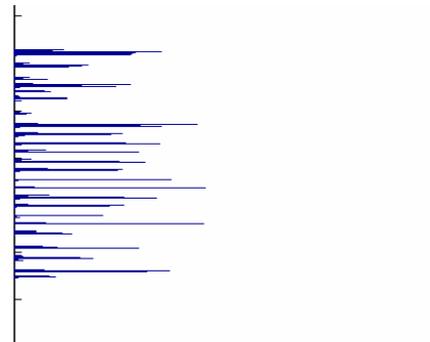
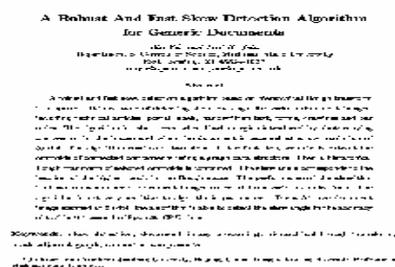
Pour le calcul de l'angle d'inclinaison, plusieurs recherches ont été proposées, parmi lesquelles, nous citons celle qui maximise la somme des largeurs des zones vides de l'histogramme [8].



(a)



(b)



(c)

(d)

Figure 1.3 : détection de l'angle d'inclinaison par la méthode de projection.

(a) : Document incliné

(b) : Histogramme horizontal

(c) : Document non incliné

(d) : Histogramme horizontal

Avantages

- Facilité d'implémentation ;
- Valable pour des documents à structure simple.

Inconvénients

- Angle d'inclinaison compris entre -10° et $+10^\circ$;
- Non valable pour des documents multi-colonnes ;
- Non adapté pour des documents contenant des graphiques ou des photographes.

1.2.3.2 La méthode de groupement des plus proches voisins :

Dans cette méthode, pour chaque composante connexe, la direction du segment qui la connecte à son plus proche voisin est calculée par :

$$\theta = \tan^{-1} \left(\frac{(M_j.y - M_i.y)}{(M_j.x - M_i.x)} \right) \quad 1.1$$

Tel que : $(M_i.x, M_i.y)$ sont les coordonnées de centre de gravité du composante connexe i.

Ces directions sont accumulées dans un histogramme d'angle et son pic représente l'angle d'inclinaison du document (voir figure 1.4) [10] [11].

Avantages

- Indépendance de l'intervalle de détection.
- Détection de plusieurs angles d'inclinaison.
- Valable pour des documents multi-colonnes.

Inconvénients

- Temps de traitement très important.

- Sensible au bruit.
- Présence des caractères attachés.
- Non adapté pour des documents contenant des graphiques ou des photographes.

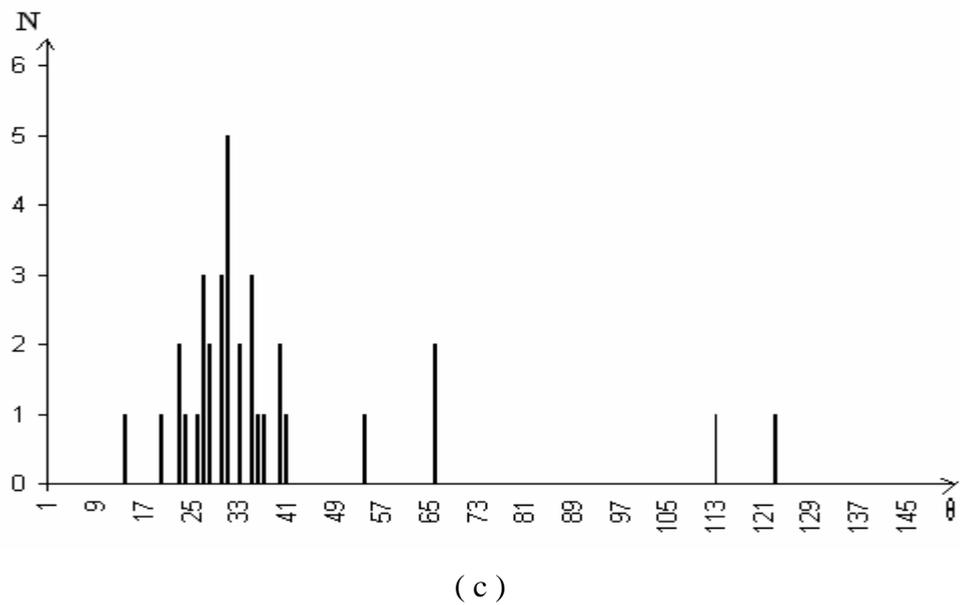
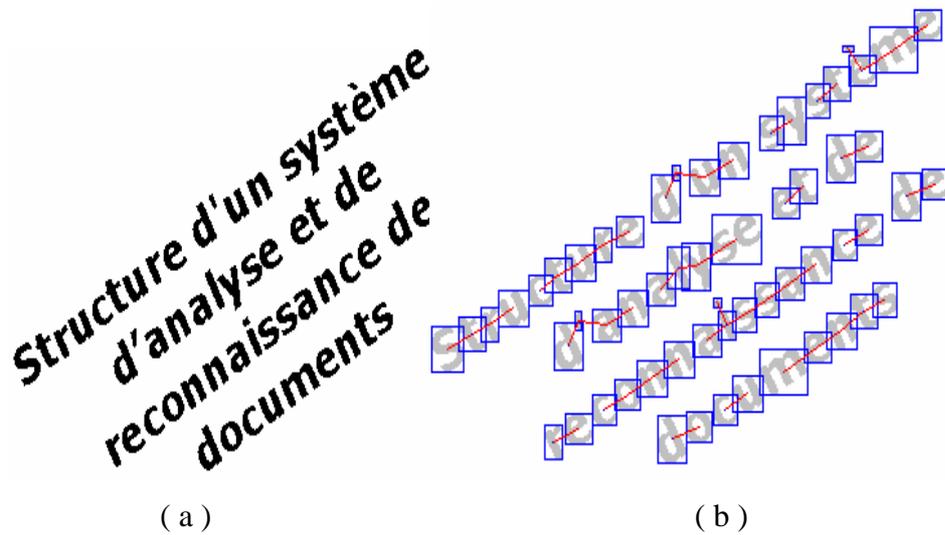


Figure1. 4. : Méthode de groupement des plus proches voisins.

(a) Texte incliné ; (b) Les composantes connexes ; (c) Histogramme des angles.

N : Etant le nombre de segments correspondant à chaque angle θ .

1.2.3.3 La transformée de Hough :

La transformée de Hough est une technique de détection des lignes et des courbes dans une image [10], (voir appendice 1).

Pour la détection d'une ligne, chaque point M_i de coordonnées (x_i, y_i) d'une droite se transforme dans l'espace de Hough en une sinusoïde d'équation :

$$\rho = x_i \cos \theta + y_i \sin \theta . \quad 1.2$$

Donc une droite sera représentée par un ensemble de sinusoïdes qui se coupent en un seul point de coordonnées polaires (ρ_0, θ_0) caractéristique de cette droite dans le plan des paramètres (voir figure 1.5).

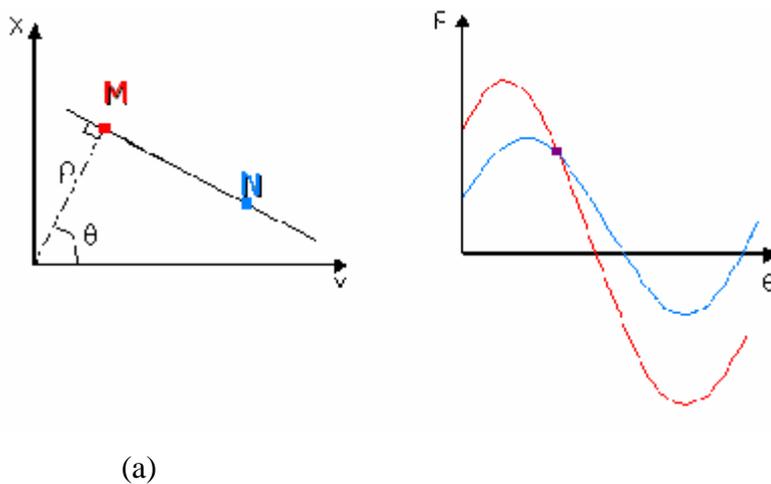


Figure 1.5 : Représentation d'une droite dans l'espace de Hough
 (a). Plan cartésien (xy) (b). Plan des paramètres (ρ, θ)

L'angle d'inclinaison est détecté par l'algorithme suivant [8] :

```

Pour chaque point (x, y)
{
Pour  $\theta_1 \leq \theta \leq \theta_2$  /*  $[\theta_1, \theta_2]$  est l'intervalle de détection */
{
Calculer  $\rho = x \cos \theta + y \sin \theta$ 
Incrémenter  $H(\rho, \theta)$ 
}
}

Trouver le maximum dans le tableau  $H(\rho, \theta)$ 
    
```

La valeur maximale dans le tableau accumulateur est associée à l'angle d'inclinaison.

Avantages

- L'exactitude peut être contrôlée en changeant la résolution de l'axe θ .
- Le temps de traitement peut être réduit en diminuant le nombre de pixels à représenter dans l'espace paramétrique (ρ, θ) .
- L'intervalle de détection est compris entre 0° et 180° .
- Valable pour des documents multi-colonnes.

Inconvénients

- Nécessité d'un espace mémoire important.
- Difficulté de détecter le vrai angle d'inclinaison lorsque cette méthode est appliquée sur des documents avec un texte clair et des régions non textuelles.
- Sensibilité au bruit.
- Détection d'un seul angle d'inclinaison.
- Temps de traitement assez important.

1.2.3.4 Les méthodes de réduction du nombre de points :

L'étude des différentes méthodes pour la détection de l'angle d'inclinaison nous a permis de déduire qu'elles ont des complications qui peut être résumer dans les points suivants :

- Problèmes liés au document multi-colonnes.
- Le temps de traitement.
- Présence des graphiques et des images.
- Les documents scannés attachés ont souvent une partie de la page opposée qui apparaît incliné avec un angle différent [12].

En vue de solutionner ces problématiques, des techniques hybrides ont été proposées pour le choix du point à traiter.

A. La méthode de codage des séries de pixels noirs :

Cette méthode permet de détecter les séries verticales de pixels noirs d'un document, et marque le point final de chaque série alors que les autres points sont mis à blanc (voir figure1.6) [13].

Avantages

- La rapidité.

Inconvénient

- Ne présente pas une précision suffisante pour l'évaluation de l'angle d'inclinaison pour les documents arabes.

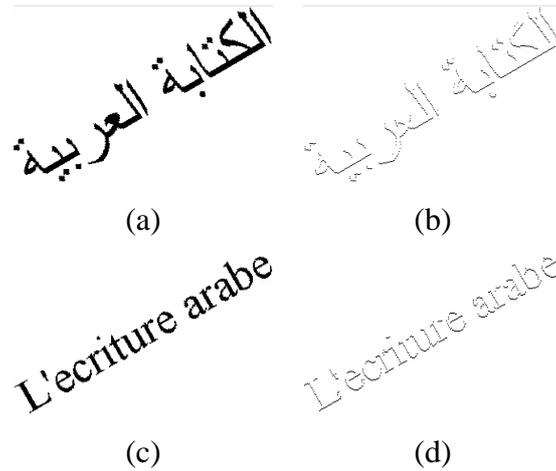
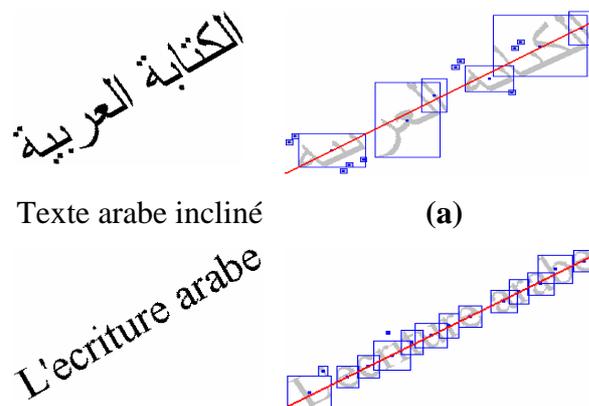


Figure 1.6. : Méthode de codage des séries de pixels noirs.

(a) Texte arabe incliné, (c) Texte latin incliné, (b, d) Application de la méthode de codage des séries de pixels noirs.

B. La méthode des composantes connexes :

Cette méthode consiste à calculer les composantes connexes du document (voir figure 1.7), ensuite seulement les points centraux de ces derniers sont utilisés pour évaluer l'angle d'inclinaison [13].



Texte latin incliné. (b)

Figure 1.7 : la méthode des composantes connexes.

(a, b) Les points centraux des composantes connexes.

Avantages

- Cette méthode donne de bons résultats pour les documents latins.

Inconvénients

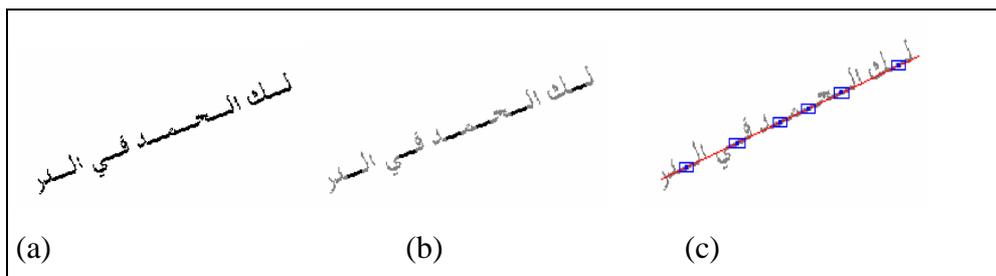
- Imprécise pour les documents arabes, car l'écriture arabe est cursive ainsi les points centraux des composantes connexes ne sont pas alignés sur la même ligne, ceci est dû au fait que ces composantes ne sont pas de taille homogène.

C. Une approche se basant sur l'extraction des liaisons :

Dans [13], une technique se basant sur les histogrammes verticaux et horizontaux afin de détecter les liaisons a été proposée.

L'étape la plus importante dans cette approche est la détection de liaisons, elle consiste à détecter les liaisons inter caractères (Figure 1.8). Elle est effectuée en calculant les histogrammes horizontaux et verticaux pour chaque composante connexe qui contient un sous mot.

En général, les liaisons sont caractérisées par une hauteur inférieure à la hauteur des caractères et leur fréquence d'apparition est très grande dans un sous-mot. Pour cela, la hauteur (largeur) la plus fréquente est calculée à partir des histogrammes, ensuite, les liaisons sont extraites, elles correspondent aux zones dont la hauteur (largeur) est inférieure à la hauteur (largeur) la plus fréquente dans l'histogramme vertical (horizontal).[13]



(a) Texte incliné, (b) Détection de liaisons, (c) Les points centraux des liaisons.

Figure 1.8 : la méthode d'extraction des liaisons

Avantages

- Réduction du nombre de point
- La précision.

Inconvénients

- Influencé par certains caractères Arabes.

D. Une approche travaillant seulement sur des sous régions de l'image :

L'utilisation de la page entière n'est pas nécessaire. Quelques lignes de texte sont suffisantes pour la détection de l'angle d'inclinaison [15].

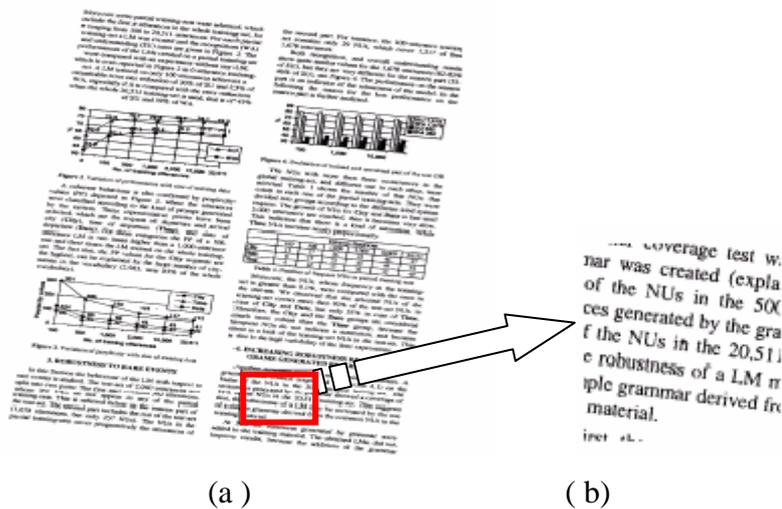


Figure. 1.9: la sélection d'une région à partir d'un document avec deux colonnes.
 (a) document avec deux colonnes (b) la région sélectionnée du document.

Cependant, le choix porte sur la partie appropriée à utiliser, pour sélectionner cette partie, nous extrayons un segment de la page centré sur les points dont les coordonnées sont les pics horizontaux et verticaux du profil de projection. De cette façon, la partie de la page dont la densité avec le maximum de pixels noirs et blancs est localisée.

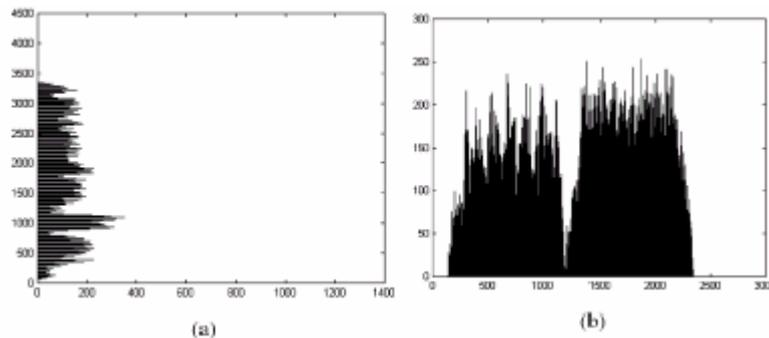


Figure.1.10 histogrammes du document.

(a) histogramme horizontal du document (b) histogramme vertical du document de la figure 1.9

Cette technique garantit la sélection du segment du texte dans la majorité des cas des documents, car la partie textuelle présente une grande intensité en pixels noirs par rapport au graphique et aux tableaux, etc.

La figure 1.9 montre un document de deux colonnes avec des tables et des graphiques. La partie sélectionnée montrée dans le rectangle de dimensions 400*400 pixels est centrée au Pixel de coordonnées (1276.1564), les figures 1.10 (a) et 1.10 (b) montrent les maxima des projections des histogrammes horizontaux.

Avantages

- cette méthode est imposée en cas de documents multi colonnes, où l'utilisation d'une partie de la page est plus sûre,
- rapide et facile.

Inconvénients

- la présence des images dans le document perturbe la technique. Dans ces cas-ci l'utilisation de la page entière est plutôt recommandée.

E. Les approches travaillant sur le document compressé:

OLEG et al ont proposé la réduction de la résolution du document [14].

Avantages

- Rapidité.
- Une basse résolution est suffisante pour le traitement.
- ne limite pas la gamme d'angle.

Inconvénients

- Il y'a seulement quelques techniques de détection de l'inclinaison fonctionnant avec les images comprimées
- Balayage de l'image deux fois (une fois pour la détection d'inclinaison et une autre fois pour l'analyse).

Par analogie a cette méthode nous nous sommes orienté à un type de compression effectué par la transformée en ondelettes.

F. L'approche proposée :

La transformée de Hough permet de détecter avec exactitude l'angle d'inclinaison toutefois, elle accuse un temps de calcul très important.

Afin de solutionner les problèmes de détection de l'angle d'inclinaison :

- Le temps de calcul élevé.
- La présence du graphique et des images.
- La sensibilité au bruit.

Nous avons effectués les étapes suivantes :

- La transformée en ondelettes (voir figure 1.11) a été introduite et seule la bande LH est utilisée, les détails concernant cette technique feront l'objet du chapitre 3, car cette bande représente le contour horizontal, ceci élimine la totalité des zones graphiques et zones d'images [16].

Le but dans la présente partie est de :

- ❖ préserver les lignes;
- ❖ enlever le bruit;
- ❖ Assurer une première réduction : le fait de travailler avec la première résolution sur un seul quadrant de l'image.



Figure 1.11. la décomposition en ondelettes d'un document image

- Un seuillage global a été défini afin d'assurer [17] :
 - une deuxième réduction du nombre de points
 - l'obtention d'une image binaire qui est une chose indispensable pour la transformée de HOUGH.
- Une transformée de Hough a été introduite pour la détection de l'angle d'inclinaison.

1.2.3.5 Correction de l'inclinaison

Une fois que l'angle d'inclinaison a été déterminé, l'image est alors corrigée. Si cette correction est faite simplement en calculant la nouvelle position de chaque pixel [10].

Pour chaque pixel (x, y) de l'image inclinée, les nouvelles coordonnées (x', y') dans l'image redressée sont calculées de la façon suivante :

$$x' = x \cos \alpha + y \sin \alpha \quad (1.3)$$

$$y' = y \cos \alpha - x \sin \alpha \quad (1.4)$$

Où $\alpha = -\theta$, avec θ est l'angle d'inclinaison de l'image du document.

1.2.4 Analyse de la structure physique

L'analyse de la structure physique est l'opération d'analyse globale qui consiste à détecter et à localiser les zones distinctes : textes, graphiques, tableaux. A l'issue de cette opération, nous soumettons les différentes zones à des traitements spécifiques. En particulier, les blocs contenant du texte sont envoyés à un module de reconnaissance de caractères (OCR).

Le rôle principal de la structure physique est de faciliter l'accès à l'information et de la lisibilité des documents [10].

L'analyse de la structure physique peut être décomposée en deux traitements : la segmentation et la classification.

1.2.4.1 Segmentation

Elle consiste à segmenter l'image du document en blocs homogènes, il existe plusieurs techniques de segmentation de documents [18]. Celles-ci seront présentées dans le chapitre suivant.

1.2.4.2 Classification

La classification [10] consiste à déterminer la nature des régions obtenues par l'étape de segmentation, le type des régions peut être : un texte, un graphique, un tableau ou une image.

La classification est effectuée en utilisant des règles qui décrivent certaines caractéristiques des blocs physiques. Les entités telles que le nombre de pixels, les découpages horizontaux et verticaux, les caractéristiques du texte et les séries de pixels sont utilisées pour calculer les caractéristiques qui permettent de classer les blocs.

1.2.5 Analyse de la structure logique

L'analyse de la structure logique d'un document permet de transmettre la sémantique du document et l'ordre de lecture des blocs, elle est associée au contenu du document et à la hiérarchie des objets logiques, tels qu'un titre, un résumé, des paragraphes, des sections, des tableaux, des figures, des renvois de page.

Généralement la structure logique est stockée dans un arbre, la racine est le document en entier, les feuilles sont les objets logiques de bases et les nœuds internes sont les objets logiques composites [7].

1.2.6 Reconnaissance optique des caractères

Une fois que les régions de textes ont été identifiées par le module de classification, elles sont appliquées à un système de reconnaissance optique de caractères (OCR) qui consiste à identifier les caractères contenus dans les blocs de texte et de les représenter par des codes ASCII ou Unicode

Un OCR consiste donc, à extraire les paramètres décrivant au mieux les caractères à reconnaître, ensuite un apprentissage est effectué pour que le système apprenne à reconnaître les propriétés des caractères. Durant la phase de décision, le système compare les paramètres du caractère à ceux enregistrés dans la phase d'apprentissage, le caractère est soit accepté, soit rejeté.

1.3 Conclusion

Ce chapitre a fait l'objet d'une introduction à l'analyse des documents et aux différentes étapes la constituant.

Parmi les techniques, les plus utilisées pour la détection de l'angle d'inclinaison, nous citons : la méthode de projection, la transformée de Hough, le groupement des k plus proches voisins et les méthodes hybrides. L'étude de ces différentes méthodes nous a permis de déduire qu'elles ont des complications qui peuvent être résumées comme suit :

- Problèmes liés au document multi colonnes [12].
- Temps considérable de traitement.
- Présence de graphiques et d'images.

Afin de solutionner le problème de détection de l'angle d'inclinaison en minimisant toutefois les coûts de calculs, en considérant seulement les zones pouvant contenir l'angle

dominant et en augmentant la précision de détection, nous avons suivi les étapes suivantes :

La transformée en ondelettes assure l'extraction de la partie textuelle, ensuite un seuillage est effectué et enfin la TH est appliquée permettant ainsi de détecter avec exactitude l'angle d'inclinaison.

Cette étape étant réalisée, nous passons ensuite à l'analyse de la structure physique du document qui est une étape primordiale de l'extraction des informations des documents, elle consiste à diviser le document en plusieurs blocs homogènes, celle-ci sera traitée dans le chapitre suivant.

CHAPITRE2

LES METHODES D'ANALYSE DE LA STRUCTURE PHYSIQUE

2.1 Introduction

Le rôle principal de la structure physique est de faciliter l'accès à l'information et la lisibilité des documents [2], en segmentant l'image du document en blocs homogènes puis en les classant.

Il existe plusieurs méthodes de segmentation et d'étiquetage logique (classification).

Dans ce chapitre, nous présentons dans l'ordre les méthodes ascendantes, les méthodes descendantes et les techniques utilisées pour la classification des documents.

2.2 Les méthodes de segmentation des documents

La segmentation a pour but de localiser à partir de l'image numérisée d'une page les blocs qui composent cette dernière. Plus concrètement, le problème de la segmentation peut se formuler comme suit :

Etant donnée l'image digitalisée d'une page, déterminer une partition géométrique de cette dernière de sorte à isoler tous les blocs qui la composent [19].

Les méthodes de segmentation des documents qui existent, peuvent être classées selon deux catégories : les méthodes descendantes et les méthodes ascendantes.

2.2.1 Les méthodes descendantes

Les méthodes descendantes prennent le document en entier, le segmentent en petits blocs, qui sont ensuite segmentés jusqu'à l'obtention des blocs suffisamment petits et isolés. Ces méthodes supposent que le document est conforme à une structure particulière. Bien qu'elle soit une limitation, la plupart des journaux et des revues ont été segmentés sans aucun problème. Parmi ces algorithmes, on peut citer : l'algorithme RSLA (Run Length Smoothing Algorithm) et l'algorithme RXYC (Recursive X-Y cut).

2.2.1.1. L'algorithme RLSA [10]:

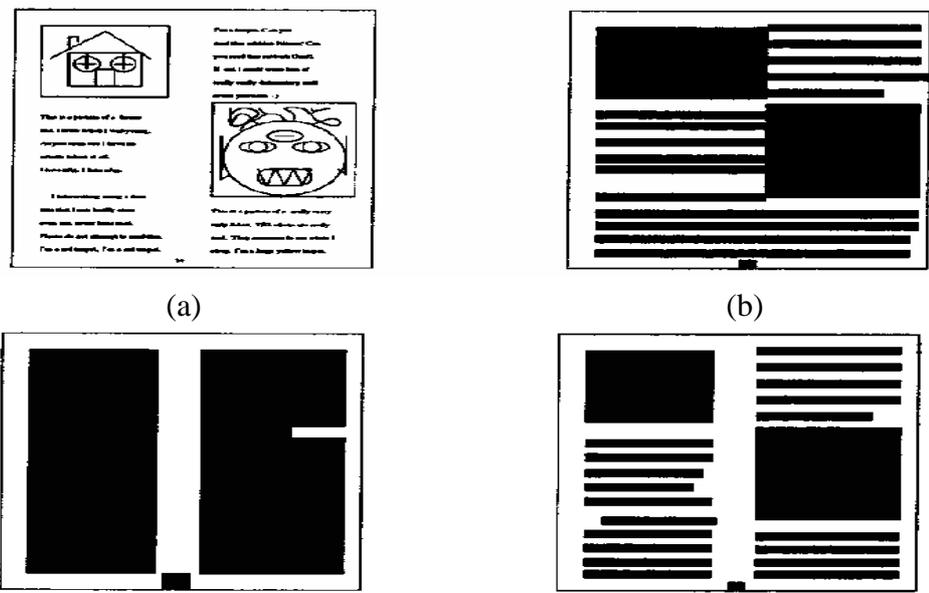
Cet algorithme est appliqué sur des images binaires contenant des suites de «0» et de «1». Les pixels blancs sont maculés et fusionnés dans des suites continues de pixels noirs si leur nombre est inférieur ou égal à un certain seuil qui représente la distance minimale en dessous de laquelle des entités de l'image sont considérées comme voisines (voir figure 2.1).

Exemple :

Avec un seuil = 3, une chaîne x est transformée en une chaîne y comme suit :

```
X = 0 1 1 0 0 1 1 1 1 1 0 0 1 1 0 1 1 1 0
Y = 0 0 0 0 0 1 1 1 1 1 0 0 0 0 0 1 1 1 0
```

RLSA est appliqué pour chaque ligne et pour chaque colonne, ce qui permet d'obtenir deux images dans lesquelles les pixels sont maculés horizontalement et verticalement. Le « Et logique » est appliqué sur ces deux images afin de délimiter les régions où des formes imprimées existent dans le document comme le montre la Figure 2.1. Les seuils 'tx' (pour le maculage horizontal) et 'ty' (pour le maculage vertical) ne sont pas nécessairement les mêmes.



(c)

(d)

Figure 2.1 : Exemple d'application de la méthode de maculage.

(a) Image originale.

(b) Application du maculage horizontal

(c) Application du maculage verticale.

(d) Le ET logique entre (b) et (c).

Avantages

- Facile à implémenter.

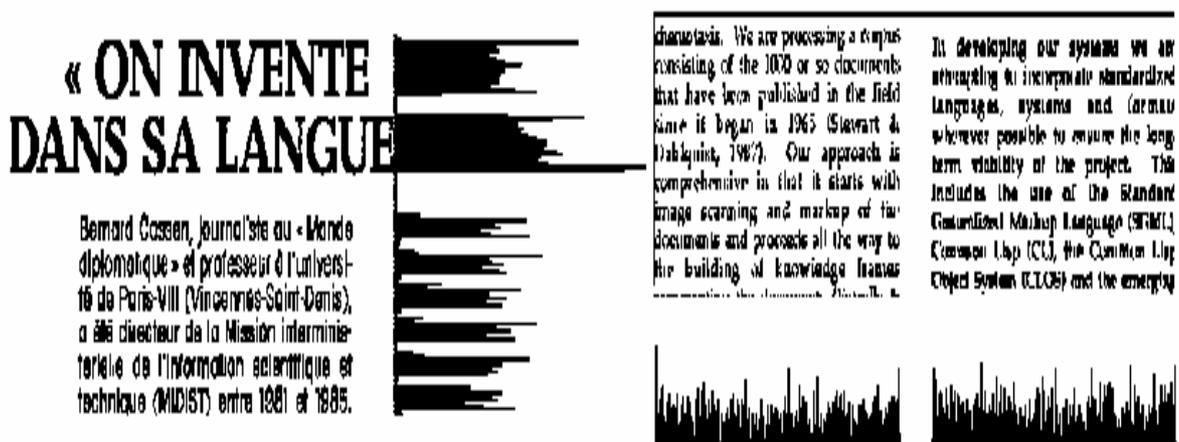
Inconvénients

- Choix des valeurs des seuils.
- Blocs obtenus ne sont pas forcément rectangulaire.
- Sensibilité à l'inclinaison.

2.2.1.2 L'algorithme « Recursive X-Y cut » [20] :

Cet algorithme divise récursivement le document en petits blocs rectangulaires. A chaque étape de récursion, les histogrammes : horizontal et vertical sont calculés. Le découpage de l'image est effectué dans les vallées de l'histogramme qui sont plus grande qu'un seuil déterminé (voir figure2.2). Ce processus est appliqué récursivement sur chaque sous région, lorsque aucun découpage ne peut être effectué, l'image est complètement segmentée en bloc.

Les histogrammes vertical et horizontal sont utilisés pour déterminer si oui ou non un découpage peut être fait et de savoir aussi où le découpage sera fait. Les vallées dans l'histogramme représentent les zones où les pixels noirs n'existent pas, comme par



exemple : l'espace entre les lignes et les paragraphes dans l'histogramme horizontal et l'espace entre les colonnes dans l'histogramme vertical.

(a)

(b)

Figure 2.2 : Exemple d'application de l'algorithme Recursive X-Y cut.

(a) histogramme horizontal ; (b) histogramme vertical d'un document.

Avantages

- Facile à implémenter.
- Adapté pour les documents décomposables par une séquence de subdivisions horizontales et verticales.

Inconvénients

- Coût de calcul des histogrammes.
- Choix du seuil qui peut être différent pour chaque niveau de récursion.
- Sensibilité à l'inclinaison.
- Nécessité d'un modèle de décomposition car cette méthode ne peut pas être directement appliquée sur des structures de page suffisamment imbriquées.

2.2.2 Les méthodes ascendantes

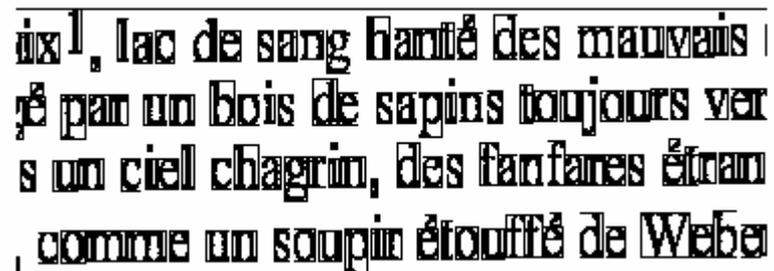
Les méthodes ascendantes considèrent les documents au niveau pixel et affinent les données par des opérations de regroupement. Ainsi les composantes connexes sont formées, puis les mots sont alors retrouvés et enfin les lignes et les paragraphes [10]. Une composante connexe est un ensemble de pixels adjacents les uns aux autres dans l'image. Parmi ces méthodes on peut citer la méthode des composantes connexes, projection des rectangles englobants (Bounding Box Projection) et Docstrum.

2.2.2.1 Les composantes connexes [10]

Cette méthode est utilisée pour retrouver les segments dans l'image. La première étape de cette méthode est de calculer les composantes connexes qui sont par la suite fusionnées en mots (voir figure 2.5). Les mots sont fusionnés en lignes et ensuite en paragraphes, l'opération de regroupement s'arrête jusqu'à ce que le niveau de segmentation soit atteint.

Le regroupement se fait par la méthode des k plus proches voisins, le 'k' est utilisé pour détecter les différents niveaux de voisinage des composantes connexes.

Une analyse 1 plus proche voisin ($k = 1$) repère toutes les lignes et une analyse 1 plus proche voisin des lignes repère les paragraphes dans le document.



ix 1, lac de sang hanté des mauvais |
é par un bois de sapins toujours ver
s un ciel chagrin, des fanfares étran
comme un sloop étouffé de Weber

Figure 2.3 : Les composantes connexes.

Avantages

- Ne nécessite pas une connaissance à priori sur la structure de documents.

Inconvénients

- Temps de calcul très important.
- Choix du paramètre k .
- Sensible à l'inclinaison.

Des erreurs de segmentation peuvent apparaître comme :

- Fusion de deux lignes qui n'appartient pas à la même colonne.
- Le non regroupement de deux mots appartenant à la même colonne.

2.2.2.2 Bounding Box Projection [10]

Pour cette technique, un algorithme de calcul de composantes connexes est appliqué au document pour générer les composantes connexes (voir figure 2.6). Le rectangle englobant de chaque composante connexe est calculé. Les caractères qui sont proches l'un de l'autre seront fusionnés pour former des mots, ensuite leurs rectangles englobants seront calculés. De même, les rectangles englobants des mots sont fusionnés pour former les rectangles englobant des lignes et qui seront par la suite fusionnées pour former les rectangles englobant des blocs de texte comme le montre la figure 2.4.

Avantages

- Ne nécessite pas une connaissance à priori sur la structure des documents.

Inconvénients

- Temps de calcul très important.

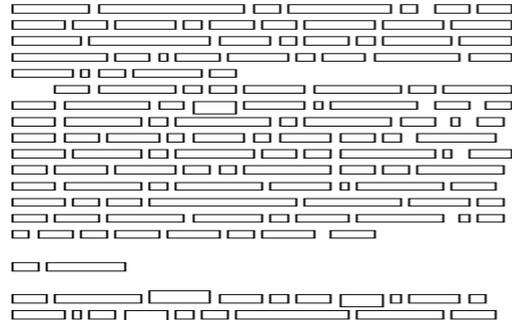
- Choix des paramètres pour grouper les composantes connexes.
- Sensible à l'inclinaison.

surfaces (see Section 2.3), the combination of local planarity and rigidity is used. For arbitrary motion, rigidity between environmental points is used to recover motion parameters from a small number of image locations (See Section 2 and Section 3.1)

The remainder of this section introduces the notation used throughout this paper. Section 2 describes how the local direction of translation is estimated from a flow field and cases of motion for which this is particularly robust. Section 3 describes how the parameters of relative sensor motion can be recovered from the estimated local directions of translation. Section 4 discusses computing the local transformational component directly from real image sequences without the initial extraction of optic flow and other areas for future work.

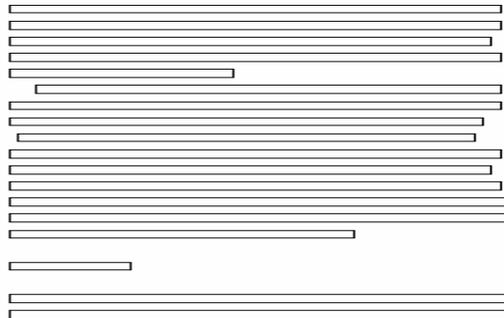
1.1 Notation

The coordinate system used in this paper is shown in Figure 1. The origin of this right-handed coordinate sys-

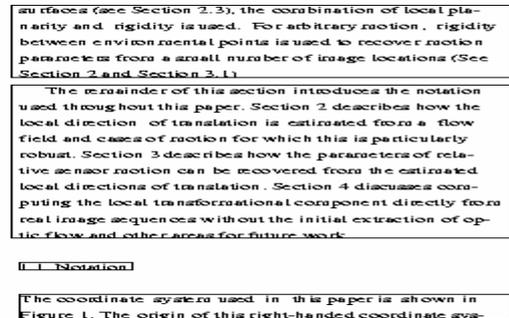


(a)

(b)



(c)



(d)

Figure 2.4 : Application de la projection des rectangles englobants sur un document.

- (a) image originale ;
- (b) Rectangles englobants des mots ;
- (c) rectangles englobant des lignes ;
- (d) rectangles englobant des blocs de texte.

2.2.2.3 Le filtrage morphologique [5]

La morphologie mathématique permet d'éroder ou de dilater des formes spécifiques dans une image. Ces opérations permettent d'isoler ou de fusionner des formes particulières plus au moins éloignées.

Une approche a été développée pour extraire les zones graphiques, elle est réalisée suivant les opérations suivantes :

- Fusion des caractères en lignes par une fermeture (dilatation suivie d'une érosion) dans la direction horizontale.
- Suppression des lignes de texte par ouverture (érosion suivie d'une dilatation) dans les directions 0° et 90° , ce qui élimine la plupart des caractères et les zones de texte

- Suppression des éléments graphiques de faible taille par fermeture suivie d'un remplissage des vides. Le résultat produit un masque binaire qui délimite les régions graphiques.
- Une différence symétrique est effectuée XOR entre l'image originale et l'image qui ne contient que les zones graphiques, permet d'obtenir les zones de texte. Filtrage des zones de textes en éliminant les petites connexités par fermeture (voir figure 2.7)

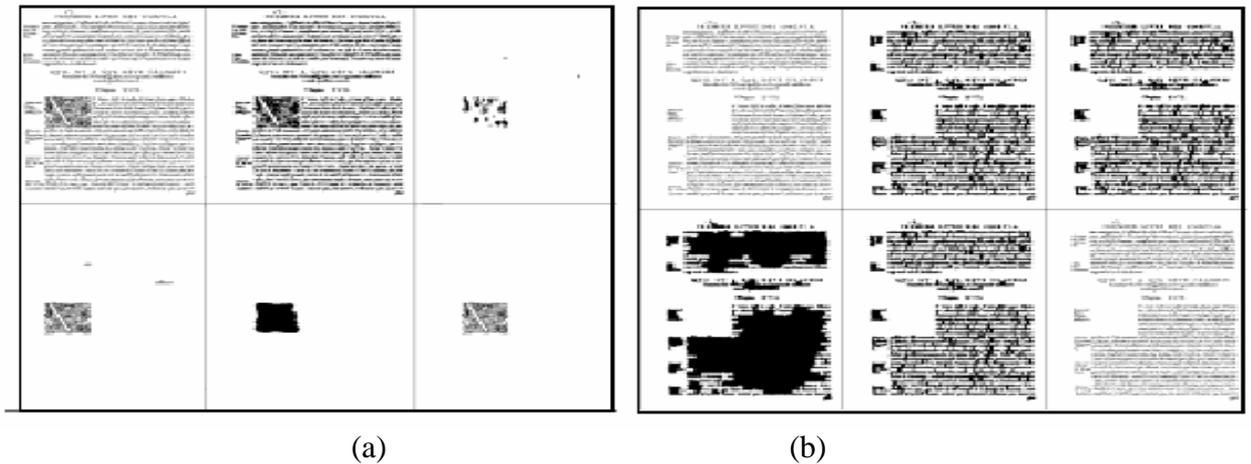


Figure 2.5 : Application d'un filtre morphologique sur un document.

(a) Détection des zones graphiques

(b) Détection des zones de texte

Avantages :

- Ne nécessite pas une connaissance a priori sur la structure des documents.

Inconvénient :

- Nécessite beaucoup de paramètres comme :
 - le nombre d'érosions et de dilations correspondant aux différents espaces entre les caractères ;
 - la hauteur minimale d'une ligne pour pouvoir l'effacer ;
 - la distance maximale entre les lignes pour pouvoir les fusionner, afin de supprimer le texte des zones graphiques restantes ;
- Le temps de traitement des opérations morphologiques augmente avec les documents de grande dimension.

2.2.2.4 L'analyse de la texture :

Une région textuelle a généralement une texture spéciale, car elle est constituée d'un caractère identique qui implique une variation périodique d'intensité horizontale due à l'alignement horizontal des caractères [2] [21].

La ligne de texte, grâce à sa texture régulière et son alignement, est l'élément le plus facile à localiser quelque soit le support. On peut définir une zone de texte comme « Une région de l'image présentant une très forte densité de traits qui forment des alignements à une échelle donnée ». Dans une image naturelle, une telle configuration est très rare. C'est la raison pour laquelle il n'est pas difficile d'extraire du texte à partir des images vidéo, d'images naturelles fixes, d'images couleurs ou à niveaux de gris.

Les approches basées sur l'analyse de la texture traitent les composantes des documents: texte, background, image et graphe comme étant différentes textures [22].

Les algorithmes de segmentation de la texture sont constitués de trois étapes principales :

- 1- Tout d'abord, il est nécessaire de transformer les données de l'image afin de les rendre pertinentes pour la recherche de texture. Cette transformation est faite par les filtres de GABOR puis par la transformée en ondelettes dans les derniers travaux [8], [21], [22].
- 2- Calcul des vecteurs caractéristiques.
- 3- Classification des vecteurs caractéristiques.

2.2.2.4.1 La transformée de GABOR :

Cette approche basée sur la transformée de Gabor permet de filtrer les fréquences orientées dans une direction à une fréquence donnée. En choisissant les directions orthogonales et obliques pour les caractères italiques associés à un filtrage passe bande de fréquence, correspondant aux fréquences produites par les traits des caractères, le filtrage de Gabor élimine les formes graphiques et met en valeur les lignes de texte[2],[8].

Avantages :

- 1- Applicable à des images à niveaux de gris.
- 2- Ne nécessite pas une connaissance à priori sur la structure des documents.
- 3- Insensible à l'inclinaison.

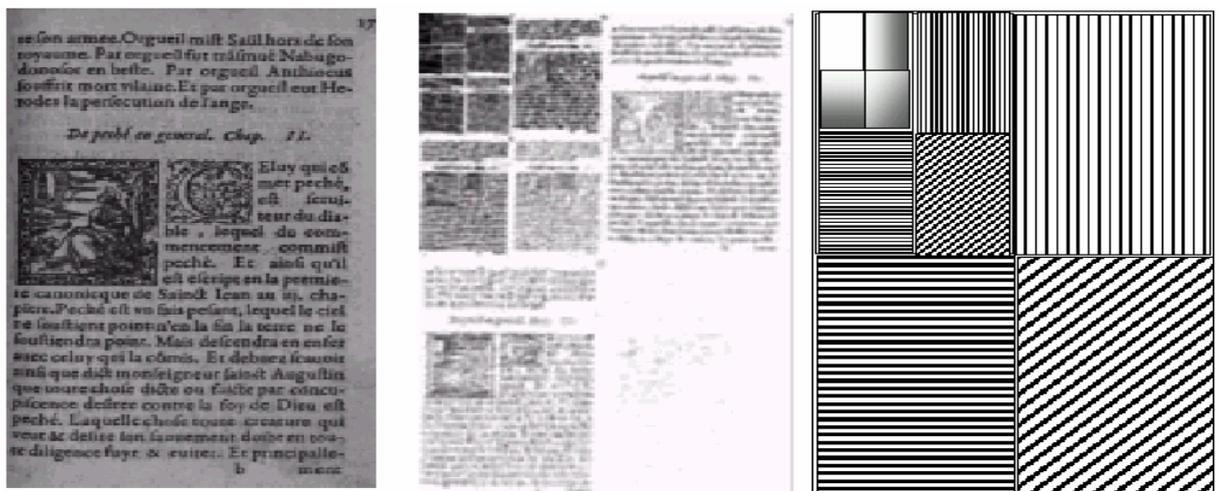
Inconvénients :

- 1- Temps de calcul très long, impossible à mettre en œuvre par toute une image.
- 2- Une complexité combinatoire élevée.
- 3- Une implémentation difficile à mettre en œuvre.

2.2.2.4.2 La transformée en ondelettes :

La transformée de Haar ou bien une autre transformée en ondelettes (cette technique fera l'objet du chapitre 3) adaptés aux images de texte étaient utilisées pour extraire les transformations à chaque échelle sur les alignements horizontaux, verticaux et obliques. Les résultats de la détection des lignes de texte par rapport aux zones graphiques peuvent être aussi bonnes avec une complexité de calcul très faible [8].

La transformée en ondelettes divise récursivement l'image en trois parties par changement de résolution et application des filtres différentiels directionnels (horizontale, verticale et oblique) du premier ordre dans chacune des parties (voir figure 2.8). Il existe donc une échelle pour laquelle les lignes apparaissent puis les blocs de lignes ainsi que les zones graphiques.



(a)

(b)

(c)

Figure 2.6 : Application de la transformée de Haar sur un document.

(a) : Image originale (b) : Transformée de Haar (c) : Orientation et résolution

Avantages :

- 1- Applicable à des images à niveaux de gris.
- 2- Ne nécessite pas une connaissance à priori sur la structure des documents.
- 3- Insensible à l'inclinaison.
- 4- Valables pour des documents contenant des images et des graphiques.

Inconvénients :

- 1- Le choix des seuils.
- 2- Choix de la taille du masque pour l'extraction de l'information statique.

2.2.3 Etude comparative des méthodes de segmentation

Les méthodes ascendantes sont plus sûres que les méthodes descendantes, car elles démarrent à partir du pixel pour former des blocs, mais elles nécessitent un temps de calcul ainsi qu'un espace mémoire très important. Par contre les méthodes descendantes sont appropriées uniquement aux documents avec une structure bien définie (par exemple les journaux, les articles et les revues) mais elles consomment moins de temps CPU par rapport aux méthodes ascendantes.

Une combinaison des deux approches peut être utilisée afin de réduire les inconvénients des méthodes ascendantes et descendantes.

2.2.4 L'approche utilisée pour la segmentation du document:

On a opté pour une méthode de segmentation basée sur l'analyse de la texture avec les ondelettes car elle est plus robuste que les méthodes traditionnelles [22].

Beaucoup d'algorithmes de segmentation utilisent *une classification par bloc* d'une certaine taille dans l'espoir que tous les pixels du bloc appartiennent à la même classe. La segmentation typique consiste alors à classer chaque bloc de pixels suivis de certains post-traitements.

Il est évident que la taille du bloc à classer est cruciale. Un large bloc augmente habituellement la fiabilité de la classification (parce que beaucoup de pixels fournissent des informations statistiquement plus riches), mais simultanément il y a risque d'avoir des pixels de différentes classes à l'intérieur du bloc. Ainsi, un large bloc produit des segmentations précises dans des grandes régions homogènes, mais une faible segmentation le long des frontières entre les régions. Un petit bloc réduit la possibilité d'avoir des classes multiples dans le bloc, mais une classification moins consistante globale. Ainsi, un petit bloc est mieux adapté aux frontières inter-régions.

Pour obtenir l'information de chaque région d'image à segmenter, les grandes échelles ainsi que les petites échelles sont utilisées pour segmenter correctement les larges

régions homogènes et les régions de frontière. Dans *la segmentation* multiechelle [23] les résultats des blocs classifiés de différentes tailles sont combinés pour obtenir une segmentation plus précise.

2.3 Les méthodes de classification régions

Les algorithmes de classification permettent de trouver la nature de chaque bloc obtenu par le processus de segmentation, chaque bloc peut appartenir à l'un des éléments suivants : texte, graphique, tableau, ... [12].

Parmi les méthodes de classifications, nous pouvons citer : la classification linéaire discriminante (Linear discriminant classifiers), l'arbre de classification binaire et les réseaux de neurones.

2.3.1 La classification linéaire discriminante

Cet algorithme a été décrit par Wang et al, il est basé sur le calcul de certaines caractéristiques des blocs obtenus par l'étape de segmentation qui est basée sur une technique de maculage. La classification permet de distinguer entre le texte et les images par le moyen d'une classification linéaire qui s'adapte avec des hauteurs différentes des caractères. Une étape supplémentaire de classification permet de résoudre les problèmes d'ambiguïtés dans la classification entre une seule ligne de texte, une ligne solide et une image[12].

2.3.2 Arbre binaire de classification (Binary classification tree)

C'est une technique statistique, proposé par Sivaramaakrishnan et al, elle permet de classer les régions homogènes extraites à partir du document. Les blocs sont supposés être rectangulaires avec des cotés alignés dans les directions : horizontal et vertical. La classification est basée sur 67 caractéristiques qui sont extraites à partir de chaque bloc[12].

2.3.3 Classification par réseau de neurone (Neural Network Classification)

Cette méthode permet de classer les blocs obtenus à partir d'une image binaire, en deux classes : le texte et le non-texte. Quatre modèles de réseaux de neurones ont été utilisés : « Back propagation », « Radial basis functions », « Probabilistic neural networks » et « Kohonen's self organizing features map », la motivation de base est

d'éliminer la sélection manuelle de plusieurs paramètres qui est une tâche critique [13] [12].

2.3.4 Classification des régions par la méthode de Jain et Yu

La méthode d'identification des régions présentée par Jain et Yu qui utilise un modèle descendant du document pendant la segmentation pour l'extraction des caractéristiques qui sont utilisées pour l'identification des régions.

Les entités appartenant à des régions de textes sont définies lors de l'obtention des lignes de textes qui sont des composantes connexes voisines. Une région de texte très petite est considérée comme une légende si elle est proche d'une région non textuelle. Ensuite, une région image est détectée après le groupement des lignes dans une région non textuelle. Enfin, en analysant les lignes extraites des régions non textuelles restantes, les tableaux peuvent être reconnus et séparés des régions de dessin [13].

2.4 Représentation de la structure physique des documents

La structure physique de chaque page d'un document est généralement stockée dans un arbre (voir figure 2.9), tel que le nœud racine est le document en entier et chaque nœud est associée à chaque bloc obtenu par la phase de segmentation.

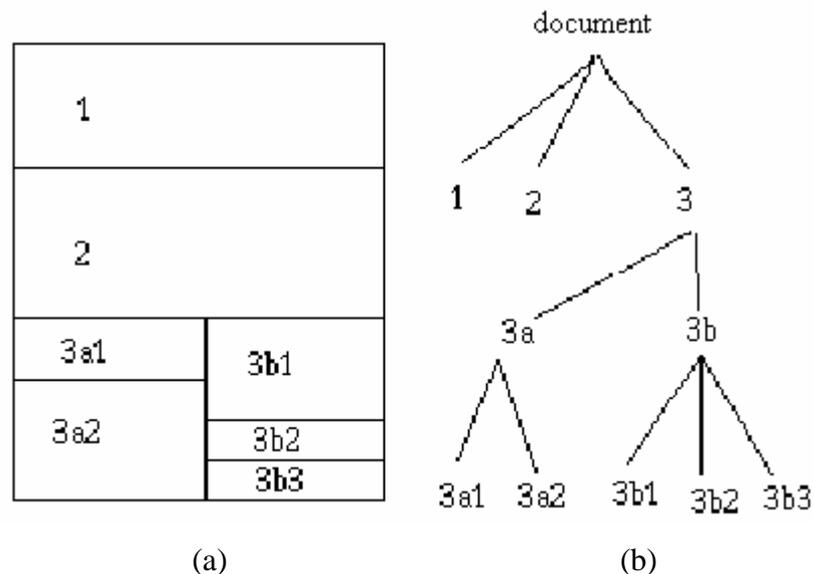


Figure 2.7 : Segmentation et structure physique

(a) Segmentation du document; (b) Représentation de la structure physique.

2.5 Conclusion

Dans ce chapitre, nous avons présenté l'analyse de la structure physique des documents, cette analyse est basée sur la segmentation, la classification et la représentation de la structure physique.

Il existe plusieurs méthodes de segmentation, nous pouvons citer les méthodes descendantes et les méthodes ascendantes. Les techniques descendantes sont très rapides lorsque la structure du document est connue au préalable. Par contre les techniques ascendantes sont plus robustes et plus exactes mais plus lentes.

Les approches basées sur l'analyse de texture traitent les composantes des documents: texte, fond, image et graphe comme étant différentes textures.

Dans notre étude, nous avons utilisé une méthode de segmentation basée sur l'analyse de la texture qui nécessite une transformation des données du document, afin de les rendre pertinentes, pour cela nous avons utilisé la Transformée en ondelettes qui donne une meilleure discrimination des zones textuelles. Cette méthode sera présentée dans le chapitre suivant.

CHAPITRE 3

INTRODUCTION A LA THEORIE DES ONDELETTES

3.1 Introduction

La transformée en ondelettes est une étape importante dans notre chaîne de traitement de l'image, son emploi doit s'intégrer parfaitement à cette chaîne. Ce chapitre présente la transformée en ondelettes et les raisons de son introduction dans le domaine du traitement du signal et de l'image (TSI), et plus particulièrement celui de l'analyse multi-résolution.

La transformée en ondelettes comble certaines lacunes de la transformée de Fourier, permettant notamment de régler la précision de l'analyse sur la fréquence ou sur le temps. D'autre part, la transformée en ondelettes est naturellement liée à l'analyse multi-résolution, et permet donc d'étudier l'information contenue dans un signal ou une image à différentes échelles, permettant de mettre en évidence des caractéristiques invisibles au premier abord [24][25].

3.2 L'analyse temps-fréquence :

3.2.1. La transformée de Fourier :

L'analyse de Fourier est une des bases majeures de la physique et des mathématiques. Elle est indissociable du traitement du signal, et ce pour deux raisons principales. La première est l'universalité du concept de fréquence sur lequel elle repose. La seconde tient à la structure même de l'analyse de Fourier qui se prête aisément à des transformations communes comme le filtrage linéaire en les traduisant de manière particulièrement simple.

$$X(\omega) = \int_{-\infty}^{+\infty} x(t) e^{-i\omega t} dt \quad 3.1$$

Le principe de la transformée de Fourier repose sur le fait que toute fonction périodique (ou pseudo-périodique) peut être représentée comme la somme d'une série de sinus et de cosinus dont on fait varier d'une part les amplitudes en les multipliant par des coefficients, et d'autre part les phases en les décalant de manière à ce qu'elles s'additionnent ou se compensent[26].



Figure 3.1 : La transformée de Fourier.

Cependant, l'analyse de Fourier, de par sa nature, montre assez vite ses limitations : son calcul nécessite la connaissance de toute l'histoire temporelle du signal . De plus, dans une transformée de Fourier, l'information sur le temps est présente (la transformée inverse est donc possible), mais elle est cachée dans les phases : il est en pratique impossible à extraire. On en est donc réduit à étudier un signal soit en fonction du temps, soit en fonction des fréquences qu'il contient, sans possibilité de conjuguer les deux analyses.

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} X(\omega) e^{i\omega t} d\omega \quad 3.2$$

L'analyse de Fourier, malgré l'introduction - et le succès - de la transformée de Fourier rapide (FFT), est donc par définition peu adaptée à l'étude des signaux non-stationnaires. Ces signaux se rencontrent beaucoup plus souvent que les signaux stationnaires. Pour leur étude, il faudra donc employer une autre méthode d'analyse qui tienne compte de l'information temporelle du signal [24][27].

3.2.2 La transformée de Fourier à fenêtre :

Pour pallier au manque d'information sur le temps dans la transformée de Fourier, une nouvelle méthode d'analyse est donc introduite : elle utilise une « fenêtre glissante ». Cette méthode, pouvant être adaptée aux signaux non-stationnaires, est très proche de l'analyse spectrale : on définit une fenêtre qui sera utilisée comme masque sur le signal, et

dans laquelle on considère que le signal est localement stationnaire, puis on décale cette fenêtre le long du signal afin de l'analyser entièrement [24][28].

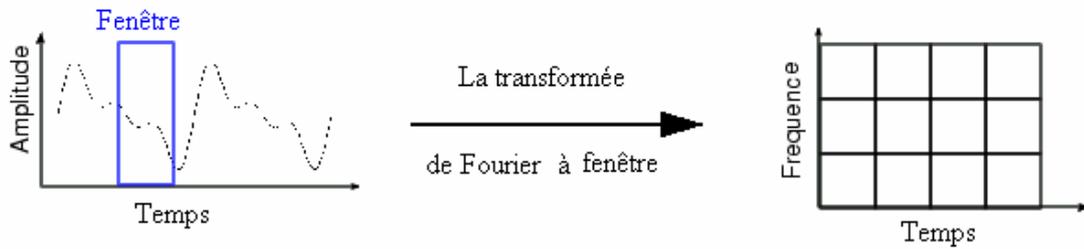


Figure 3.2 : La transformée de Fourier à fenêtre.

Gabor reprendra cette méthode dans les années 40 ; sa fenêtre est représentée par une fonction gaussienne [27]:

$$g_{a,b} = e^{iat} g(t - b) \quad a, b \in \mathfrak{R} \quad 3.3$$

où a représente le facteur d'échelle, et b le facteur de translation.

On constate que le membre $g(t - b)$ de l'équation 3.3 est indépendant de a , ce qui signifie que l'enveloppe de la fenêtre glissante sera constante : on aura donc une résolution fixe sur toute la durée du signal.

Ainsi, l'étude d'un signal avec la transformée de Gabor permet d'obtenir à la fois une information sur le temps et sur la fréquence, mais la résolution d'analyse est fixée par le choix de la taille de l'enveloppe [27]: si la fenêtre est trop petite, les basses fréquences n'y seront pas contenues, et si la fenêtre est trop grande, l'information sur les hautes fréquences est noyée dans l'information concernant la totalité de l'intervalle contenu dans la fenêtre.

3.3. La transformée en ondelettes

3.3.1. L'idée de la transformée en ondelettes et l'analyse multi-résolution :

La méthode de Gabor permet donc de décomposer un signal donné en une combinaison linéaire temps-fréquence judicieusement choisie. Mais l'on aimerait pouvoir faire varier la résolution d'analyse en fonction du signal afin de l'adapter à celui-ci : on n'a pas besoin d'un très haut niveau de résolution lorsque le signal est constitué uniquement

de basses fréquences, et il serait dommage d'étudier un signal comportant beaucoup de hautes fréquences avec une résolution trop basse.

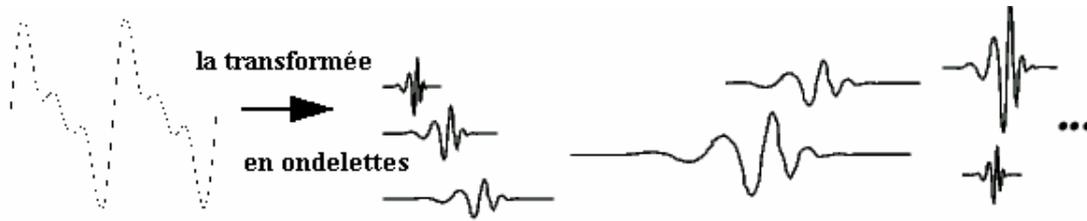


Figure 3.3 : La transformée en ondelettes.

Ainsi, plutôt que de choisir à l'avance une résolution adaptée à un type de signal donné, il serait préférable de disposer d'une méthode d'analyse dont la résolution, aussi bien en temps qu'en fréquence, s'adapte au signal en fonction de ses caractéristiques. Plutôt que de conserver une enveloppe fixe dans laquelle le nombre d'oscillations varie[26].

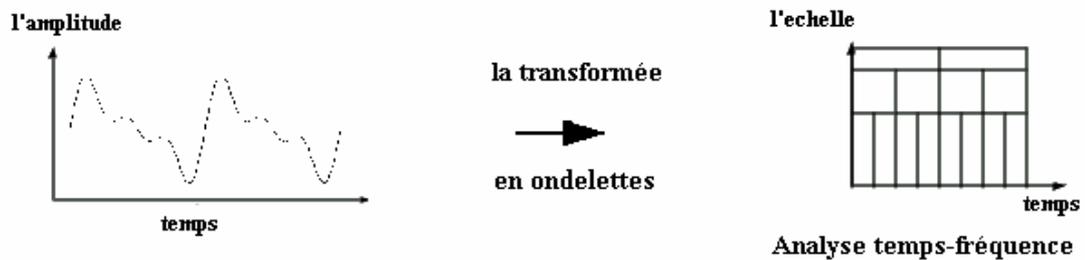


Figure 3.4 :L'analyse temps fréquence

On conserve un nombre d'oscillations constant dans une enveloppe que l'on peut contracter et dilater à volonté : c'est l'analyse multi-résolution, dont la figure 3.4 décrit le principe [27].

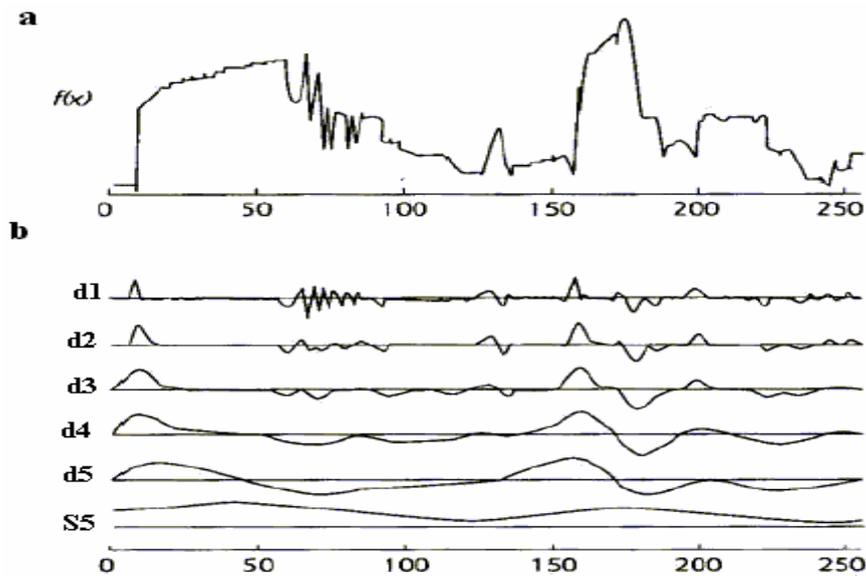


Figure 3.5: La transformée en ondelettes agit comme un *zoom* mathématique

Le signal d'origine (a) est décomposé en un ensemble de sous-signaux (b) selon la procédure suivante :

La fonction $f(x)$ est décomposée en un signal de détail **d1** et un signal d'approximation **S1** (qui n'est pas représenté ici). Ce signal **S1** est à son tour décomposé en un détail **d2** et une approximation **S2**, et ainsi de suite jusqu'à obtenir les cinq signaux de détails **d1** à **d5** correspondant aux détails des 5 premières résolutions, et le signal d'approximation à la résolution 5, **S5**.

Ainsi, l'analyse multi-résolution agit comme un *zoom mathématique* sur le signal à analyser: on peut en faisant varier l'échelle d'analyse du signal, extraire les détails présents à telle ou telle résolution.

3.3.2. La transformée en ondelettes continue :

La transformée en ondelettes utilise des translations et des dilatations d'une fonction fixe, l'ondelette mère $\psi \in L^2(\mathfrak{R})$ [27]. Dans le cas de la transformée continue, les paramètres de translation et de dilatation varient de manière continue. En d'autres mots, la transformée utilise les fonctions :

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right) \quad 3.4$$

Avec $a, b \in \mathfrak{R}, a \neq 0$, où a sert à *dilater* (compresser ou étendre) la fonction, et b sert à la *translater* (la déplacer selon l'axe des temps).

Quand on analyse un signal $f(x)$ avec ces ondelettes, on le transforme en une fonction de deux variables (le temps et l'échelle d'analyse du signal) qu'on peut appeler $W(a, b)$ qui représente les coefficients correspondant au signal f projeté sur la base d'ondelettes $\psi_{a,b}$:

$$W(a, b) = \langle f, \psi_{a,b} \rangle \quad 3.5$$

que l'on peut également noter :

$$W(a, b) = \frac{1}{\sqrt{|a|}} \int f(x) \psi_{a,b} dx \quad 3.6$$

Modèle graphique du calcul des coefficients d'ondelettes [28]:

L'équation 3.5 donne ce que l'on appelle le coefficient d'ondelette. Le calcul de ces coefficients et leur comportements peuvent être représentés graphiquement grâce aux figures 3.5 et 3.6.

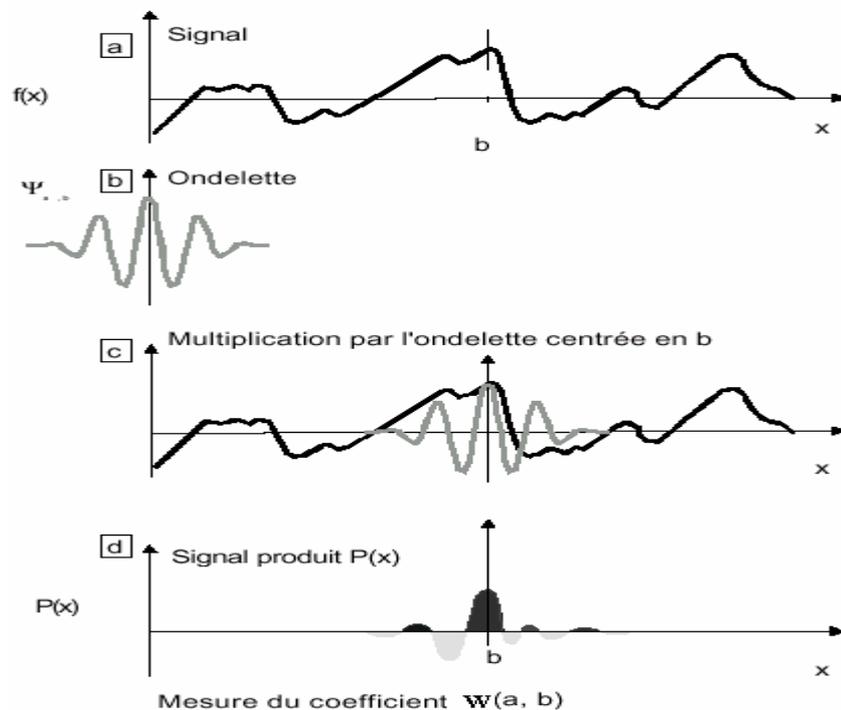


Figure 3.6: Modèle graphique du calcul du coefficient d'ondelette $W(a, b)$.

(a) Soit $f(x)$ un signal à analyser.

(b) On prend une ondelette analysante (ici l'ondelette de Morlet $y(t)$ caractérisée par une "fréquence" $1/a$).

(c) On centre l'ondelette sur une valeur b de l'espace et on multiplie le signal $f(x)$ par les valeurs de $y(x)$.

(d) On calcule l'aire du signal produit $P(x)$ ainsi obtenu.

L'aire est comptée positivement pour les parties de la courbe situées au-dessus de l'axe des abscisses (gris foncé) et négativement dans le cas opposé (gris clair). Cette aire est égale au coefficient d'ondelette $W(a, b)$ correspondant à la valeur de la "fréquence" $1/a$ et à la position b .

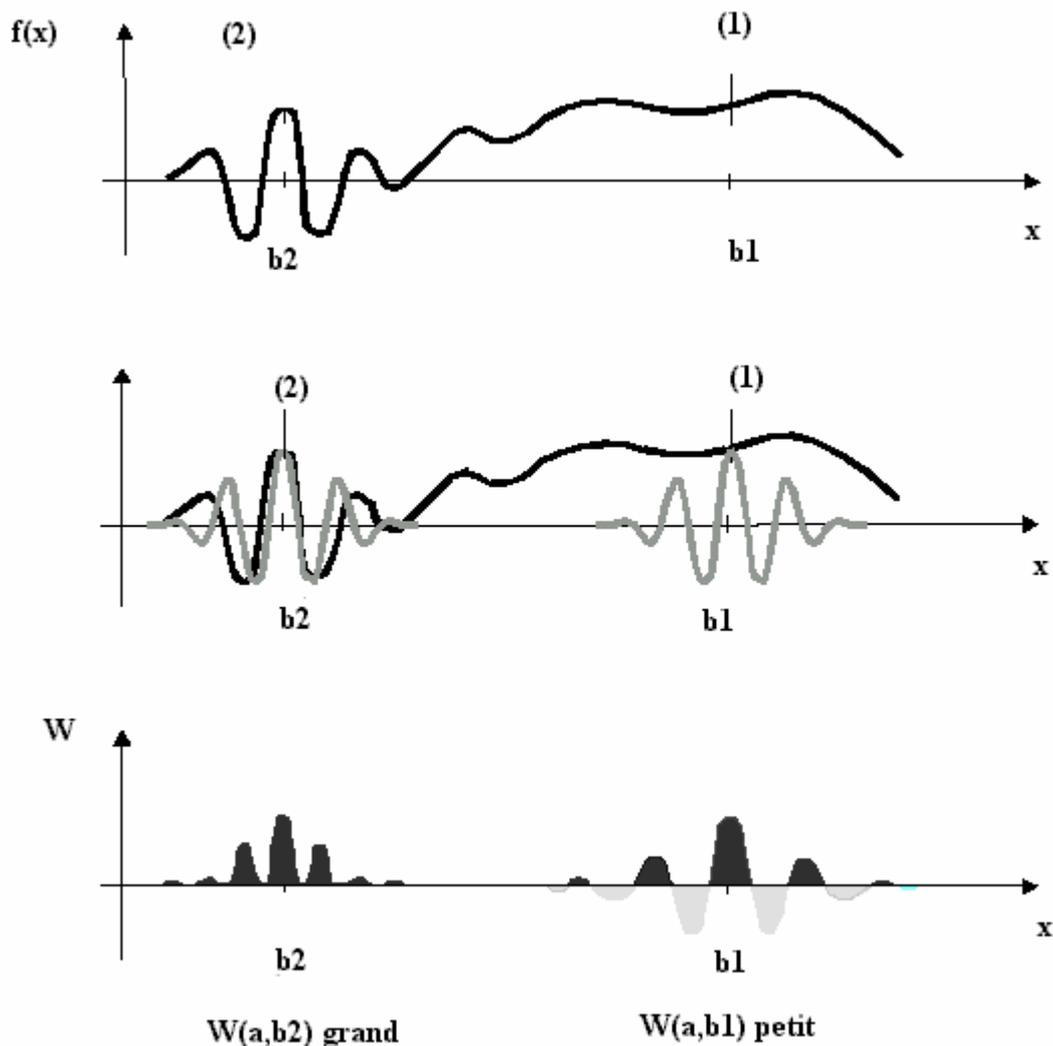


Figure 3.7 : Comportement du coefficient d'ondelette.

Le coefficient d'ondelette $W(a, b)$ d'un signal $f(x)$ dépend de la forme de celui-ci au voisinage de la position b . Quand $f(x)$ est à peu près constant (1) autour d'une position ($b = b_1$), le produit de $f(x)$ par $y(x)$ (d'aire nulle) a une aire très petite, donc $W(a, b_1)$ est très petit.

Quand au contraire la portion de signal autour de $b = b_2$ est irrégulière (2) et que ses variations sont de "fréquence" comparable à celle de l'ondelette, l'aire du produit signal-ondelette, donc $W(a, b_2)$, est généralement beaucoup plus grande.

Commentaire

Cette transformation est en théorie infiniment redondante, puisque l'ondelette est translatée de manière continue ; cependant il existe des méthodes pour diminuer cette redondance : l'une de ces méthodes consiste en l'emploi de la transformée en ondelettes discrète.

3.3.3 La transformée en ondelettes discrète

Contrairement à la transformée continue, dans laquelle l'ondelette est dilatée et translatée de manière continue, la transformée en ondelettes discrète translate et dilate l'ondelette selon des valeurs discrètes[27].

Ces coefficients a et b seront discrétisés de la manière suivante :

$$a = a_0^m \text{ et } b = n \cdot b_0 \cdot a_0^m, \text{ avec } a_0 > 1 \text{ et } b_0 > 0 \text{ fixés et appartenant à } \mathbb{Z}.$$

Les ondelettes sont alors définies de la manière suivante :

$$\psi_{m,n}(x) = \frac{1}{\sqrt{a_0^m}} \psi \left(\frac{1}{a_0^m} x - n b_0 \right) \quad 3.7$$

Nous avons vu que la transformée en ondelettes continue est infiniment redondante. La transformée en ondelettes discrète permet de diminuer fortement cette redondance sans l'annuler entièrement : les signaux sont décrits par un nombre fini d'échantillons. Lorsque que l'on voudra obtenir un signal transformé aussi concis que possible (en pratique, un signal transformé qui contient autant d'échantillons que le signal d'origine) on utilisera alors un cas particulier de la transformée en ondelettes discrète : la transformée en ondelettes orthogonale.

La transformée dyadique

Lorsque l'on pose $a_0 = 2$ et $b_0 = 1$, on parle de transformée dyadique [25].

$$\psi_{m,n}(x) = \frac{1}{\sqrt{2^m}} \psi \left(\frac{1}{2^m} x - n \right) \quad 3.8$$

3.3.4 l'analyse multi-résolution

Une autre manière de décrire la transformée en ondelettes est de dire qu'elle consiste à obtenir une approximation d'un signal f par projection de celui-ci sur l'espace d'approximation V_j grâce à un opérateur A_j [27]. Cette approximation faisant perdre une partie de l'information, la reconstruction du signal nécessite d'associer à l'opérateur A_j un opérateur D_j projetant le même signal sur l'espace de détail W_j tel que $A_{j+1}f = A_j f + D_j f$. Les espaces d'approximation et de détail sont construits respectivement à partir des fonctions de base $\varphi(x)$ et $\psi(x)$.

L'analyse multi-résolution permet de rapprocher l'analyse en variable d'espace avec l'analyse en variable de fréquence. Elle formalise l'idée intuitive, selon laquelle tout signal peut être construit par raffinements successifs, c'est-à-dire par l'ajout de détails lorsque l'on passe d'une résolution à la suivante. D'une manière plus précise, une analyse multi-résolution de $L^2(\mathfrak{R})$ est définie comme une suite de sous-espaces fermés V_j de $L^2(\mathfrak{R})$, $j \in \mathbb{Z}$, ayant les propriétés suivantes [27]:

1. $V_j \subset V_{j+1}$ l'approximation à la résolution a_{j+1} contient toutes les informations nécessaires pour calculer le même signal à la résolution inférieure a_j .
2. $v(x) \in V_j \Leftrightarrow v(2x) \in V_{j+1}$, si $v(x)$ appartient à V_j , la même fonction dilatée d'un facteur 2 appartient à V_{j+1} .
3. $v(x) \in V_0 \Leftrightarrow v(x-k) \in V_0 \Rightarrow$ si $v(x)$ appartient à V_j , la même fonction traduite d'un facteur quelconque appartient aussi à V_j .
4. $\bigcup_{j=-\infty}^{+\infty} V_j$ est dense dans $L^2(\mathfrak{R})$ et $\bigcap_{j=-\infty}^{+\infty} V_j = \{0\} \Rightarrow$ chaque fonction $f(x)$ d'énergie finie ($f \in L^2$) peut être approximée avec une précision arbitraire par une fonction de V_j d'une hauteur j adéquate.

5. Il existe une *fonction d'échelle* $\varphi \in V_0$, ayant une intégrale non nulle, telle que l'ensemble $\{\varphi(x-k) | k \in \mathbb{Z}\}$ est une base de Riesz de V_0 . \Rightarrow il existe une fonction φ telle que $\{\varphi(x-k)\}$ soit une base orthonormée de V_0 . La fonction φ et toutes ses translatées d'un facteur entier forment une base orthonormée.

Cette dernière propriété de l'analyse multi-résolution introduit une nouvelle fonction, la fonction d'échelle.

a. Fonction d'échelle et sous-espaces

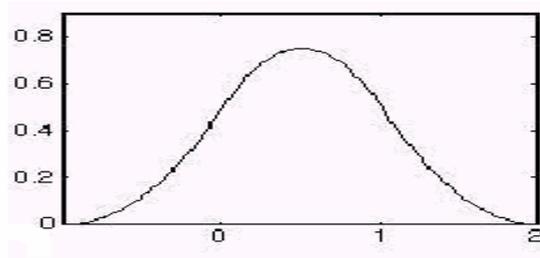


Figure 3.8 : Une fonction d'échelle B-spline

Nous pouvons faire quelques observations à propos de la définition précédente.

Comme $\varphi \in V_0 \subset V_1$, il existe une suite $(h_k) \in l^2(\mathbb{Z})$ telle que la fonction d'échelle satisfasse :

$$\varphi(x) = \sqrt{2} \sum_k h_k \varphi(2x-k) \tag{3.9}$$

On constate que la définition des fonctions d'échelles est similaire à celle des ondelettes. Elles sont construites par dilatation et translation d'une fonction unique, appelée parfois *père des ondelettes*. Chaque fonction de V_0 peut être écrite comme une combinaison linéaire des fonctions de base $\varphi(x) = \sqrt{2} \sum_k h_k \varphi(2x-k)$ de V_1 .

Cette équation fonctionnelle porte différents noms dont les plus utilisés sont «*équation de dilatation*» ou «*équation de raffinement*».

b. Fonction ondelette et espaces de détail

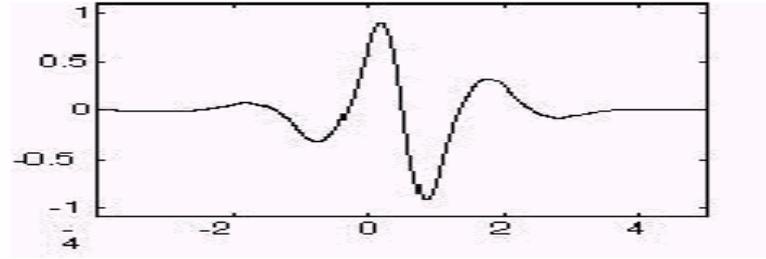


Figure 3.9 : l' ondelette B-spline

Nous avons vu plus haut que $A_{j-1}f = A_j f + D_j f$. Ceci revient à dire :

$$V_{j+1} = V_j \oplus W_j \quad 3.10$$

En d'autres termes, chaque élément de V_{j+1} peut être écrit de manière unique comme la somme d'un élément de W_j et d'un élément de V_j . Notons que les espaces W_j eux-mêmes ne sont pas nécessairement uniques ; il peut y avoir plusieurs moyens de compléter V_j en V_{j+1}

L'espace W_j contient l'information de détail nécessaire pour passer d'une approximation à la résolution j à la résolution $j + 1$:

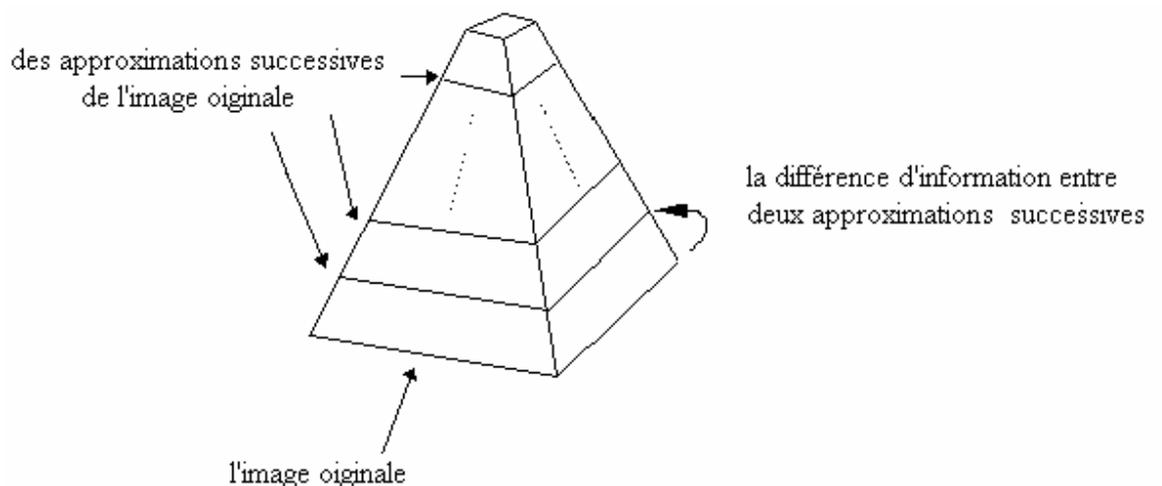
$$\oplus W_j = L^2(\mathfrak{R}) \quad 3.11$$

Une fonction ψ est une *ondelette* si l'ensemble des fonctions $\{\psi(x-l) | l, j \in \mathbb{Z}\}$ est une base de Riesz de $L^2(\mathfrak{R})$. La définition de $\psi_{j,l}$ est similaire à celle de $\varphi_{j,l}$ de la section précédente. Ainsi, de la même manière que précédemment, comme une ondelette est également un élément de V_0 , il existe une suite $\{g_k\}$ telle que :

$$\psi(x) = \sqrt{2} \sum_k h_k \psi(2x-k) \quad 3.12$$

3.3.5 Les algorithmes pyramidaux

Si on ne se préoccupait pas du temps de calcul, on pourrait décomposer le signal en ondelettes en le comparant, à chaque échelle, aux ondelettes de taille appropriée (fonction échelles et analyse **multi-échelle**). Le retour, à chaque étape, au signal initial est très lente. Il est plus astucieux de se servir du travail déjà fait. Afin de calculer plus vite, on analyse le signal par "tranches de résolutions consécutives" (fonctions ondelettes et analyse **multi-résolution**) [29].



Ondelette: traduit la "différence" d'information entre 2 approximations successives de l' image

Figure3.10: Pyramide multi-résolution

La première étape de ces algorithmes consiste à séparer le signal en deux composantes, une composante lisse qui donne l'allure générale du signal et une composante qui en fournit les détails. L'image lisse est le signal tel qu'on le voit à la moitié de la résolution la plus fine, avec deux fois moins d'échantillons. On obtient cette image à l'aide d'un filtre passe-bas qui correspond à la fonction échelle. Les détails s'obtiennent en utilisant un filtre passe-haut. Ils permettent de reconstituer le signal initial à partir de l'image lisse.

La deuxième étape consiste à répéter la procédure sur le signal à une résolution moitiée. On sépare ce signal lissé en deux parties : un signal encore plus lissé (vu à un

quart de la résolution du signal initial) et de nouveaux détails deux fois plus grands que les précédents. Pour cela, on dilate d'un facteur 2 la fonction échelle et l'ondelette. On réitère ensuite la procédure.

3.3.6. Algorithme de décomposition –reconstruction de S Mallat

Le principe de cet algorithme pour les ondelettes est présenté dans la figure 3.11. Soit $x(n)$ un signal échantillonné correspondant au signal d'origine. Ce signal est décomposé sur plusieurs niveaux de résolutions en deux bandes de fréquences (passe-haut et passe bas) de la manière suivante, approximation et détails.

3.3.6.1 Décomposition/reconstruction.

L'opération de décomposition passe par plusieurs itérations [30]. Une itération est représentée par le schéma suivant :

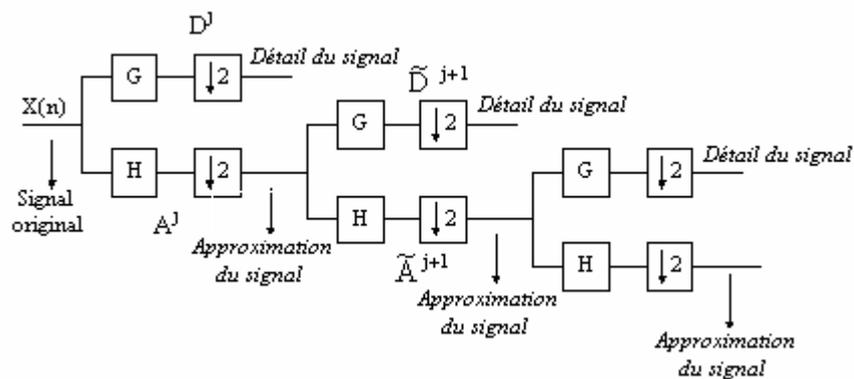


Figure 3.11: Schéma de L'algorithme de Décomposition de S.Mallat

A^j : signal d'approximation de niveau j

\tilde{A}^{j+1} : signal issu de la convolution de A^j avec le filtre H

D^j : signal de détail de niveau j .

\tilde{D}^{j+1} : signal issu de la convolution de A^j avec le filtre G

Avec :

$$\tilde{A}^{j+1}(n) = \sum_k h(k-n)A^j$$

$$\tilde{D}^{j+1}(n) = \sum_k g(k-n)A^j \quad 3.14$$

Décimation :

L'opération de décimation consiste à prendre un échantillon sur deux d'un signal [31] d'où :

$$A^{j+1}(n) = \tilde{A}^{j+1}(2n) = \sum_k h(k-2n)A^j(k) \quad 3.15$$

$$D^{j+1}(n) = \tilde{D}^{j+1}(2n) = \sum_k g(k-2n)A^j(k) \quad 3.16$$

Le signal A^{j+1} représente une version lissée du signal A^j

D^{j+1} représente la différence d'information entre A^j et A^{j+1} .

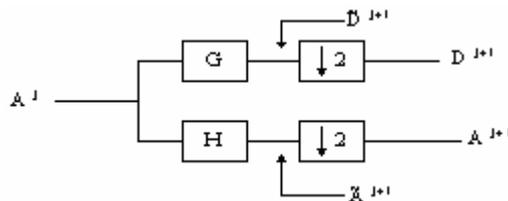


Figure 3.12 : Schéma de la décimation

3.3.6.2 Reconstruction.

La reconstruction est l'étape inverse de la décomposition avec une interpolation

Le schéma de reconstruction du signal original est le suivant [30] :

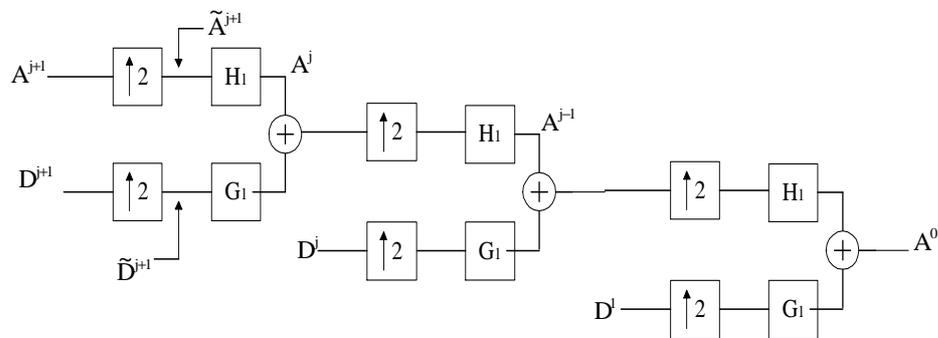


Figure 3.13 : Processus de reconstruction du signal

Interpolation par des zéros :

L'interpolation par des zéros consiste à insérer un zéro tous les deux échantillons successifs [31].

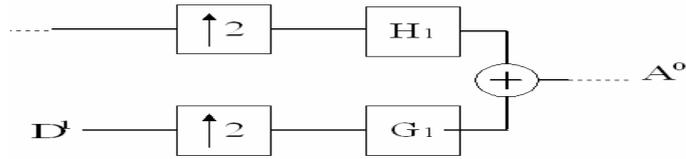


Figure 3.14 : Processus d'interpolation du signal

3.3.5.4. Application aux images

En deux dimensions, la représentation en ondelettes peut être vue de la même manière qu'une représentation en ondelettes sur chacun des axes x et y . On utilise donc une extension de l'algorithme pyramidal à une dimension. A chaque étape on décompose

$$A^d_{2^{j+1}} f \quad \text{en :} \quad \left\{ \begin{array}{l} A^d_{2^j} f, \\ D^1_{2^j} f, \\ D^2_{2^j} f, \\ D^3_{2^j} f, \end{array} \right. \quad \text{où :}$$

L'approximation :

$$A^d_{2^j} f = ((f(x, y) * \phi_{2^j}(-x) \phi_{2^j}(-y))(2^{-j}n, 2^{-j}m)) \text{ pour } (m, n) \text{ appartenant à } \mathbb{Z}^2 \quad 3.17$$

Les coefficients horizontaux (C^H) :

$$D^1_{2^j} f = ((f(x, y) * \phi^j(-x) \psi^j(-y))(2^{-j}n, 2^{-j}m)) \text{ pour } (m, n) \text{ appartenant à } \mathbb{Z}^2 \quad 3.18$$

Les coefficients verticaux (C^V) :

$$D^2_{2^j} f = ((f(x, y) * \psi_{2^j}(-x) \phi_{2^j}(-y))(2^{-j}n, 2^{-j}m)) \text{ pour } (m, n) \text{ appartenant à } \mathbb{Z}^2 \quad 3.19$$

Les coefficients (C^D) :

$$D^3_{2^{\wedge}j} f = ((f(x, y) * \psi_{2^{\wedge}j}(-x) \psi_{2^{\wedge}j}(-y))(2^{-j}n, 2^{-j}m)) \text{ pour } (m, n) \text{ appartenant à } \mathbb{Z}^2 \quad 3.20$$

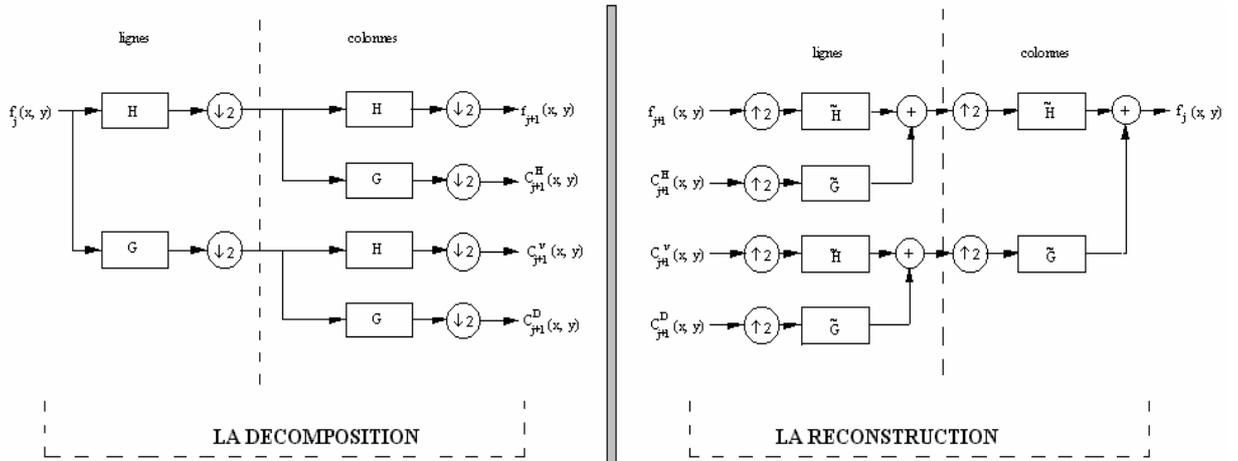


Figure3.15 schéma de décomposition/reconstruction multi-résolution

(cas bidimensionnel)

On convolue d'abord les colonnes de $A^d_{2^{\wedge}j+1} f$ avec un filtre à une dimension, on garde toutes les colonnes impaires, on convolue l'image résultante avec un autre filtre à une dimension, on retient les lignes impaires autres colonnes. Les filtres utilisés dans cette décomposition sont les filtres miroir en quadrature H ($L : low$) et G ($H : height$). On itère le procédé ci-dessus en faisant varier j . Ceci correspond à une décomposition en filtres miroirs conjugués séparables. Les coefficients d'ondelettes ainsi obtenus ont une grande amplitude au voisinage des contours et dans les textures selon une orientation spatiale donnée.

Les expressions (3.17)-(3.20) montrent qu'en deux dimensions $A^d_{2^{\wedge}j} f$ et $D^k_{2^{\wedge}j} f$ sont calculés à l'aide de filtres séparables selon les abscisses et les ordonnées. La décomposition en ondelettes peut alors être vue comme la décomposition d'un signal en un ensemble de bandes de fréquences orientées dans l'espace. $\phi(x)$ peut être considéré comme un filtre passe-bas parfait et ψ comme un filtre passe-bande.

La décomposition de $A^d_{2^{\wedge}j+1} f$ est telle que:

- $A^d_{2^{\wedge}j} f$: correspond aux basses fréquences LL,
- $D^1_{2^{\wedge}j} f$: donne les hautes fréquences verticales (contours horizontaux) LH,
- $D^2_{2^{\wedge}j} f$: donne les hautes fréquences horizontales (contours verticaux) HL,
- $D^3_{2^{\wedge}j} f$: donne les hautes fréquences diagonales (contours diagonaux) HH,

Cette décomposition est illustrée dans la figure ci-dessous :

LL3	LH3	LH2	LH1
HL3	HH3		
HL2		HH2	
HL1			HH1

Figure 3.16 : Matrices d'une transformation de niveau 3.



Figure 3.17 : Une image (à gauche) et ses coefficients d'ondelettes (Matrices de droite).

3.4 Intérêt en traitement d'image

L'emploi de l'analyse multi-résolution dans le traitement d'images fournit un certain nombre d'avantages selon l'application envisagée :

➤ *Compression :*

Le taux de compression de l'image pourra facilement être adapté à la qualité désirée. En effet, en éliminant les détails contenus dans certains niveaux de résolution, jugés peu importants, on pourra fortement augmenter le taux de compression, en conservant uniquement l'information nécessaire. Si l'on désire conserver une bonne qualité d'image, il suffira alors de conserver tous les détails ;

➤ *Détection de contours :*

La détection de contours est une tâche ardue lorsque les images traitées présentent des variations brusques dans des zones inintéressantes ; en pratique, on voudrait pouvoir ignorer certains contours et ne conserver que les plus représentatifs. L'analyse d'un arbre par exemple présente des contours principaux, ceux de l'arbre lui-même, ainsi que des contours plus précis (ceux du feuillage). En fonction de l'application, on voudrait pouvoir conserver uniquement les contours principaux (l'arbre), ou bien conserver l'aspect de texture (le feuillage).

Ce type d'analyse est permis par la multi-résolution. En analysant l'image à une résolution grossière, après avoir éliminé les détails, les informations sur le feuillage auront disparu. *A contrario*, l'analyse à une résolution plus précise nous donnera tous les contours présents dans l'image.

➤ *Détection/reconnaissance de texture :*

L'analyse multi-résolution apporte un avantage considérable dans le domaine de la reconnaissance de texture, puisque l'échelle est prise en compte en plus des paramètres habituels de détection comme les motifs.

Bien évidemment, la transformée en ondelettes ne se limite pas aux applications mentionnées ci-dessus.

On peut également citer la reconnaissance de visage, la détection de mouvements, et d'une manière générale la plupart des applications reposant sur l'analyse d'images.

D'autre part, un avantage non négligeable de la transformée en ondelettes est qu'elle n'est pas liée à une fonction prédéfinie comme l'est la transformée de Fourier qui utilise les fonctions sinus et cosinus exclusivement. Ainsi, le choix de l'ondelette utilisée pour l'analyse pourra dépendre de l'application envisagée, voire même du type de donnée traitée.

3.5 Conclusion

Nous avons vu dans ce chapitre que la transformée en ondelettes présente de nombreux avantages dans le domaine du traitement du signal et de l'image. C'est un outil puissant de transformation du signal qui permet d'adapter celui-ci à un traitement ultérieur dont le principal intérêt est de «préparer» le signal dans le but de faciliter le traitement envisagé. Ainsi, elle permet de réduire la redondance pour améliorer la compression d'une image ; elle peut également extraire les informations importantes (texture, contours, etc.) contenues dans une image. Le tableau 3.1 résume les différences entre

les différentes transformées présentées : transformée de Fourier, transformée de Fourier à fenêtre (transformée de Gabor) et transformée en ondelettes.

	Transformée de Fourier	Transformée de Gabor	Transformée en ondelettes
Fonctions utilisées	sinus cosinus	sinus, cosinus gaussienne	ondelette fonction d'échelle
Paramétrable	non	en partie	oui
Précision	aucune	fixe	variable

Tableau 3.1: Comparaison entre la transformée en ondelettes, la transformée de Gabor (Fourier à fenêtre), et la transformée de Fourier

CHAPITRE 4 CONCEPTION ET MISE EN ŒUVRE

4.1 Introduction

Dans le chapitre précédent, nous avons vu les différentes techniques de la segmentation des documents. L'étude de ces différentes techniques, nous a permis de trouver leurs limites. L'exploitation de ces limites peut nous aider à réaliser un système plus fiable et robuste.

L'élaboration d'un tel système est très difficile sachant qu'il existe plusieurs facteurs que nous devons prendre en considération, parmi eux, nous pouvons citer :

- La qualité de numérisation des documents.
- La structure du document est différente d'un document à l'autre.
- L'inclinaison des documents.

Afin de réaliser ce système, nous avons utilisé la segmentation multiechelle pour l'extraction du texte. Ainsi que la transformée de Hough pour la détection de l'angle d'inclinaison. Un tel choix a été fait pour les raisons suivantes :

- ❖ La méthode multiechelle est une méthode de segmentation ascendante qui permet de segmenter le document en blocs homogènes, cette méthode :
 - ne nécessite pas un modèle de décomposition.

- et elle peut être appliquée sur n'importe quel type de document.
- N'est pas influencé par le fond.
- N'est pas influencé par l'angle d'inclinaison.

Mais elle nécessite la fixation des seuils utilisés dans la classification des blocs.

- ❖ La transformée de Hough permet de détecter des formes particulières dans une image (droite, courbe,...).

Dans ce qui suit, nous présentons l'architecture générale de notre système qui consiste à segmenter des documents inclinés multi fontes, la segmentation est basée essentiellement sur une analyse multiechelle, mais des traitements supplémentaires sont indispensables pour améliorer la fiabilité de cette analyse.

4.2 Système de segmentation de document

Notre système a pour but de segmenter les images des documents en séparant les régions texte des régions graphiques, image et du fond, afin de reproduire la version électronique. Nous avons choisi une analyse multiechelle qui nous a permis de calculer des paramètres statiques des 3 bandes hautes fréquences LH, HH, HL qui caractérisent les blocs pour faciliter la discrimination entre eux.

Des pré traitements sont nécessaires pour éliminer quelques problèmes dues à l'estimation des seuils, ils seront détaillés dans ce qui suit.

L'architecture générale de notre système est montrée par la figure 4.1 :

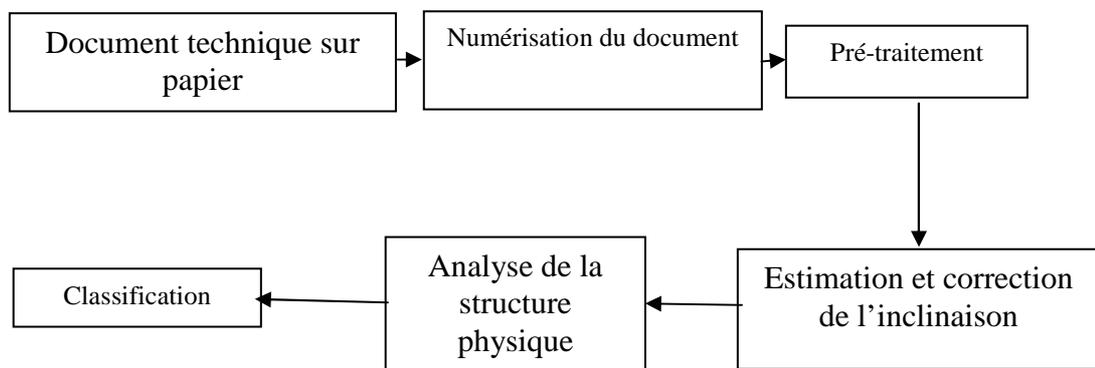


Figure 4.1 : Structure d'un système d'analyse des documents.

4.2.1 Architecture du système de segmentation de documents.

4.2.1.1 Acquisition et prétraitement

L'acquisition d'un document est faite à l'aide d'un scanner. Généralement, les images acquises nécessitent, avant la phase de segmentation, quelques opérations de prétraitement et qui comprennent.

L'élimination du bruit, l'estimation de l'angle d'inclinaison, le redressement de l'image.

4.2.1.1 Estimation de l'angle d'inclinaison

L'inclinaison des images est due essentiellement au mauvais positionnement des documents lors de l'étape de numérisation. L'estimation de l'angle d'inclinaison est nécessaire dans certaines techniques de segmentation comme celle que nous allons utiliser.

Pour détecter l'angle d'inclinaison, nous avons utilisé la transformée de Hough qui permet de détecter l'angle d'inclinaison, nous avons utilisé la transformée de Hough qui permet de détecter les droites dans les images, ce choix est justifié par le fait que cette méthode possède un intervalle de détection qui varie entre 0° et 180° . Malgré qu'elle possède des inconvénients comme :

- La sensibilité au bruit
- Intervalle de détection réduit pour les documents contenant des images, des tableaux et des graphiques.
- Temps de traitement très important.

Pour palier à ces inconvénients, nous avons suivi les étapes suivantes :

- La transformée en ondelettes a été introduite et seule la bande LH est utilisée, car cette bande représente le contour horizontal, ceci élimine la totalité des zones graphiques et zones d'images [15].

Le but dans la présente partie est de :

- ❖ Préserver les lignes;
- ❖ Enlever le bruit;
- ❖ Assurer une première réduction : le fait de travailler avec la première résolution sur un seul quadrant de l'image.

- Un seuillage global a été défini afin d'assurer [9] :
- Une deuxième réduction du nombre de points
- L'obtention d'une image binaire qui est indispensable pour la transformée de HOUGH.
- Une transformée de Hough a été introduite pour la détection de l'angle d'inclinaison.

Cette méthode a été testée sur des documents latins et arabes et elle a donné de bons résultats concernant la détection de l'angle d'inclinaison et la réduction du temps de traitement. Nous avons testé cette méthode sur des documents arabes, les résultats seront présentés dans le cinquième chapitre.

L'algorithme de la transformée de Hough

```

Pour chaque point (x, y)
{
    Pour  $\theta_1 \leq \theta \leq \theta_2$  /*  $[\theta_1, \theta_2]$  est l'intervalle de détection */
    {
        Calculer  $\rho = x \cos\theta + y \sin\theta$ 
        Incréments  $H(\rho, \theta)$ 
    }
}

Trouver le maximum dans le tableau  $H(\rho, \theta)$ 

```

La valeur maximale dans le tableau accumulateur est associée à l'angle d'inclinaison.

4.2.1.2 Correction de l'angle d'inclinaison

La correction de l'inclinaison s'effectue par rotation de toute l'image en fonction de l'angle d'inclinaison ($\alpha = -\theta$), ainsi les nouvelles coordonnées (x', y') de chaque point (x, y) de l'image seront calculées comme suit :

Pour chaque point (x, y)

$$\left\{ \begin{array}{l} x' = x \cos \alpha + y \sin \alpha \\ y' = y \cos \alpha - x \sin \alpha \end{array} \right.$$

4.2.2 La Segmentation

L'étape de segmentation est très importante dans un système d'analyse et de reconnaissance de documents. Elle permet de localiser les blocs d'information homogènes.

Pour segmenter notre document scanné en trois classes: texte, image et fond, l'image est divisée en blocs de taille de $N \times N$, et chacun de ces blocs est classifié (voir figure 4.2).

Si les blocs $N \times N$ ne peuvent pas être classifiés, la longueur du bloc est changée pour avoir une segmentation plus précise.

Chaque bloc de $N \times N$ est transformé en utilisant la transformée en ondelettes à un niveau, afin de former un vecteur des coefficients de l'ondelette, en prenant les deux bandes LH et HL.

Pour classifier chaque bloc, sa variance est calculée :

- 1- Si la variance est proche de zéro, le bloc est classifié comme fond.
- 2- Si la variance approche des valeurs extrêmes, le bloc est classifié comme texte ou image;
- 3- Aautrement, le bloc est marqué comme indéterminé.

Le bloc indéterminé est alors divisé en quatre sous blocs $N/2 \times N/2$, qui sont à leur tour classifié individuellement en tant que texte, fond, image ou indéterminé.

- Si le sous bloc est indéterminé, l'information de contexte est utilisée pour classifier le sous blocs.
- Si le sous bloc a des blocs adjacents comme étant texte ou image, le bloc est localisé sur une frontière.

- La classification ne peut pas être améliorée en fusionnant le bloc avec les blocs d'entourage, puisque le grand bloc fusionné contiendrait les types mélangés. le bloc est supposé inclassable parce que la dimension de l'échantillon est trop petite.
- La fusion des huit blocs $N \times N$ d'entourage avec le bloc concerné est effectuée pour former un bloc de taille $3N \times 3N$, le sous bloc initial de $N/2 \times N/2$ est alors classifié selon la variance du bloc $3N \times 3N$.

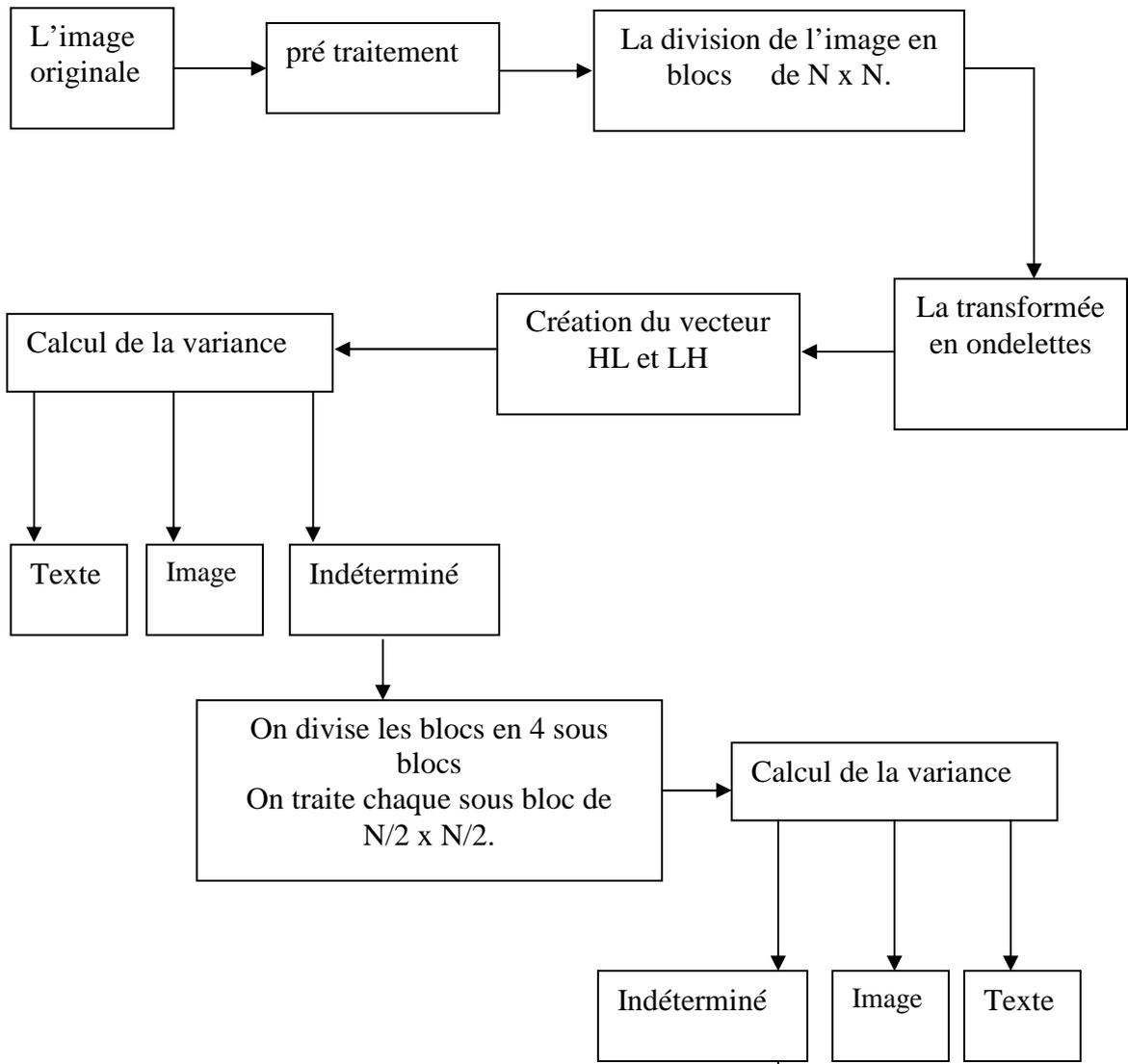


Figure4.2 : Les étapes de segmentation

4.2.2.1 Étape 1 : Segmentation par bloc :

Les pixels de l'image sont divisés en blocs dyadiques, cette étape présente les avantages suivants :

- Segmentation simple.
- Pas de redondance d'information.

La taille du bloc varie suivant les régions

Quelques exemples :

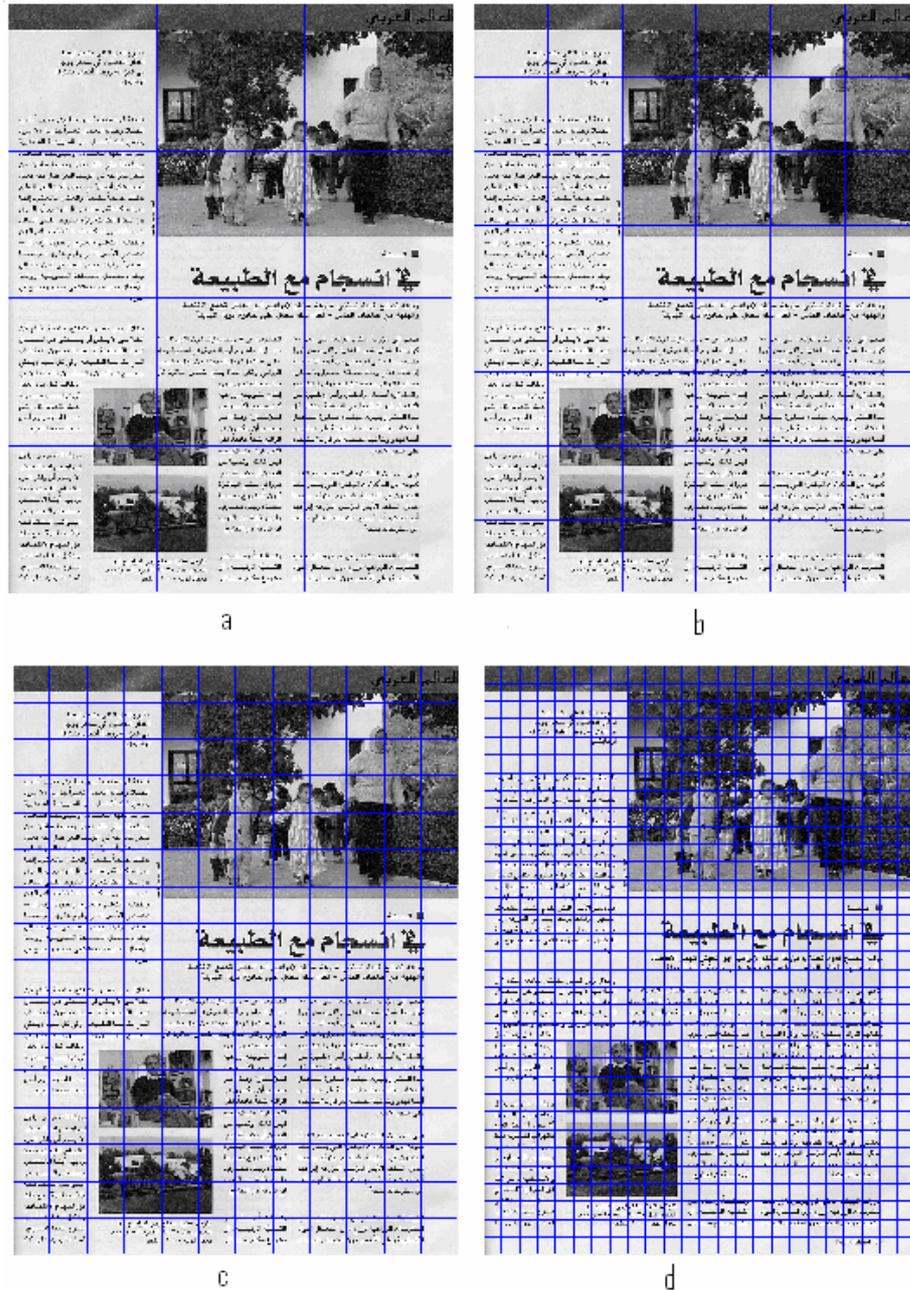


Figure 4.3: Segmentation en blocs de différentes tailles
 (a) 256*256 (b) 128*128 (c) 64*64 (d) 32*32

Remarque :

On remarque plus que le bloc a une taille grande, plus que il y aura le risque d'avoir des classes mixtes (voir figure 4.3).

4.2.2.2 Etape2 : La transformée en ondelettes

L'utilisation de la transformée d'ondelettes pour décomposer l'image (appliquée a chaque bloc de l'étape précédente):

- ❖ Donne des approximations successives à l'image par décimation après le passage par un filtre passe bas.
- ❖ A la capacité de détecter les contours après le passage par un filtre haut (détaille).

Ceci peut être facilement vu dans la décomposition d'image comme illustre la figure 4.4, montre l'image originale et sa décomposition au premier niveau.

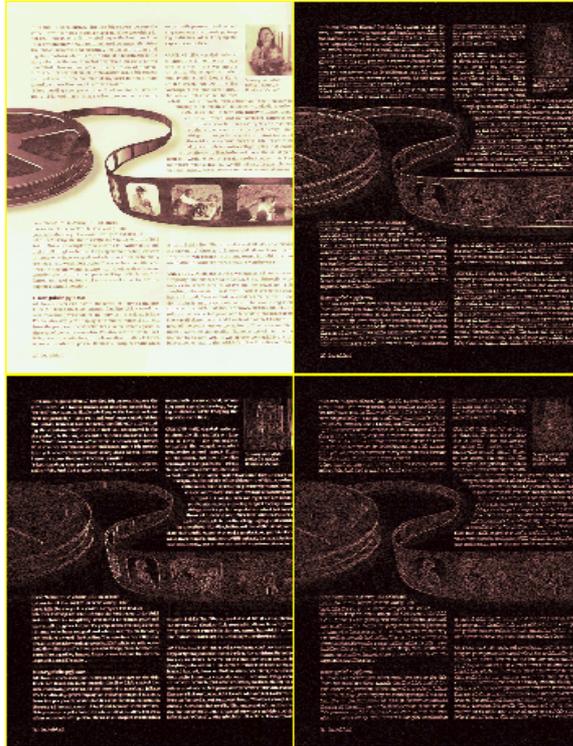


Figure4.4 : La décomposition en ondelette d'un document image

On remarque que les régions de type texte montrent l'activité la plus élevée dans les trois hautes sous-bandes de fréquence (High-Low HL, Low-High, LH, High-High HH).

Ceci est due à leur nature locale, seulement les ondelettes qui sont localisées proche du Contour donnent les grands coefficients, rendant ainsi les régions de type texte discernables dans les sous-bandes à haute fréquence.

Les filtres de db2 sont comme suit :

$$h_k = -0.1294 \quad 0.2241 \quad 0.8365 \quad 0.4830$$

$$g_k = -0.4830 \quad 0.8365 \quad -0.2241 \quad -0.1294$$

Pour une image $I(x, y)$ représenté comme

$$I(x, y) = \begin{bmatrix} i_{0,0} & i_{0,1} & \cdot & \cdot & \cdot & i_{0,2N-1} \\ i_{1,0} & i_{1,1} & \cdot & \cdot & \cdot & i_{1,2N-1} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ i_{2N-1,0} & i_{2N-1,1} & \cdot & \cdot & \cdot & i_{2N-1,2N-1} \end{bmatrix}_{2N \times 2N}$$

Nous pouvons employer l'algorithme du Mallat [17] pour avoir la transformée en deux dimensions de db2 de $I(x, y)$.

Pour quoi db2 ?

Le choix de l'algorithme des ondelettes dépend de l'application. L'algorithme d'ondelette de Haar a l'avantage d'être simple pour calculer et plus facile à comprendre. Mais l'algorithme de Daubechies permet de reprendre les détails qui ne sont pas détectés par l'algorithme d'ondelette de Haar.

Exemple :

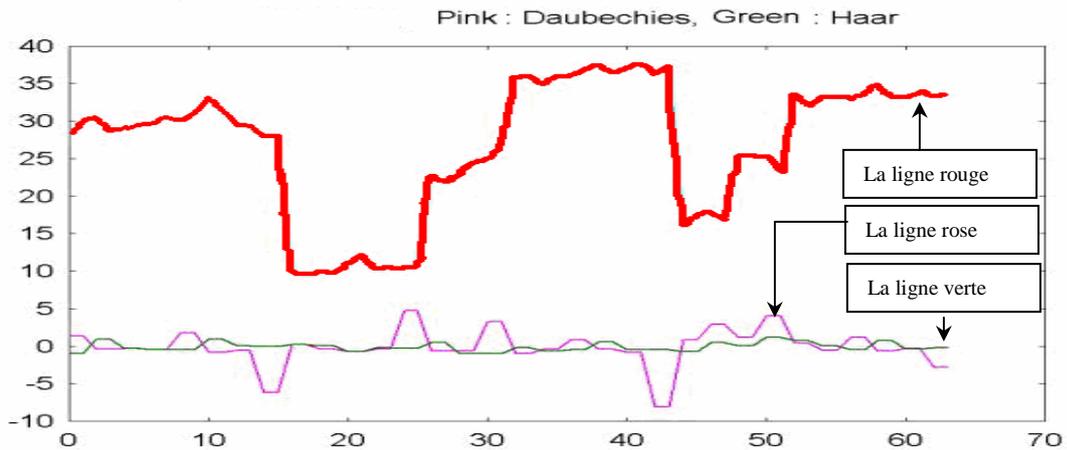


Figure 4.5 : comparaison entre Haar et db2

Dans la figure 4.5 :

- La ligne rouge : montre un signal avec des variations (ressemble à une ligne scannée, voir appendice 2).
- La ligne verte : montre les coefficients d'ondelette de Haar
- La ligne rose : montre les coefficients d'ondelette de Daubechies

A. La distribution des coefficients d'ondelettes dans les bandes à haute fréquence

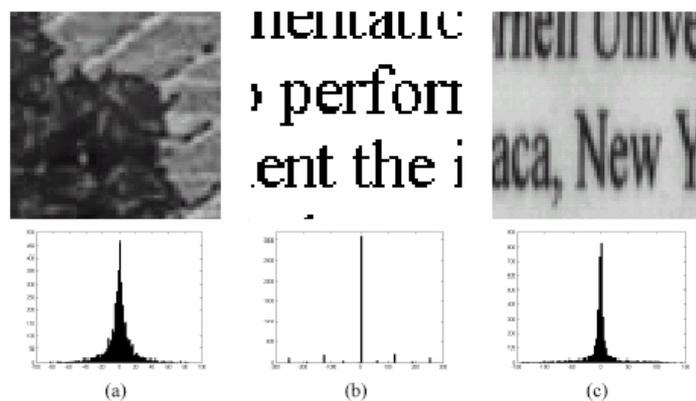


Figure4.6 : Histogramme des coefficients de Haar

(a) bloc d'une photo; (b) bloc d'un texte généré par micro ; (c) bloc d'un texte scannée

La première ligne : l'image originale ;, deuxième ligne : l'histogramme

On a observé que pour la photographie, les coefficients d'ondelettes des bandes hautes fréquence : LH HL HH comme montrées dans figure 4.6 (a) tendent à suivre une distribution Laplacienne. Cette approximation est controversée dans certaines applications. Donc c'est un moyen de distinction des images de tonalité continue du graphique et du texte.

Par exemple, l'histogramme des coefficients d'ondelette de Haar dans la bande LH pour une image de photographie, une image de graphique, et une image des textes sont illustrés dans la figure 4.6. Une autre différence importante à noter est la continuité de la distribution observée.

Les histogrammes des coefficients de l'image du graphique et de l'image des textes suggèrent que les valeurs soient fortement concentrées sur quelques valeurs discrètes, mais l'histogramme pour l'image de photographie montre une continuité bien meilleure de la distribution [32].

Dans le cas spécial montré dans la figure 4.6(b). L'histogramme de l'image des textes contient cinq barres indépendamment loin l'un de l'autre parce que l'image des textes a des intensités à deux niveaux. Pour l'image de graphique, on remarque qu'un pourcentage très élevé des données localisées autour de zéro. Bien que la concentration des données autour du zéro soit absolue. On remarque que la quantité de données non nulle non négligeable.

Pour l'image de photographie, il n'existe aucune valeur couvrant la quantité suffisante de données à concentrer d'une façon fortement convaincante. L'histogramme a une crête au zéro et s'atténue rudement d'une façon exponentielle.

Nous précisons que dans la pratique les histogrammes des trois types ne sont habituellement pas aussi précis que ceux de l'exemple de la figure 4.6. L'ambiguïté se produit le plus avec la classe graphique parce que c'est une classe intermédiaire entre la photographie et le texte.

Les caractéristiques du Laplacien donnent des bons résultats pour les images générées par l'ordinateur, mais des faibles performances une fois appliquées aux images scannées. Comme montré dans la figure 4.6 (c). C'est parce que le texte des document scannées ne varie pas abruptement au loin du noir au blanc mais descend par à coup, passant par beaucoup de nuances de gris.

En conséquent, un mélange des coefficients d'ondelettes est produit, pas simplement quelques valeurs concentrées. Ceci est connu sous le nom de « texte sale ».

Ainsi l'aptitude du dispositif Laplacien de distinction de l'image est régit par le problème du texte sale, au point que l'erreur de classification est autour de 25%. Pour remédier à ce problème on passe à la résolution suivante.

Jusqu'ici, nous avons présenté la transformée d'ondelettes 2d de Haar et son exécution pratique.

B. La variation de l'histogramme des coefficients d'ondelettes en fonction de la taille du bloc :

B.1. Bloc homogène :

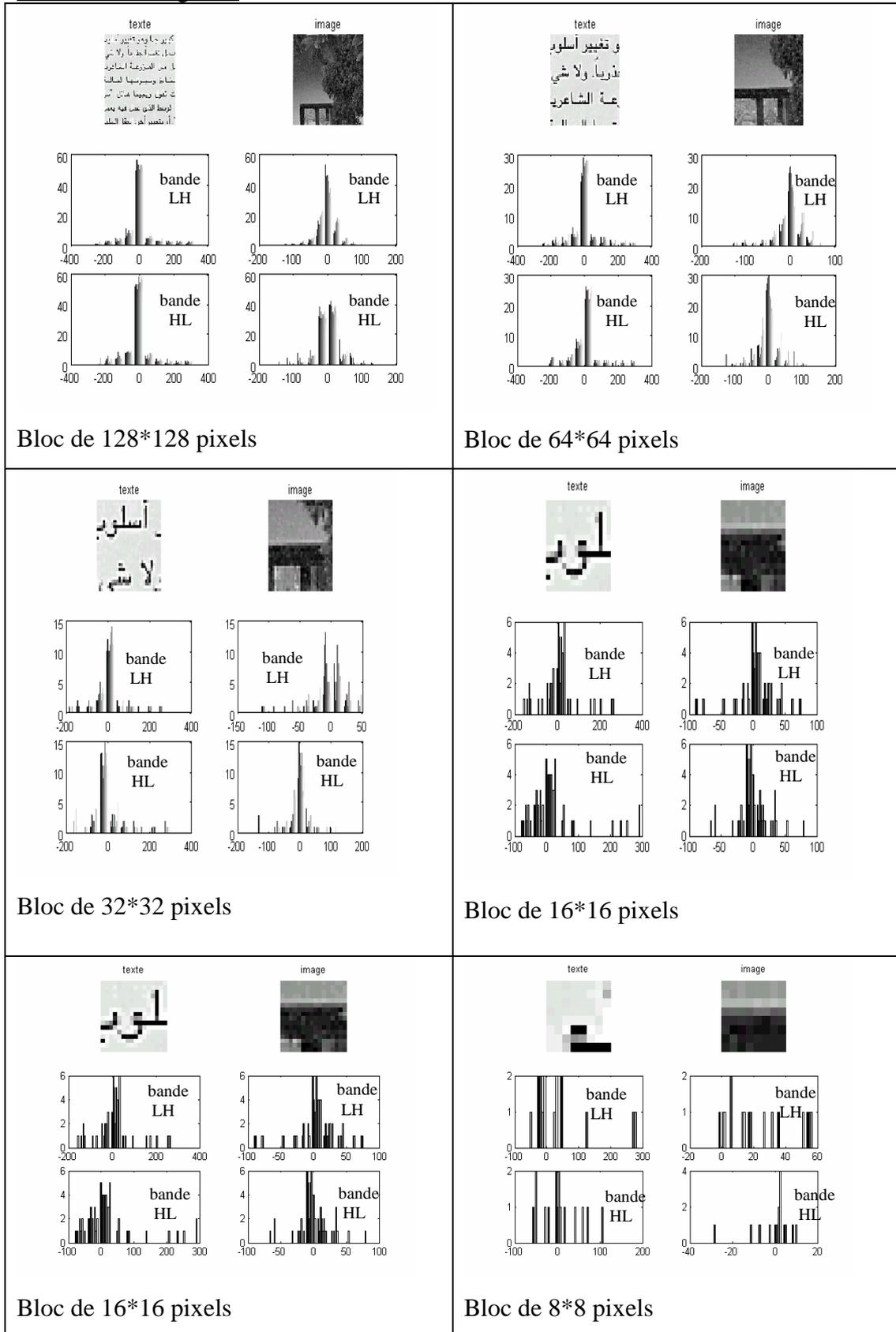


Tableau 4.1 : Variation des histogrammes selon avec la taille du bloc homogène.

Remarque :

On remarque que plus la taille diminue, la discrimination n'est plus évidente

Plus la taille est grande, plus on aura le risque d'avoir des blocs mixtes (hétérogène)

B.2. Bloc hétérogène :

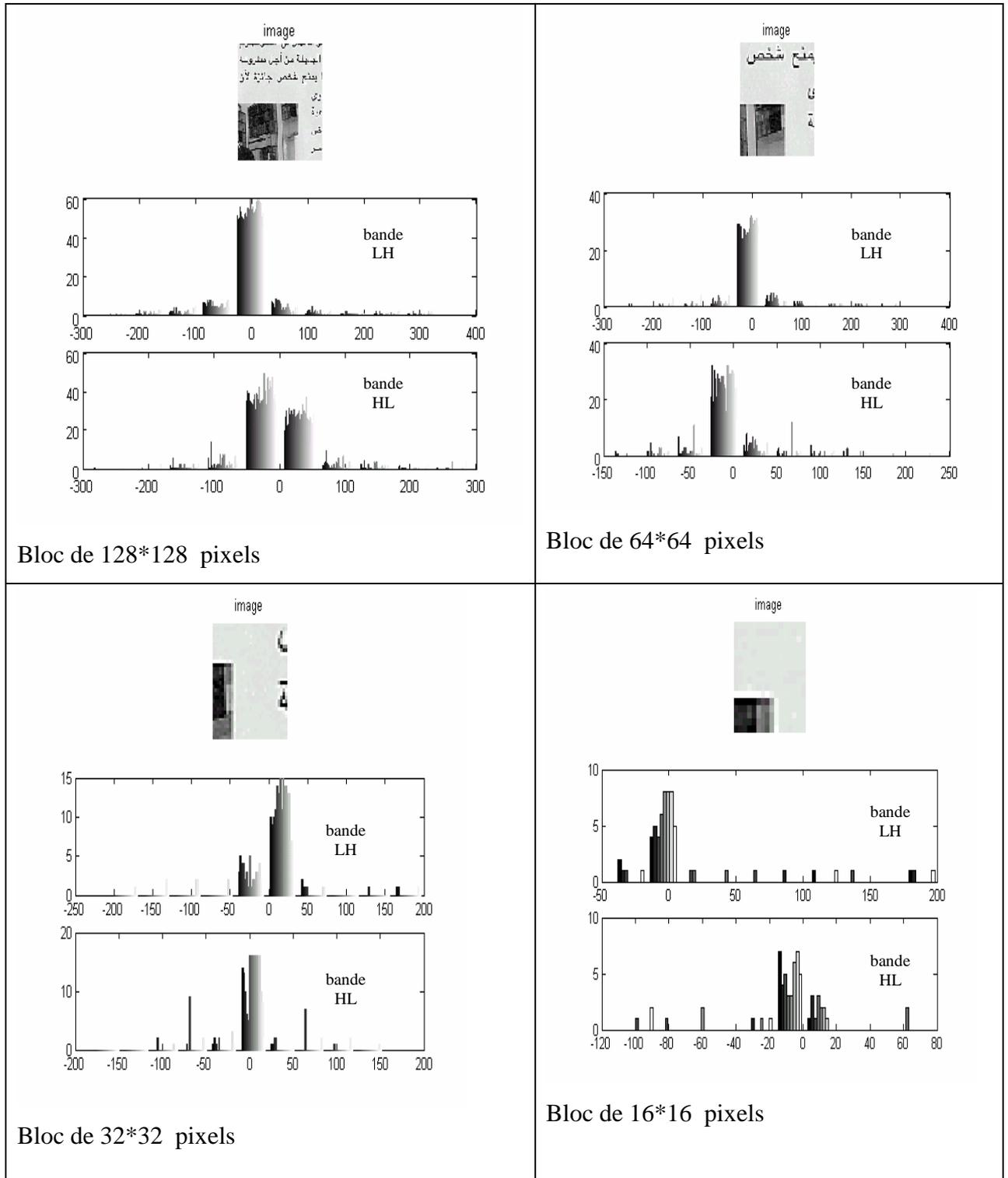


Tableau 4.2: Variation des histogrammes avec la taille du bloc hétérogène.

Commentaire :

Pour avoir une bonne discrimination, on doit faire un compromis sur la taille du bloc.

4.2.2.3 Etape3 : *Calcul du vecteur caractéristique*

Nous extrayons le vecteur caractéristique à partir de la décomposition d'ondelettes de l'image (chaque bloc tout seul). Nous utilisons la variance comme vecteur caractéristique pour les blocs d' $I(x, y)$, Nous calculons la variance par :

$$\text{var} = \mu_2(I) = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (I(i, j) - M(I))^2 \quad 4.1$$

$$\text{avec :} \quad M(I) = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} I(i, j) \quad 4.2$$

La déviation standard des coefficients d'ondelettes (σ) dans un bloc d'image est un dispositif faisable pour la classification des images scannées [29]. Il peut être pris comme dispositif primaire pour identifier les blocs de fond.

4.2.2.4 Etape 4 : *Classification pour la première résolution*

Si la variance est proche de zéro, le bloc est classifié comme fond. Si la variance approche des valeurs extrêmes (suivant certains seuils) le bloc est classifié comme texte ou image ; autrement le bloc est marqué comme indéterminé.

4.2.2.5 Etape5 : *Classification pour la deuxième résolution*

Le bloc indéterminé est alors divisé en quatre sous blocs 16x16, qui sont à leur tour classés individuellement en tant que texte, fond, image ou indéterminé.

4.2.2.6 Etape 6 : *Rregroupements des blocs adjacents en croissance*

Si le sous bloc est indéterminé, l'information de contexte est utilisée pour classifier le sous bloc. Si le sous bloc a des blocs adjacents comme étant texte ou image, le bloc est localisé sur une frontière. La classification ne peut pas être améliorée en fusionnant le bloc avec les blocs d'entourage, puisque le grand bloc fusionné contiendrait les types mélangés. le bloc est supposé inclassable parce que la dimension de l'échantillon est trop

petite. La fusion des huit blocs 32×32 d'entourage avec le bloc concerné est effectuée pour former un bloc de taille 96×96 , le sous bloc initial de 16×16 est alors classifié selon la variance du bloc 96×96 .

4.2.3 Le post traitement :

C'est l'étape finale dans la chaîne de notre traitement, et il est constitué de deux opérations :

a. Operation1 : Suppression des blocs isolés :

Pour bien améliorer le résultat on a supprimé les régions (des blocs) indésirables qui ont une taille proportionnelle à la résolution utilisée.

b. Opération 2 : Dilatation :

On a remarqué aussi qu'il existe des trous au niveau des zones texte, pour résoudre ce problème on a utilisé une dilatation avec un élément structurant de taille proportionnelle à l'application.

Dilatation :

La **dilatation** d'un objet X par un élément structurant B est l'ensemble des positions u telles que le translaté de B en u touche X . [33]

La dilatation d'une image par un disque (voir figure 4.7):

- connecte les objets quand ils sont proches ;
- comble les trous étroits présents dans les objets ;
- élargit les objets d'une taille correspondant au rayon de l'élément structurant.

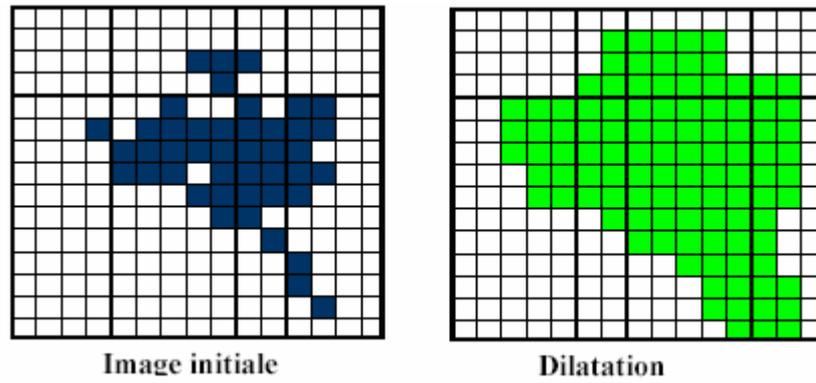


Figure 4.7 : Exemple de dilatation

4.3 Conclusion :

Dans ce chapitre, nous avons présenté notre approche. Pour avoir une bonne idée sur les performances de notre système, différents tests ont été réalisés, ils seront présentés dans le prochain chapitre.

CHAPITRE 5 : TESTS ET RESULTATS

5.1 Introduction :

Dans ce chapitre, notre objectif est l'évaluation des différentes étapes qui contribuent à la segmentation de documents développés dans le chapitre précédent. Afin de réaliser cette évaluation, nous avons effectué des tests sur plusieurs documents.

5.2 Environnement d'évaluation :

Nos développements ont été réalisés avec le langage MATLAB 6.5, sur une machine dotée d'un microprocesseur Celeron à une fréquence de 1.7Ghz et une mémoire de 128 Mo. Pour l'évaluation, nous avons testé notre méthode sur deux bases de données:

- Base 1 : Une base de données de documents contient : 100 documents latin et 100 documents arabes de types article de la même revue. Ces documents ont été scannés à l'aide d'un scanner « Epson Expression 1600 Pro » avec une résolution de 100 dpi. Les images ont été sauvegardées sous le format bmp.
- Base 2 : Quelques échantillons de la base de données de documents de l'université de Oulou de Finland. Cette base contient 512 documents de différents types (article, cheque bancaire, courrier, dictionnaire, formulaire, formules mathématiques, journaux, annuaire téléphonique, programmes). Ces documents ont été scannés à l'aide d'un scanner « Hewlett Packard ScanJet » avec une résolution de 400 dpi. Les images ont été sauvegardées sous le format TIFF. Cette base de données a été utilisée par différents chercheurs afin de tester la fiabilité de leur système.



Document latin



Document Arabe

Figure 5.1 : Exemple de quelques documents de la base 1.



Journaux



Chèque bancaire



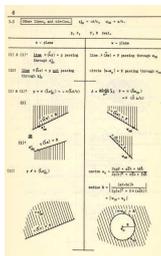
Courier



Dictionnaire



Formulaire



Mathématiques



Note musicale



Programme



Article



Annuaire
téléphonique

Figure 5.2 : Exemple de quelques documents de la base de données de Oulu base 2.

5.3 Démarche d'expérimentation :

Nous avons testé notre système sur 100 documents de la base 1 et sur 100 documents de la base 2.

Pour chaque partie de notre système, comme décrit dans le chapitre précédent, nous avons effectué une série de tests afin d'évaluer les performances de notre système.

5.3.1 Estimation et correction de l'angle d'inclinaison :

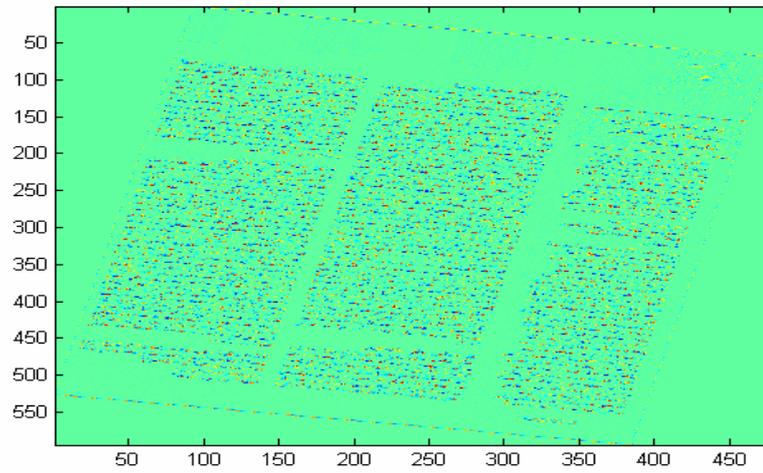
Pendant l'étape de numérisation, les documents ont été disposés d'une manière inclinée afin d'obtenir des images mal redressées.

Pour l'évaluation, nous avons testé notre méthode sur une base qui a été utilisée dans l'article [17] afin de comparer les résultats :

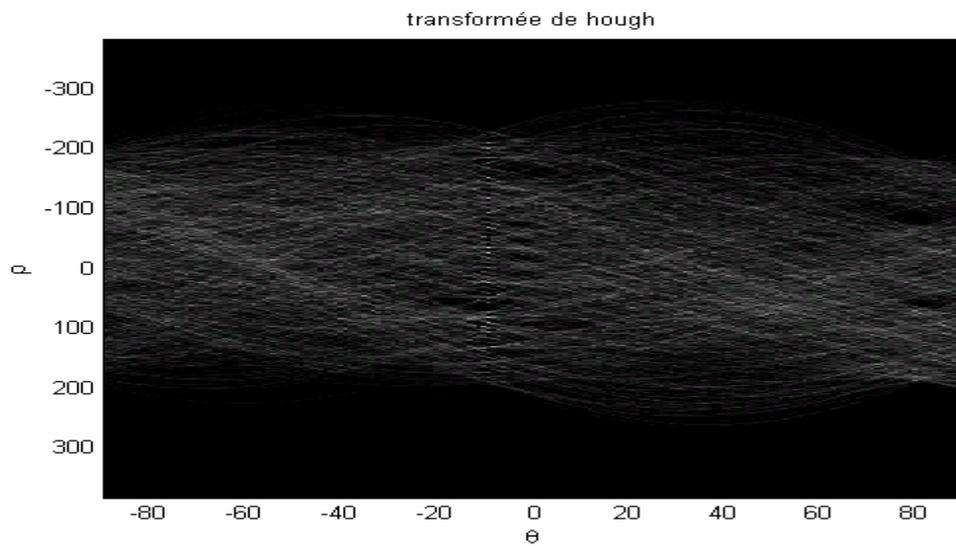


Figure 5.3 : Exemple d'un document incliné utilisé pour le test.

Dans notre système, avant d'appliquer la transformée de HOUGH sur le contour horizontal (la bande LH) du document, un seuillage est effectuée afin de réduire le temps de traitement.



Etape1 : Le contour horizontal;



Etape3 : La transformée de HOUGH;



Etape4 : Correction d'inclinaison

Figure 5.4 : Les différentes étapes de la détection/correction d'inclinaison

Evaluation :

Afin de tester l'exactitude de l'algorithme de détection proposé dans notre mémoire, nous l'avons appliqué sur la même base utilisée dans l'article [13], qui contient cent images de documents arabes imprimés. Ces dernières ont été obtenues en scannant des livres, des lettres et des revues avec un scanner «Epson Expression 1600 Pro»; la résolution de numérisation était de 100, 150 et 200 dpi. Les images sont sauvegardées sous format BMP (noir et blanc). Quelques images sont présentées dans le tableau 5.1.

Les tests ont été établis en effectuant des rotations sur les documents scannés; l'angle de rotation testé lors des expériences varie entre 0° et 180° . Le déplacement dans l'espace de Hough était de un pour les axes ρ et θ ($\Delta\rho=1, \Delta\theta=1$).

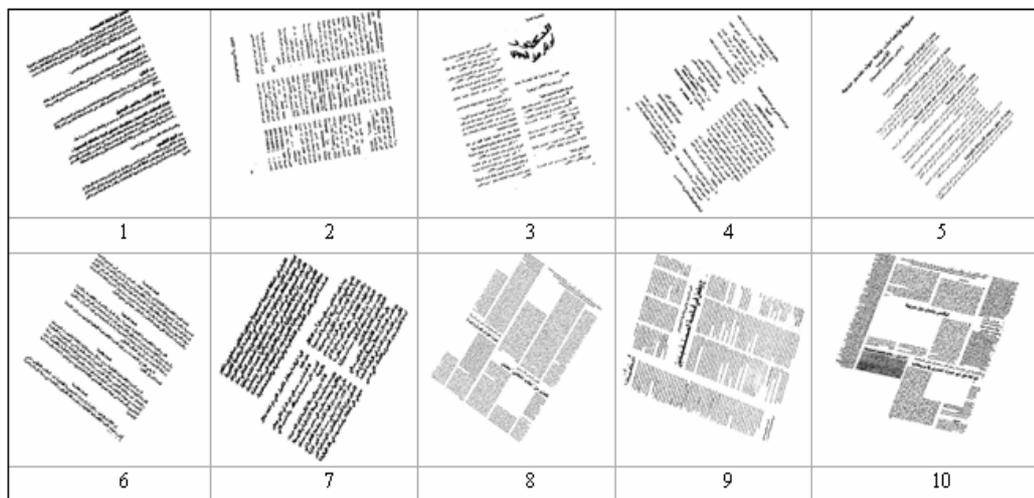


Tableau 5.1 : Exemple de quelques documents inclinés

Le tableau 5.2 illustre l'efficacité de notre algorithme sur un échantillon de 10 images choisies arbitrairement sur l'ensemble des 100 images traitées. Afin de comparer notre méthode avec la méthode originale de Hough, et la transformée de Hough appliquée sur les points centraux des liaisons, nous mentionnons dans ce tableau la taille de l'image, la résolution et le nombre de points traités.

Comme nous pouvons le constater sur ces exemples, que notre algorithme permet de réduire le temps d'exécution qui est due à la réduction du nombre de points utilisés, ce nombre est très petit par rapport au nombre de points dans la première méthode, mais il est proche du nombre de points dans la deuxième méthode.

Image	Taille	Résolution	Nombre de points utilisés		
			(2)	(1)	(3)
1	714*781	100	38627	445	205
2	1237*1429	150	124168	809	357
3	1801*2045	200	189136	862	511
4	1577*1617	100	87130	586	588
5	2149*2149	200	158332	884	1372
6	945*909	100	43859	551	221
7	567*590	100	28895	458	238
8	2709*2673	200	344065	2025	912
9	2595*2286	200	372100	4229	903
10	2014*1575	150	279376	3220	662

Tableau 5.2 : tableau comparatif du nombre de point utilisé

- (1) La transformée de Hough appliquée sur tous les points de l'image.
- (2) La transformée de Hough appliquée sur les points centraux des liaisons.
- (3) La transformée de Hough appliquée sur les points du contour horizontale seuillé.

Afin de valoriser notre travail concernant la TH, nous avons simulé sur 100 images à niveau de gris différents angles d'inclinaison (5, 10, 50, 85, 120), afin de les détecter et de déterminer avec quel degré d'exactitude l'inclinaison est évaluée.

Les différents taux de reconnaissance de l'angle d'inclinaison est présenté dans le tableau suivant :

L'angle L'erreur tolérée	L'angle				
	5°	10°	50°	85°	120°
0°	87 %	97 %	52 %	72 %	78 %
1°	96 %	98 %	61 %	80 %	93 %
2°	96 %	99 %	63 %	83 %	93 %

Tableau 5.3 : Taux de reconnaissance de l'angle d'inclinaison

Nous remarquons que l'intervalle de confiance est entre (50 < angle < 120).

Commentaire :

Si nous utilisons tous les points de l'image, le temps d'exécution sera très important.

D'autres exemples: illustrant l'efficacité de notre approche sur des documents complexes

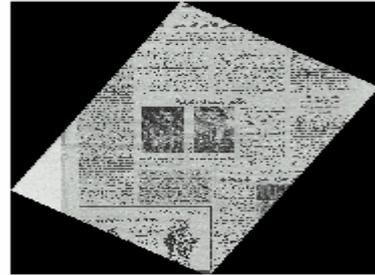
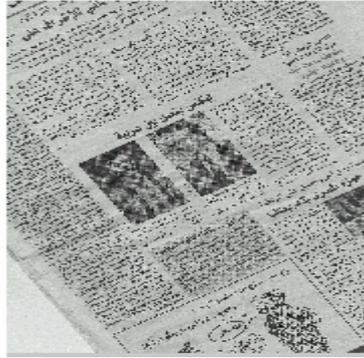


Figure5.5 : Document multicolonne avec des images



Figure 5.6: Documents avec tableau



Figure5.7 : Document avec la partie image dominante

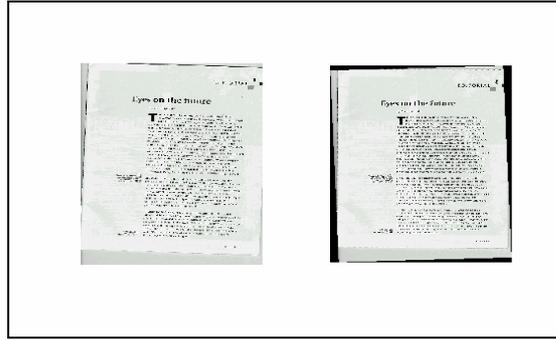


Figure5.8 : Document latin

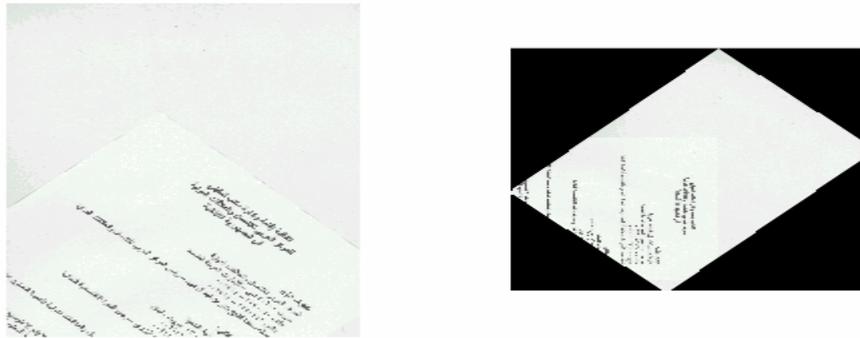


Figure5.9 : Exemple qui présente une fausse détection

Afin de valider notre travail, nous avons fait un comparatif avec les travaux réalisés au sein du laboratoire d'architecture des systèmes.

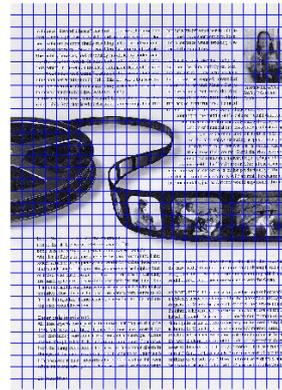
Les deux travaux ont été effectués sur 100 documents, les différents taux obtenus par cette équipe présente une meilleure précision (89% pour détection exacte et 96% pour une erreur de $\pm 1^\circ$ pour différents angles, en enlevant les images et les graphiques au préalable. Toutefois notre travail présente une rapidité de traitement sans enlever du document traité les zones graphiques et les images.

5.3.2. Exemple détaillé de test :

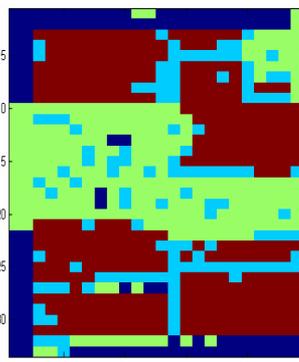
Nous allons présenté dans ce qui suit les différentes étapes de segmentation traitées par notre système sur deux exemples de documents, l'un est pris de la base 1, l'autre est de la base 2. Ces deux documents sont caractérisés par un contenu textuel.



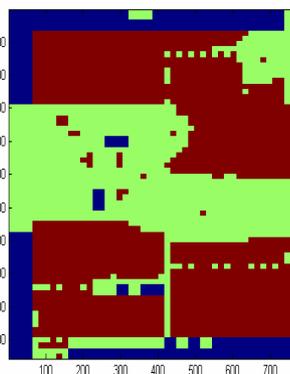
a



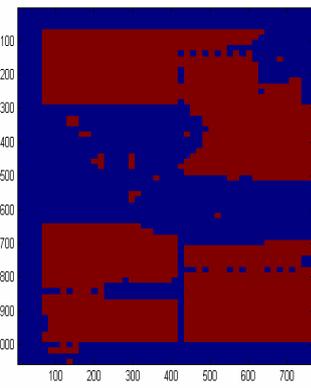
b



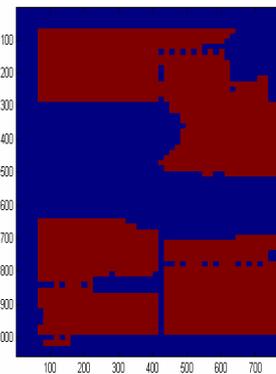
c



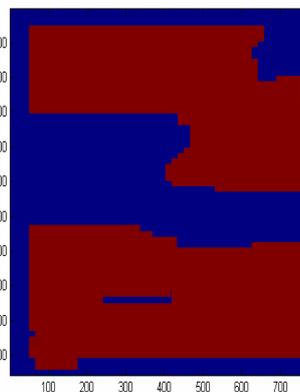
d



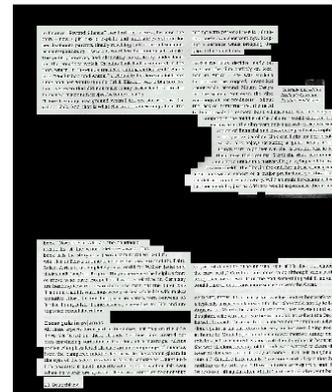
e



f



g



h

a : image originale

b : segmentation par bloc

c : Classification pour la première résolution

d : classification pour la deuxième résolution

e : extraction de la partie texte

f : élimination des régions isolés

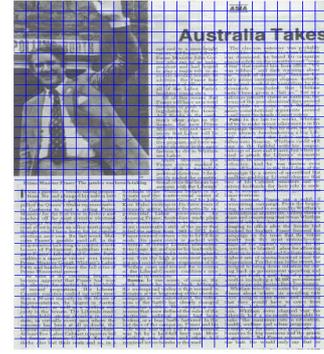
g : dilatation

h : image finale

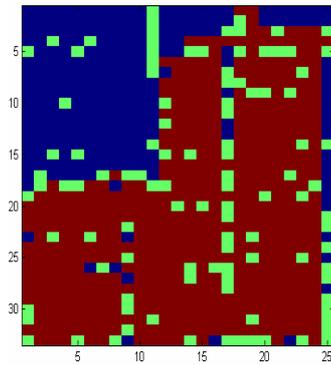
Figure 5.10: Les différentes étapes de segmentation d'un exemple de la première base



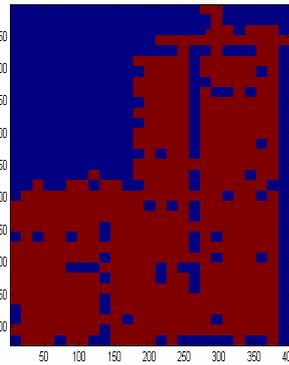
a



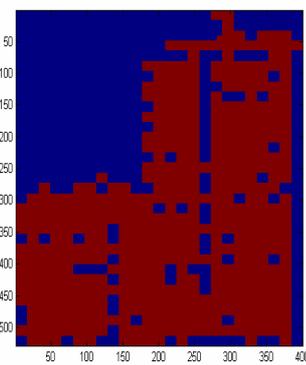
b



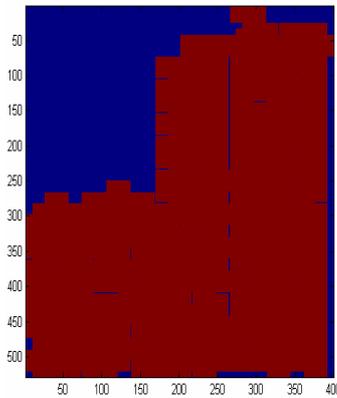
c



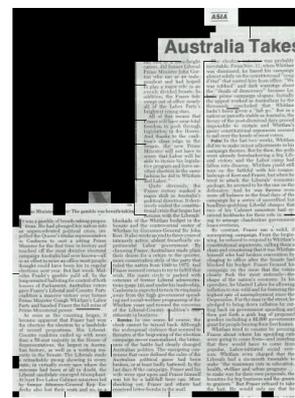
d



e



f



g

a : image originale

b : segmentation par bloc

c : Classification pour la première résolution

d : classification pour la deuxième résolution

e : extraction de la partie texte

f : élimination des régions isolés

g : dilatation

h : image finale

Figure 5.11: Les différentes étapes de segmentation d'un exemple de la deuxième base

5.3.3 Estimation des seuils :

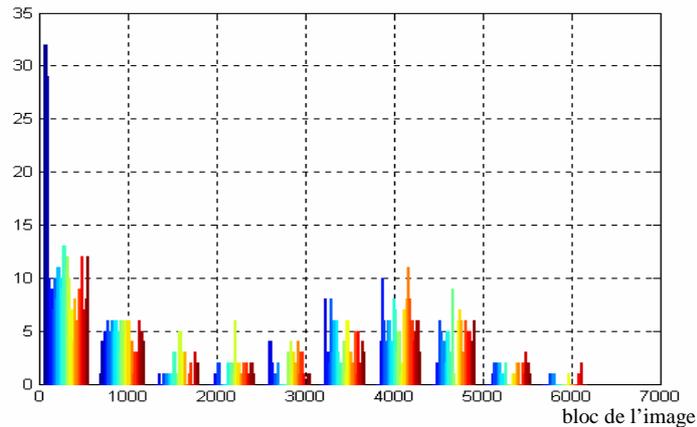


Figure5.12 : Histogramme de la variance des blocs pour un document complet

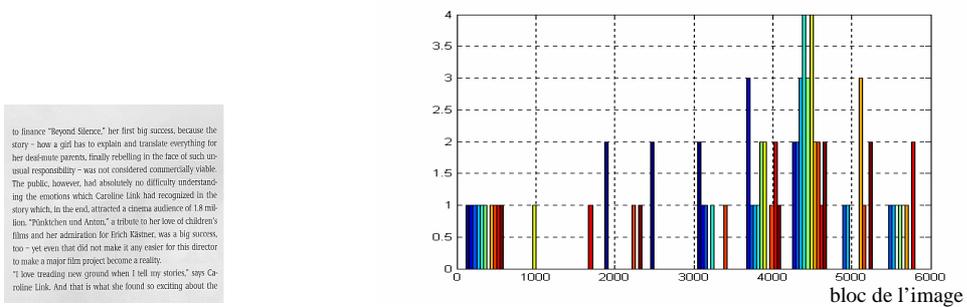


Figure5.13 : Histogramme de la variance des blocs pour une partie texte du document

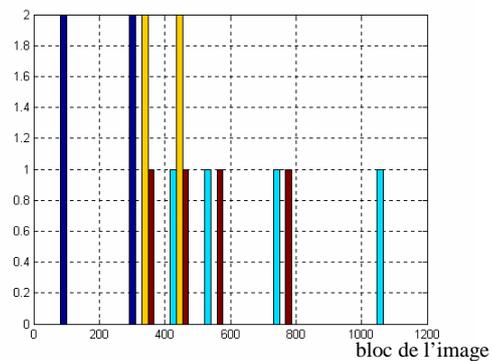


Figure5.14 : Histogramme de la variance des blocs pour une partie photo du document

Commentaire :

Ces graphes ne sont pas toujours les représentants pour les zones textes et les zones images, d'autre cas défavorable sont mentionné dans la figure 5.15.

5.3.4 Des cas défavorables:

Ce sont des cas où le contenu des images a beaucoup de variation qui ressemble à celle du texte qui empêche la discrimination entre les zones.

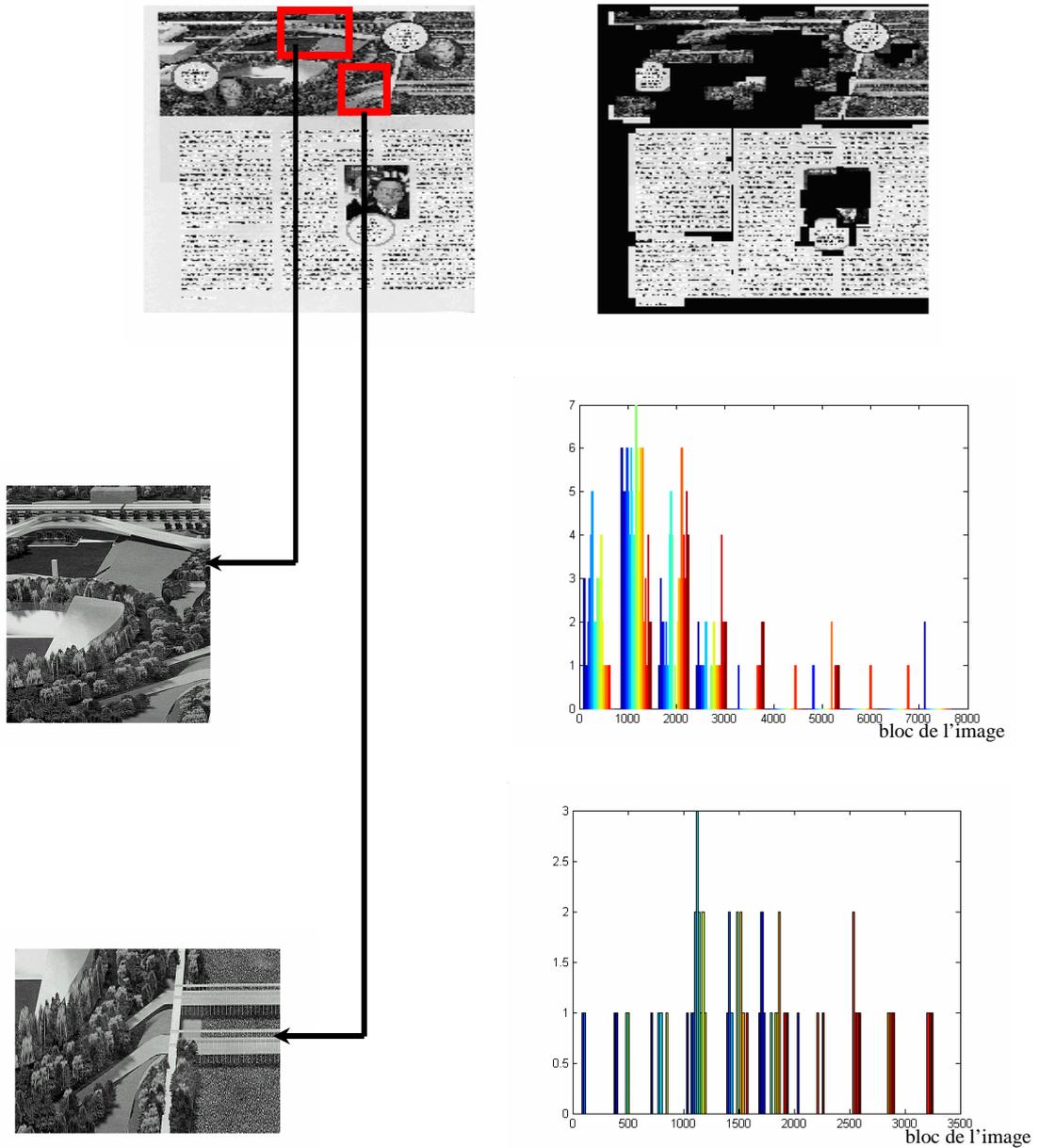


Figure 5.15 : Exemple d'un document présentant un cas défavorable.

5.4 Évaluation :

Afin de mesurer l'efficacité de notre système, nous avons calculé le temps d'exécution de chaque étape de segmentation, ils sont présentés dans le tableau suivant :

Type	Doc	Etape1	Etape2	Etape3	total	Rapport1	Rapport2	Rapport3
Base 1	Doc1.1	6.05	2.09	2.64	10.78	56.12%	19.38%	24.48%
	Doc1.2	8.02	3.62	4.12	15.76	50.88%	22.96%	26.14%
	Doc1.3	5.66	2.91	1.81	10.38	54.52%	28.03%	17.43%
	Doc1.4	8.40	2.36	3.62	14.38	58.41%	16.41%	25.17%
	Doc1.5	5.77	3.35	1.93	11.05	52.21%	30.31%	17.46%
Base 2	Doc2.1	4.23	1.42	0.99	6.64	63.70%	21.38%	14.90%
	Doc2.2	4.17	1.54	0.93	6.64	62.80	23.19	14.00%
	Doc2.3	5	0.99	0.94	6.93	72.15%	14.28%	13.56%
	Doc2.4	4.17	1.70	0.99	6.86	60.78%	24.78%	14.43%
	Doc2.5	4.39	12.75	1.54	18.68	23.50%	68.25%	8.24%

Tableau 5.4 : Temps de calcul (second) et le rapport entre le temps de calcul total et le temps de calcul des différentes étapes de la segmentation.

5.5 Discussion critique

Les systèmes courants pour la transformation de document du papier au format électronique présentent plusieurs limitations. Ils sont conçus

- en fonction des applications spécifiques, (pour cela on a choisi une revue comme base de données pour valider l'application).
- Il est difficile d'ajuster un système à une application différente de l'originale (la base d'OULOU est une base qui contient différent types).

Ce qui apparaît souvent est le fait que les documents contiennent avec l'information appropriée (le texte), des éléments étrangers (bruit ou fond), ces éléments peuvent déranger l'extraction de dispositifs des textes et compromettre le processus de compréhension du document entier.

D'autres faiblesses ont été remarquées :

- les documents dégradés n'ont pas été rapporté, bien qu'ils soient souvent présents dans des applications pratiques ou puissent représenter l'entrée exclusive en traduisant de vieilles archives de papier en format électronique.

- En outre, nous avons observé que quelques types de régions sont négligés dans la phase de décomposition de page : équations mathématiques, formules chimiques, tables.

5.6 Conclusion :

L'emploi des méthodes proposées pour l'analyse de structure (cité dans les premiers chapitres) en beaucoup d'applications est rendu impraticable par leur coût informatique élevé.

Le fond non uniforme et le texte avec différentes directions, deviennent fréquents dans le monde de documents pour une variété de raisons :

- pour obtenir un regard visuel attrayant (journaux, magazine et les couvertures de livre),
- pour donner l'information d'un coup d'oeil (annonces) ou pour éviter la falsification (contrôles).

Notre système d'analyse d'image de document a satisfait :

- l'efficacité informatique ;
- la gestion automatique du document avec :
 - ❖ fond non uniforme (voir figure 5.16) ;



Figure 5.16 : Exemple d'un document avec fond non uniforme

- ❖ présence de niveau gris ou d'information de couleur (voir figure 5.17) ;



Figure 5.17 : Exemple d'un document en niveau de gris

- ❖ page structurée avec les dispositions complexes (non rectangulaires ou même blocs librement formés) (voir figure 5.18) ;

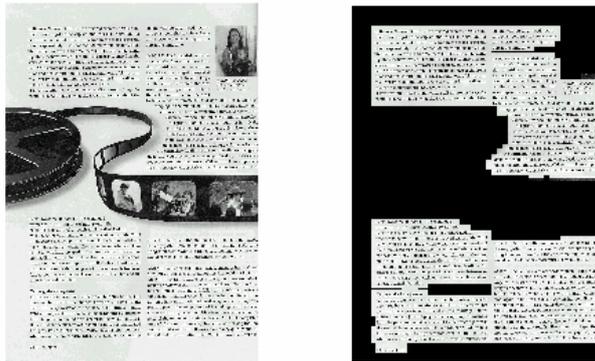


Figure 5.18 : Exemple d'un document structuré.

- ❖ texte avec des tailles, des modèles et des polices très différentes dans la même page (voir figure 5.19) ;



Figure 5.19 : Exemple d'un document avec différents fond et fontes

- ❖ régions des textes avec différentes orientations dans la même page (voir figure 5.20) ;

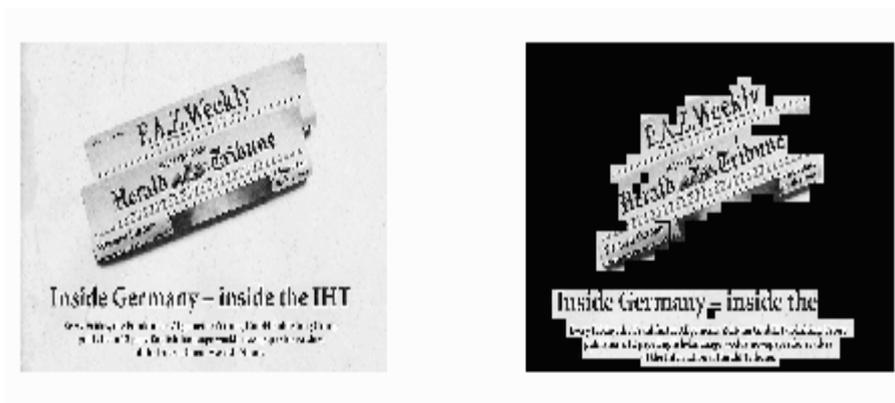


Figure 5.20 : Exemple d'un document avec différentes orientation

- ❖ texte et texte inverse dans la même page ou, plus généralement, texte (voir figure 5.21) ;

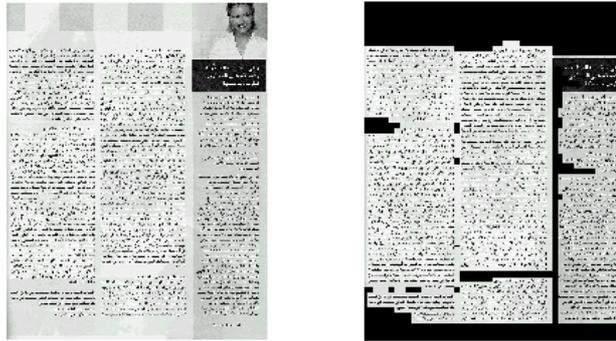


Figure 5.21 : Exemple d'un document avec texte de différentes couleurs

- ❖ fond avec différentes couleurs dans la même page (voir figure 5.22) ;

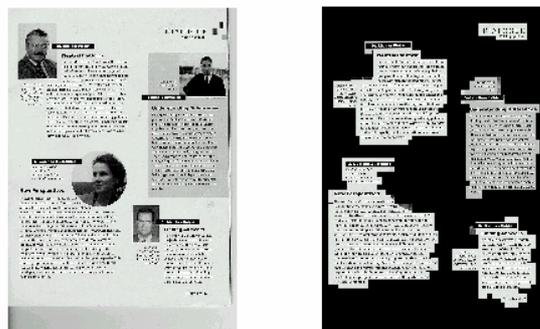


Figure 5.22 : Exemple d'un document avec fond de différentes couleurs

- robustesse en ce qui concerne les documents test dégradés (voir figure 5.23) ;

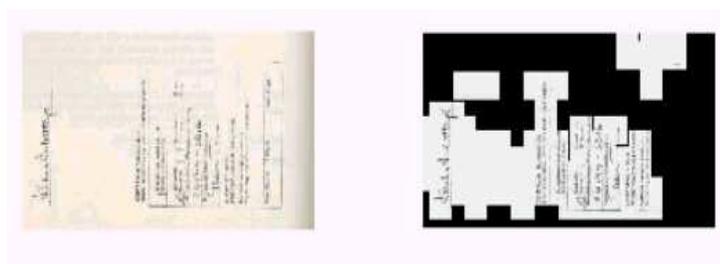


Figure 5.23 : Exemple d'un document dégradée

- ❖ texte inclus dans les graphiques ou les figures (voir figure 5.24) ;

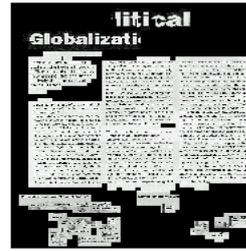


Figure 5.24 : Exemple d'un document avec texte inclut dans les graphiques

CONCLUSION

L'une des majeures problématiques liées à l'analyse des documents concerne la segmentation efficace des documents, essentiellement ceux contenant différentes textures comme le texte, les graphiques, les images, ...

En effet, les différents algorithmes existants tendent à segmenter le document en régions classées selon leur contexte, toutefois, ces algorithmes perdent en précision de segmentation lorsque la structure du document n'est pas connue préalablement ou contient des fonds ou des fontes non uniformes.

La valeur ajoutée par notre mémoire se distingue par l'utilisation des ondelettes afin de solutionner ces différents problèmes, en plus de la prise en charge de l'inclinaison du texte dans le document.

Nous avons adopté la méthodologie suivante qui se résume par une étude comparative, des méthodes de détection de l'angle d'inclinaison, et les critères du choix par l'utilisation de la transformée de Hough qui présente une bonne précision et un large intervalle de détection, ensuite, nous avons présenté les différentes méthodes de segmentation et nous avons montré la prépondérance des ondelettes sur ces méthodes.

Les différents tests que nous avons réalisés, nous ont montré la fiabilité de la méthode adoptée avec des taux de segmentation bien meilleure aux ceux utilisés dans d'autres littératures.

Nous préconisons comme perspective à ce travail :

- Un système permettant de segmenter les mots en caractères isolés afin de faire une reconnaissance optique des caractères et de générer la version électronique des documents en jouant sur l'échelle d'analyse.
- Ainsi qu'un système de classification permettant de trouver la nature de chaque bloc, qui peut être un texte, un tableau, une photo, un graphique, une formule mathématique, ...etc.
- Et enfin arriver à détecter différentes inclinaisons contenues dans une ou différentes zones textuelles.

APPENDICE A LA TRANSFORMEE DE HOUGH

a. Principe de la transformée de Hough

Une droite est décrite dans le plan cartésien (xy) par l'expression suivante :

$$y - ax - b = 0 \tag{1}$$

Sachant que a est la pente et b l'ordonnée à l'origine des abscisses. Etant donné un ensemble de points contours d'un objet représenté par un ensemble de points discrets M qui font partie d'une courbe dont les paramètres a et b restent à définir. **Hough** puis **Rosenfeld** ont proposé une méthode pour détecter les droites à l'aide des points du plan (xy) . Son principe est de calculer pour chaque point M_i de coordonnées (x_i, y_i) de contour d'un objet, l'ensemble des paramètres a qui vérifient l'équation $f(y_i, x_i, a, b) = 0$ avec b fixe [34].

Pour chaque point $M_i(x_i, y_i)$ de l'image, il y a un ensemble de valeurs possibles pour les paramètres a et b . Cet ensemble forme une droite d'équation $b = -ax + y$ dans l'espace des paramètres (ab) appelé espace de Hough. Deux points p_i et p_j de coordonnées (x_i, y_i) et (x_j, y_j) respectivement, appartenant à la même droite, forment des droites dans l'espace des paramètres (a, b) , qui se coupent au point N de coordonnées (a', b') . De cette façon tous les points qui appartiennent à la même droite forment des droites dans le plan des paramètres (ab) qui se coupent au même point. Ce concept est illustré dans les figures (Figure1.a) et (Figure2.b).

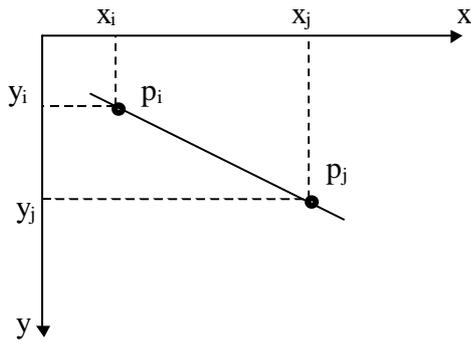


Figure 1.a. Plan cartésien (xy)

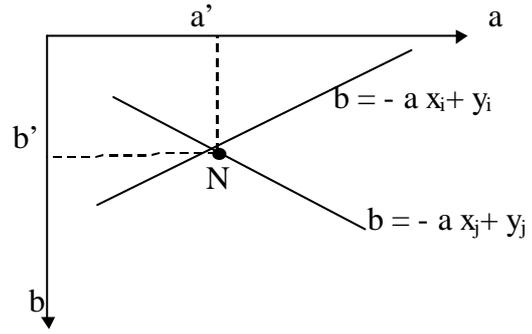


Figure 1.b. Plan des paramètres (ab)

Le traitement de Hough consiste en une quantification du plan des paramètres en cellules accumultrices (Figure 2) où (a_{\min}, a_{\max}) et (b_{\min}, b_{\max}) sont les valeurs limites de l'intervalle de la pente a et de l'ordonnée à l'origine des abscisses b .

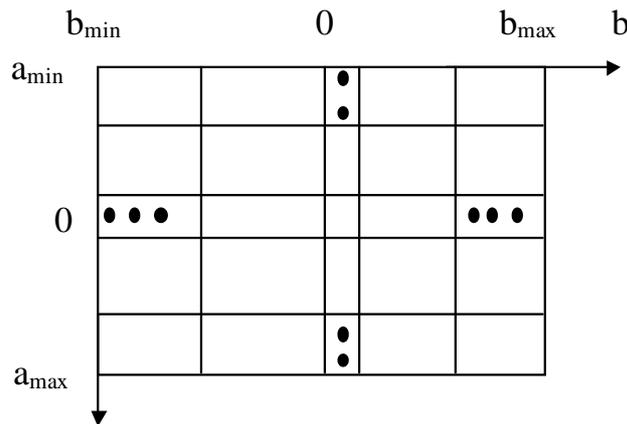


Figure 2. Quantification du plan des paramètres ab

Chaque cellule de coordonnées (i, j) a une valeur accumulée $A(i, j)$ et correspond à la cellule de coordonnées (a_i, b_j) dans le plan des paramètres (ab) . Initialement, ces cellules sont mises à zéro. Pour chaque point de l'image de coordonnées (x_k, y_k) , on calcule pour chaque valeur de a quantifiée à la valeur a_p sur l'axe des a , son correspondant b en utilisant l'équation suivante : $b = -a x_k + y_k$. La valeur résultante résultat b est arrondi à la valeur la plus proche de b quantifiés b_q sur l'axe des b . Si on obtient une valeur b_q suite à a_p choisie, on incrémente la valeur de la cellule correspondante: $A(p, q) = A(p, q) + 1$

A la fin de cette procédure, la valeur n de $A(i, j)$ dans une cellule (i, j) correspond à n points dans le plan (xy) qui vérifient l'équation $y=a_ix+b_j$, donc il existe n points qui appartiennent à la droite de pente a_i et de l'ordonnée à l'origine des abscisses b_j . L'inconvénient majeur de cette procédure réside dans son incapacité à détecter les droites verticales.

Pour remédier à ce problème, un paramétrage polaire (ρ, θ) est plus satisfaisant. Ce paramétrage est illustré dans la figure 3.

Une droite est alors définie par l'équation suivante :

$$f(x, y, \rho, \theta) = \rho - x \cos \theta - y \sin \theta = 0 \quad (2)$$

avec ρ la distance perpendiculaire à la droite de l'origine du plan (xy) et θ l'angle entre cette distance et l'axe des x .

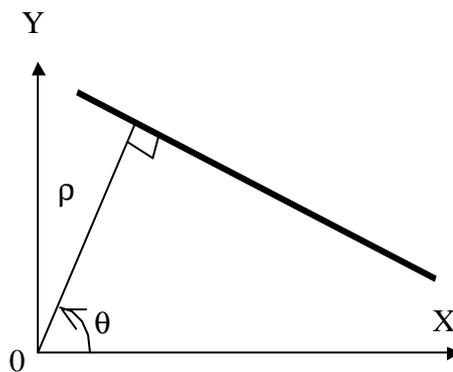


Figure 3 Paramétrage polaire d'une droite

L'utilisation de cette représentation dans la construction du tableau accumulateur est identique à celle développée précédemment (représentation ab) (voir figure 4), on précisera que le choix de quantification de l'espace des paramètres $(\rho\theta)$, doit porter sur les trois objectifs essentiels suivants :

- Garantir une précision de détection aussi bonne que possible ;
- Diminuer la mémoire nécessaire au stockage des accumulateurs ;

- Accélérer les calculs.

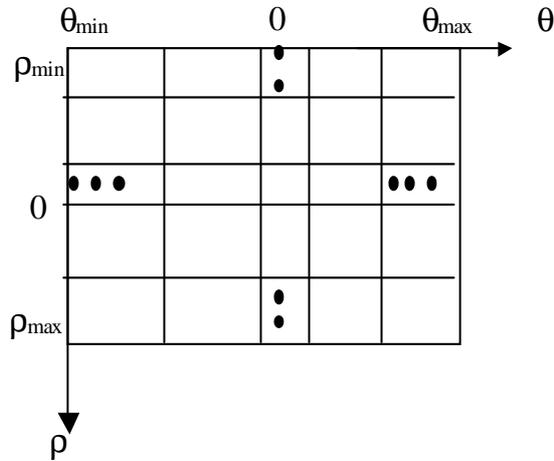


Figure 4. Quantification du plan des paramètres (ab)

Chaque point M_i de coordonnées (x_i, y_i) d'une droite se transforme dans le plan des paramètres (ab) en une sinusoïde d'équation : $\rho = x_i \cos \theta + y_i \sin \theta$. Donc une droite sera représentée par un ensemble de sinusoïdes qui se coupent en un seul point de coordonnées polaires (ρ_0, θ_0) caractéristique de cette droite dans le plan des paramètres voir (figure 5.a) et (figure 5.b).

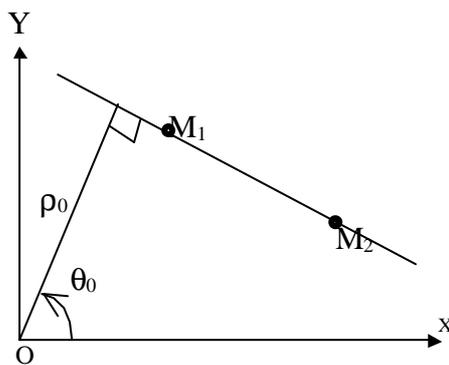


Figure 5.a. Plan cartésien (xy)

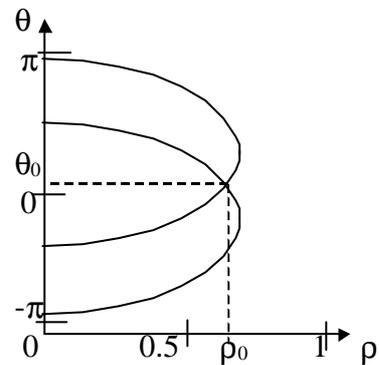


Figure 5.b. Plan des paramètres (ab)

b. Propriétés de la Transformée de Hough

- ❖ Un point du plan cartésien correspond à une sinusoïde dans le plan des paramètres.

- ❖ Un point du plan des paramètres correspond à une droite dans le plan cartésien.
- ❖ Les points appartenant à la même droite dans le plan cartésien correspondent au point d'intersection des courbes du plan des paramètres.
- ❖ Les points appartenant à la même courbe du plan des paramètres correspondent à des droites du même point du plan cartésien.

APPENDICE B

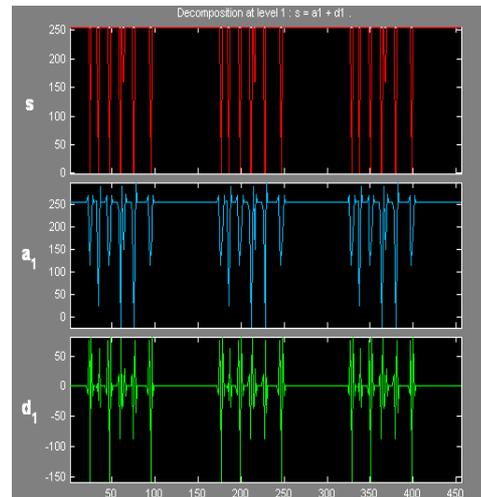
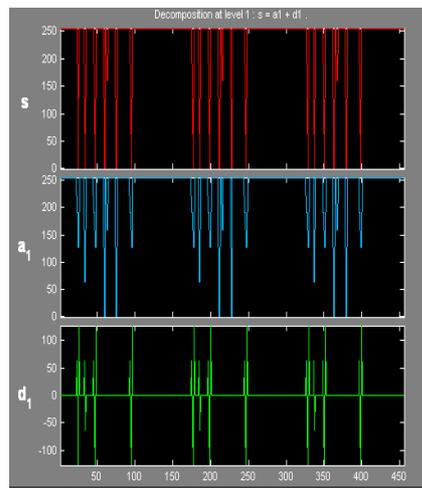
Des exemples d'analyse en ondelette en utilisant :

l'ondelette de haar pour les graphes : (a) ; (c) ; (e) ; (g).

et l'ondelette de db2 pour les graphes : (b) ; (d) ; (f) ; (h).

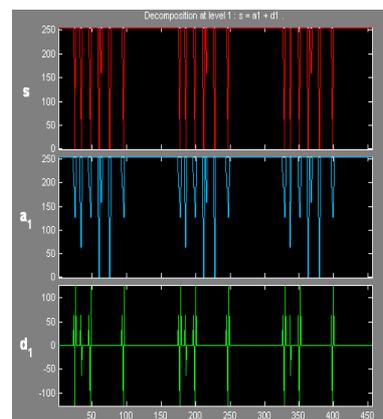
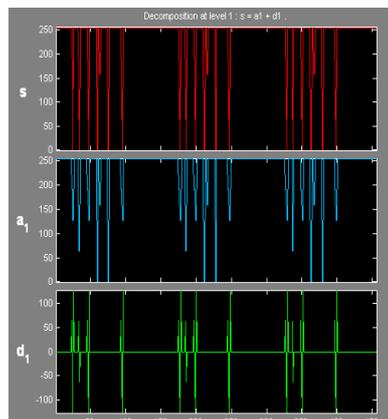
Exemple1 : texte non scannée arabe

مصالحة وطنية ... تنمية مستدامة



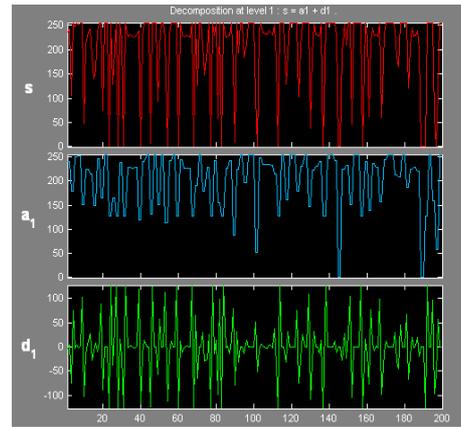
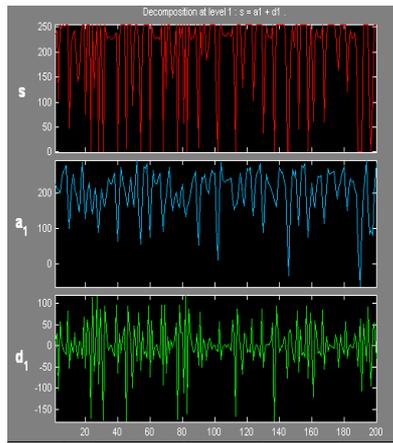
Exemple2 : texte non scannée latin

institut d'electronique



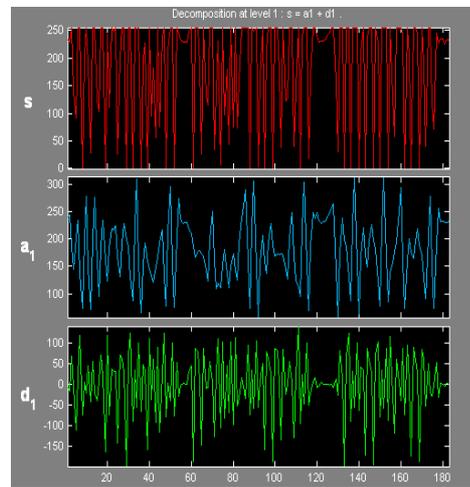
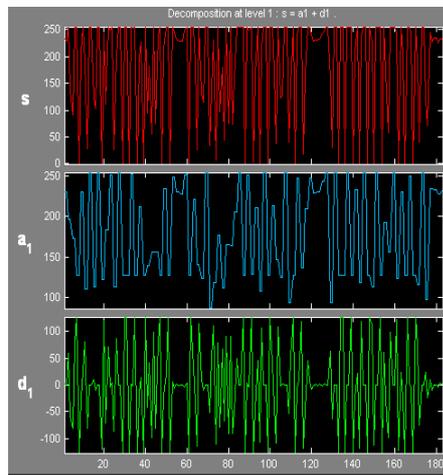
Exemple3 : texte scannée arabe

فئة لها هدف كبير جدا وهو تغيير اسلوب

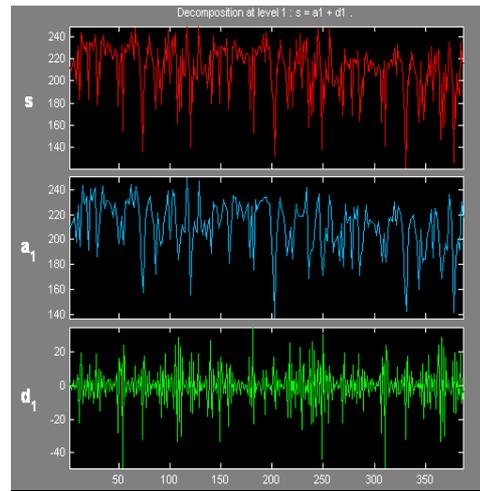
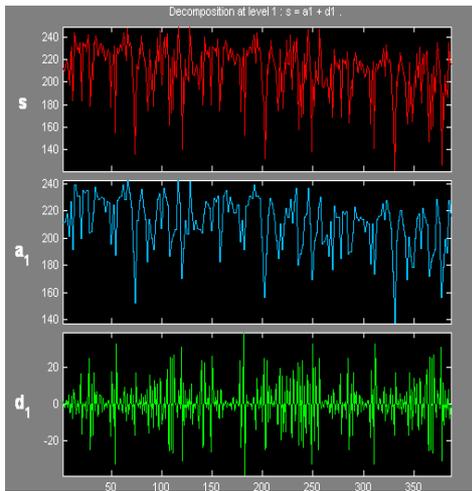


Exemple4 : texte scannée latin

unknown newcomers alongside



Exemple5 : texte scannée latin base de OULOU



APPENDICE C

Des exemples de détection d'inclinaison pour différents angles

Angle=15



Angle=5



Angle=1

a



b



Angle=45

a

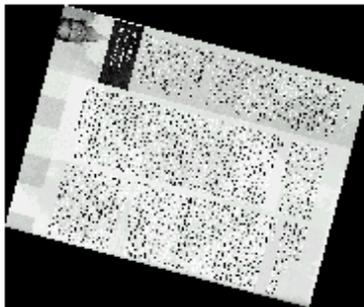


b



Angle=75

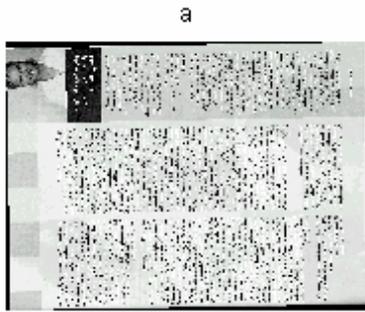
a



b



Angle=91



Angle=170



Angle=10



REFERENCES

1. C. Faure, « Traitement automatique de l'écriture et du document, présentation du domaine », URA CNRS 820 - Télécom Paris, (1993).
2. H.Banke, PS.Wang and HS.Baird, « Document image analysis », Serie in machine Perception Artificiel intelligence, vol 16, ISBN : 981-02-2046-4, USA. (1997).
3. J. Girardot, « Documents Numériques : une introduction », Ecole des Mines de Saint- Etienne, (février 2004).
4. A.Lawrence Spitz, A.Dengel, « Document Analysis Systems », Serie in machine Perception Artificiel intelligence, vol 14, ISBN : 981-02-2122-4, USA.(1995).
5. B. Epron, « Guide pour les revues numériques », France, (janvier 2002)
6. F.Le Bourgeois, H.Empto, E.Trinh, F.Muge, C.Pinto and I.Granado, « Description du matériel et logiciel de traitement d'images pour la numérisation des collections et leur interprétation », DEBORA, Telematics Applications Programme, project no.5608, Institut national des sciences Appliquées de Lyon, France, (2000).
7. A.Belaïd, « Cours 4 », Cours de DEA,PRTAL7, France, (2003).
8. R.Cattoni, T. Coianiz, S.Messelodi, C.M. Modena, "Geometric layout analysis techniques for document image understanding: a review". Rapport technique, ITC-IRST, Via Sommativa, I-38050 Povo, Trento Italy, (Janvier 1998).
9. S.Mao ; A.Rosenfield and T.Kanungo, "Document structure Analysis Algorithms, A literature survey", Center for automation research, University of Maryland, IBM Alamden research center, USA.,(2003)
10. L. Mezai « Analyse et reconnaissance de documents », Mémoire de Magister I.N.I, (juin 2004).
11. H.Bunke, D.Widmer-kljajo, J.Xiaoyi, « Skew Detection of document images by focused a nearest-neighbor clustering », Proceeding of the Fifth International Conference On, ICDAR, 1999, pages: 629-632.
12. D. Bloomberg, « Analysis of Document Skew », revue Leptonica N° 1, (2002).
13. A.Sehad, L.Mezai, M.T.Laskri et M.Chriet, "Estimation de l'inclinaison des documents arabes imprimés », CIFED, France, (2004).

14. O.Okun ; M. P. Ainen and J. Sauvola, “Robust Skew Estimation on Low-Resolution Document Images”, CIMO, (1999).
15. E. Kavallieratou, N. Fakotakis, G. Kokkinakis “Skew angle estimation for printed and handwritten documents using the Wigner–Ville distribution», Wire Communications Laboratory, Dept. of Electrical Computer Engineering Published by Elsevier Science, (2002).
16. D.Xi and S.Lee, “Reference line extraction from form documents with complicated background”, Computer Society, IEEE proceedings , ICDAR, (2003).
17. J. Yu, « Parallel Line Detection Using a Cascaded Hough Transform Method with a Possible Application to Image Alignment», Course at Purdue University, (May 2004).
18. P.K. Das, S.Sural, «A document image analysis system on parallel processors», fourth international conference on high performance computing Bangalore, India (1997).
19. A.S.Azokly « Une approche uniforme pour la reconnaissance de la structure physique de documents composites fondée sur l’analyse des espaces», Thèse de Ddoctorat, Université de Fribourg Suisse, (1995).
20. H.Jaekye, R.Haralick, T.Phillips, “Recursive X-Y cut using bounding boxes of connected components”, Proceeding of the Third International Conference On, ICDAR, 1995,vol 2 pages: 952-955.
21. C.W. Liang and P.Chen « DWT Based Text Localisation», International journal of applied science and Engineering, Chaoyang University of Ttechnology, ISSN 1727-2394,(2004).
22. H. Cheng, C.A. Bouman, and J.P. Allebach “Multiscale document segmentation” IS&T 50th Annual conference, pp 417-425, Cambridge, (1997).
23. H. Choi and R. G.Baraniuk «Multisale image segmentation using wavelet-Domain Hidden Markov Models», IEEE transaction on image processing, (1999).
24. E. B. Bouchereau « Analyse d’image par transformées en ondelettes : Application aux images sismiques» Tthèse de Doctorat, L’université Joseph FOURIER, Grenoble I, France (Mars 1997).
25. B.Barras, « Estimateur de la régularité à l’aide de l’analyse en ondelettes » travail de diplôme, Ecole polytechnique Fédérale de Lausanne, (février 2000).

- 26 Matlab de Mathworks, Wavelet toolbox ; release 13; 2004
27. C.DIOU, « Contribution à l'intégration sur silicium de la transformée en ondelettes
Application au traitement d'images», Université de Montpellier II, (Décembre 2000).
28. Richard barond, «Contribution a l'étude des réseaux d'ondeletes», Thèse de doctorat, Ecole normale supérieure de Lyon, (1997).
29. Kristina Bouchitte, Souad Guemghar, Olivier Magneau, Julie Vandebussche, Marylène Warthmann , « Projet Traitement Numérique des images», Projet n° 19 : Bases d'ondelettes, (03/2000).
- 29 Kush R.Varshney, « Block-segmentation and classification of grayscale postal images» school of electrical and computer engineering, Cornell university, Ithacan, NY 14853, (2004).
- 30 A. Lewis, G. Knowles «Image Compression Using the 2-D Wavelet», *IEEE Transactions on Image Processing*, Vol. 1 No. 2, pp. 244-250, (1992).
- 31 Y.Y.Tang, L.H.Yang, J.Liu, H.Ma, « Wavelet Theory and Its Application to Pattern Recognition», series in machine perception and artificial intelligence, vol 36, (2000).
- 32 Kush R.Varshney, « Block-Segmentation and classification of grayscale postal images» school of electrical and computer engineering, Cornell university, Ithacan, NY 14853, (2004).
- 33 Yoann Juge et Patrick Horain , « La morphologie mathématique», ERNEST, (1998).
- 34 K.Messaoudi , L.Rabea, «L'implémentation de la transformée de Hough sur un circuit FPGA de xilinx», Mémoire d'ingénieur, USTHB, (2003).

