

الجمهورية الجزائرية الديمقراطية الشعبية
REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
وزارة التعليم العالي و البحث العلمي
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE

جامعة سعد دحلب البليدة
Université SAAD DAHLEB de BLIDA



كلية الهندسة – دائرة الإلكترونيك
FACULTE DES SCIENCES DE L'INGENIEUR
DEPARTEMENT D'ELECTRONIQUE

**MEMOIRE DE
PROJET DE FIN D'ETUDES
POUR L'OBTENTION DU DIPLOME DE
MASTER EN ELECTRONIQUE**

Option : Traitement d'information et systèmes électroniques TISE

Redressement et reconnaissance optique des
documents scannés à plusieurs blocs inclinés

Présenté par :

Mr. ZIDANE Hamza

Proposé par :

Mr. KHORISSI.N

Session: Octobre 2011



Dédicaces

Je dédie ce modeste travail

A mes chers parents, mes chers frères et ma chère sœur

A Mes chères familles.

A tous mes amis

A toute la promotion

A tous ceux qui m'ont aidé durant ma vie universitaire



Hamza



Remerciements

Avant tous je tiens à remercier « ALLAH ».

A travers ce modeste travail, je tiens à exprimer mes remerciements vivement ma promoteur Mr: N-KHORISSI pour l'intéressante documentation qu'il a mise à ma disposition, pour ses conseils précieux et pour toutes les commodités et aisances qu'il m'a apportées durant mon étude et réalisation de ce projet.

Mon remerciements les plus vifs s'adressent aussi aux messieurs le président et les membres de jury d'avoir accepté d'examiner et d'évaluer notre travail.

J'exprime également mon gratitude à tous les professeurs et enseignants qui ont collaboré à ma formation depuis mon premier cycle d'étude jusqu'à la fin de mon cycle universitaire. Sans omettre bien sur de remercier profondément tous ceux qui ont contribué de près ou de loin à réalisation du présent travail.

Enfin, que me chers parents et familles, et bien avant tout, trouvent ici l'expression de mon remerciements les plus sincères et les plus profonds en reconnaissance de leurs sacrifices, aides, soutien et encouragement afin d'assurer cette formation de Master dans les meilleures conditions.



Résumé :

Notre travail consiste à redresser les documents imprimés scannés à plusieurs blocs inclinés en utilisant la méthode de profil de projection et la méthode de hough. Le document redressé sera reconnu par un système OCR

mots clés : angle d'inclinaison, redressement, profil de projection, transformé de hough, transformé en ondelette, OCR.

Abstract :

Our work consists in straightening the scanned printed documents with several sloped blocks by using projection profile method and Hough method. The straightened document will be recognized by an OCR system

Key words: skew angle , alignment , projection profile , hough transform, wavelet transform, ocr.

ملخص :

يتمثل عملنا هذا في تعديل الوثائق المطبوعة المتعددة الانحرافات و المصورة بواسطة الماسح الضوئي بطريقة الإسقاط الأفقي الجانبي , و باستعمال تحويل هوق. يتم التعرف على الوثيقة المعدلة باستعمال نظام التعرف الضوئي للحروف.

الكلمات المفتاحية : زاوية الإنحراف , تعديل , الإسقاط الأفقي الجانبي, تحويل هوق , محول الموجات, تعرف ضوئي على الحروف.

chapitre I

Généralités sur l'analyse et le pré -traitement de document

I.1 LE DOCUMENT

I.1.1 Définition d'un document

Le document est un moyen de communication de la pensée, de la connaissance, de l'information et du savoir.

I.1.2 Format d'un document

Il se présente sous la forme de textes, tableaux, dessins, photo, ...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 [1]

I.1.3 Le document électronique

Dans cette section nous parlerons du document électronique, des documents à structures complexes et de la production de documents avec les moyens électroniques.

I.1.3.1 Définition document électronique

Un document électronique est la représentation d'un document, sous la forme d'une structure de données stockées en mémoire ou sur un support informatique, transmissible entre ordinateurs. Dans un système informatique, cette structure de données est représentée dans un fichier sous forme d'une séquence d'octets. Un document électronique peut avoir plusieurs représentations, d'où la notion de format de fichiers.

I.1.3.2 Caractéristique document électronique

La principale caractéristique d'un document électronique est sa facilité de modification. Tout document électronique peut être copié dans un système informatique, d'un support à un autre, sous forme d'un fichier. Tout document électronique est modifiable ; différentes opérations y sont applicables ; parmi celles-ci, nous citons l'impression, et l'édition. En ce qui concerne l'impression, elle consiste en la matérialisation du document électronique sous forme de papier. Par contre, les opérations d'édition (ajout, modification et suppression) modifient le contenu du document électronique. La suppression physique détruit de façon permanente le document électronique.

I.1.4 L'image synthétique

Une image synthétique peut être générée à partir d'un document électronique structuré. L'image numérisée est obtenue en numérisant le document papier. L'image synthétique et l'image numérisée représentent toutes les deux un aperçu du document électronique. Certes, elles sont toutes les deux des images électroniques ; cependant l'image numérisée est déformée et comporte du bruit alors que l'image synthétique en est dépourvue.

I.1.5 Documents à structures complexes

Les documents électroniques diffèrent entre eux du point de vue du contenu et du point de vue organisationnel. En effet, trois structures sont possibles : structures linéaires, structures hiérarchiques simples et structures complexes. Les premières sont représentées par exemple par les œuvres littéraires tels que les romans. Les deuxièmes sont représentées par les articles scientifiques ou les livres. Elles possèdent une organisation en chapitres, sections, articles et paragraphes que l'on peut représenter sous forme d'arbres. Les troisièmes sont représentées par les journaux, les magazines et les dépliants publicitaires. Elles possèdent une typographie riche et ne sont pas composées uniquement

de texte mais d'une combinaison, selon une disposition variable, de textes, de graphiques et d'images.

De ce fait, nous pouvons définir les documents à structures complexes comme étant des documents possédant une structure de pages assez complexe et variable. Ce type de documents n'est pas régi par des règles claires et définies.

Les documents à structures complexes à l'instar des journaux sont construits de la manière suivante : l'éditeur en chef du journal et ses collaborateurs se réunissent ensemble pour décider de l'ensemble d'articles à insérer dans le journal. Une fois arrivé à consensus, le maquettiste et le graphiste interviennent pour le mettre en forme. Cette mise en forme est effectuée en utilisant un logiciel de PAO. Il incombe au maquettiste et au graphique de se soumettre aux exigences aussi bien du logiciel que de la structure de pages du journal. La variabilité de l'information à insérer dans le journal et les exigences du logiciel génèrent une variabilité entre l'édition du jour et celle du lendemain, connue sous la notion de variabilité intra-classes. Chaque éditeur de journal essaie de se distinguer de ses concurrents en donnant une empreinte à son journal par le biais d'une typographie et d'une représentation spécifique qui engendre une variabilité entre plusieurs éditeurs de journaux connue sous la notion de variabilité inter-classes.

La figure (I.1) illustre un exemple de documents à structures complexes et la variabilité de la structure physique intra-classes.



Figure I.1 : Exemple de document à structure complexe d'un journal.

Afin de faciliter l'identification des différents documents à structures complexes, nous utilisons le principe des classes. En effet, nous regroupons dans une classe un ensemble de documents ayant des caractéristiques similaires. Par exemple, la classe "journaux" regroupe tous les journaux de différents éditeurs, la classe "magazines" regroupe tous les magazines, idem pour la classe "dépliants".

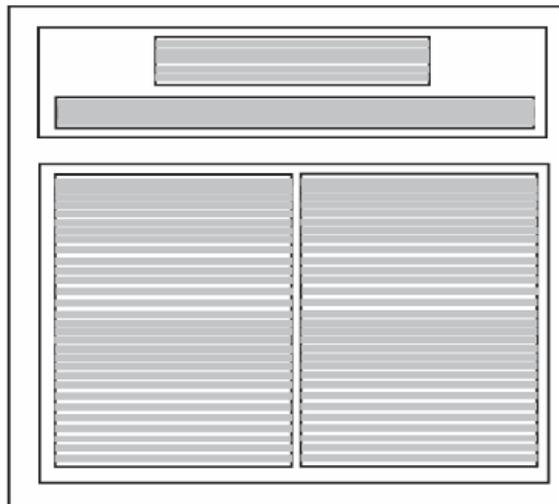
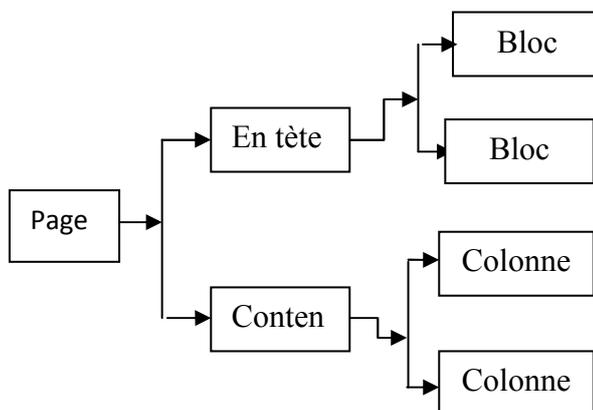
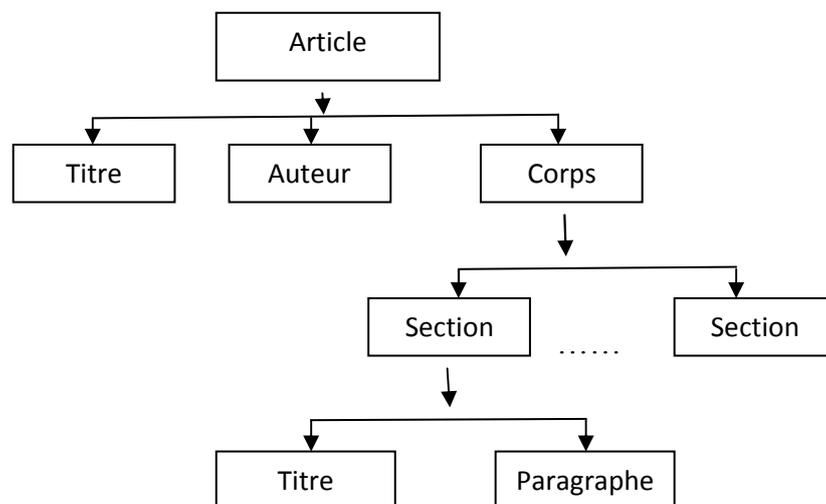
I.1.6 Production de documents avec les moyens électroniques

Dans l'édition professionnelle, le processus d'élaboration d'un document, depuis sa conception jusqu'à son impression, comporte plusieurs étapes faisant intervenir un certain nombre de personnes : l'auteur, le correcteur, le maquettiste et le typographe. Les systèmes de production de documents structurés retiennent essentiellement deux acteurs : l'auteur et le typographe.

Un document électronique commence par être saisi soit avec un logiciel de traitement de texte, soit avec un logiciel de publication assistée par ordinateur (PAO). Une fois cette étape achevée, le document est formaté. L'étape suivante consiste en la restitution du document au moyen d'un format de fichier permettant d'assurer une bonne qualité d'impression. La dernière étape se résume à l'impression du document électronique sous forme papier.

I.1.6.1 Structure logique et structure physique d'un document électronique

Le document électronique passe par différentes formes : structure logique, structure physique, image et papier. La structure logique reflète le point de vue de l'auteur; elle permet de représenter l'organisation du document en entités telles que chapitres, sections, paragraphes, etc. Il est à noter que le niveau de structuration utilisé est fonction de l'application visée. La structure physique permet de représenter la structuration du document en vue de son impression ; de ce fait certains critères, comme par exemple la découpe en page et la répartition des espaces, sont nécessaires. La figure (I.2) illustre un exemple de structure physique et de structure logique.

**Structure physique****Structure logique****Figure I.2 : Exemple de structures.**

L'image est obtenue par le biais d'une conversion de la structure physique. Finalement la représentation sur papier résulte de l'impression de l'image du document.

I.1.6.2 Les étapes de production et les différentes formes intermédiaires d'un document

Les différentes étapes de production sont schématisées sur l'organigramme de la figure (I.3)

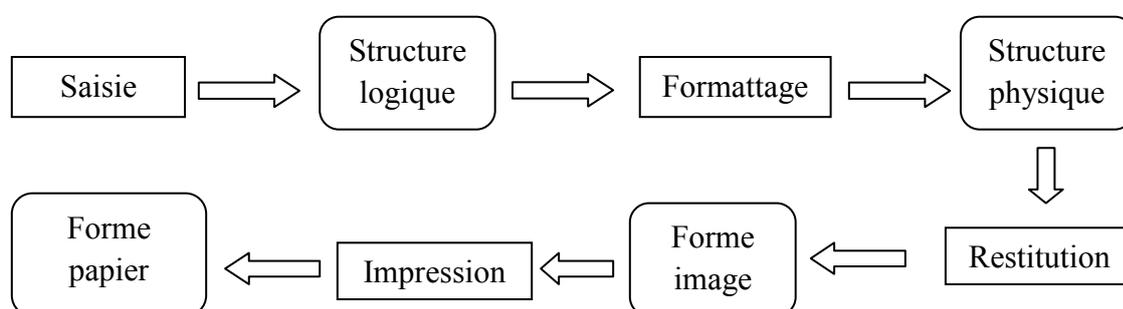


Figure I.3 : *Etapes de production et les différentes formes d'un document.*

I.1.6.3 Types de formats de fichiers

Il existe plusieurs types de formats de fichiers : les formats structurés éditables, les formats d'échange et d'impression, et les images.

I.1.6.3.1 Les formats de fichiers structurés éditables

Les formats de fichiers structurés éditables se subdivisent en deux catégories : ceux avec balises et ceux WYSIWYG (What You See Is What You Get). Parmi les formats de fichiers structurés avec balises nous trouvons LATEX et HTML. Le premier est très répandu dans la communauté scientifique pour la rédaction des articles et des thèses. Le second est le langage de publication de pages web sur Internet. Pour les formats de fichier WYSIWYG, ils sont générés par des logiciels qui offrent une interface

d'édition permettant de visualiser immédiatement le résultat final obtenu. Parmi ces logiciels nous citons Microsoft Word, Quark Xpress, Adobe Indesign, Adobe Framemaker et Adobe Pagemaker.

I.1.6.3.2 Les formats d'échange et d'impression

Les formats d'échange et d'impression quant à eux sont des formats non éditables. Parmi ces formats nous trouvons PS, EPS et PDF [2,3]. Ils ont été créés pour permettre la visualisation et l'impression sur toutes les plateformes d'une manière identique à l'original. Les fichiers en format PS et EPS sont, en réalité, des fichiers texte ASCII et ils représentent un programme de description de page. Cependant avec l'essor d'Internet, PDF est devenu le format de diffusion de documents électroniques sur la toile car il est plus compact ; mais aussi le format le plus prisé pour les imprimeurs. Les fichiers en format PDF sont des fichiers binaires.

I.1.6.3.3 Les images

Les images permettent aussi de représenter un document électronique. Nous distinguons les images non compressées et les images compressées. Les premières sont très gourmandes en espace de stockage. Parmi ces formats, nous citons : BMP, PSD (format propriétaire d'Adobe) et TIFF [4] non compressé. Les deuxièmes ont été introduites dans le but d'accélérer leurs échanges sur la toile. En effet, les images compressées prennent moins de place en espace de stockage que celles non compressées et par conséquent leur transfert se trouve accélérer. Les formats d'images compressées comprennent ceux avec perte d'information et ceux sans perte d'information. Parmi les formats d'images compressés avec perte d'information nous citons le JPEG. Ce dernier est très utilisé dans le web mais aussi comme format de stockage dans les appareils photos numériques. En revanche il existe une multitude de formats d'images compressées sans perte d'information, tels que : GIF, PNG, et TIFF compressé.

I.1.6.3.4 La structure générique

Il arrive que certaines règles de structuration régissent un ensemble de documents. Ces règles sont communément appelées structures génériques. La structure générique définit le mode de construction des structures spécifiques. Chaque élément de la structure spécifique appartient à une classe générique. Les classes génériques sont définies par un ensemble de règles grammaticales.

Par exemple, la structure générique d'un **mail** exprime certaines règles. Parmi celles-ci nous citons celle relative à l'ordre des éléments : un **mail** comprend l'adresse du destinataire, un sujet, un contenu, une signature et des pièces jointes. Le contenu est composé d'un ensemble de paragraphes suivi optionnellement d'une signature. Les pièces jointes sont facultatives.

I.2 TRAITEMENT D'IMAGE

I.2.1 Introduction

Les premiers traitements d'image par moyens technique remontent au 18^{ème} siècle, à l'époque des pionniers de la photographie. La comparaison entre les premiers résultats de noircissement du nitrate d'argent par la lumière et ceux que l'on obtient aujourd'hui avec les films et les appareils disponible sur le marché montre le progrès accompli dans ce domaine. [5]

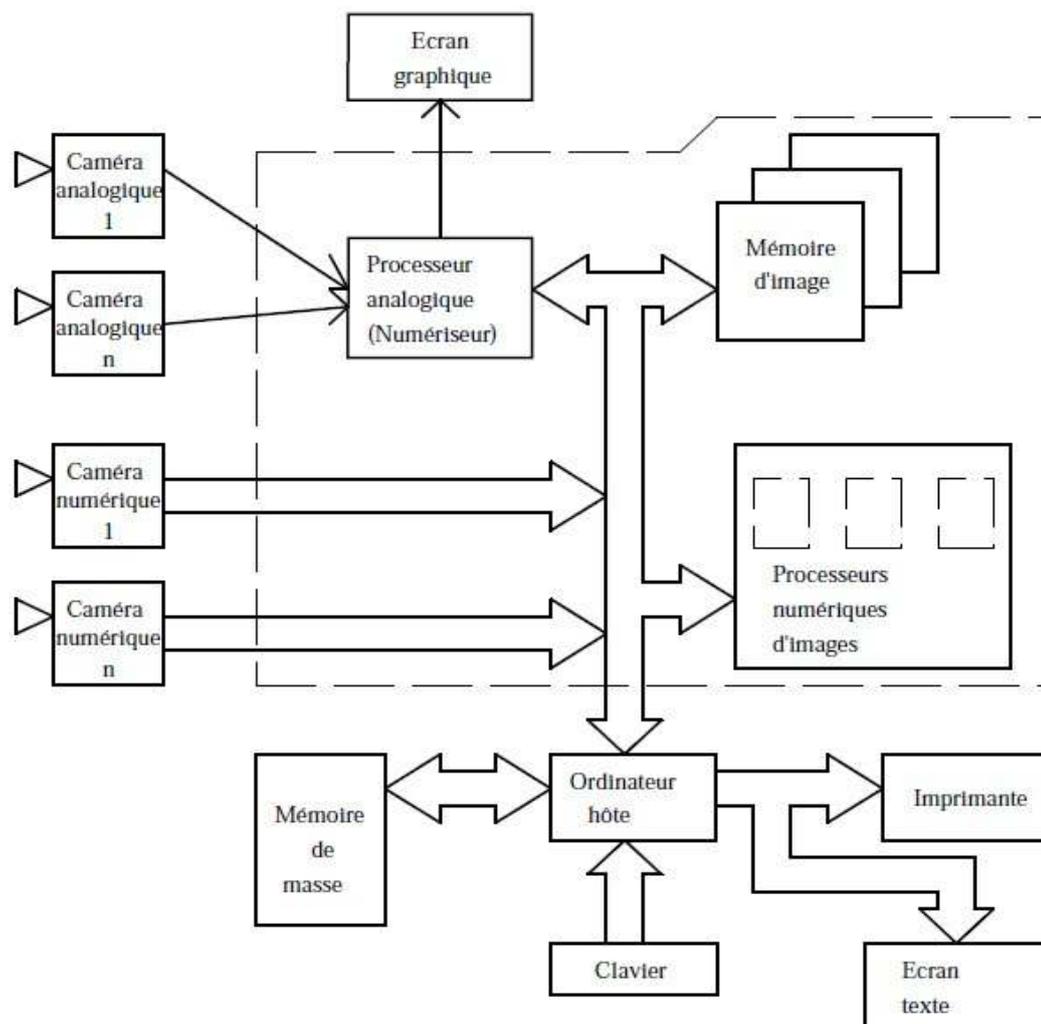


Figure I.4 : Architecture matérielle d'un système complet de traitement d'image.

Le traitement d'images désigne une discipline des mathématiques appliquées qui étudie les images numériques et leurs transformations, dans le but d'améliorer leur qualité ou d'en extraire de l'information. Il s'agit donc d'un sous-ensemble du traitement du signal dédié aux images et aux données dérivées comme la vidéo (par opposition aux parties du traitement du signal consacrées à d'autres types de données : son et autres signaux monodimensionnels notamment), tout en opérant dans le domaine numérique (par opposition aux techniques analogiques de traitement du signal, comme la photographie ou la télévision traditionnelles). [6]

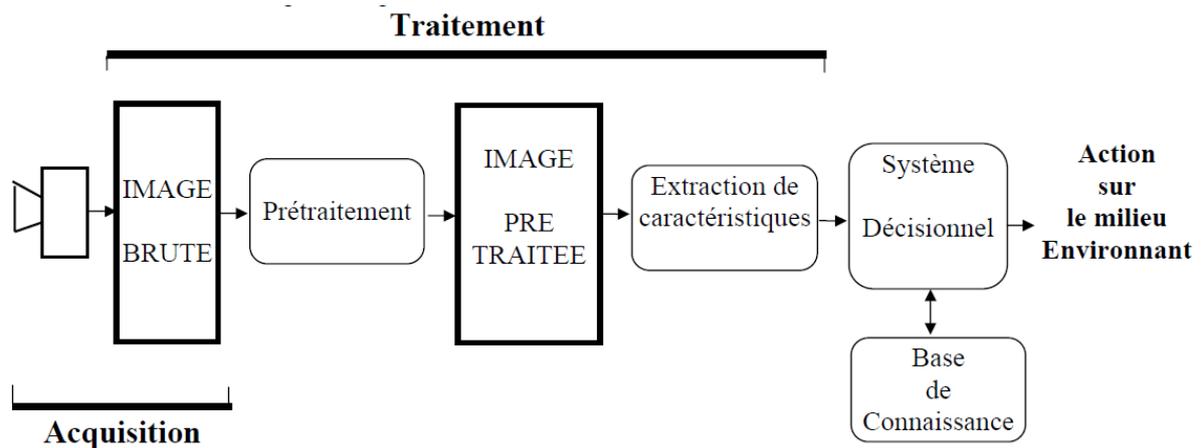


Figure I.5 : La chaîne de vision assistée par ordinateur.

I.2.2 Domaines d'application du traitement d'image

- Imagerie aérienne et spatiale (images satellites, cartographie, météorologie ...)
- techniques biomédicales (scanner, échographie ...)
- robotique (assemblage, contrôle de qualité ...)
- astronomie
- physique nucléaire
- ...

I.2.3 Notions fondamentales en traitement d'images

I.2.3.1 Définition d'une image

L'image est une représentation bidimensionnelle de structure simple dans son organisation mais complexe dans son contenu, et comme étant aussi la projection d'objet tridimensionnelle sur un plan. Sa projection est une matrice de points portant chacun une information différente de l'autre, c'est aussi un ensemble structuré d'informatique qui, après affichage sur écran, ont une signification pour l'œil humain. [7]

I.2.3.2 Acquisition d'une image

L'étude de cette étape passe inmanquablement par le système d'acquisition qui fait référence : l'œil. On peut utiliser des webcams, des appareils photo numériques, des caméras industrielles, des caméras infrarouges... En médecine, on utilise des imageurs IRM, TEP, scanner X, échographie Doppler, échographie, scintigraphie, Tomographie, etc.

Tous ces systèmes peuvent être comparés à des capteurs. Il ne faut pas oublier qu'il y a une étape de conversion analogique/numérique. C'est souvent cette étape qui limite la résolution de l'image. L'une des caractéristiques intéressantes de ces capteurs est la taille du plus petit élément (pixel), mais aussi l'inter-corrélation de deux éléments voisins : plus cette inter-corrélation est faible, meilleure est l'image. [6]

I.2.3.3 Définition d'une image numérique

Une image numérique est un signal numérique composé d'unités élémentaires (appelées pixels) qui représentent chacun une portion de l'image. [8]

I.2.3.4 Définition du PIXEL

(Picture Élément) est le plus petit élément qui constitue l'image.

Un pixel est décrit par [9]:

- Ses coordonnées dans l'image.
- Sa valeur qui représente sa couleur ou son niveau gris.

I.2.3.5 Les images binaires (noir ou blanc)

Un pixel peut prendre uniquement les valeurs noir ou blanc. C'est typiquement le type d'image que l'on utilise pour scanner du texte quand celui ci est composé d'une seule couleur.

I.2.3.6 Les images en niveaux de gris

En général, les images en niveaux de gris renferment 256 teintes de gris. Par convention la valeur zéro représente le noir (intensité lumineuse nulle) et la valeur 255 le blanc (intensité lumineuse maximale). Le nombre 256 est lié à la quantification de l'image. En effet chaque entier représentant un niveau de gris est codé sur 8 bits. Il est donc compris entre 0 et $2^8-1=255$. C'est la quantification la plus courante. On peut coder une image en niveaux de gris sur 16 bits ($0 \leq n \leq 2^{16}-1$) ou sur 2 bits : dans ce dernier cas le « niveau de gris » vaut 0 ou 1 : il s'agit alors d'une image binaire (Noir et Blanc).

I.2.3.7 Les images couleurs

L'espace couleur est basé sur la synthèse additive des couleurs, c'est à dire que le mélange de trois composantes (par exemple (R, V, B)) donne une couleur. On garde l'information couleur, ou intensité lumineuse et chromaticité. Un pixel est codé par trois valeurs numériques. La signification de ces valeurs dépend du type de codage choisi. Le plus utilisé pour le maniement des images numériques est l'espace couleur Rouge, Vert, Bleu (R,V,B) (RGB en anglais). La restitution des couleurs sur écran utilise cette représentation (synthèse additive).

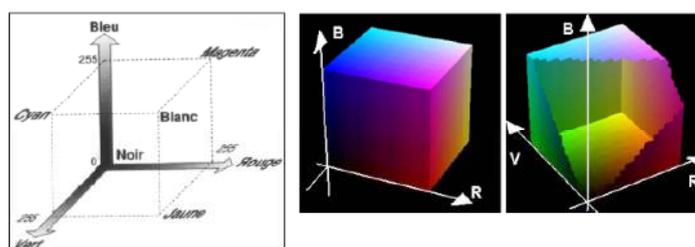


Figure I.6 : L'espace couleur Rouge Vert Bleu (RGB)

Il en existe d'autres : CMJ (ou CMY en anglais), TSL (ou HSL en anglais), YUV, YIQ, Lab, XYZ .

I.2.3.8 Bruit dans une image

Le bruit dans une image est défini comme étant un phénomène de brusque variation d'un pixel isolé par rapport à ses voisins, il affecte la qualité de l'image.

Ce bruit se présente par :

- Sur-éclairage ou sous-éclairage.
- Saturation ou mauvais rapport S/B de la caméra.
- Distorsion optique : aberrations, astigmatisme, problème de prise de vue , mise au point.
- Bruit électronique lié au capteur (bruit thermique des composants).
- Défauts du système de numérisation.

I.2.3.9 L'histogramme

L'histogramme d'une image est l'histogramme de la série de données correspondant aux niveaux de gris des pixels. C'est une fonction définie par :

$\forall P \in \{0, \dots, 255\}$ $h(p)$ = Nombre des pixels ayant p pour niveau de gris.

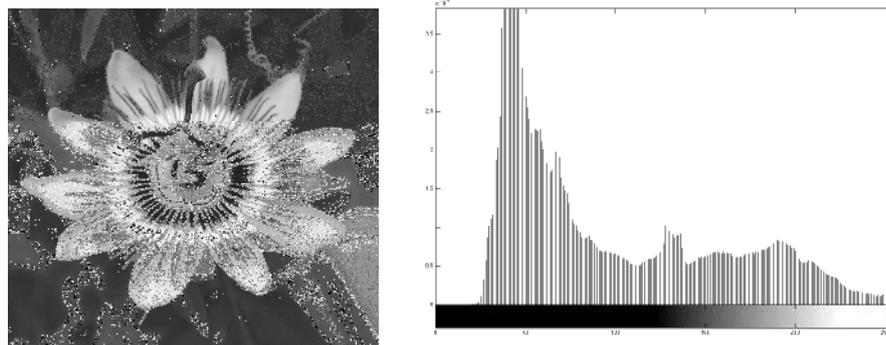


Figure I.7: *Histogramme.*

L'histogramme donne une indication de la dynamique de l'image (répartition des niveaux de gris) mais n'est, en aucun cas, une caractéristique de l'image.

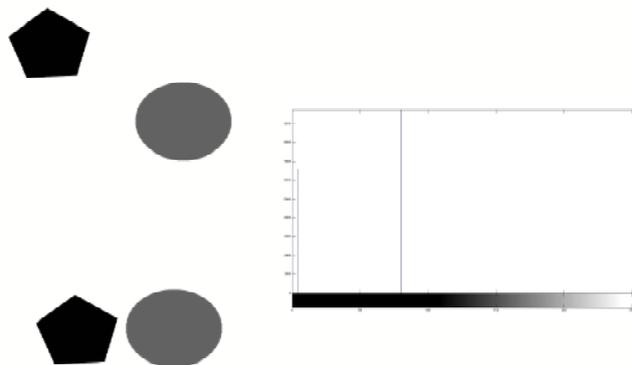


Figure I.8: *Deux images différentes peuvent avoir le même histogramme.*

I.2.3.10 Segmentation d'une image

La segmentation d'image vise à rassembler les pixels de l'image entre eux suivant des critères prédéfinis. [10]

Les méthodes de segmentation des documents peuvent être classées selon deux catégories :

A. Les méthodes ascendantes :

Les méthodes ascendantes considèrent les documents au niveau pixel et affinent les données par des opérations de groupement. Ainsi les composantes connexes sont formées, puis les mots sont alors retrouvés et enfin les lignes et les paragraphes. [11]

B. Les méthodes descendantes :

Les méthodes descendantes prennent en entier, le segment en petits bloc, 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(RunLengthSmoothingAlgorithm) et l'algorithme RXYC (Recursive X-Y Cut).

I.2.3.11 Classification

La classification 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.

I.3 CONCLUSION

Ce chapitre nous a permis de faire un aperçu global sur l'analyse et le prétraitement de document, c'est une étape très importante et même décisive pour une bonne reconnaissance optique de caractères avec un meilleur taux de réussite.

chapitre II

Généralités sur le redressement de document

II.1 INTRODUCTION

Lors de la numérisation d'un document, une inclinaison du texte par rapport à l'axe horizontal de l'image se produit très souvent, en raison d'un mauvais positionnement des pages ou d'une erreur de manipulation, ou un document enclore une sous structure physique inclinée. Une phase d'estimation de l'orientation d'un document est donc indispensable avant de passer aux étapes ultérieures de segmentation et de reconnaissance.

En effet la plupart des algorithmes de reconnaissance ne fonctionne bien que si le document est bien redressé. D'où l'intérêt du redressement [12]

La l'angle β est la l'angle d'inclinaison entre la ligne horizontale du document et l'orientation des lignes du texte (voir figure II.1).

Notre objectif est trouver β et faire une rotation de texte de $-\beta$.



Figure II.1 : un document incliné

Plusieurs méthodes d'évaluation de l'angle d'inclinaison ont été proposées. On peut les classer suivant l'approche adoptée :

- techniques basées sur la projection.
- techniques basées sur la transformée de Hough.
- techniques basées sur les k plus proches voisins.
- techniques basées sur la Transformation de Fourier.
- Cross corrélation

II.2 ALGORITHMES DE DETECTION D'ORIENTATION

II.2.1 Méthodes 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 ou long d'une ligne (voir figure II.2). 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 (voir figure II.4).

Pour le calcul de l'angle d'inclinaison, plusieurs recherches ont été proposées, parmi les quelles, nous citons celle qui maximise la somme des largeurs des zones vides de l'histogramme [1].

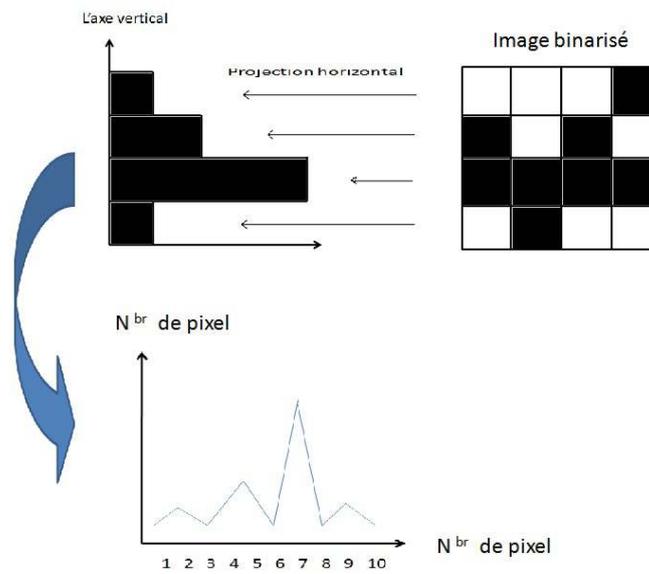
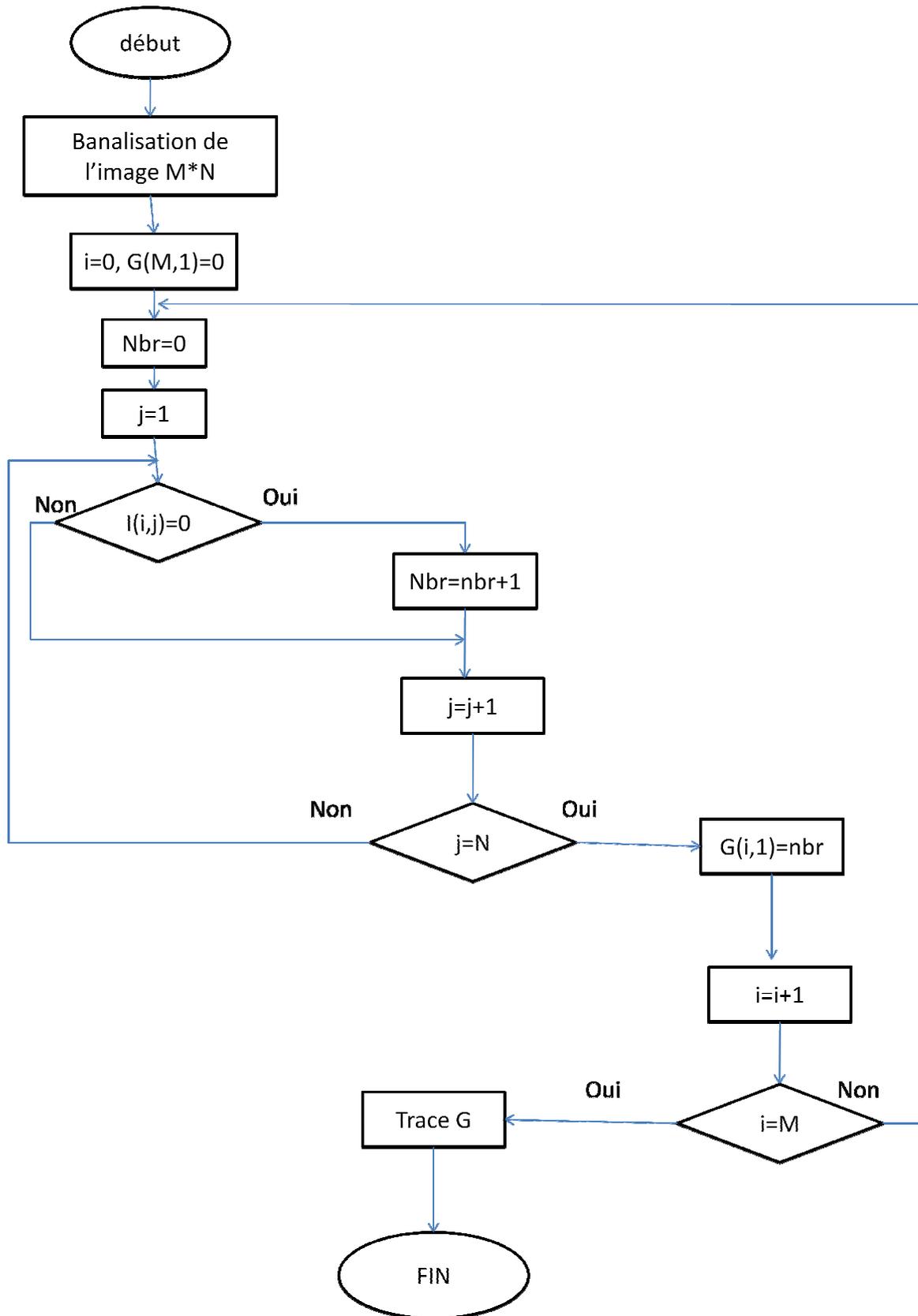


Figure II.2 : Illustration du calcul l'histogramme horizontal.

Pour extraire l'histogramme horizontal à partir de n'importe qu'elle image binaire, on propose l'algorithme suivant [9] :



$I(i,j)$: pixel de l'image binaire.

i : l'ordre de ligne.

j : l'ordre de colonne.

G : vecteur colonne qui compte le nombre de pixel noir dans chaque ligne.

nbr : nombre de pixel noir trouvé dans chaque ligne

II.2.1.1 Analyse du profil de projection

Une méthode pour l'estimation de l'angle d'inclinaison, utilise le profil de la projection horizontale, parce que les textes dans la plupart des images du document sont alignés horizontalement. Quand le profil de la projection horizontale est appliqué sur une image numérique $M \times N$ binarisée, un vecteur colonne de dimension $M \times 1$ est obtenue. Les éléments de ce vecteur sont la somme des valeurs d'un niveau de gris considéré (noir ou blanc) de chaque ligne de la matrice de l'image numérique, un exemple du profil de projection d'une image d'un document incliné est montré dans figure II.3.

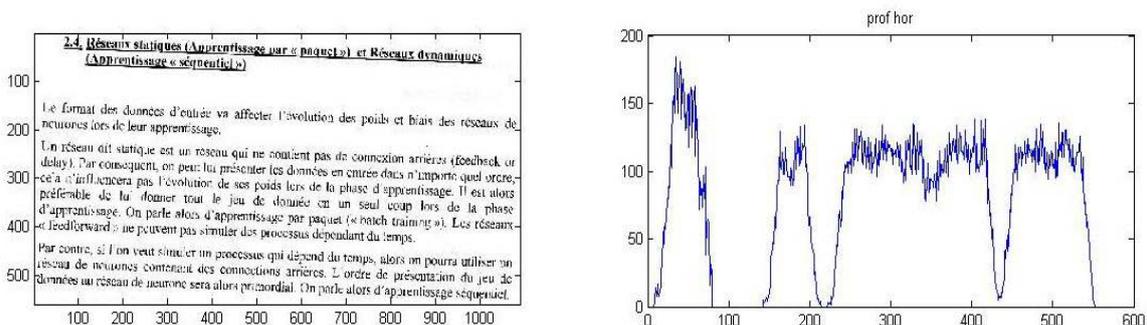


Figure II.3 : Image inclinée et sa profil horizontal.

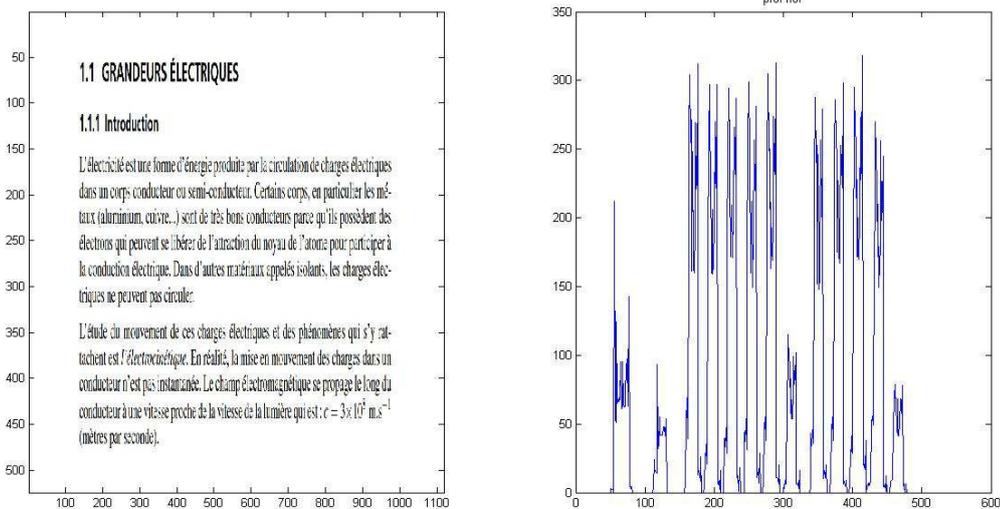


Figure II.4 : le profil de projection après le redressement.

II.2.2 Transformation de Hough

Les transformations de Hough sont très communes en détection d'orientation. Hough proposa une façon très simple de reconnaître des lignes spécifiques à partir d'un ensemble de points caractéristiques. Pour chaque point (x,y) du plan, il existe une infinité de droites auquel ce point appartient. Ces droites sont caractérisées par un coefficient directeur a et une ordonnée à l'origine b , la droite étant définie par l'équation : $Y = a * X + b$.

Ainsi pour chaque point du plan, on peut associer dans l'espace de Hough une droite selon l'équation : $b = -a * X + Y$.

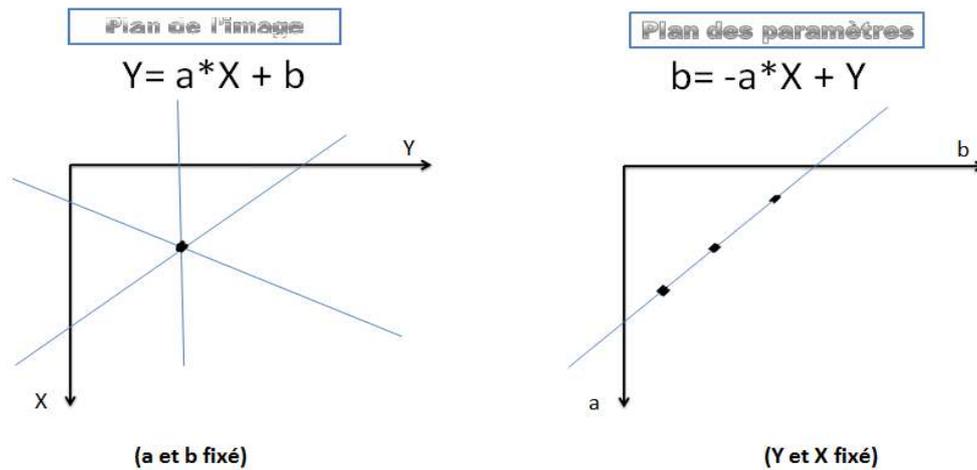


Figure II.5 : Représentation d'une droite dans un plan.

Cependant ce choix de paramétrage est maladroit car les paramètres a et b ne varient pas de façon uniforme. Par exemple, on ne peut pas représenter une droite verticale dans cet espace (car le coefficient directeur est infini).

Donc on va utiliser les coordonnées polaires, l'équation polaire d'une droite passant par le point $(x; y)$ étant :

$$\rho = x.\cos(\theta) + y.\sin(\theta)$$

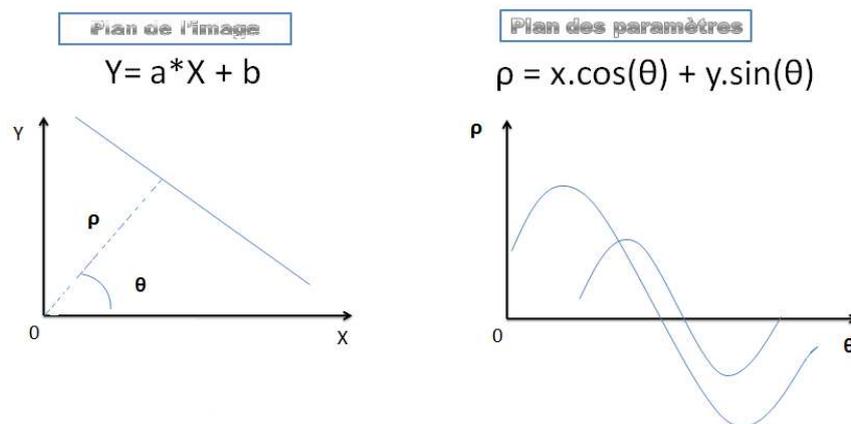


Figure II.6 : Représentation d'une droite dans l'espace de Hough.

II.2.2.1 L'algorithme proposé :

L'algorithme proposé est basé sur la transformé en ondelette (figure II.5) et Seuillage du contour horizontal et algorithme Hough.

La décomposition de l'image par transformée en ondelette (DWT) (voir figure 1 de l'annexe) nous permet d'avoir des approximations successives de l'image par décimation après le passage par un filtre passe bas, et aussi de détecter les contours après le passage par un filtre passe haut. Cette décomposition peut être achevée on appliquant simplement une convolution de l'image avec les masques de HAAR : [15]

Le Seuillage du contour horizontal est une opération de réduction du nombre de points, est une binarisation de l'image indispensable pour le calcul de la TH

L'algorithme de détection de l'angle d'inclinaison se résume comme suit :

Début

Initialiser le tableau $H(\rho, \theta)$ à 0

Calculer la transformée en ondelettes

Seuillage du contour horizontale

Pour chaque point de la matrice du contour seuil faire

{

 Soit (x_c, y_c) point noir

 Pour $0 \leq \theta \leq \pi$

 {

 Calculer $\rho = x_c \cos\theta + y_c \sin\theta$

 Incrémenter $H(\rho, \theta)$

 }

 }

Trouver les coordonnées (ρ_m, θ_m) de la valeur maximale du tableau $H(\rho, \theta)$

θ_m Correspond à l'angle d'inclinaison

Fin

Avantages

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

Inconvénients

- Nécessité d'un espace mémoire important.
- Sensibilité de bruit.
- 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.
- Détection d'un seul angle d'inclinaison.

II.2.3 Regroupement des plus proches voisins

Cette méthode possède la particularité d'être capable de détecter une orientation quelconque (de 0° à 360°).

Pour chaque composante connexe, la direction du segment qui la connecte à son plus proche voisin est calculée par :

$$\theta = \tan^{-1} \left(\frac{(M_{jy} - M_{iy})}{(M_{jx} - M_{ix})} \right)$$

Tel que : (M_{ix}, M_{iy}) sont les coordonnées 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 [1] [11].

Avantage

- 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.

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 figure II.7) [13].



Figure II.7 : 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 du document (voire figure II.8), ensuite seulement les points centraux de ces derniers sont utilisés pour évaluer l'angle d'inclinaison [13].

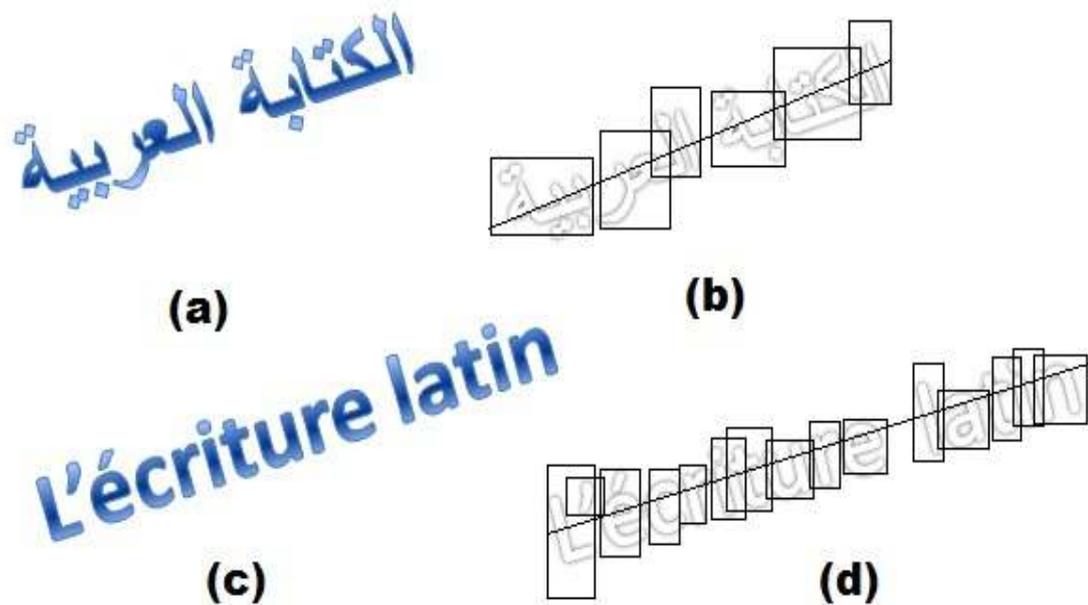


Figure II.8 : La méthode des composantes connexes, (a) : Texte arabe incliné, (c) : Texte latin incliné, (b, d): les points centraux des composantes connexes.

Avantage

- ✓ 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 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

Cette technique basée sur les histogrammes verticaux et horizontaux afin de détecter les liaisons a été proposée [13].

L'étape la plus importante dans cette approche est la détection de liaisons, elle consiste à détecter les liaisons inter caractères (figure II.9). Elle est effectuée en calculant les histogrammes verticaux et horizontaux pour chaque composante qui contient un sous mot.

En général, les liaisons sont caractérisées par une hauteur à la hauteur des caractères et leur fréquence d'apparition 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]

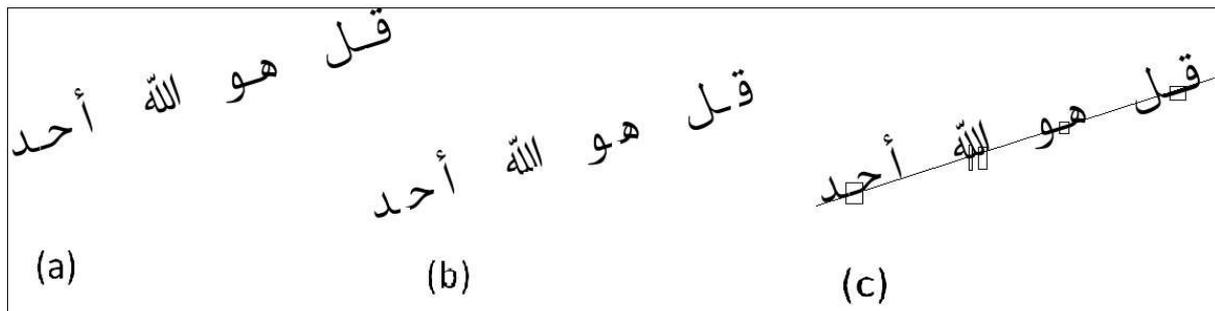


Figure II.9 : La méthode d'extraction des liaisons.

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

Avantages

- ✓ Réduction du nombre de point.
- ✓ La précision.

Inconvénients

- ✓ Influence 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 [1]. (figure II.10)

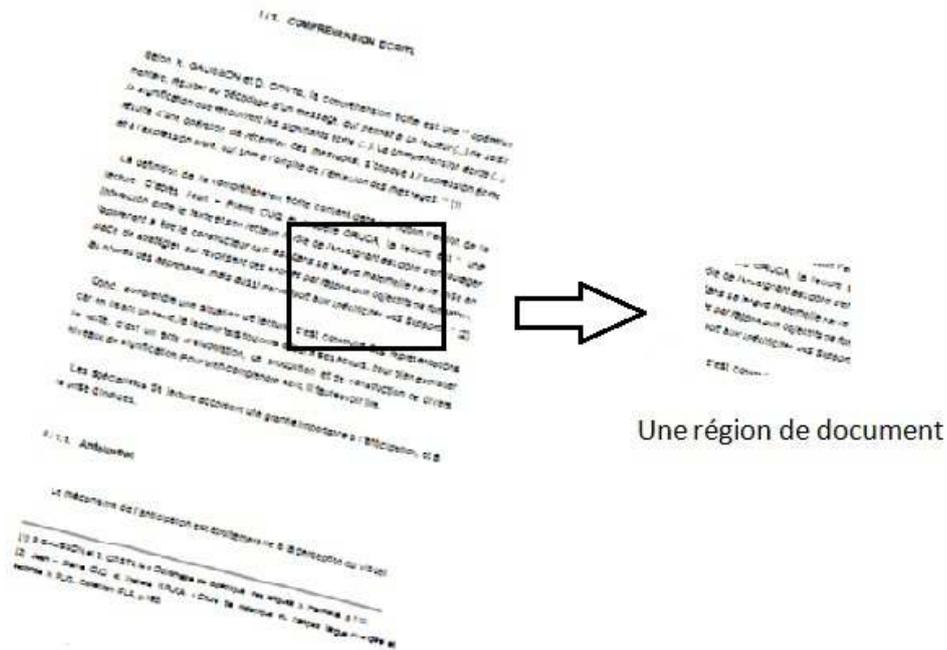


Figure II.10 : La sélection d'une région à partir d'un document.

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.

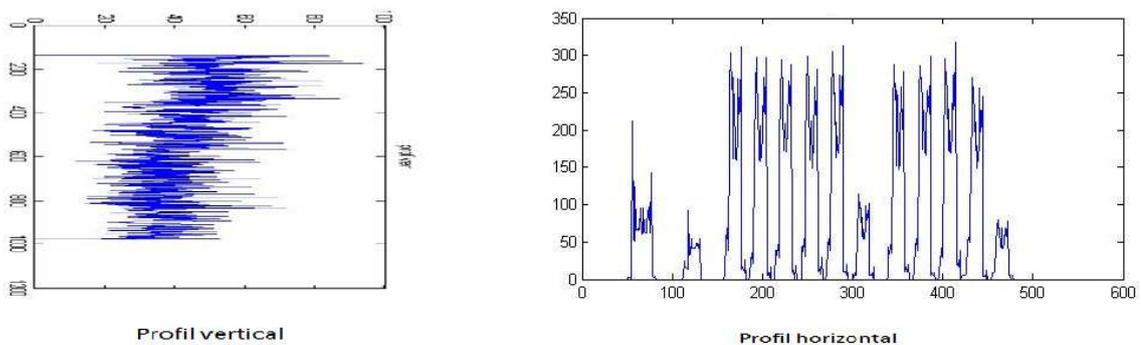


Figure II.11 : Histogramme du document.

Cette technique garantit la section 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.

Avantages

- ✓ Cette méthode est imposée en cas de documents multi colonne.
- ✓ Rapide et facile.
- ✓ L'utilisation d'une partie de la page est plus sûre.

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 de réduction de la résolution du document [14]

Avantage

- ✓ Rapidité.
- ✓ Pas de limite pour l'angle de détection.
- ✓ Une basse résolution est suffisante pour le traitement.

Inconvénients

- Il y'a seulement quelque technique de détection de l'inclinaison fonctionnant avec les images comprimées.
- Balayage de l'image deux fois (le premier pour la détection de l'inclinaison et l'autre pour l'analyse).

II.3 CORRECTION DE L'INCLINAISON

Une fois que l'angle d'inclinaison à été déterminé, l'image sera corrigé, on applique pour chaque pixel (x, y) de l'image incliné, les nouvelle coordonné (x', y') dans l'image redressée qui sont calculées de façon suivant :

$$\begin{cases} x' = x \cos(-\theta) + y \sin(-\theta) \\ y' = y \cos(-\theta) - x \sin(-\theta) \end{cases} \quad / \text{ Avec } \theta \text{ est l'angle d'inclinaison de du document.}$$

II.4 CONCLUSION

Le redressement est une opération fréquente en analyse de documents, plusieurs algorithmes ont été proposés dans la littérature. Deux paramètres très importants sont retenus pour déterminer leurs performances : le temps de détection et de correction de l'inclinaison ainsi que l'exactitude de l'angle estimé.

chapitre III

RECONNAISSANCE DE DOCUMENT

III RECONNAISSANCE DE DOCUMENT

III.1. INTRODUCTION

La reconnaissance de document consiste à convertir un document papier en document électronique en se basant sur l'analyse et l'interprétation du document [16].

La technologie OCR permet de transformer des images de caractères d'impression machine en caractères susceptibles d'être interprétés ou reconnus par un ordinateur. [17]

On appelle reconnaissance de documents le passage de la forme papier ou d'une forme image d'un document, en une forme électronique codée ou structurée (figure III.1). [18]

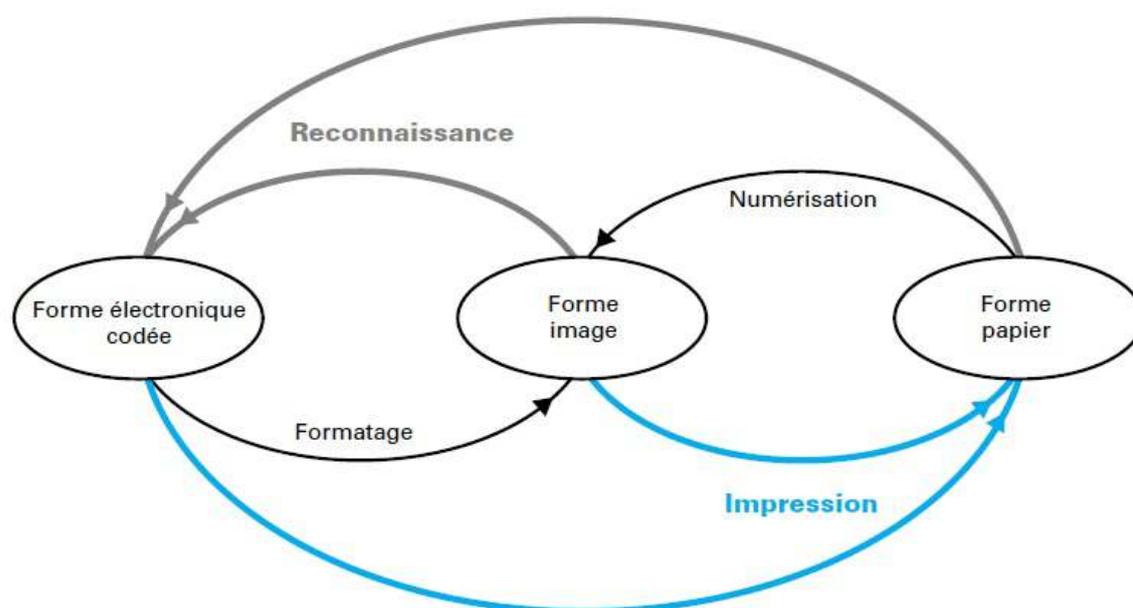


Figure III.1 : Les formes des documents et la reconnaissance

III.2. AVANTAGES DES FORMES ELECTRONIQUES CODEES PAR RAPPORT AUX FORMES IMAGE

Les avantages des formes électroniques codées sont multiples [18] :

➤ **Compacité**

Une page de 3 000 caractères occupe sous forme codée un peu plus de 3 000 octets, alors que son image numérique, selon sa nature et le type de compression, a une taille de 60 kilooctets à 2 mégaoctets, voire plus pour des images à niveaux de gris ou en couleurs. Ce gain, d'un facteur 20 à 1 000, prend toute son importance lorsqu'il s'agit de stocker et surtout de transmettre à distance les documents (coûts et délais de transmission).

➤ **Modification facile du contenu**

Cet avantage est encore plus important que le précédent. Le fichier texte d'une forme électronique codée peut être corrigé : tous les traitements de texte génèrent des documents dont le format a été conçu pour permettre leurs modifications. Par contre, un texte sous forme image est une entité figée : on ne peut que le consulter de façon passive, et les possibilités de modifications sont minimales.

➤ **Accès potentiel à la sémantique du texte**

Seule la forme électronique codée permet l'accès au contenu (potentiellement au sens) du texte, à l'aide des techniques de traitement automatique du langage naturel (TALN) : recherche d'informations, analyse syntaxique des phrases, mise en correspondance d'informations, text mining, extraction de terminologies, classement et classification, diffusion ciblée, représentation de la signification du texte, résumé automatique et traduction assistée...

➤ **Structuration de l'information et hypertextes**

Bien plus, les documents électroniques codés peuvent être structurés, c'est-à-dire organisés selon un plan précis, avec des renvois. Il devient ainsi possible de les consulter non plus linéairement, mais selon leur logique interne, et d'y « naviguer » lorsqu'ils sont munis de liens hypertextuels, comme les documents HTML.

III.3. COMPOSANTES D'UN SYSTEME DE RECONNAISSANCE

Un système de reconnaissance se décompose généralement en quatre étapes ayant une importance variable suivant la catégorie du document : la numérisation, le prétraitement, l'analyse et la reconnaissance [19].

La figure III.2 indique la suite des traitements qui doivent être effectués pour aboutir à la reconnaissance complète d'une page.

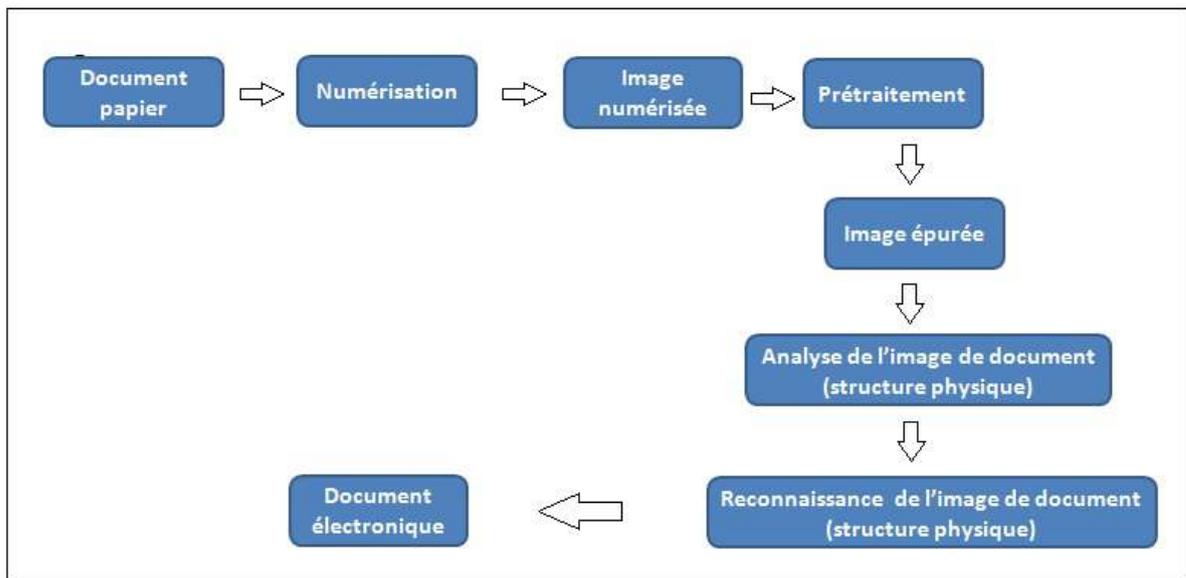


Figure III.2 : Les étapes du processus de reconnaissance de document.

III.3.1 Acquisition ou numérisation

La première opération consiste à acquérir l'image numérique de la page avec un scanner. Cependant, les documents à traiter sont de contenu et de qualité divers ; il est donc quelquefois nécessaire d'acquérir les images en niveaux de gris. Pour éviter l'encombrement de la mémoire, des images bi-niveaux sont préférables. Les résolutions utilisées pour l'OCR sont au minimum de 300 dpi (dots per inch), voire 400 dpi pour les petits caractères.

III.3.2 Redressement

Les scanners d'exploitation ont un débit de 60 à 120 pages à la minute. Les feuilles, lors de leur défilement rapide sur la glace d'exposition, sont susceptibles de s'incliner de plusieurs degrés par rapport à la direction de balayage des images. Pour une bonne exploitation des documents, il faut détecter cet angle et redresser les images, de telle sorte que les lignes de texte soient parallèles aux bords.

III.3.3 Binarisation

La binarisation permet de passer d'une image de niveaux de gris à une image binaire composée de 2 valeurs 0 et 1 [20], plus simple à traiter. Et est la plupart du temps réalisée électroniquement dans les scanners. Elle entraîne dans certains cas une perte importante d'informations, et selon la manière dont elle est effectuée, l'image bi-niveau résultante est exploitable ou non pour des traitements ultérieurs.

III.3.4 Segmentation des pages

On dispose maintenant en mémoire d'une image bi-niveau. Avant de passer à une analyse plus fine, il est nécessaire de donner au système l'équivalent de la vision globale d'un lecteur humain. Au niveau des pages, on appelle segmentation l'opération d'analyse globale qui consiste à détecter et à localiser les zones distinctes : textes, graphismes par traits, tableaux, photos... etc. A l'issue de cette opération, on soumet les différentes zones à des traitements spécifiques. En particulier, les blocs contenant du texte sont envoyés à un module de reconnaissance de caractères (OCR).

III.3.5 Reconnaissance des caractères

Après la phase de repérage des zones de texte, il faut identifier les caractères, les coder et les regrouper en mots : c'est en cela que consiste la reconnaissance de caractères. Cette opération fait appel à des techniques diverses ; la complexité du problème provient

de la grande diversité des formes et des tailles des caractères imprimés, ainsi que des défauts d'impression et de numérisation.

III.3.6 Reconnaissance de la structure Logique

La plupart du temps, on se contente de récupérer le texte. Il existe cependant des applications (exemple des revues de presse, des notices bibliographiques...) dans lesquelles on alimente une base de données à partir du texte reconnu. Dans ce cas, certaines parties du document d'origine vont jouer un rôle particulier dans l'alimentation de la base : titre de l'article, chapeau, nom de l'auteur, date de parution, etc. Il est alors important de connaître la nature ou le rôle de l'information, en plus de son contenu. Cela nécessite la reconnaissance de la structure logique du document. Elle est fondée sur la reconnaissance d'éléments de mise en page, eux-mêmes en relation avec cette structure logique.

III.3.7 Transcodage

Il ne suffit pas de reconnaître un document : encore faut-il en donner une traduction dans un format adéquat, exploitable par la ou les applications qui vont utiliser ce document. Par exemple, si on reconnaît un texte pour l'insérer dans un document en cours de création avec un traitement de texte d'usage courant, on va souhaiter le récupérer dans un format RTF, qui s'importe facilement dans MS-Word, Wordperfect ou tout autre traitement de texte ou logiciel de PAO. Dans ce cas précis, on préfère généralement récupérer le texte, car la mise en page va dépendre plutôt du nouveau document que de l'ancien.

III.4. RECONNAISSANCE DES CARACTERES (OCR)

On se limite à la reconnaissance des caractères dactylographiés ou imprimés de l'alphabet latin.

III.4.1 Structure des caractères et variété

Plusieurs éléments de style peuvent différer d'une fonte à l'autre, ce qui explique les multiples variétés existantes. Citons le gras, la stature (rapport hauteur-largeur), l'italique (inclinaison), le contraste entre les pleins et les déliés, l'aplatissement des rondeurs, la présence ou l'absence d'empattements, la nuance anglaise (arrondissement des empattements). La classification Vox, retenue par l'Association typographique internationale, regroupe les fontes en une dizaine de classes (figure III.3).

Quelques valeurs numériques sont nécessaires pour situer l'ampleur du problème : en dactylographie, on recense facilement entre les divers constructeurs plus de 200 polices différentes, chacune dans un seul corps ; en imprimerie, le catalogue d'un seul fournisseur, comme la Monotype, comporte plus de 8 000 fontes, avec tous les corps possibles pour chacune (échelle de 6 à 100). [18]

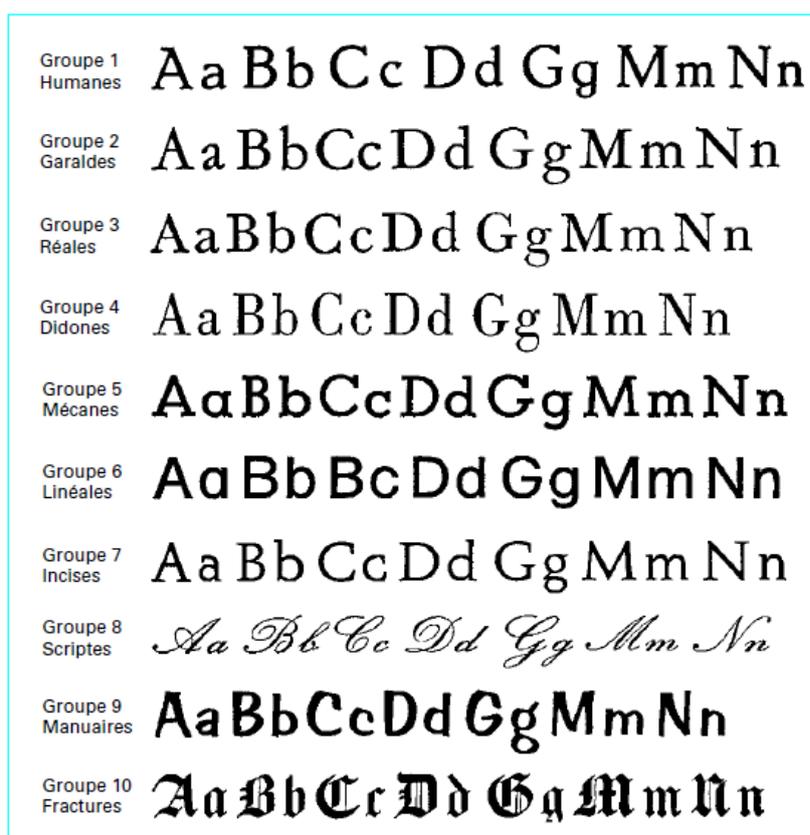


Figure III.3 : Exemples de la classification Vox.

III.4.2 Déformations dues à l'impression, la reproduction, la numérisation

En fait, les logiciels de reconnaissance sont confrontés à des millions de formes possibles, car aux variantes de fontes viennent s'ajouter les différences dues aux qualités d'impression et de reproduction des documents. Une impression ou une photocopie de mauvaise qualité engendrent des caractères collés, coupés, ou du bruit de fond. Il faut tenir compte en plus des dégradations dues à l'acquisition : problèmes d'échantillonnage (résolution), de numérisation et de seuillage déjà évoqués. Cela accentue les phénomènes de déformation, collage ou cassure des caractères (figure III.4).

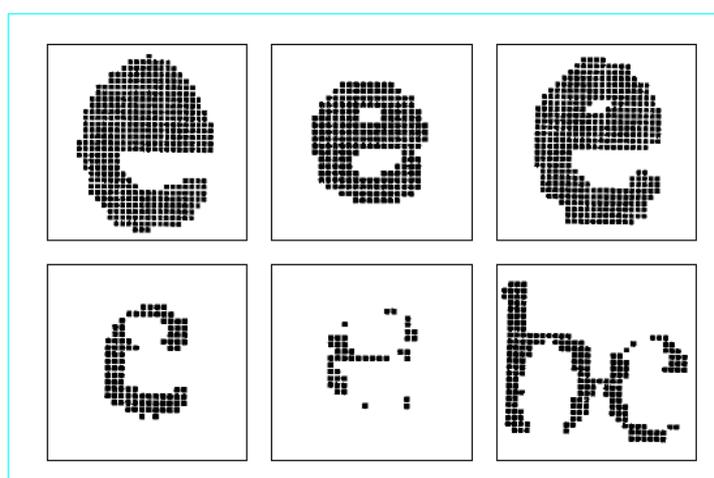


Figure III.4 : Exemples d'images de caractères rencontrés.

III.4.3 Erreurs de reconnaissance et mesure des performances

Les défauts sur les caractères engendrent pour les logiciels d'OCR des erreurs de reconnaissance. L'évaluation de la qualité de ces logiciels nécessite de caractériser puis d'évaluer quantitativement les erreurs.

III.4.3.1 Types d'erreurs

On classe les erreurs en quatre catégories :

- Rejet : un caractère n'est pas reconnu, le logiciel le détecte cependant et le signale par un code spécial : \$, #. C'est le cas le moins grave ;
- substitution : un caractère est pris pour un autre. C'est le cas le plus ennuyeux ;
- omission : un caractère est perdu (cas des caractères collés) ;
- ajout : il y a un caractère supplémentaire. Dans ce cas, une tache est prise pour un caractère, ou un caractère est coupé en deux.

Les erreurs d'omission et d'ajout sont le plus souvent liées à une substitution ; par exemple, un « m » coupé est reconnu comme « rn » : substitution + ajout, ou « ri » collés vont être reconnus comme « n » : omission + substitution. Les cas les plus fréquents de substitutions ont lieu par exemple entre les lettres « I », « l », et le chiffre « 1 » ou entre la lettre « O » et le chiffre « 0 », qui ne se distinguent que par le contexte.

III.4.3.2 Taux d'erreur

On définit le nombre total d'erreurs de reconnaissance comme la somme des erreurs de chaque type :

$$Err = Rejet + Substitution + Omission + Ajout$$

Si le texte reconnu comporte N caractères, le taux d'erreur est simplement le rapport des deux nombres :

$$Taux\ d'erreur = \frac{Err}{N}$$

On définit de même le taux de rejet et le taux de confusion :

$$Taux\ de\ rejet = \frac{Rejet}{N}$$

$$Taux\ de\ confusion = \frac{Substitution + Omission + Ajout}{N}$$

Le taux de reconnaissance est le complément à 1 du taux d'erreur :

$$\text{Taux de reconnaissance} = \frac{N - \text{Err}}{N}$$

Comme ordre de grandeur pour ces taux, il faut savoir qu'au-dessus de 1 à 1,5 % de taux d'erreur, un logiciel d'OCR est considéré comme difficilement exploitable. Pour obtenir l'erreur sur les mots, il faut en effet multiplier le taux d'erreur sur les caractères par un facteur proche de 7 (en considérant qu'un mot comporte en moyenne 7 caractères) : un taux d'erreur de 1,5 % implique donc qu'un relecteur doit corriger 1 mot sur 10 dans le texte, ce qui devient vite impraticable. Le taux d'erreur d'un logiciel de reconnaissance n'est pas affichable a priori : il dépend de la qualité du document reconnu. Des logiciels récents comme Fine Reader ou PRASAD ont des taux d'erreurs de 0,2 % ou moins sur des documents originaux de bonne qualité.

III.4.4 Définition de la reconnaissance de caractères

Avant de poursuivre, il est nécessaire de préciser la notion de « reconnaissance de caractères ». Ce terme recouvre deux concepts.

Le premier que nous désignerons par reconnaissance de bas niveau, en anglais pattern matching, consiste à décider de l'identité quasi point à point, d'une forme-échantillon avec une forme-modèle prédéfinie (gabarit). Par exemple la forme #, constituée de quatre traits parallèles deux à deux, peut servir de modèle. Si maintenant, nous retrouvons ce même signe # imprimé dans un document, nous le reconnaissons, c'est-à-dire nous constatons son identité visuelle avec le signe imprimé sur la ligne précédente.

Le deuxième que nous désignerons par reconnaissance symbolique consiste à analyser une forme-échantillon, pour lui attribuer une étiquette, un code ou un nom symbolique, qui caractérise un modèle générique et abstrait. Ce problème est beaucoup plus complexe car un modèle symbolique correspond à une multitude de formes-modèles qui se distinguent les unes des autres par des variantes topologiques et métriques. Par exemple, les formes suivantes : f, f, f, f correspondent toutes à la sixième lettre de l'alphabet, et sont nommées « caractère f », bien qu'elles soient graphiquement assez différentes, par leur taille comme par d'autres caractéristiques.

Les logiciels d'OCR qui se sont succédé sur le marché pendant les 30 dernières années ont évolué d'une reconnaissance de bas niveau vers une reconnaissance symbolique.

III.4.5 Processus mis en œuvre

La reconnaissance, qu'elle soit de bas niveau ou symbolique, est une tâche complexe qui recouvre plusieurs étapes quasiment identiques dans les deux cas. Globalement, il s'agit de comparer des caractères-échantillons en entrée, à des modèles (formes-modèles ou modèles symboliques) décrivant des classes prédéfinies, et de décider à laquelle des classes existantes l'échantillon appartient. C'est donc un problème de classement.

III.4.5.1. Etapes de reconnaissance

- Segmentation des caractères
- Extraction des caractéristiques
- Décision ou reconnaissance proprement dite

III.4.5.2 Synthèse

La figure III.5 synthétise ces trois étapes qui sont globalement identiques pour la reconnaissance de bas niveau ou symbolique.

Certains logiciels mettent en œuvre successivement les deux types de reconnaissance, ce qui donne un modèle global plus complexe : il est possible en effet, soit d'effectuer la reconnaissance symbolique d'un caractère directement après sa segmentation, soit de commencer par constituer une bibliothèque de formes-modèles des caractères présents dans la page analysée (reconnaissance de bas niveau), puis de réaliser la reconnaissance symbolique des formes-modèles.

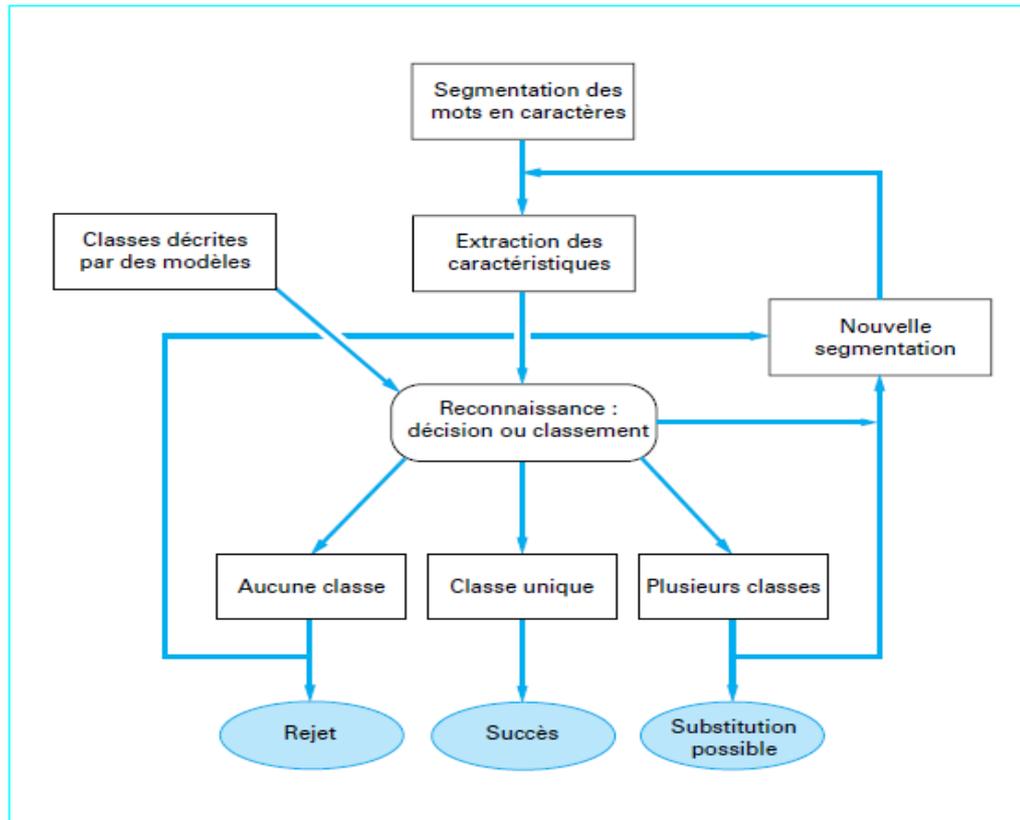


Figure III.5 : Processus mis en œuvre dans la reconnaissance de caractères.

III.4.5.3 Chronologie des étapes de fonctionnement d'un logiciel d'OCR

Nous résumerons ainsi la chronologie et la logique des étapes de constitution et fonctionnement d'un logiciel de reconnaissance :

- choix manuel ou automatique (par exemple dans le cas d'un réseau de neurones) de caractéristiques qui permettent de modéliser les caractères, par des formes-modèles proches des images, ou des modèles symboliques ;
- apprentissage sur une base de référence, qui nécessite une classification, manuelle ou automatique, des images des caractères de la base ; celles-ci sont modélisées, puis réparties en classes homogènes, et enfin labellées ;
- reconnaissance de caractères : prise de décision, qui consiste en un classement des échantillons de caractères par rapport aux classes définies lors de l'apprentissage.

III.5 METHODES DE RECONNAISSANCE

Les méthodes de reconnaissances existe sont :

- Arbre de décision
- Discrimination fonctionnelle linéaire
- Méthodes bayésiennes
- Programmation dynamique et automates
- Réseaux de neurones
- Modèles de Markov
-

La typologie des méthodes de reconnaissance sont représenté dans le tableau suivant :

Décision	Extraction des caractéristiques							
	Méthodes de bas niveau				Méthodes structurales			
	Comparaison globale	Masques partiels	Descripteurs géométriques	Intersection de droites	Grille ajustable	Projections partielles	Analyse des zones blanches	Description topologique
Arbre de décision	I	I						
Discrimination fonctionnelle linéaire	I	I						
Méthodes bayésiennes			II	II	II			
Programmation dynamique et automates			II	II	III	III	III	III
Réseaux de neurones			II	II	III	III	III	III
Modèles de Markov							III	III

Tableau III.1 : Typologie des méthodes de reconnaissance avec: I : 1^{er} génération, II : 2^{ème} génération et, III : 3^{ème} génération.

III.6. CONCLUSION

Convertir un document papier en document électronique est l'objectif des systèmes de reconnaissances de documents. Cette tâche s'avère une nécessité pour offrir une manipulation plus aisée des données : archivage, indexation, recherche, etc.

IV.1 INTRODUCTION

L'objectif de notre travail est de tester les méthodes proposées de détection et de correction des différentes inclinaisons dans les documents imprimés scannés. Comme critères de performance des méthodes proposées, nous examinerons :

- Le temps de détection et de correction de l'inclinaison
- La précision de détection
- Le taux de reconnaissance du module OCR

IV.2 L'ENVIRONNEMENT D'EVALUATION

Les différents programmes ont été écrits avec le langage MATLAB R2010a, et exécutés sur un ordinateur doté d'un microprocesseur intel®Core™ i3-370M, et une mémoire RAM de capacité 2 GB DDR3. Les documents de la base de données ont été scannés à l'aide d'un scanner epson perfection 1250 avec une résolution de 300dpi. **Les documents sont organisés en différentes catégories :**

- ❖ Document arabe (imprimé)
- ❖ Document latin (imprimé et manuscrit)
- ❖ Document chinois
- ❖ Document à structure complexe (texte avec : tableau, graphique, image...).

Chaque document peut contenir jusqu'à 4 sous-régions inclinées. Quelques exemples de la base de données sont illustrés sur la figure IV .1

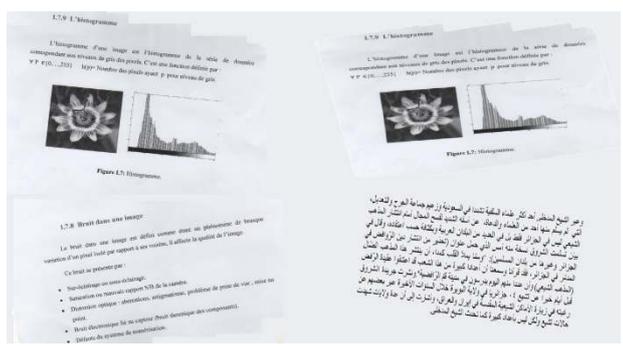
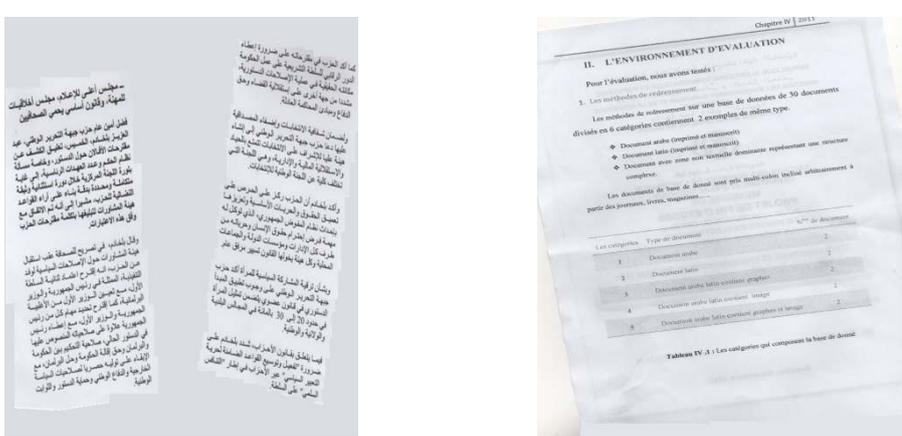


Figure IV.1 : quelques échantillons de la base de données.

IV.3 Présentation de l'interface graphique du logiciel :

Cette interface est représentée sur la figure IV.2

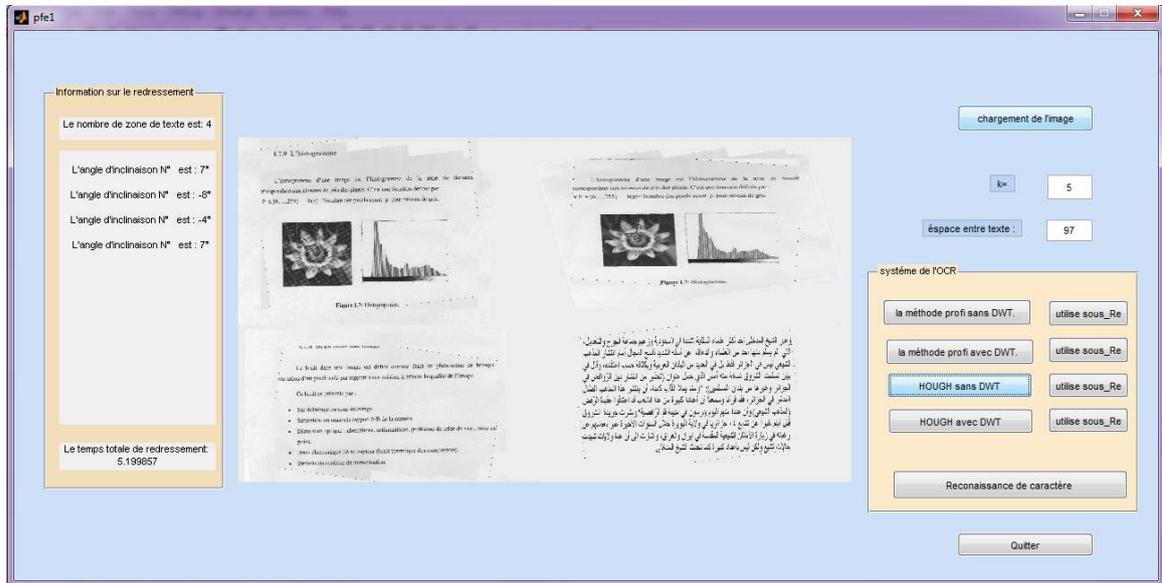


Figure IV.2 : interface graphique

Chargement de l'image :

Cette commande nous permet de choisir une image de la base de données

Bloc correction de l'inclinaison et OCR :

Dans ce bloc nous avons la possibilité de faire un choix parmi les méthodes de détection et de correction des différentes inclinaisons du document en question puis de faire la reconnaissance optique de caractères

Bloc résultats :

Dans cette partie nous trouverons les informations concernant le nombre de sous-régions inclinées, les valeurs des angles détectés, et enfin le temps de détection et de correction.

Dans cette interface nous avons aussi la possibilité de choisir le nombre de sommets K du profil de projection et le rayon R du disque de détection des sous-régions.

VI.4 RESULTAT ET DISCUSSION

IV.4.1 Résultats obtenus pour le redressement des documents :

Les différents résultats sont résumés dans les tableaux suivants :

1^{er} catégorie (2 blocs inclinées avec texte arabe)

	Angle réel	Méthode Profil de projection sans DWT	Méthode Profil de projection avec DWT	Méthode Hough sans DWT	Méthode Hough avec DWT
1 ^{er} angle détecté (en degré)	-8	-8	-8	-7	-7
2 ^{eme} angle détecté (en degré)	10	9	9	10	10
Temps global de détection(s)	-	10.41	3.41	0.432	0.38

Tableau VI.1 : résultat obtenu avec la 1^{ère} catégorie

2^{eme} catégorie (4 blocs inclinées : 2 textes arabes et 2 textes chinois)

	Angle réel	Méthode Profil de projection sans DWT	Méthode Profil de projection avec DWT	Méthode Hough sans DWT	Méthode Hough avec DWT
1 ^{er} angle détecté (en degré)	5	4	3	-89	5
2 ^{eme} angle détecté (en degré)	-10	-10	-9	-10	-10
3 ^{eme} angle détecté (en degré)	3	4	3	-89	5
4 ^{eme} angle détecté (en degré)	-11	-10	-9	-9	-11
Temps global de détection(s)	-	99.39	28.64	2.92	3.27

Tableau VI.2 : résultat obtenu avec la 2^{eme} catégorie

3^{ème} catégorie (2 blocs inclinés : texte manuscrit latin)

	Angle réel	Méthode Profil de projection sans DWT	Méthode profil de projection avec DWT	Méthode Hough sans DWT	Méthode Hough avec DWT
1^{er} angle détecté (en degré)	4	5	5	6	6
2^{ème} angle détecté (en degré)	7	7	7	8	8
Temps global de détection(s)	-	127.28	36.76	2.14	2.68

Tableau VI.3 : résultat obtenu avec la 3^{ème} catégorie**4^{ème} catégorie (1 angle d'inclinaison : document avec texte latin et 1 tableau)**

	Angle réel	Méthode Profil de projection sans DWT	Méthode Profil de projection avec DWT	Méthode Hough sans DWT	Méthode Hough avec DWT
Angle (degré)	-8	-9	-9	-8	-8
Temps global de détection(s)	-	127.28	46.63	4.62	3.56

Tableau VI.4 : résultat obtenu avec la 4^{ème} catégorie

5^{ème} catégorie (1 angle d'inclinaison : document avec texte latin et une image)

	Angle réel	Méthode Profil de projection sans DWT	Méthode Profil de projection avec DWT	Méthode Hough sans DWT	Méthode Hough avec DWT
Angle (en degré)	7	7	7	8	8
Temps global de détection(s)	-	22.47	6.69	0.50	0.54

Tableau VI.5 : résultat obtenu avec la 5^{ème} catégorie

6^{ème} catégorie (2 blocs inclinés : texte latin et texte latin+image+graphique)

	Angle réel	Méthode Profil de projection sans DWT	Méthode Profil de projection avec DWT	Méthode Hough sans DWT	Méthode Hough avec DWT
1 ^{er} angle détecté	-10	-9	-9	-8	-8
2 ^{ème} angle détecté	6	6	6	7	7
Temps global de détection(s)	-	128.48	36.95	2.64	3.17

Tableau VI.6 : résultat obtenu avec la 6^{ème} catégorie

7^{ème} catégorie (3 blocs inclinées : texte latin, texte latin+image+graphique et texte latin+tableau)

	Angle réel	Méthode Profil de projection sans DWT	Méthode Profil de projection avec DWT	Méthode Hough sans DWT	Méthode Hough avec DWT
1 ^{er} angle détecté (en degré)	-9	-9	-9	-8	-8
2 ^{ème} angle détecté (en degré)	6	6	6	7	7
3 ^{ème} angle détecté (en degré)	-8	-9	-8	-8	-8
Temps(s) globale	-	210.69	59.20	5.08	6.02

Tableau VI.7 : résultat obtenu avec la 7^{ème} catégorie

8^{ème} catégorie (4 blocs inclinées : 2 textes latin+image+graphique ; texte latin ; texte arabe)

	Angle réel	Méthode Profil de projection sans DWT	Méthode Profil de projection avec DWT	Méthode Hough sans DWT	Méthode Hough avec DWT
1 ^{er} angle détecté (en degré)	6	6	6	7	7
2 ^{ème} angle détecté (en degré)	-9	-9	-9	-8	-8
3 ^{ème} angle détecté (en degré)	7	7	7	7	7
4 ^{ème} angle détecté (en degré)	-4	-4	-4	-4	-4
Temps(s) globale	-	250.14	70.26	5.09	5.93

Tableau VI.8 : résultat obtenu avec la 8^{ème} catégorie

- Les erreurs de détermination des angles d'inclinaison pour les différentes méthodes sont exprimées en degré dans le tableau IV.9 :

Numéro de catégorie de document	Numéro d'angle	Méthode Profil de projection sans DWT	Méthode Profil de projection avec DWT	Méthode Hough sans DWT	Méthode Hough avec DWT
1	1	0	0	1	1
	2	1	2	84	0
2	1	1	2	84	0
	2	0	1	0	0
	3	1	2	84	0
	4	1	2	2	0
3	1	1	1	2	2
	2	0	0	1	1
4	1	1	1	0	0
5	1	0	0	1	1
6	1	1	1	2	2
	2	0	0	1	1
7	1	0	0	1	1
	2	0	0	1	1
	3	1	0	0	0
8	1	0	0	1	1
	2	0	0	1	1
	3	0	0	0	0
	4	0	0	0	0

Tableau VI.9 : Les erreurs de détermination des angles d'inclinaison en degré.

Discussion :

- Toutes les méthodes sont efficaces, l'erreur variant entre 0° et 2° est à la limite acceptable.
- La méthode du Profil de projection est très lente par rapport à la méthode de Hough.
- La transformé en ondelettes est considérée comme une amélioration efficace à la méthode de profil de projection, en effet, on constate une réduction de temps d'environ $\frac{3}{4}$.
- la transformé en ondelette n'a pas un effet significatif sur la méthode de hough.

Pour réduire le temps de détection nous appliquons les méthodes précédentes non pas au bloc entier mais juste à une sous région du bloc. Les testes ont portés sur 3 catégories de documents, les résultats sont présentés dans les tableaux suivants :

	Angle réel	Méthode Profil de projection sans DWT	Méthode Profil de projection avec DWT	Méthode Hough sans DWT	Méthode Hough avec DWT
angle (degré)	-8	-9	-9	-8	-8
Temps(s) globale	-	5.09	4.63	2.77	2.68

Tableau IV.10 : résultats de détection (méthode de sous région de bloc appliquée au document de la 4ème catégorie)

	Angle réel	Méthode profil de projection sans DWT	Méthode Profil de projection avec DWT	Méthode Hough sans DWT	Méthode Hough avec DWT
Angle (degré)	7	8	8	-82	-82
Temps(s) globale	-	1.16	0.98	0.42	0.48

Tableau IV.11 : résultats de détection (méthode de sous région de bloc appliquée au document de la 5ème catégorie)

	Angle réel	Méthode Profil de projection sans DWT	Méthode Profil de projection avec DWT	Méthode Hough sans DWT	Méthode Hough avec DWT
1 ^{ère} angle détecté (en degré)	-9	-9	-5	-8	-10
2 ^{ème} angle détecté (en degré)	6	30	9	7	7
3 ^{ème} angle détecté (en degré)	-8	0	-13	-8	-8
Temps(s) globale	-	7.69	7.52	4.47	4.62

Tableau IV.12 : résultats de détection (méthode de sous région de bloc appliquée au document de la 7ème catégorie)

Discussion :

- On remarque que la méthode de sous région est efficace dans le cas où le document est textuel ou contenant des tableaux. La méthode n'est pas adaptée aux documents contenant des images.
- La réduction du temps de détection est importante. En effet, nous constatons dans le tableau IV.10 que le temps de détection passe de **127.28 s** à **5.09 s** pour le cas du profil de projection.

IV. 4 .2 Effet des méthodes de détection sur le taux de reconnaissance des caractères du document

Pour vérifier l'efficacité de notre programme de reconnaissance optique de caractères, nous avons commencé par faire des tests sur 4 blocs alignés comportant des textes écrits avec 4 polices différentes : Arial, Times New Roman, Cambria, Calibri . Les taux de reconnaissances obtenus sont représentés dans le tableau IV.13

Types d'écritures	Arial	Times New Roman	Cambria	Calibri
Taux de reconnaissance (%) Caractères majuscules	95.27	10.65	91.64	94.06
Taux de reconnaissance (%) Caractères minuscules	87	46	83	91

Tableau IV.13 : Taux de reconnaissance dans le cas de document aligné

Nous obtenons des taux supérieurs à 90% pour les polices Arial, Cambria et Calibri dans le cas des caractères majuscule, et des taux acceptables pour les polices Arial et Calibri dans le cas des caractères minuscules.

Pour comparer les performances des méthodes d'alignement proposées nous avons testé le taux de reconnaissance d'un document comportant 4 blocs inclinés écrits avec 4 polices différente. Les résultats sont représentés sur les tableaux IV.14

Police d'écriture	Arial	Times New Roman	Cambria	Calibri
Angles estime par profil de projection	-7	10	-11	14
Angles estime par Hough	-6	11	-10	15
Taux de reconnaissance (redressement par profil de projection) (%)	78	72	64	50
Taux de reconnaissance (redressement par Hough) (%)	74	75	67	54
Taux de reconnaissance (%) (cas aligné)	87	46	83	91

Tableau IV.14 : Taux de reconnaissance dans le cas de document multi-inclinés et aligné

IV.5 CONCLUSION

Les différents taux de reconnaissances obtenus après pré-traitement sont inférieurs à ceux obtenus pour les documents initialement alignés. Ces taux sont d'autant plus élevés que la précision de détection est meilleur. Les taux sont pratiquement voisins pour les deux méthodes (profil de projection et hough). La précision de détection est aussi la même pour les deux méthodes. Par contre, le temps de détection diffère d'une méthode à l'autre. En effet, l'application de la DWT réduit le temps de détection pour le cas du profil de projection. Nous constatons aussi que la méthode de hough est nettement plus rapide que la méthode de profil de projection. La méthode de profil de projection appliquée à une sous région de bloc nous a permis de faire une réduction très significative du temps de détection.

Bibliographie :

- [1]- F.ABDAT, « L'analyse des documents par la méthode des ondelettes », mémoire de magister, université SAAD DAHLEB Blida département d'électronique, 2005
- [2] Adobe Systems Incorporated. "Adobe Portable Document Format Reference", <http://partners.adobe.com/asn/tech/pdf/specifications.jsp>.
- [3] Adobe Systems Incorporated. "PostScript Language", Addison-Wesley, 1986.
- [4] Adobe Systems Incorporated. "Tagged Image File Format Specification Revision 6.0", <http://partners.adobe.com/asn/tech/tiff/specification.jsp>.
- [5] GILLES BUREL, « introduction au traitement d'images ».
- [6] P. Bernas, « *Traitements d'images pour la sécurité globale* », Optics Valley, 2006.
- [7] BENCHERCHALI MOHAMED et LOURCI MOHAMED AMINE, « reconnaissance automatique de plaque d'immatriculation de véhicules », PFE, département d'électronique, université SAAD DAHLEB Blida, 2008
- [8] M. Bergounioux, « Méthodes mathématiques pour le Traitement d'Image ».document numérique, 2009.
- [9] TAKILALT Abdllatif et MEHRAD Belkacem , « Redressement de documents scannés » , PFE, département d'électronique , université SAAD DAHLEB Blida, 2010.
- [10] Mélanie Lemaitre, « Approche markovienne bidimensionnelle d'analyse et de reconnaissance de documents manuscrits », DOCTORAT, UFR de Mathématiques et Informatique, Université Paris 5 René Descartes, 2007
- [11] L. Mezai " Analyse et reconnaissance de documents", Mémoire de Magister, I.N.I, (juin 2004).
- [12] M.Belaid, « Cour 4 », Cour de DEA, PRTAL7, France, (2003).
- [13] A.Shad, L.Mezai, M.T Laskri et M.Cheriet, « Estimation de l'inclinaison des documents arabes imprimés », CIFED, France, 2004.
- [14] O.Okun ; M, P. Ainen and J.Sauvola, « robbust Skew Estimation on Low-Resolution Document Image », CIMO, (1990).

[15] N. Khorissi, F. Abdat, A. Mellit, A. Guessoum, M. Bendekhis ‘Methodology for fast skew angle detection and correction by using the Wavelet and the Hough transform: Application for Arabic printed documents’ WSEAS TRANSACTIONS on SIGNAL PROCESSING, Issue 7, Volume 3, July 2007, ISSN1790-5022, <http://www.wseas.org>

[16] DERDOUR Khedidja, « Reconnaissance de formes du chiffre arabe imprimé : Application au code à barre d'un produit », Magistère, Département Génie Industriel, Faculté des Sciences de l'Ingénieur, Université HADJLAKHDAR –BATNA, 2008/2009

[17] <http://www.eldish.net/hp/automat/matlab.htm>

[18] Philippe LEFÈVRE, « Reconnaissance de l'imprimé », Techniques de l'Ingénieur, traité Informatique.

[19] Dalel Ketata et Maher Khemakhem, « Un survol sur l'analyse et la reconnaissance de documents : imprimé, ancien et manuscrit », publié dans "Colloque International Francophone sur l'Écrit et le Document (CIFED2010), sousse : Tunisie (2010)", page(1), 2010.

[20] Abdel Belaïd , « Reconnaissance automatique de l'écriture et du document » ,document numérique.

Introduction générale

Depuis l'avènement de l'écriture, aux environs du III^{ème} millénaire avant Jésus Christ (JC) en Mésopotamie, l'être humain n'a cessé d'améliorer ce moyen de communication. En effet, plusieurs civilisations ont apporté leur savoir faire dans le domaine de l'écriture. Il y a eu d'abord l'écriture des hiéroglyphes des pharaons, l'écriture chinoise, l'écriture grecque, et puis l'écriture arabe et romaine. Certaines de ces écritures ont disparu avec la destruction de leur civilisation alors que d'autres sont encore d'actualité.

Malheureusement parmi les 3000 langues dénombrées dans le monde, seulement une centaine sont dotées d'un système d'écriture. Mais la percée majeure de l'écriture a atteint son apogée au 15^{ème} siècle après JC avec l'invention de l'imprimerie par Gutenberg. Et c'est à partir de cette époque que l'écriture est entrée dans une nouvelle ère, à savoir l'ère du document imprimé.

La notion de document est très générale et il existe une panoplie de définitions. Une définition appropriée d'un document est la suivante : un document est le support physique pour conserver et transmettre de l'information. Selon le support choisi, un document peut être textuel, graphique, multimédia (sonore, vidéo). Le document imprimé a eu un essor en deux temps grâce à l'introduction de l'imprimerie et de l'informatique. En effet, de nos jours, avec l'apport de l'informatique, un grand nombre de documents actuels, qui existent de par le monde, est en format numérique. Ces documents sont soit conçus et réalisés, dès le départ, par ordinateur, soit numérisés au moyen d'un scanner. Il est à noter que le style des documents s'est enrichi de nouvelles typographies et de nouveaux styles grâce aux nouveaux logiciels permettant leur création. Ceci a permis de rendre les documents beaucoup plus attractifs et plus ergonomiques.

L'être humain arrive avec aisance à reconnaître n'importe quel document. Cette faculté est assez basique pour un être humain ; en revanche elle pose jusqu'à ce jour encore des problèmes pour l'ordinateur. L'analyse et la reconnaissance d'images de documents

englobent un ensemble de techniques informatiques avec comme but la reconstitution du contenu du document sous la forme de documents structurés.

Le but de notre projet est de redresser les différentes zones de texte inclinées d'un document imprimé scanné, puis de faire la reconnaissance optique de son écriture.

Se projet fait partie du travail de recherche, de mon directeur de mémoire, dans le domaine de l'analyse, le traitement, et la reconnaissance de document de l'équipe RDF du laboratoire LATSI.

Notre mémoire est organisée comme suit :

Le chapitre un sera consacré aux généralités sur L'analyse et le pré -traitement de document. Dans le deuxième chapitre nous présenterons quelques généralités sur le redressement. La reconnaissance optique de caractères fera l'objet du troisième chapitre. Nous présenterons dans le quatrième chapitre les différents résultats expérimentaux. Enfin, nous terminons notre travail par une conclusion générale.

Annexe

La transformé en ondelette:

La figure 1 explique les étapes de transformées en ondelette

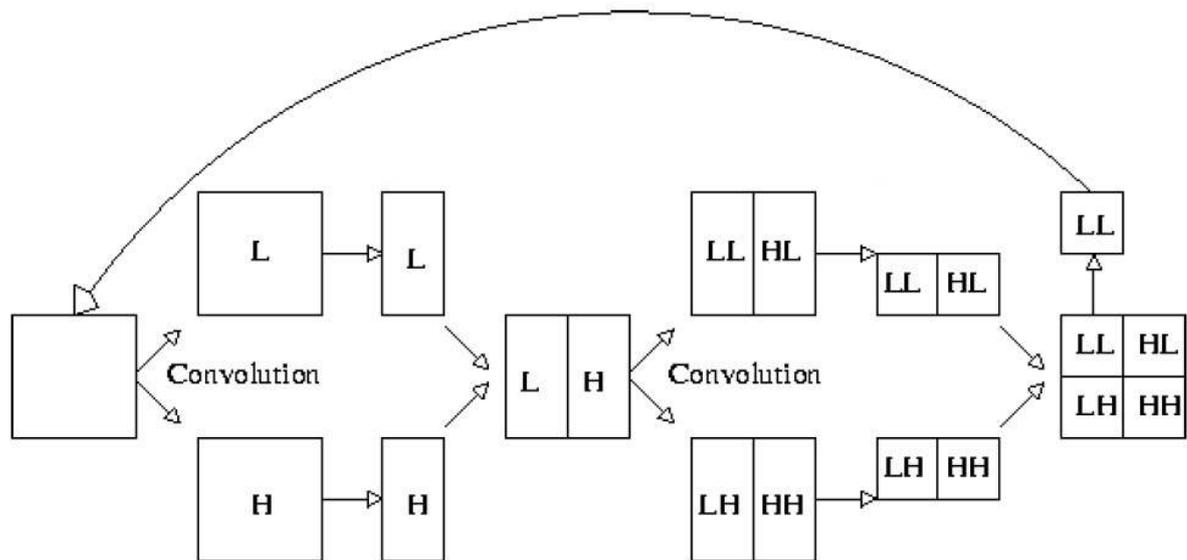


Figure 1 : schéma général de traitement d'images par ondelettes

1. Si l'image du document de l'entrée n'est pas au niveau de gris, transformez-le et noté I_g
2. Décomposer I_g avec 2D-DWT (transformé en ondelette discret de deux dimensions). Alors, quatre sous-bandes fréquentiel (LL, LH, HL, HH) sont obtenues. (ici), la sous bande LL est sélectionnée parce qu'il conserve la structure de l'image du document.

3. Noter la matrice formée par les valeurs absolues des coefficients de sous bande $|LL|$, qui est tournée à travers $[-\alpha, \alpha]$ et un profil de projection horizontal est calculé à chaque angle.
4. L'angle d'inclinaison est estimé grâce à la détermination des sommets dans chaque profil de projection. L'algorithme est décrit sur la figure II.6.

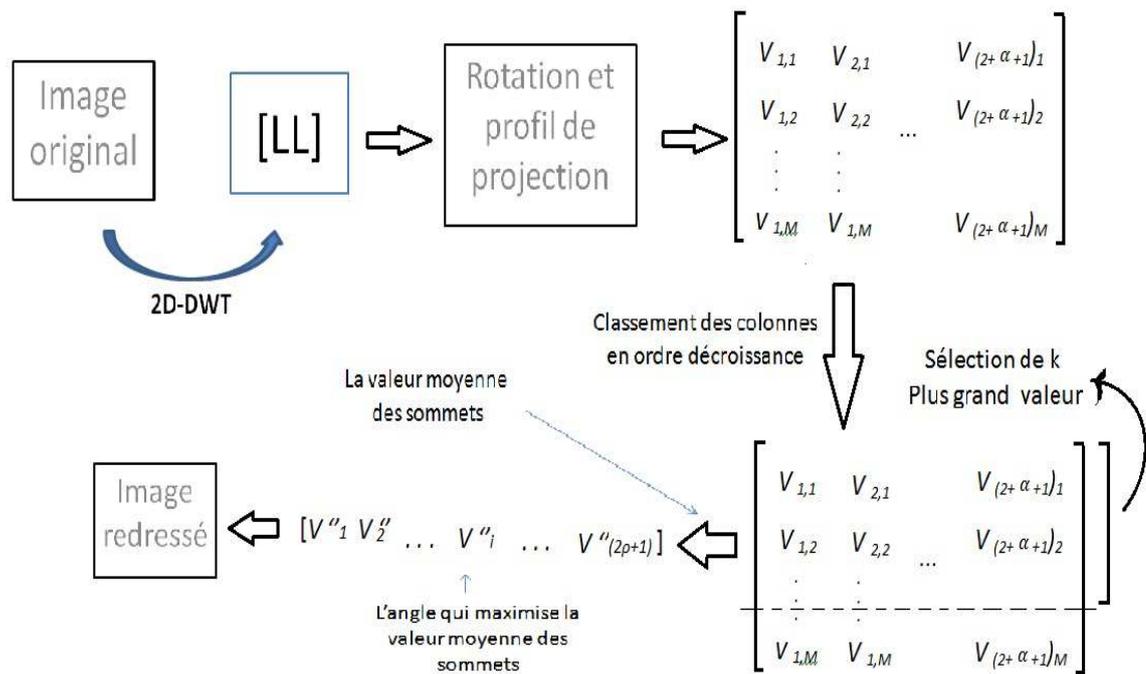
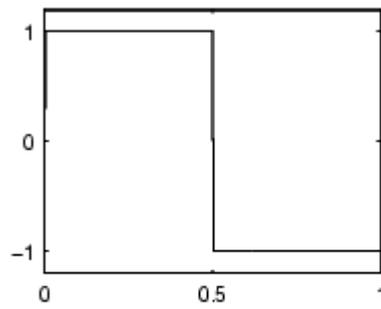
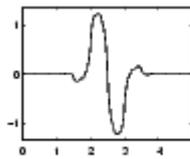


Figure 2: Schéma de la méthode de profil par projection.

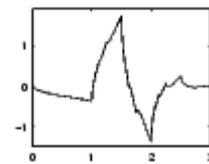
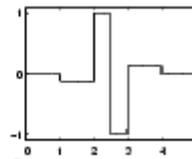
Quelque exemple des bases d'ondelette



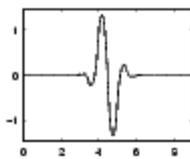
Haar



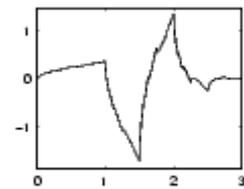
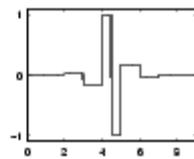
bior1.3



db2



bior1.5



sym2

sommaire

Remerciements

Dédicaces

Résumé

Sommaire

Liste des figures

Liste des tableaux

Introduction générale1

Chapitre I : Généralités sur l'analyse et le prétraitement de document

I.1 LE DOCUMENT.....4

I.1.1 Définition d'un document.....4

I.1.2 Format d'un document.....4

I.1.3 Le document électronique.....4

I.1.3.1 Définition document électronique.....4

I.1.3.2 Caractéristique document électronique.....5

I.1.4 L'image synthétique5

I.1.5 Documents à structures complexes.....5

I.1.6 Production de documents avec les moyens électroniques.....8

I.1.6.1 Structure logique et structure physique d'un document électronique.....8

I.1.6.2 Les étapes de production et les différentes formes intermédiaires d'un document10

I.1.6.3 Types de formats de fichiers10

I.2 TRAITEMENT D'IMAGE12

I.2.1 Introduction.....12

I.2.2 Domaines d'application du traitement d'image.....14

I.2.3 Notions fondamentales en traitement d'images.....14

I.2.3.1 Définition d'une image14

I.2.3.2 Acquisition d'une image.....15

I.2.3.3 Définition d'une image numérique15

I.2.3.4 Définition du PIXEL.....	15
I.2.3.5 Les images binaires (noir ou blanc)	15
I.2.3.6 Les images en niveaux de gris	16
I.2.3.7 Les images couleurs	16
I.2.3.8 Bruit dans une image.....	17
I.2.3.9 L’histogramme	17
I.2.3.10 Segmentation d’une image.....	18
I.2.3.11 Classification.....	19
I.3 CONCLUSION.....	19
Chapitre II : Généralités sur le redressement de document	
II.1 INTRODUCTION	21
II.2 ALGORITHMES DE DETECTION D’ORIENTATION.....	22
II.2.1 Méthodes de projection	22
II.2.1.1 Analyse du profil de projection	25
II.2.2 Transformation de Hough	26
II.2.2.1 L’algorithme proposé :	28
II.2.3 Regroupement des plus proches voisins.....	29
A. La méthode de codage des séries de pixels noirs.....	30
B. La méthode des composantes connexes.....	31
C. Une approche se basant sur l’extraction des liaisons.....	32
D. Une approche travaillant seulement sur des sous régions de l’image.....	33
E. Les approches travaillant sur le document compressé	34
II.3 CORRECTION DE L’INCLINAISON	35
II.4 CONCLUSION	35
Chapitre III : Reconnaissance de document	
III.1.INTRODUCTION	37
III.2. AVANTAGES DES FORMES ELECTRONIQUES CODEES PAR RAPPORT AUX FORMES IMAGE.....	38
III.3. COMPOSANTES D’UN SYSTEME DE RECONNAISSANCE.....	39
III.3.1 Acquisition ou numérisation	40
III.3.2 Redressement	40
III.3.3 Binarisation	40
III.3.4 Segmentation des pages	40
III.3.5 Reconnaissance des caractères.....	40

III.3.6 Reconnaissance de la structure Logique	41
III.3.7 Transcodage	41
III.4. RECONNAISSANCE DES CARACTERES (OCR)	41
III.4.1 Structure des caractères et variété	42
III.4.2 Déformations dues à l'impression, la reproduction, la numérisation	43
III.4.3 Erreurs de reconnaissance et mesure des performances	43
III.4.3.1 Types d'erreurs	44
III.4.3.2 Taux d'erreur.....	44
III.4.4 Définition de la reconnaissance de caractères.....	45
III.4.5 Processus mis en œuvre	46
III.4.5.1. Etapes de reconnaissance	46
III.4.5.2 Synthèse	46
III.4.5.4 Chronologie des étapes de fonctionnement d'un logiciel d'OCR.....	47
III.5 METHODES DE RECONNAISSANCE.....	48
III.6. CONCLUSION.....	49
Chapitre IV : Partie expérimentale	
IV.1 INTRODUCTION.....	51
IV.2 L'ENVIRONNEMENT D'EVALUATION.....	51
IV.3. PRESENTATION DE L'INTERFACE GRAPHIQUE DU LOGICIEL.....	53
IV.4. RESULTAT ET DISCUSSION.....	54
IV.4.1 Résultats obtenus pour le redressement des documents.....	54
IV.4.2 Effet des méthodes de détection sur le taux de reconnaissance des caractères du document.....	61
VI CONCLUSION.....	62
Conclusion générale.....	63
Bibliographie	
Annexes	

CONCLUSION GENERALE

L'objectif de notre travail était la détection et l'estimation des différentes inclinaisons dans un document imprimé scanné. Pour cela, nous avons expérimenté deux approches qui existent déjà dans la littérature : le profil de projection et la transformée de hough. Ces deux approches ont été amélioré grace à l'introduction de la DWT et au traitement restreint à une sous-région d'un bloc du document. Ainsi, les résultats obtenus sont très prometteurs

Nous avons retenu trois critères de performances pour ces méthodes :

- La précision de détection des différentes inclinaisons
- Le temps de détection
- Le taux de reconnaissance

Ces critères sont très décisifs lorsqu'il s'agit de faire un choix parmi les logiciels de traitement de documents existants dans le commerce.

Comme perspectives il serait très intéressant de comparer nos résultats avec ceux obtenus par les autres approches, à savoir : la méthode basée sur les k plus proches voisins.

Enfin, nous espérons que ce modeste projet de fin d'étude puisse apporter un plus aux travaux futurs dans le domaine de l'analyse et le traitement de document

liste des figures

Chapitre I : Généralités sur l'analyse et le pré-traitement de document

Figure I.1 : Exemple de deux documents à structures complexes d'un même journal.....	7
Figure I.2 : Exemple de structures.....	9
Figure I.3 : Etapes de production et les différentes formes d'un document.....	10
Figure I.4 : Architecture matérielle d'un système complet de traitement d'image.....	13
Figure I.5 : La chaîne de vision assistée par ordinateur.....	14
Figure I.6 : L'espace couleur Rouge Vert Bleu (RGB).....	16
Figure I.7 : Histogramme	17
Figure I.8 : Deux images différentes peuvent avoir le même histogramme.....	18

Chapitre II : Généralités sur le redressement de document

Figure II.1 : un document incliné.....	21
Figure II.2 : Illustration du calcul l'histogramme horizontal.....	23
Figure II.3 : Image incliné et se profil horizontal.....	25
Figure II.4 : Le profil de projection après le redressement.....	26
Figure II.5 : Représentation d'une droite dans un plan.....	27
Figure II.6 : Représentation d'une droite dans l'espace de Hough.....	27
Figure II.7 : 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.....	30
Figure II.8 : La méthode des composantes connexes, (a) : Texte arabe incliné, (c) : Texte latin incliné, (b, d): les points centraux des composantes connexes.....	31
Figure II.9 : La méthode d'extraction des liaisons.....	32
Figure II.10 : La sélection d'une région à partir d'un document.....	33
Figure II.11 : Histogramme du document.....	33

Chapitre III : Reconnaissance de document

Figure III.1 : Les formes des documents et la reconnaissance.....	37
--	----

Figure III.2 : Les étapes du processus de reconnaissance de document.....	39
Figure III.3 : Exemples de la classification Vox.....	42
Figure III.4 : Exemples d'images de caractère rencontrées.....	43
Figure III.5 : Processus mis en œuvre dans la reconnaissance de caractère.....	47
Chapitre IV : Partie expérimental	
Figure IV.1 : Quelques échantillons de la base de données.....	52
Figure IV.2 : Interface graphique.....	53
Annexe :	
Figure 1 : schéma général de traitement d'images par ondelettes.	
Figure 2 : Schéma de la méthode de profil par projection.	

Liste des tableaux

Chapitre III : Reconnaissance de document

Tableau III.1 : Typologie des méthodes de reconnaissance avec: I : 1 ^{er} génération, II : 2 ^{ème} génération et, III : 3 ^{ème} génération.....	48
---	----

Chapitre IV : Partie expérimental

Tableau VI.1 : résultat obtenu avec la 1 ^{ère} catégorie.....	54
Tableau VI.2 : résultat obtenu avec la 2 ^{ème} catégorie.....	54
Tableau VI.3 : résultat obtenu avec la 3 ^{ème} catégorie.....	55
Tableau VI.4 : résultat obtenu avec la 4 ^{ème} catégorie.....	55
Tableau VI.5 : résultat obtenu avec la 5 ^{ème} catégorie.....	56
Tableau VI.6 : résultat obtenu avec la 6 ^{ème} catégorie.....	56
Tableau VI.7 : résultat obtenu avec la 7 ^{ème} catégorie.....	57
Tableau VI.8 : résultat obtenu avec la 8 ^{ème} catégorie.....	57
Tableau VI.9 : Les erreurs de détermination des angles d'inclinaison en degré.....	58

Tableau IV.10 : résultats de détection (méthode de sous région de bloc appliquée au document de la 4ème catégorie).....	59
Tableau IV.11 : résultats de détection (méthode de sous région de bloc appliquée au document de la 5ème catégorie).....	60
Tableau IV.12 : résultats de détection (méthode de sous région de bloc appliquée au document de la 7ème catégorie).....	60
Tableau IV.13 : Taux de reconnaissance dans le cas de document aligné.....	61
Tableau IV.14 : Taux de reconnaissance dans le cas de document multi-inclinés et aligné.....	62