

الجمهورية الجزائرية الديمقراطية الشعبية

République Algérienne Démocratique et Populaire

Ministère de L'Enseignement Supérieur et de la Recherche Scientifique



**UNIVERSITÉ FERHAT ABBAS – SETIF 1**

**FACULTÉ DE TECHNOLOGIE**

**THÈSE**

**Présentée au Département d'électronique**

**Pour l'obtention du diplôme de**

**DOCTORAT**

**Domaine : Sciences et Technologie**

**Filière: Electronique**

**Option: Systèmes Embarqués et  
Technologie**

**Par**

**KHITAS Mehdi**

**THÈME**

**Implantation d'application d'aide de décision avec un  
serveur temps réel pour images dégradées via des moyens  
de communication**

**Soutenue le 24/07/2019 devant le Jury:**

<b>AMARDJIA Nouredine</b>	<b>Professeur</b>	<b>Univ. F. Abbas Sétif 1</b>	<b>Président</b>
<b>ZIET Lahcene</b>	<b>Professeur</b>	<b>Univ. F. Abbas Sétif 1</b>	<b>Directeur de thèse</b>
<b>CHEMALI Hmimi</b>	<b>Professeur</b>	<b>Univ. F. Abbas Sétif 1</b>	<b>Examineur</b>
<b>BOUMEHRAZ Mohamed</b>	<b>Professeur</b>	<b>Univ. M. Khider Biskra</b>	<b>Examineur</b>
<b>BOUGUERRA Abderrahmene</b>	<b>M.C.A.</b>	<b>Univ. M. Boudiaf Msila</b>	<b>Examineur</b>

## DEDICACES

Je dédie ce travail à la mémoire de mon père

« Pour toi qui n'a pas vu l'aboutissement de mon travail mais je sais que tu en aurais été très fier »

## REMERCIEMENTS

Je remercie tout d'abord **ALLAH** de m'avoir prêté santé et volonté pour mener à terme cette thèse de doctorat.

Je remercie chaleureusement mon directeur de thèse Monsieur **ZIET LAHCENE**, Professeur à l'université de Sétif, pour son intérêt et son soutien, sa grande disponibilité et ses nombreux conseils durant la réalisation de ma thèse.

Je tiens également à remercier vivement les membres de jury pour avoir accepté d'évaluer cette thèse : Monsieur **AMARDJIA Noureddine**, Professeur à l'Université de Sétif, d'avoir accepté de juger mon travail et de présider le jury de soutenance de cette thèse. Messieurs **CHEMALI Hmimi** Professeur à l'Université de Setif, **BOUMEHRAZ Mohamed**, Professeur à l'Université de Biskra et **BOUGUERRA Abderrahmene**, Maître de conférences classe A à l'Université de Batna, pour m'avoir fait l'honneur d'accepter d'être les examinateurs de cette thèse.

Je tiens à exprimer mes plus vifs remerciements à Monsieur **BOUGAZEL Saad** Professeur à l'université de Sétif pour son soutien scientifique mais également pour sa bonne humeur et sa disponibilité au cours de l'élaboration de cette thèse.

Un grand merci aux membres de **Ma Famille** et **Belle Famille** pour leurs amours et leurs soutiens, en particulier mes grands-parents, mon frère et ma sœur dont je suis très fier.

Merci à ma petite sœur, pour ses conseils et son précieux aide au cours de l'élaboration de cette thèse.

Merci à ma chère épouse, mon plus grand soutien. Merci d'avoir cru à ce travail, merci de l'avoir défendu.

Merci à ma chère mère '**FATIMA**', elle mérite tout le respect et la reconnaissance pour tous ce qu'elle a du faire pour moi et pour ce qu'elle fait chaque jour Que Dieu fait.

Enfin, pour tous ceux qui, de près ou de loin, ont rendu ce travail possible.

# TABLE DES MATIERES

Dédicaces .....	i
Remerciments .....	ii
Table des matières .....	iii
Listes des figures .....	vi
Liste des tableaux .....	ix
Liste des abreviations .....	x
Introduction générale .....	1

## CHAPITRE 1

NUMERISATION DES DOCUMENTS, ANALYSE DES CONTENUS DE DOCUMENT ET BIBLIOTHEQUES ELECTRONIQUES .....	5
1.1. Introduction .....	5
1.2. Le processus de digitalisation.....	5
1.2.1. Acquisition et scanning .....	6
1.2.2. Traitement des images .....	8
1.2.3. Restauration.....	8
1.2.4. Reconnaissance et reconstruction.....	9
1.2.5. Contrôle de la qualité.....	9
1.2.6. Hébergement web .....	10
1.3. Analyse des images document / cas des bibliothèques historiques électroniques.....	10
1.3.1. Acquisition d'images .....	10
1.3.2. Analyse de la mise en page et extraction des métadonnées. ....	11
1.3.3. Reconnaissance de l'écriture manuscrite pour DL .....	12
1.3.4. Bibliothèques électroniques multi-langues.....	13
1.3.5. Analyse de forme multi-langue .....	14
1.3.6. Analyse du contenu .....	14
1.4. Conclusion.....	15
Bibliographie .....	16

## CHAPITRE 2

APERÇU SUR LES METHODES DE BINARISATION, IMPLEMENTATION ET COMPARAISON .....	19
2.1. Introduction .....	19

2.2.	Binarisation .....	19
2.3.	Les méthodes de binarisation .....	19
2.3.1.	Méthodes de binarisation globale .....	20
2.3.2.	Méthodes de binarisation locale .....	25
2.3.3.	Méthodes de binarisation hybride .....	27
2.4.	Implémentation d'algorithmes de calcul du seuil de binarisation .....	32
2.4.1.	Implémentation d'une méthode simple de calcul du seuil .....	32
2.4.2.	Implémentation méthode globale de calcul du seuil .....	33
2.4.3.	Implémentation de la méthode locale de calcul du seuil .....	36
2.5.	Résultats de binarisation sur quelques images vieux documents .....	37
2.6.	Conclusion.....	43
	Bibliographie.....	45

## CHAPITRE 3

### BINARISATION PAR ESTIMATION DU BACKGROUND ET SELECTION

	AUTOMATIQUE DE FENETRE DE TRAITEMENT .....	47
3.1.	Introduction .....	47
3.2.	Binarisation des documents ancien par utilisation du filtre médian pour l'estimation du background.....	47
3.2.1.	Détection du type du document.....	48
3.2.2.	Estimation de l'arrière-plan.....	50
3.2.3.	Compensation du contraste de l'image.....	52
3.2.4.	Estimation du seuil globale .....	52
3.2.5.	Étape de poste traitement.....	52
3.3.	Bibliothèque et outils d'évaluations des images .....	53
3.3.1.	Bibliothèque des images .....	54
3.3.2.	Evaluation des méthodes de binarisation .....	54
3.4.	Expérimentale et résultat .....	58
3.4.1.	Test sur les différentes bases de données DIBCO.....	59
3.4.2.	Teste sur la bibliothèque « Bickley Diary » .....	67
3.5.	Conclusion.....	68
	Bibliographie.....	70

## CHAPITRE 4

IMPLEMENTATION HARDWARE D'ALGORITHMES DE BINARISATION D'IMAGES DE DOCUMENTS ANCIENS SUR FPGA .....	73
4.1. Introduction .....	73
4.2. Plateforme et outils de conception d'un système SoPC .....	73
4.2.1. Carte de développement .....	74
4.2.2. Logiciels utilisés.....	76
4.3. Système proposé.....	79
4.3.1. Le processeur Nios II.....	80
4.3.2. Module histogramme.....	85
4.3.3. Module Otsu .....	86
4.3.4. Transmission internet de l'image .....	88
4.4. Résultats .....	90
4.5. Conclusion.....	95
Bibliographie.....	96
Conclusion generale .....	98
ANNEXE A .....	99

# LISTE DES FIGURES

Figure 1 : Exemple d'image dégradée.....	1
Figure 2 : Système de processus de préservation numérique.....	2
Figure 3 : (a) et (c) Images originales. (b) et (d) Images après traitement. ....	3
Figure 4 : (a) Image scannée (b) Image binarisée après normalisation de l'intensité.....	4
Figure 1.1 : Vue générale du processus de la numérisation de la bibliothèque.....	6
Figure 1.2 : Processus de scannage, de traitement d'image et diagramme de reconnaissance OCR.....	7
Figure 1.3 : Contrôle de la qualité et de la performance.....	10
Figure 2.1 : Allures idéales d'un histogramme. (a): Unimodal, (b): Bimodal, (c) Multimodal.....	21
Figure 2.2 : Illustration du seuillage adaptatif. (a) Image originale avec un fond non uniforme, (b) Image segmentée avec un seuil global, (c) Image originale découpée en sous images, (d) Image segmentée avec un seuillage adaptatif.....	25
Figure 2.3 : Résultat de binarisation sur la base de données DIBCO 2009. (a) Image originale, (b) Image de référence, (c) Méthode d'Otsu, (d) Méthode d'ISODATA, (e) Méthode de Bernsen, (f) Méthode de Niblack.....	41
Figure 2.4 : Résultat sur la base de données DIBCO 2010. (a) Image originale, (b) Image de référence, (c) Méthode d'Otsu, (d) Méthode d'ISODATA, (e) Méthode de Bernsen, (f) Méthode de Niblack..	41
Figure 2.5 : Résultat sur la base de données DIBCO 2011. (a) Image originale, (b) Image de référence, (c) Méthode d'Otsu, (d) Méthode d'ISODATA, (e) Méthode de Bernsen, (f) Méthode de Niblack..	42
Figure 2.6 : Résultat sur la base de données DIBCO 2012. (a) Image originale, (b) Image de référence, (c) Méthode d'Otsu, (d) Méthode d'ISODATA, (e) Méthode de Bernsen, (f) Méthode de Niblack..	42
Figure 2.7 : Résultat sur la base de données DIBCO 2013. (a) Image originale, (b) Image de référence, (c) Méthode d'Otsu, (d) Méthode d'ISODATA, (e) Méthode de Bernsen, (f) Méthode de Niblack.....	43
Figure 3.1: Type d'Histogramme : (a) Histogramme bimodale, (b) Histogramme multimodale.....	48
Figure 3.2: (a) Image originale. (b) Estimation du background par le lissage polynomial ligne par ligne [2]. (c) Estimation du background par le lissage polynomial colonne par colonne [2]. (d) Estimation du background par la méthode proposée. ....	51
Figure 3.3: Compensation du contraste: (a) Image originale. (b) Background estimé par la méthode proposée. (c) Image compensée.....	52
Figure 3.4 : Evaluation d'un algorithme de binarisation.....	53
Figure 3.5 : Comment évaluer une méthode de binarisation.....	54
Figure 3.6 : Exemple pour le calcul de la F-measure.....	56

Figure 3.7 : Image originale de la collection DIBCO 2009 Handwritten avec l'image de <i>vérité-terrain</i> correspondante. ....	60
Figure 3.8 : Image originale de la collection DIBCO 2009 Printed avec l'image de <i>vérité-terrain</i> correspondante. ....	61
Figure 3.9 : Image originale de la collection DIBCO 2011 Handwritten avec l'image de <i>vérité-terrain</i> correspondante. ....	61
Figure 3.10 : Image originale de la collection DIBCO 2011 Printed avec l'image de <i>vérité-terrain</i> correspondante. ....	61
Figure 3.11 : Image originale de la collection DIBCO 2013 Printed avec l'image de <i>vérité-terrain</i> correspondante. ....	61
Figure 3.12 : Image originale de la collection DIBCO 2013 Printed avec l'image de <i>vérité-terrain</i> correspondante. ....	62
Figure 3.13 : Résultat de la binarisation par différentes méthodes (a) Image originale DIBCO 2009, (b) Otsu, (c) Sauvola,(d) Niblack,(e) Bernsen,(f) Gatos,(g) BE,(h) Méthode proposée, (i) Image de référence .....	62
Figure 3.14 : Résultat de la binarisation par différentes méthodes (a) Image originale DIBCO 2009, (b) Otsu, (c) Sauvola,(d) Niblack,(e) Bernsen,(f) Gatos,(g) BE,(h) Méthode proposée, (i) Image de référence.....	63
Figure 3.15 : Résultat de la binarisation par différentes méthodes (a) Image originale DIBCO 2011, (b) Otsu,(c) Sauvola,(d) Niblack,(e) Bernsen,(f) Gatos,(g) BE,(h) LELO,(i) Su SNUS , (j) Howe,(k) Méthode proposée,(l) Image de référence. ....	64
Figure 3.16 : Résultat de la binarisation par différentes méthodes (a) Image originale DIBCO 2011, (b) Otsu,(c) Sauvola,(d) Niblack,(e) Bernsen,(f) LELO,(g) Su SNUS,(h) Howe, (i) Méthode proposée,(j) Image de référence .....	64
Figure 3.17 : Résultat de la binarisation par différentes méthodes (a) Image originale DIBCO13, (b) Otsu,(c) Sauvola,(d) Niblack,(e) Bernsen,(f) 1 <sup>st</sup> Rank DIBCO 2013,(g) 2 <sup>nd</sup> Rank DIBCO 2013,(h) Méthode proposée,(i) Image de référence.. ....	65
Figure 3.18 : Résultat de la binarisation par différentes méthodes (a) Image originale DIBCO13, (b) Otsu,(c) Sauvola,(d) Niblack,(e) Bernsen,(f) 1 <sup>st</sup> Rank DIBCO 2013,(g) 2 <sup>nd</sup> Rank DIBCO 2013,(h) 3 <sup>rd</sup> Rank DIBCO 2013,(i) 4 <sup>th</sup> Rank DIBCO2013,(j) 5 <sup>th</sup> Rank DIBCO 2013,(k) Méthode proposée,(l) Image de référence .....	66
Figure 3.19 : Résultat de la binarisation par différentes méthodes (a) Image originale DIBCO13, (b) Otsu,(c) Sauvola,(d) Niblack,(e) Bernsen,(f) 1 <sup>st</sup> Rank DIBCO 2013,(g) 2 <sup>nd</sup> Rank DIBCO 2013,(h) Méthode proposée,(i) Image de référence. ....	66

Figure 3.20 : Résultats de la binarisation d'une image dégradé de la bibliothèque Bickley diary obtenue par différentes méthodes de binarisation. (a) Otsu. (b) Sauvola. (c) Niblack. (d) Bernsen. (e) Gatos. (f) LMM. (g) BE. (h) Su. (i) Méthode proposée. (j) Image de référence. ....	68
Figure 4.1 : Carte DE2 .....	76
Figure 4.2 : Périphérique de Cyclone FPGA II 2C35 .....	77
Figure 4.3 : IDE Quartus II .....	78
Figure 4.4 : Flot de conception.....	79
Figure 4.5 : Flot de conception mixte.....	79
Figure 4.6 : Interface Qsys et mapping mémoire.....	80
Figure 4.7 : Système de binarisation et de transfert.....	81
Figure 4.8 : Conception du système embarqué proposé sur Altera DE2.....	82
Figure 4.9 : Architecture interne du processeur Nios II.....	83
Figure 4.10 : Etape de conception Nios II .....	83
Figure 4.11 : Interface Avalon.....	84
Figure 4.12 : Brochage de la carte SD.....	85
Figure 4.13: Schéma bloc d'IP SD carte.....	85
Figure 4.14 : Schéma bloc d'IP UART .....	86
Figure 4.15 : Exemple de calcul d'histogramme .....	87
Figure 4.16 : Schéma block du module histogramme.....	87
Figure 4.17 : Schéma block du module calcul de seuil .....	88
Figure 4.18 : Schéma block du calcul probabilités moyennes et l'écart type.....	89
Figure 4.19 : Schéma block du calcul de la variance interclasse.....	89
Figure 4.20 : Schéma block de la maximisation de la variance interclasse.....	90
Figure 4.21 : Schéma du mécanisme de transmission.....	90
Figure 4.22 : Brochage NodeMCU 1.0 ESP8266.....	91
Figure 4.23 : Exemple d'une trame UART .....	91
Figure 4.24 : Format d'une trame WiFi.....	92
Figure 4.25 : Système complet sur Qsys partie-1 .....	92
Figure 4.26 : Système complet sur Qsys partie-2 .....	93
Figure 4.27 : Système complet sur Qsys partie-3 .....	93
Figure 4.28 : block diagramme du système .....	94
Figure 4.29 : Rapport de compilation du système .....	94
Figure 4.30 : État de système avant d'implémentation.....	95
Figure 4.31 : Affichage VGA le cas d'une RAM vide avant transfert du programme.....	95
Figure 4.32 : Résultat de binarisation : (a,c,e) Images originales, (b,d,f) Images binarisées.....	96
Figure 4.33 : Résultats de comparaison.....	97

# LISTE DES TABLEAUX

Tableau 1.1 : Paramètres normalisés pour le format image TIFF .....	9
Tableau 3.1: Répartition d'images de documents dans les collections publiques de DIBCO .....	60
Tableau 3.2. Résultat de mesure DIBCO 2009.....	62
Tableau 3.3. Résultat de mesure DIBCO 2011 .....	63
Tableau 3.4. Résultat de mesure DIBCO2013.....	65
Tableau 3.5. Résultat de mesure Tableau 3.4. Résultat de mesure DIBCO2013. ....	68

# LISTE DES ABREVIATIONS

## Chapitre 1

AIIM	Association for Information and Image Management
DIA	Digital Image Analysis
DL	Digital Library
dpi	dot per inch
HLS	Hue, Lightness, and Saturation
HMM	Hidden Markov Models
LIS	Library Integrated System
MXY	Modifie X-Y
NPES	Association for Suppliers Publishing and Converting
OCR	Optical Character Recognition
PDA	Personal Digital Assistant
RMSC	Rochester Museum & Science Center
R&D	Rresearch and development
SNR	Signal to Noise Ratio
UI	User Interface

## Chapitre 3

DIBCO	Document Image Binarization Contest
DRD	Distance-Reciprocal Distortion metric
FM	F-measure
H-DIBCO	Handwritten Document Image Binarization Competition
MPM	Misclassification penalty metric
NRM	Negative Rate Metric
PSNR	Peak Signal to Noise Ratio
DMA	Direct Memory Access
EDS	Embedded Design Suite
FPGA	Field Programmable Gate Array
HAL	Hardware Abstraction Layer
IP	Intellectual property
RISC	Reduced instruction Set Computer
SD	Secure Digital

## Chapitre 4

DMA	Direct Memory Access
EDS	Embedded Design Suite
FPGA	Field Programmable Gate Array
HAL	Hardware Abstraction Layer
IP	Intellectual property
RISC	Reduced instruction Set Computer
SD	Secure Digital
SoC	System On a Chip
SoPC	System On Programmable Chip
SPI	Serial Peripheral Interface
SS	Slave Select
UP	University program
VHDL	Very high speed integrated circuit Hardware Description Language
WiFi	Wireless Fidelity

# INTRODUCTION GENERALE

## Introduction générale

Les documents historiques précieux, les manuscrits anciens, les vieux livres et journaux accumulés au fil du temps sont généralement conservés dans les bibliothèques et les archives. Cependant, la majorité de ces documents anciens subissent plusieurs dégradations et détériorations dues, principalement, à l'usure de longues années de stockage dans des endroits moins conservatifs en raison de nombreux facteurs environnementaux telles que la chaleur, l'humidité, la poussière, la mauvaise manipulation et la mauvaise qualité du papier. Ces documents anciens sont numérisés et mis à disposition par le biais d'interfaces logicielles telles que des bibliothèques en ligne.

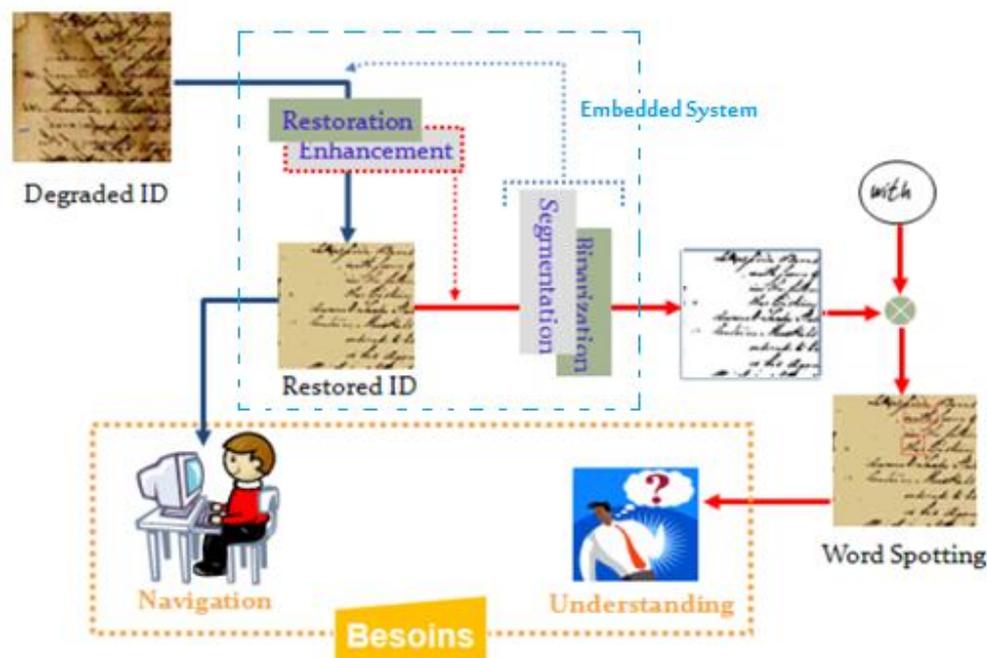
Les images numérisées des documents originaux sont généralement affichées en niveaux de gris ou en couleur. Aujourd'hui, le développement dans le domaine du traitement numérique des images, qui est l'opération la plus efficace pour éliminer les effets indésirables des facteurs environnementaux sur les documents anciens. Certains de ces effets indésirables tels que la transparence, la rétroversion, les taches d'humidité, décoloration, l'encre fanée et les tâches d'encre etc.... (Figure.1).



**Figure 1** : Exemple d'image dégradée

Le processus de préservation numérique évolue à partir de la création des images de haute résolution jusqu'à la génération de différentes versions de ces images qui présentent différents aspects comme la qualité et la taille. Une caractéristique commune entre toutes ces images, c'est qu'elles montrent la détérioration existante dans les documents originaux. La restauration et le rehaussement numérique, dans le contexte d'images de documents anciens dégradés, peuvent être vus comme des processus de transformation qui retrouvent l'aspect

original de l'image du document ancien qui montre un certain état de dégradation. Ces traitements sont nécessaires non seulement pour améliorer la qualité de l'image du document, mais aussi pour améliorer les résultats des opérations ultérieures de segmentation et de reconnaissance. Ainsi, par le nettoyage des dégradations dans les images de documents dégradés, le rapport de l'interprétation du contenu sera plus important et plus significatif. Le traitement numérique des documents anciens a pour but de partager et de diffuser l'information entre plusieurs types d'utilisateurs. L'information à diffuser sera utilisée selon des objectifs bien précis (Figure.2), d'où la nécessité d'accélérer le traitement par des applications matérielles embarquées spécialisées et l'utilisation d'outils très performants .



**Figure 2 :** Système de processus de préservation numérique

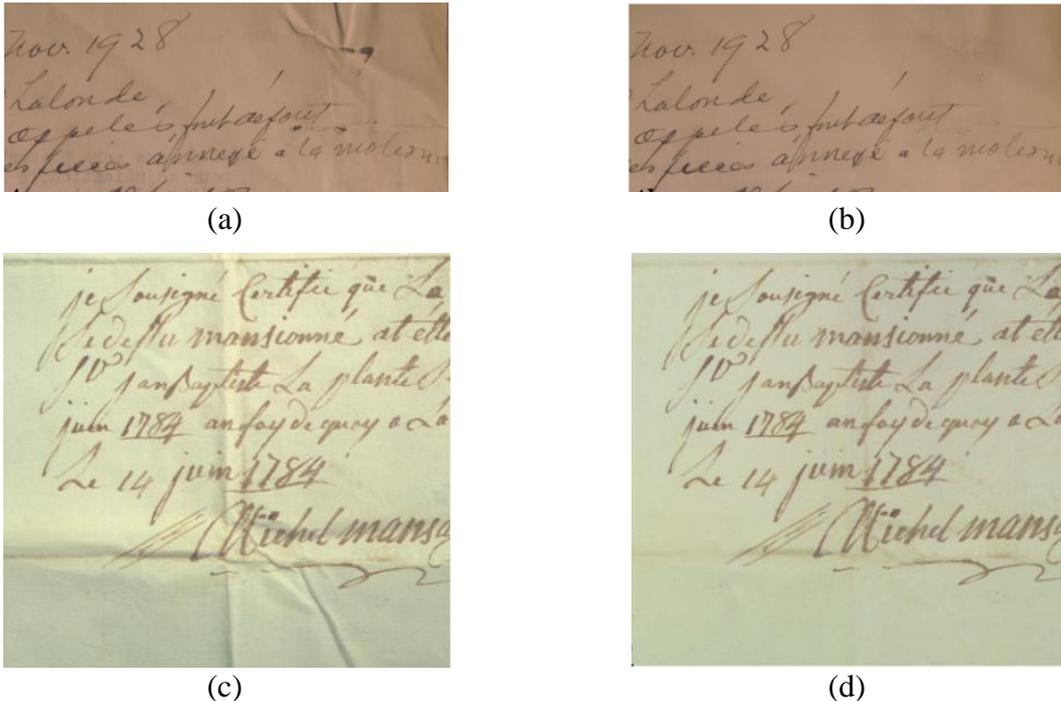
Deux problèmes majeurs seront ciblés par cette thèse :

- Restauration d'images de documents anciens dégradés pour améliorer leur lisibilité et extraction de l'information textuelle ou graphique pour offrir une meilleure accessibilité aux méthodes d'analyse de haut niveau tels que les OCR (Optical Caractère Recognition) et interprétation.
- Accélération du traitement par l'implémentation hardware des méthodes pour des systèmes embarqués.

Dans la première partie de la thèse, nous étudierons à un thème très intéressant qui est le remplissage des zones perdues de l'image de document dégradé par des techniques

basées sur le contexte et sur l'information à priori du document, etc. Nous parlons ici des techniques de « Inpainting » spécifiquement la binarisation des images de documents anciens. Plusieurs techniques ont été proposées dans la littérature.

L'objectif ultime est donc de définir des modèles de restauration ayant la capacité de remplacer les zones dégradées de l'image de document historique par des informations significatives rendant leur lisibilité meilleure et leur apparence proche de l'originale (Fig. 3).

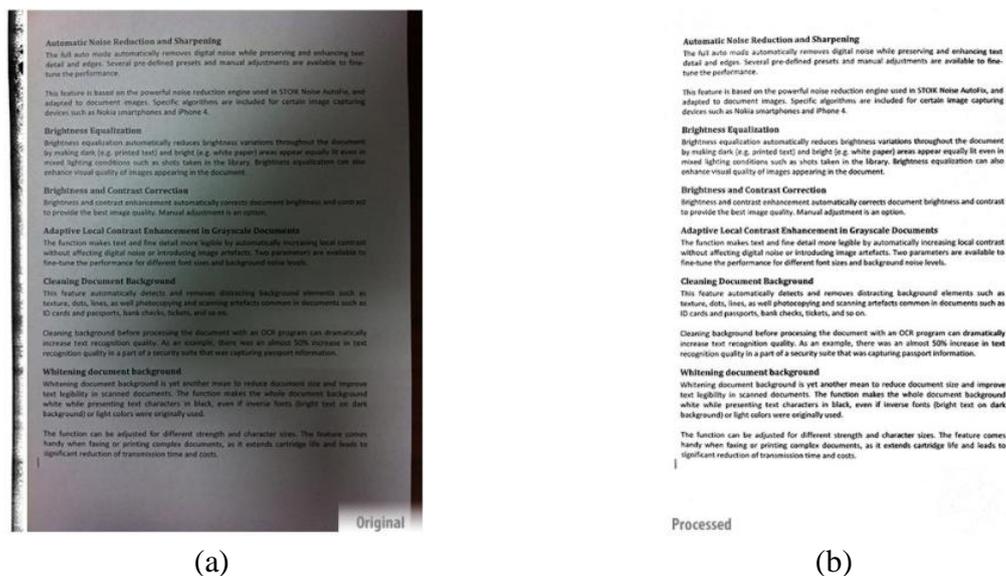


**Figure 3 :** (a) et (c) Images originales. (b) et (d) Images après traitement.

Nous nous intéressons au problème d'extraction (segmentation et binarisation) de l'information textuelle ou graphique de l'image de document avec plus de précision. Augmenter la précision de l'extraction de l'information textuelle offre une meilleure accessibilité aux méthodes d'analyse ultérieures de reconnaissance et d'interprétation.

La deuxième partie consiste à implémenter des applications qui s'exécutent en temps réel avec des serveurs informatiques via des moyens de communication tel que l'internet. À titre d'exemple, un utilisateur peut capturer une image d'un document ou d'une carte géographique, par son cellulaire, une partie du traitement peut s'exécuter sur le cellulaire par des circuits ou applications spécialisés embarqués (par exemple : un simple rehaussement, une segmentation ou extraction d'une partie du texte), puis l'émission de cette information à un serveur dédié à compléter l'analyse et à prendre des décisions. L'acquisition par téléphone

cellulaire ou iPad par exemple, offre de nouvelles perspectives de recherche telles que la calibration géométrique de l'image et la normalisation de l'intensité de l'image (Figure.4).



**Figure 4 :** (a) Image scannée (b) Image binarisée après normalisation de l'intensité

La thèse est organisée en quatre chapitres.

Le premier chapitre donne un aperçu sur les processus de digitalisation et les défis rencontrés dans le domaine d'analyse de documents dans le contexte des bibliothèques numériques, qui regroupe les conférences, entretiens et présentations d'enseignants, document historique etc.

Le deuxième chapitre est dédié à la présentation détaillée de l'état de l'art ainsi que les formulations mathématiques des différents algorithmes de binarisation les plus connus (globaux, locaux et hybrides), notamment des brefs aperçus d'implémentation de quelques méthodes.

Dans le troisième chapitre nous présentons l'approche proposée dans le contexte de la binarisation. Nous allons donner les descriptions détaillées des différentes mesures d'évaluation. Des résultats de comparaison sur des références de bases de données réputés mondialement (DIBCO's BICKLEY DIALY) seront exposés.

Le quatrième chapitre traite la mise en place d'un système embarquant un exemple d'application dédié à la binarisation d'images sur une plateforme reconfigurable FPGA où nous avons adopté une approche de conception mixte.

Des perspectives sur les futurs travaux seront données en conclusion.

# CHAPITRE 1

NUMERISATION DES DOCUMENTS, ANALYSE DES CONTENUS DE  
DOCUMENT ET BIBLIOTHEQUES ELECTRONIQUES

## 1.1. Introduction

Lorsqu'un utilisateur final accède à des images, à des fichiers PDF, des audio, des vidéo ou à tout autre document multimédia via internet, cela signifie que l'objectif principal de tout l'effort de numérisation est atteint. Du début à la fin, la production de documents multimédias numériques est un processus de fabrication et de livraison qui doit être assimilé à une chaîne de montage. À ce jour, le processus de numérisation dans de nombreuses bibliothèques s'est concentré sur les entrées (numérisation) et la réalisation via internet et sur le logiciel de gestion de contenu. Ce qui n'a pas fait l'objet d'une attention suffisante, c'est l'automatisation, le suivi et la gestion de l'ensemble du flux de travail, en mettant l'accent sur ce qui se passe entre analyse et livraison.

Le système de numérisation, que l'on peut appeler laboratoire de numérisation DL (Digital Library), nécessite des opérations de digitalisation efficaces et hautement intégrées comprenant des processus matériels, logiciels et de gestion du flux de travail pouvant exploiter pleinement les capacités uniques du laboratoire numérique. Plusieurs expériences avec de grands projets de numérisation nous ont montré la nécessité d'un système hautement intégré qui gère l'ensemble du processus de numérisation avec ses phases, les utilisateurs du système, la gestion des exceptions, l'historique de suivi des actions, le déplacement des fichiers, l'archivage et l'intégration au LIS (Library Integrated System) et au référentiel numérique de la bibliothèque.

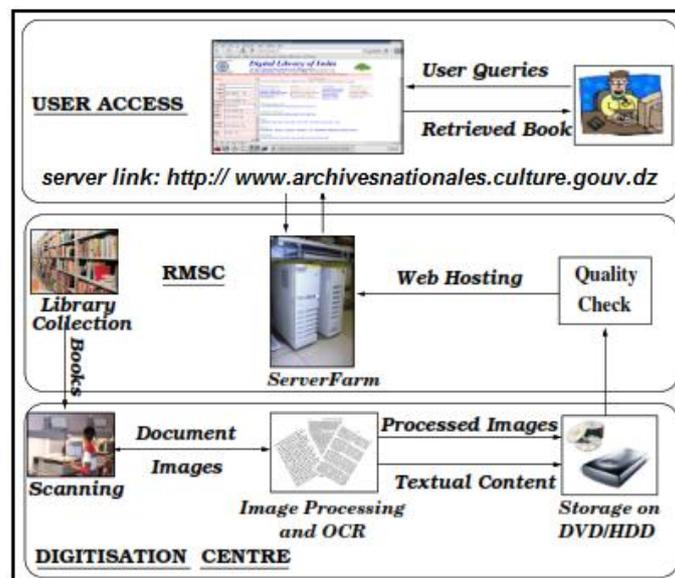
Un tel système devrait être suffisamment souple pour gérer simultanément plusieurs projets avec une diversité de matériels couvrant des livres, des revues, des journaux, des manuscrits, des matériaux non reliés, des supports audio, vidéo et des diapositives. Le système devrait également alimenter de manière transparente le contenu dans le référentiel numérique des bibliothèques afin de garantir la conservation du contenu pendant de nombreuses années.

## 1.2. Le processus de digitalisation

Comme tout flux de travail, celui de numérisation est une description d'un processus technique avec suffisamment de détails pour qu'il puisse être exécuté directement par un système de gestion de flux de travail [1]. Un processus de numérisation est composé d'une séquence de phases. Les phases sont entreprises par les ressources du laboratoire numérique, telles que les opérateurs et les appareils de laboratoire numérique (scanners ou serveurs de

codage). Les données de production sont des objets d'information. Par exemple, les fichiers TIFF, PDF, DJVU, ZIP ou tous les fichiers numériques, dont l'existence ne dépend pas de la gestion du flux de travail. Plusieurs outils sont utilisés pour exécuter des activités élémentaires au cours d'une phase de numérisation, tels que le traitement d'image et les combinaisons de reconnaissance optique des caractères OCR (Optical Character Recognition).

Le processus de numérisation est réparti en différentes étapes logiques au sein d'un même centre de numérisation. Comme un environnement distribué convient au projet, les différentes phases de travail peuvent être réparties sur différents emplacements. Un flux de processus visant, la réalisation d'une configuration hautement automatisée et la création d'un environnement distribué, la figure suivante décrit les étapes du processus de traitement.



**Figure 1.1:** Vue générale du processus de la numérisation de la bibliothèque

Chacune des composantes du flux de travail est expliquée ci-dessous. Le processus de numérisation commence par la procuration du document. Un bibliothécaire crée par exemple les métadonnées des livres, ce qui permet de vérifier la possibilité de les dupliquer par rapport à des livres numérisés auparavant.

### 1.2.1. Acquisition et scanning

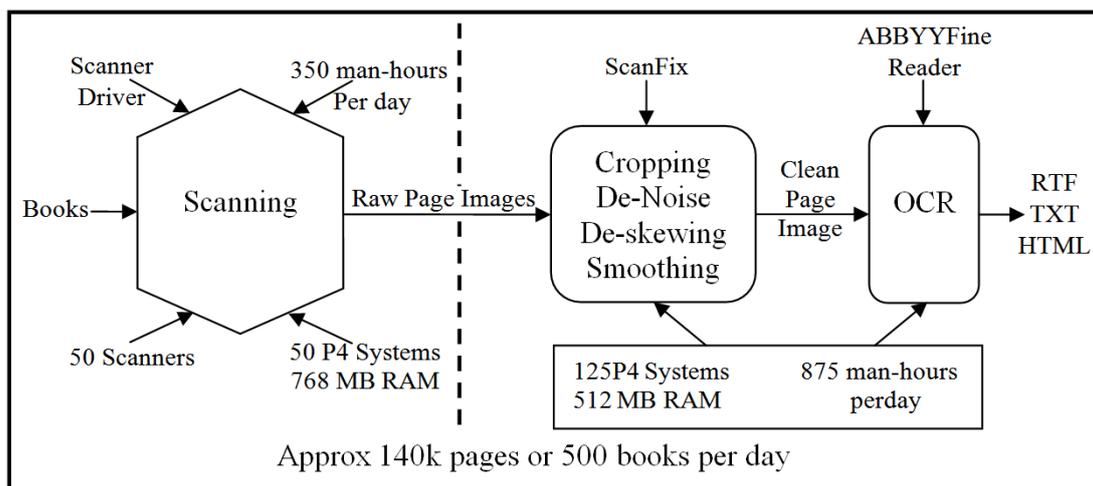
Le processus de numérisation commence par la numérisation des documents. Les scanners sont des dispositifs de capture à grande vitesse qui sont utilisés pour convertir des documents en images numériques correspondantes. Les résolutions de numérisation (fréquence d'échantillonnage spatial) des documents textuels se situent généralement entre 300 et 400 dpi (dot per inch). La fréquence de numérisation et la capacité de stockage sur disque augmentent

la résolution à 600 dpi. Pour quelles utilisations en aval ces directives approximatives sont-elles suffisantes? Les possibilités de recherche sont nombreuses, de ce type général, un régime de numérisation particulier pour les livres et les documents imprimés modernes par exemple : 300 dpi couleur 24 bits fournit de manière fiable des images (au moins du texte) garantissant la meilleure précision de reconnaissance possible à l'avenir, à mesure que les méthodes de traitement des images s'améliorent. Où devrions-nous en tant que communauté de recherche, aider à élaborer des normes d'analyse plus strictes.

Une activité conjointe entre AIIM (Association for Information and Image Management) et NPES (Association for Suppliers Publishing and Converting) examine une norme internationale qui est « PDF-Archive » visant à définir l'utilisation du PDF pour l'archivage et la conservation des documents.

Les cibles de test pour évaluer les scanners comprennent:

- IEEE Std 167A-1987, une cible d'essai pour télécopieur produite par photographie photographique en tons continus, avec motifs et marques pour une large gamme de mesures de précision modérée.
- AIIM Scanner Target, une cible imprimée en demi-teinte sur papier collé.
- RIT Process Ink Gamut Chart, un graphique en quadrichromie (cyan, magenta, jaune et noir) imprimé en demi-teinte permettant de déterminer la sensibilité des couleurs avec une faible précision.



**Figure 1.2:** Processus de scannage, de traitement d'image et diagramme de reconnaissance OCR

La figure 1.2 résume la numérisation et les étapes de traitement de l'image. Le côté gauche de la ligne en pointillé représente la phase de numérisation. Plusieurs scanners

fonctionnent par tranche de quelques heures par jour. Le débit maximal obtenu touche une centaine de milliers de pages par jour.

### **1.2.2. Traitement des images**

Les images brutes générées à partir de la numérisation doivent être traitées pour améliorer la qualité. L'image numérisée doit être recadrée pour supprimer l'arrière-plan. Le contenu textuel d'une page contient généralement un certain nombre d'artefacts sous la forme de points et de taches sur la page en raison du vieillissement du papier ou de la morsure, de la déchirure ou de la coupure de la page, ainsi que de caractères érodés ou incomplets. Diverses opérations de traitement d'image, tels que le recadrage, le désalignement, le débruitage, le lissage, etc., sont effectuées pour débarrasser les images de telles imperfections.

L'opérateur identifie les étapes les plus appropriées pour supprimer les artefacts autant que possible. Il définit également divers paramètres pour les procédures de traitement d'image. Des outils sont nécessaires pour orienter les pages de manière à ce que le texte soit au premier plan, pour le redressement, la correction de la lecture et la réduction du *dwarping*, ainsi que pour la suppression des artefacts sombres dans les bords des images. La prise en charge logicielle des fonctions de bureau telles que la numérotation et le classement des pages, ainsi que la collecte de métadonnées, sont également essentielles au maintien d'un débit élevé.

Les images traitées sont stockées dans le format de fichier TIFF à l'aide d'algorithmes de compression tel que le CCITT 4 Fax. Ce système s'est révélé très efficace. Une image de résolution  $4300 \times 2900$ , ne contenant que quelques mots, a une taille de fichier de 1 Ko et un document de la même résolution avec le texte intégral occupe 75 Ko, alors qu'il faudrait environ 150 Ko au format PNG. Une page avec des images stockées au format binaire serait de 120 Ko au format TIFF, contre environ 230 Ko au format PNG.

### **1.2.3. Restauration**

La restauration d'images de document peut faciliter la lecture rapide et sans effort, l'OCR pour le contenu textuel, le DIA (Digital Image Analysis) pour une expérience utilisateur améliorée (conservation du format, par exemple) et la caractérisation du document (âge, source, etc.). À ces fins, des méthodes ont été développées pour l'amélioration du contraste et de la netteté, la rectification (y compris la correction de l'asymétrie et du cisaillement), la super-résolution et la reconstruction de forme [2], mais il semble en avoir assez de problèmes.

### 1.2.4. Reconnaissance et reconstruction

Après le recadrage et le nettoyage d'une image de page, une reconnaissance optique est utilisée pour en extraire le texte (OCR disponible avec ABBYY FineReader). Le texte produit par l'OCR est stocké aux formats RTF, TXT et HTML, dans des dossiers distincts. Le texte extrait est utilisé pour indexer les pages et les documents, pour permettre la recherche et la récupération pour les utilisateurs.

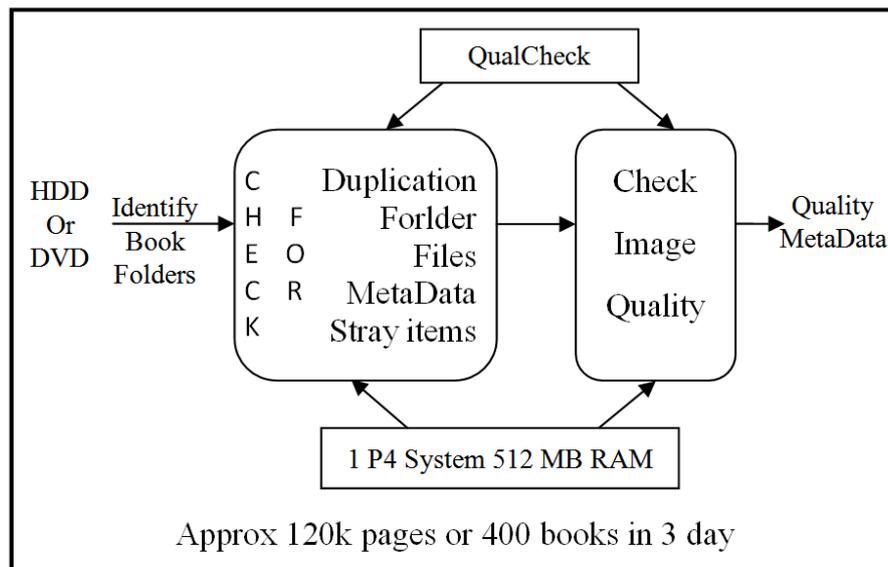
### 1.2.5. Contrôle de la qualité

Pour s'assurer que le system de numérisation maintient une qualité correcte, la qualité du document converti numériquement est contrôlée par le RMSC (Rochester Museum & Science Center) [3]. Un ensemble de normes fixes a été adopté comme indiqué par exemple dans le tableau 1. La qualité des documents soumis est vérifiée pour chacun de ces aspects et les documents contenant plus de 80% du contenu ayant des erreurs sont re-scannés et/ou re-traités.

**Tableau 1.1** : Paramètres normalisés pour le format image TIFF

Paramètre	Spécification
Dimensions	Même ligne x colonne
dpi	600 ou plus
Compression	CCITT 4 Facsimile
Marge	300 pixels de toutes les faces
Inclinaison	<2°
Page vide	Identifier et notifier

Un outil automatique, est utilisé pour scruter de manière récursive des documents sur un disque dur ou DVD et rechercher automatiquement tout le contenu requis. Il vérifie ensuite chacun des paramètres définis pour les images et génère un rapport XML pour chacune des erreurs voir figure 1.3. Les erreurs généralement observées sont les fichiers manquants dans un ou plusieurs des dossiers, les marges non uniformes dans les images traitées et l'utilisation d'un algorithme de compression différent de celui spécifié.



**Figure 1.3 :** Contrôle de la qualité et de la performance

### 1.2.6. Hébergement web

Mettre en ligne des documents numérisés consiste à mettre en place sur le web un système de recherche documentaire et une interface proposant des fonctionnalités et des services permettant la consultation de documents et leur réappropriation par le public. Ce type de projet entre dans les missions de valorisation des bibliothèques en plaçant les collections au sein d'un réseau.

Les documents numériques sont hébergés sur les serveurs de Téraoctets, qui sont mis en cluster en tant que champs de données. Un opérateur exécute le processus de métadonnées post-analyse en créant des métadonnées structurales [4], qui permettent de naviguer facilement dans les documents. Une copie du document est créée et stockée en tant que sauvegarde en cas de panne du disque dur. Ces livres sont hébergés sur le serveur et dupliqués pour les serveurs locaux.

## 1.3. Analyse des images document / cas des librairies historiques électroniques

### 1.3.1. Acquisition d'images

La capture d'images à partir d'artefacts historiques nécessite une manipulation spéciale pour contrer les défauts du vieillissement des documents et les contraintes physiques de numérisation. Une approche orientée DIA est suggérée pour augmenter efficacement la résolution de numérisation, ainsi que pour assurer la préservation des documents de numérisation et le contrôle de la qualité [5], [6].

Dans plusieurs recherches, le rapport signal sur bruit SNR (Signal to Noise Ratio) et d'autres mesures montrent les pertes de résolution par rapport aux données dans les formats de compression d'image où le stockage est en vrai couleur ou à 256 niveaux de gris. Les conservateurs devraient être informés des besoins de la technologie de la librairie digitale DL et des inconvénients des pertes observés dans les formats de fichiers tels que JPEG. En outre, les lumières non UV et la pagination automatique sont utilisées pour sauvegarder les originaux pendant le balayage, les erreurs sont contrées en utilisant l'inclinaison, l'éclairage, la correction de courbure pour les fixations de livres et la couleur de réduction de profondeur pour les vieux documents.

La reconstruction des caractères est suggérée pour restaurer les écritures cassées dans les documents anciens. La numérisation continue est suivie d'un recadrage automatique, ce qui constitue une procédure rapide pour générer des images à partir de microfilm [7]. La transformée de Fourier-Mellin est utilisée pour corriger les erreurs de rotation/cisaillement, d'échelle et de translation [3]. Les opérations morphologiques d'analyse de la luminosité et de la saturation HLS (Hue, Lightness, and Saturation) dans les données d'image et l'analyse en composantes connexes sont utilisées pour éliminer les zones de papier reconstruit [8].

### **1.3.2. Analyse de la mise en page et extraction des métadonnées.**

L'analyse de la mise en page et l'extraction métadonnées est une étape cruciale dans la création d'une base d'information pour les librairies numériques DL historiques. Alors que les chercheurs gagnent du terrain sur la reconnaissance complète du document texte historique, des systèmes pratiques ont été construits en utilisant uniquement l'étape d'analyse de la mise en page [5], [6], [9].

La disponibilité des images permet de fournir une récupération d'image basée sur le contenu, en utilisant même des caractéristiques structurelles telles que la couleur et la mise en page. Marinai et coll. [10] ont créé une structure d'arborescence MXY (Modifie X-Y) pendant la segmentation du document, puis ont utilisé la similarité de présentation en tant que caractéristique pour vérifier des documents par exemple.

Une DL historique doit compléter le contenu avec des métadonnées décrivant les caractéristiques textuelles (date, auteur, lieu) et des informations géométriques (paragraphe, zones d'image,). Couasnon et coll. utilisent un système automatisé basé sur le web pour la collecte, recueillant des annotations des archives françaises [11]. Le système associe une

analyse automatique de la présentation à une annotation assistée par l'homme dans une interface web.

La transcription des documents historiques mappe le texte ASCII aux mots dans l'image du document. Cette solution vise à contourner le manque de reconnaissance optique de caractères OCR parfaite pour les styles d'écriture anciens [5], [12], [13].

### **1.3.3. Reconnaissance de l'écriture manuscrite pour DL**

Bien que les produits commerciaux soient disponibles pour le texte dactylographique, la reconnaissance d'écriture manuscrite n'a réussi que dans des domaines spécialisés. Les systèmes de reconnaissance de modèles de caractères basés sur HMM (Hidden Markov Models) sont utilisés dans la reconnaissance d'adresses postales à partir d'images postales [14], [15]. Ce système s'appuie sur des informations de contexte liées aux adresses.

Pour la création de transcriptions à partir de documents historiques, les systèmes de cartographie utilisent la reconnaissance de l'écriture manuscrite. Les moteurs OCR utilisés dans ces applications ne peuvent pas répondre en temps réel aux demandes de reconnaissance. Les systèmes de classification automatique [16] utilisent plusieurs étapes de binarisation suivie par l'identification des documents à l'aide des caractéristiques des caractères. Pour les scripts Hanja (coréen), les techniques OCR et UI (User Interface) [15] intègrent une normalisation non linéaire, des formes de direction de contour et des modules de reconnaissance basés sur Mahalanobis afin de générer des transcriptions pour les documents Hanja.

Un module de reconnaissance basé sur HMM pour les grands lexiques est examiné pour indexer les documents historiques [5]. Le système utilise le partage de sous-chaîne, où une arborescence de préfixe est créée du lexique. Les entrées qui partagent le même préfixe partagent également son calcul sans invoquer le module de reconnaissance. Les contraintes de durée sur les états des caractères, l'élagage à choix et le décodage parallèle permettent une accélération de 7 fois.

Zhang et al. [13] combinent la reconnaissance de modèles de mots et le mappage de transcriptions pour créer des bases de données manuscrites. Lavrenko et al. [17] suggèrent une technique de reconnaissance holistique dans laquelle des images de mots normalisés sont utilisées en tant qu'entrées dans un HMM. Les caractéristiques scalaires du profil sont extraites des images et un document historique entier est modélisé comme un HMM, avec des mots constituant la séquence d'état.

Pour un document écrit par un auteur donné, les probabilités de transition de l'état sont obtenues en calculant la moyenne des probabilités de mots à deux grammes recueillies à partir de textes contemporains et écrits précédemment transcrits de l'auteur cible. Pour un document écrit par un auteur donné, les probabilités de transition d'état sont obtenues en faisant la moyenne des probabilités de mots à deux grammes recueillies à partir de textes contemporains et des écrits précédemment transcrits de l'auteur cible.

### **1.3.4. Bibliothèques électroniques multi-langues**

D'excellents progrès ont été développés dans les DL latine, mais les recherches dans d'autres scripts tels que l'anglais, l'arabe, le bengali, le devanagari, le telugu, le chinois et le coréen n'ont pas reçu le même intérêt. L'accès numérique aux contenus de documents dans ces scripts est difficile en termes de conception d'interface utilisateur UI, d'analyse de présentation et de reconnaissance OCR.

Un système DL multilingue devrait prendre en charge le stockage, la saisie et l'affichage simultanés des données dans de nombreux scripts. Beaucoup de scripts non latins ont un jeu de caractères compliqués et nécessitent un système de codage séparé [18]. L'affichage et la saisie de ces langues nécessitent de nouvelles polices [19], [20] et de nouveaux systèmes de saisie de caractères. De plus, pour assurer la compatibilité et l'indépendance des données des plates-formes, une DL ne doit pas recourir à des solutions personnalisées sans examiner en profondeur les normes existantes.

En termes de codage de caractères, le consortium Unicode vise à fournir un schéma de codage fiable pour tous les scripts du monde [18] et il est accepté comme norme de système par de nombreux chercheurs et fabricants de logiciels [11], [12], [19], [20]. Bien que d'autres schémas aient été suggérés [21], ils n'ont pas la compatibilité et la globalité d'acceptation du système Unicode. Sur le plan du stockage, le XML est en train de devenir un outil polyvalent, il est préféré pour les projets DL [12], [22], [23].

En ce qui concerne les techniques de saisie et d'affichage, des schémas multicouches pour des scripts phonétiques [24] sont suggérés pour les systèmes de saisie basés sur un stylet ou un clavier par exemple les PDA's (Personal Digital Assistant). Les systèmes de cartographie au clavier (INSCRIPT pour les scripts Indic) mappent la touche d'un clavier QWERTY standard sur les caractères d'un script cible [21].

Ce système de clavier est fonctionnel, mais a une courbe d'apprentissage difficile. De plus, chaque clavier doit être étiqueté physiquement avant qu'un utilisateur puisse associer les

touches aux caractères appropriés. Le logiciel TrueViz [11] utilise un clavier graphique pour la saisie de script en russe. Kompalli et al. [12] proposent un système de translittération, dans lequel les caractères devanagari sont entrés par des chaînes équivalentes phonétiques en anglais.

### **1.3.5. Analyse de forme multilingue**

La variation dans l'ordre d'écriture des scripts et la présence de constructions spécifiques à une langue, nécessitent différentes approches pour l'analyse de la mise en page. Par exemple, les lacunes ne peuvent pas être utilisées pour identifier des mots en chinois et en arabe. Les techniques d'identification des scripts varient de l'identification des scripts de mots individuels dans un document multilingue [25] à ceux permettant de déterminer les scripts de lignes [26] et de blocs de texte entiers [27], [28]. Une fois qu'un script est identifié, des algorithmes de séparation de mots et de lignes spécifiques peuvent être utilisés [29].

### **1.3.6. Analyse du contenu**

L'analyse et la reconnaissance du contenu des images de document exigent bien entendu la gamme complète des réalisations de DIA R&D (research and development) analyse de la mise en page, séparation texte/non-texte, séparation dactylographiée/manuscrite, reconnaissance de texte, étiquetage des blocs de texte par fonction, indexation automatique lien, tableau et graphiques, etc. La plupart des ouvrages de la DIA sont consacrés à ces sujets.

Cependant, il convient de noter que les images trouvées dans les DL, (représentent de nombreux pays, cultures et périodes historiques), ont tendance à poser des défis particulièrement graves aux méthodes DIA actuelles, et en particulier à l'architecture des systèmes DIA, qui ne sont pas robustes face au texte multilingue et aux systèmes et scripts non occidentaux, polices obsolètes, anciennes mises en page et qualité d'image faible ou variable. La grande variété d'images de documents qui sont rapidement mises en ligne menace de submerger les capacités des systèmes DIA de pointe; ce fait, pris isolément, suggère qu'une recherche fructueuse pour la R&D de la DIA consiste à rechercher des outils capables d'effectuer de manière fiable des tâches spécifiques, peut-être étroitement définies, sur toute la gamme de documents naturels.

Ceux-ci pourraient inclure:

- Une image qui contient un texte imprimé ou manuscrit.
- Une image qui contient un long passage de texte (par exemple, 50 de mots ou plus).
- L'isolation toutes les régions textuelles, en les séparant du non-texte et de l'arrière-plan.

- L'identification/Segmentation du manuscrit du texte imprimé par la machine.
- Identification du script (système d'écriture) et la langue des régions de texte.

Cela pourrait être appelé une stratégie DIA de premier plan ou de polyvalence d'abord. La plupart de ces tâches ont, bien sûr, reçu une certaine attention dans la littérature. Ce qui est nouveau, peut-être, est l'accent mis sur la réalisation d'un certain niveau de compétence (peut-être pas toujours élevé) à travers les ordres de grandeur avec plus de types d'images de documents que ce qui avait été tenté jusqu'à présent.

#### **1.4. Conclusion**

Dans ce chapitre, nous avons donné un aperçu sur la numérisation des documents et sur certains défis auxquels sont confrontés les constructeurs de systèmes d'analyse de documents dans le contexte des bibliothèques numériques. Bien qu'il puisse sembler décourageant de découvrir que de nombreux problèmes importants restent non résolus, il n'en est fait aucun doute que le DIA et les communautés DL ont beaucoup à s'offrir mutuellement. En tant que banc d'essai pratique pour les techniques d'analyse de documents et une application réelle d'énormes données culturelles, nous prévoyons que les bibliothèques numériques fourniront un axe de travail précieux dans ce domaine.

Nous pourrions nous attendre à voir tous les documents culturels et historiques conservés pour les générations futures et disponibles dans le monde entier. Conférences, entretiens et présentations d'enseignants et de chercheurs des meilleures écoles du monde seraient à la disposition des étudiants de chaque école, ce qui donnerait une impulsion importante à l'éducation et aux activités de recherche.

# *Bibliographie*

- 
- [1] N. Russell, A. H. M. ter Hofstede, D. Edmond, et W. M. P. van der Aalst, « Workflow Data Patterns: Identification, Representation and Tool Support », in *Conceptual Modeling – ER 2005*, vol. 3716, L. Delcambre, C. Kop, H. C. Mayr, J. Mylopoulos, et O. Pastor, Éd. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 353-368.
- [2] R. P. Loce et E. R. Dougherty, *Enhancement and restoration of digital documents: statistical design of nonlinear algorithms*. Bellingham, Wash., USA: SPIE Optical Engineering Press, 1997.
- [3] L. A. D. Hutchison et W. A. Barrett, « Fast registration of tabular document images using the Fourier-Mellin transform », in *First International Workshop on Document Image Analysis for Libraries, 2004. Proceedings.*, Palo Alto, CA, USA, 2004, pp. 253-267.
- [4] H. Bunke et A. L. Spitz, Éd., *Document Analysis Systems VII*, vol. 3872. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006.
- [5] H. S. Baird, « Difficult and urgent open problems in document image analysis for libraries », in *First International Workshop on Document Image Analysis for Libraries, 2004. Proceedings.*, Palo Alto, CA, USA, 2004, pp. 25-32.
- [6] F. Le Bourgeois, E. Trinh, B. Allier, V. Eglin, et H. Emptoz, « Document images analysis solutions for digital libraries », in *First International Workshop on Document Image Analysis for Libraries, 2004. Proceedings.*, Palo Alto, CA, USA, 2004, pp. 2-24.
- [7] W. Barrett, L. Hutchison, D. Quass, H. Nielson, et D. Kennard, « Digital mountain: from granite archive to global access », in *First International Workshop on Document Image Analysis for Libraries, 2004. Proceedings.*, Palo Alto, CA, USA, 2004, pp. 104-116.
- [8] A. Antonacopoulos et D. Karatzas, « Document image analysis for world war II personal records », in *First International Workshop on Document Image Analysis for Libraries, 2004. Proceedings.*, Palo Alto, CA, USA, 2004, pp. 336-341.
- [9] G. Harit, S. Chaudhury, et H. Ghosh, « Managing document images in a digital library: an ontology guided approach », in *First International Workshop on Document Image Analysis for Libraries, 2004. Proceedings.*, Palo Alto, CA, USA, 2004, pp. 64-92.
- [10] S. Marinai, E. Marino, F. Cesarini, et G. Soda, « A general system for the retrieval of document images from digital libraries », in *First International Workshop on Document Image Analysis for Libraries, 2004. Proceedings.*, Palo Alto, CA, USA, 2004, pp. 150-173.
- [11] T. Kanungo, C. H. Lee, J. Czorapinski, et I. Bella, « TRUEVIZ: a groundtruth/metadata editing and visualizing toolkit for OCR », présenté à Photonics West 2001 - Electronic Imaging, San Jose, CA, 2000, pp. 1-12.
- [12] S. Setlur, S. Kompalli, V. Ramanaprasad, et V. Govindaraju, « Creation of data resources and design of an evaluation test bed for Devanagari script recognition », in *Proceedings. Seventeenth Workshop on Parallel and Distributed Simulation*, Hyderabad, India, 2003, pp. 55-61.
- [13] Bin Zhang, C. Tomai, S. Srihari, et V. Govindaraju, « Construction of handwriting databases using transcript-based mapping », in *First International Workshop on Document Image Analysis for Libraries, 2004. Proceedings.*, Palo Alto, CA, USA, 2004, pp. 288-298.
-

- 
- [14] S. Setlur, A. Lawson, V. Govindaraju, et S. Srihari, « Large scale address recognition systems Truthing, testing, tools, and other evaluation issues », *Int. J. Doc. Anal. Recognit.*, vol. 4, n° 3, pp. 154-169, mars 2002.
- [15] S. N. Srihari et E. J. Kuebert, « Integration of hand-written address interpretation technology into the United States Postal Service Remote Computer Reader system », in *Proceedings of the Fourth International Conference on Document Analysis and Recognition*, Ulm, Germany, 1997, vol. 2, pp. 892-896.
- [16] I. B. Yosef, K. Kedem, I. Dinstein, M. Beit-Arie, et E. Engel, « Classification of hebrew calligraphic handwriting styles: preliminary results », in *First International Workshop on Document Image Analysis for Libraries, 2004. Proceedings.*, Palo Alto, CA, USA, 2004, pp. 299-305.
- [17] V. Lavrenko, T. M. Rath, et R. Manmatha, « Holistic word recognition for handwritten historical documents », in *First International Workshop on Document Image Analysis for Libraries, 2004. Proceedings.*, Palo Alto, CA, USA, 2004, pp. 278-287.
- [18] J. Aliprand et Unicode Consortium, Éd., *The Unicode standard*, Version 4.0. Boston: Addison-Wesley, 2003.
- [19] B. Couasnon, J. Camillerapp, et I. Leplumey, « Making handwritten archives documents accessible to public with a generic system of document image analysis », in *First International Workshop on Document Image Analysis for Libraries, 2004. Proceedings.*, Palo Alto, CA, USA, 2004, pp 270-277.
- [20] J. Mike, C. Peter, et H. Nicholas, « Windows glyph processing for OpenType fonts, part 1 », *Windows glyph processing for OpenType fonts, part 1*, 20-oct-2017. [En ligne]. Disponible sur: <https://docs.microsoft.com/fr-fr/typography/opentype/processing-part1>.
- [21] D. P. Madalli et D. Patel, « Issues in Indian languages computing in particular reference to search and retrieval in Telugu language », *Libr. Hi Tech*, vol. 27, n° 3, pp. 450-459, sept. 2009.
- [22] R. B. Allen et J. Schalow, « Metadata and data structures for the historical newspaper digital library », in *Proceedings of the eighth international conference on Information and knowledge management - CIKM '99*, Kansas City, Missouri, United States, 1999, pp. 147-153.
- [23] S. J. Simske et M. Sturgill, « A ground-truthing engine for proofsetting, publishing, re-purposing and quality assurance », in *Proceedings of the 2003 ACM symposium on Document engineering - DocEng '03*, Grenoble, France, 2003, pp. 150.
- [24] S. Shanbhag, D. Rao, et R. K. Joshi, « An intelligent multi-layered input scheme for phonetic scripts », in *Proceedings of the 2nd international symposium on Smart graphics - SMARTGRAPH '02*, Hawthorne, New York, 2002, pp. 35-38.
- [25] A. M. Namboodiri et A. K. Jain, « Online script recognition », in *Object recognition supported by user interaction for service robots*, Quebec City, Que., Canada, 2002, vol. 3, pp. 736-739.
- [26] U. Pal et B. B. Chaudhuri, « Script line separation from Indian multi-script documents », in *Proceedings of the Fifth International Conference on Document Analysis and Recognition. ICDAR '99 (Cat. No.PR00318)*, Bangalore, India, 1999, pp. 406-409.
- [27] J. Hochberg, L. Kerns, P. Kelly, et T. Thomas, « Automatic script identification from images using cluster-based templates », in *Proceedings of 3rd International Conference*
-

- on Document Analysis and Recognition*, Montreal, Que., Canada, 1995, vol. 1, pp. 378-381.
- [28] T. N. Tan, «Rotation invariant texture features and their use in automatic script identification », *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, n° 7, pp. 751-756, juill. 1998.
- [29] V. Govindaraju, S. Khedekar, S. Kompalli, F. Farooq, S. Setlur, et R. Vemulapati, «Tools for enabling digital access to multi-lingual indic documents », in *First International Workshop on Document Image Analysis for Libraries, 2004. Proceedings.*, Palo Alto, CA, USA, 2004, pp. 122-133.

## CHAPITRE 2

APERÇU SUR LES METHODES DE BINARISATION, IMPLEMENTATION ET  
COMPARAISON

## 2.1. Introduction

Notre problématique de seuillage s'inscrit dans le contexte général de traitement de documents anciens. Dans cette partie de thèse, nous présenterons les différents types de binarisation cités dans la littérature. Nous nous intéressons dans cette étude à la binarisation des documents en niveaux de gris, puisque la plupart des documents en couleurs peuvent être convertis fidèlement en niveaux de gris.

## 2.2. Binarisation

La binarisation est une opération de prétraitement d'images qui sert à convertir une image en couleur ou niveau de gris en une image binaire, où l'avant plan est en couleur noir et l'arrière-plan en blanc (ou l'inverse).

La binarisation peut se faire par plusieurs techniques, tels que le seuillage, la classification, etc. Elle représente donc un outil très utilisé dans la segmentation d'images pour extraire des objets de leurs fonds en fonction d'un certain seuil.

Techniquement parlant, soit une image  $I$  de taille ( $M$  (hauteur) x  $N$  (largeur)),  $I(x, y)$  la valeur représentant le niveau de gris du pixel au point coordonnées  $(x, y)$ . Étant donné un seuil  $T$  :

- Nous définissons les pixels de l'objet, ceux ayant le niveau de gris inférieur à  $T$ , et les pixels de fond, ceux ayant le niveau de gris supérieur à  $T$ .
- Donc, l'image binarisée  $g$  est déterminée comme suit :

$$g(x, y) = \begin{cases} 1 & \text{if } f(x, y) > T \\ 0 & \text{if } f(x, y) \leq T \end{cases} \quad (2.1)$$

## 2.3. Les méthodes de binarisation

Les méthodes de binarisation ont été classées selon les informations qu'elles exploitent, en six catégories [1] :

- méthodes basées sur la forme de l'histogramme de l'image.
- méthodes basées sur le clustering.
- méthodes basées sur l'entropie.
- méthodes basées sur les attributs d'objet.
- méthodes spatiales.
- méthodes locales.

Selon plusieurs travaux de recherche [2], les techniques de binarisation par seuillage d'images en niveaux de gris peuvent être classées en trois catégories :

- Seuillage global [3]:

Consistant à partitionner l'image en deux classes grâce à un seuil optimal qui sera calculé à partir d'une mesure globale sur toute l'image. L'histogramme est une mesure utilisée le plus souvent dans les méthodes de seuillage. Dans ce cas, le seuil attendu est celui qui distingue le plus possible les deux classes : fond et objet.

- Seuillage local [4]:

Le seuil est calculé localement pour chaque pixel centré dans une fenêtre définie préalablement, se basant sur l'information se trouvant dans son voisinage.

- Seuillage hybride [5] :

Consiste à combiner des informations globales et locales pour attribuer les pixels à l'une des deux classes.

### 2.3.1. Méthodes de binarisation globale

Les méthodes de seuillage global reposent sur l'exploitation de l'histogramme de toute l'image  $I$ . L'histogramme est une courbe monodimensionnelle qui caractérise la distribution des niveaux de gris, il est décrit par la fonction  $h(k)$ :

$$h(k) = \sum_{i=1}^N \sum_{j=1}^M (I(i, j) == k) \quad (2.2)$$

Où :

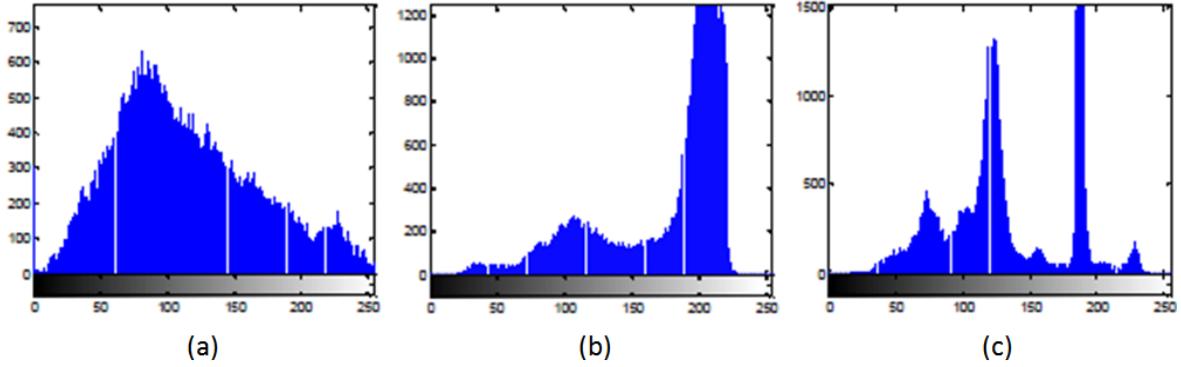
$h(k)$  Étant le nombre de pixels ayant le niveau de gris  $k \in [0:255]$

$N \times M$  le nombre total de pixels dans l'image.

A partir de l'allure de l'histogramme, on peut déduire le nombre de classes ainsi que la position relative des seuils. L'histogramme est uni-modal, s'il est formé d'un seul pic représentant les pixels de l'objet ou ceux du fond. Il est bimodal s'il est caractérisé par deux modes séparés par une vallée, il indique l'existence d'un objet sur un fond.

L'histogramme multimodal comporte plus de deux modes séparés par des vallées, il indique la présence de plusieurs classes dans l'image (chaque mode correspond à une classe). Les seuils doivent être localisés dans les vallées.

En pratique, il est rare de trouver un histogramme qui présente des modes bien distincts. En effet, la plupart des images présentent des histogrammes bruités caractérisés par des modes non discernables.



**Figure 2.1** : Allures idéales d'un histogramme. (a): Unimodal, (b): Bimodal, (c) Multimodal

L'application d'un seuil particulier  $t$  est équivalent à classer chaque pixel comme étant soit une partie de l'arrière-plan (fond) ou au premier plan (objet). Ainsi, l'ensemble de tous les pixels de l'image est divisé en deux ensembles disjoints  $C_1$  et  $C_2$  tels que:

$$C_1 = \{0,1,\dots,T\} \quad \text{et} \quad C_2 = \{T+1,\dots,255\} \quad (2.3)$$

Avant de donner le principe des différents algorithmes de binarisation, passons en revue quelques faits élémentaires au sujet des informations qui peuvent être dérivées de l'histogramme de l'image par les équations suivantes :

- La probabilité des niveaux de gris telle que :

$$p(i) = \frac{h(i)}{N} \quad (2.4)$$

- La probabilité moyenne de l'image telle que :

$$P(i) = \sum_{i=0}^k p(i) \quad (2.5)$$

- L'écart type de l'image tel que :

$$\mu(i) = \frac{\sum_{i=0}^k p(i) \times i}{P(i)} \quad (2.6)$$

Si on fait un seuillage à un niveau  $T$ , le nombre de pixels assignés à chaque classe  $C_1$  et  $C_2$  (background et foreground) est:

$$n_1(T) = |C_1| = \sum_{i=0}^T h(i) \quad \text{et} \quad n_2(T) = |C_2| = \sum_{i=T+1}^k h(i) \quad (2.7)$$

Donc :

$$n_1(T) + n_2(T) = |C_1 \cup C_2| = N \times M \quad (2.8)$$

La probabilité moyenne ainsi que l'écart type correspondant à chaque classe  $C_1$  ou  $C_2$  peuvent être calculés à partir de l'histogramme pour chaque seuil  $T$  où :

$\mu_1, \mu_2$  désignent respectivement l'écart type de chaque classe objet et fond de l'image tel que :

$$\mu_1 = \frac{\sum_{k=0}^T p(k)k}{P_1} \quad (2.9)$$

$$\mu_2 = \frac{\sum_{k=T+1}^{255} p(k)k}{P_2} \quad (2.10)$$

$P_1, P_2$  représentent respectivement les probabilités moyennes de chaque classe objet et fond telles que :

$$P_1 = \sum_{k=0}^T p(k) \quad (2.11)$$

$$P_2 = \sum_{k=T+1}^{255} p(k) \quad (2.12)$$

$p$  représente la probabilité des niveaux de gris telles que :

$$p(i) = \frac{h(i)}{N} \quad (2.13)$$

### ➤ Méthode d'Otsu [6]

Elle est considérée comme la méthode de référence dans le domaine du seuillage d'histogrammes. Dans cette méthode, l'opération de seuillage est vue comme une séparation (ou partitionnement) des pixels d'une image en deux classes « fond et objet » à partir d'un seuil  $T$  qui minimise la variance intra-classe et maximise la variance interclasse. La classe « fond » regroupe tous les pixels ayant un niveau de gris supérieur au seuil  $T$  alors que la classe « objet » contient tous les pixels de niveau de gris inférieur à  $T$ .

Soient  $\delta_W^2$  la variance d'une intra-classe,  $\delta_B^2$  la variance interclasse et  $\delta_T^2$  la variance totale telle que :

$$\delta_B^2 = P_1 P_2 (\mu_2 - \mu_1)^2 \quad (2.14)$$

$$\delta_w^2 = P_1(i - \mu)^2 \quad (2.15)$$

$$\delta_T^2 = \delta_w^2 + \delta_B^2 = \sum_{i=1}^T P_1(i - \mu_1)^2 + \sum_{i=T+1}^{255} P_2(i - \mu_2)^2 \quad (2.16)$$

Dans la littérature la variance interclasse est la plus utilisée et cela pour des raisons de simplicité. Le seuil optimal  $T$  est celui qui maximise cette variance :

$$T = \arg \max_{0 \leq T \leq L-1} \delta_B^2 = \arg \max_{0 \leq T \leq L-1} (P_1 P_2 (\mu_2 - \mu_1)^2) \quad (2.17)$$

➤ **Méthode de Kittler [7]**

La méthode J.Kittler et J.Illingworth, considère que l'image a un histogramme bimodale de niveaux de gris qui représente la densité de probabilité  $p(i)$  d'un mélange de populations formées des niveaux de gris des deux classes objets et du fond, formons une distribution normale avec un écart type  $\mu$ , une variance  $\sigma$ . Cet algorithme permettra de repérer le seuil optimal qui minimise la probabilité d'erreur de classification. Donc le seuil optimal  $T$  est le niveau de gris qui minimise le critère  $J(t)$  suivant :

$$J(t) = 1 + 2[P_1(t)\log(\sigma_1(t)) + P_2(t)\log(\sigma_2(t))] - 2[P_1(t)\log(P_1(t)) + P_2(t)\log(P_2(t))] \quad (2.18)$$

➤ **Méthode d'ISODATA [8]**

Le binarisation par ISODATA consiste à trouver un seuil en séparant l'histogramme en deux classes itérativement avec la connaissance a priori des valeurs associées à chaque classe. Cette méthode commence par la division de l'intervalle de valeurs non-nulles de l'histogramme en deux parties équidistantes, et ensuite prendre  $p_1, p_2$  comme les probabilités moyennes arithmétiques des deux classes. Répéter jusqu'à la convergence, le calcul du seuil optimal  $T$  comme l'entier le plus proche de  $\frac{P_1 + P_2}{2}$ , et mettre à jour les deux probabilités moyennes  $P_1$  et  $P_2$ .

➤ **Méthode de Pun [9]**

Pun a proposé que la maximisation de l'entropie est un critère optimal pour la binarisation de l'image. Le seuil optimal  $T$  est choisi comme le niveau de gris qui maximise la somme de l'entropie de l'objet  $H_1$  et l'entropie du fond  $H_2$ .

$$S_{optimal} = \max_t (H_1(t) + H_2(t)) = \max_t \left( - \sum_{i=0}^t (p(i) \times \log p(i)) - \sum_{i=t+1}^{255} (p(i) \times \log p(i)) \right) \quad (2.19)$$

➤ **Méthode de Kapur** [10]

La méthode proposée par Kapur est une extension de la méthode de Pun qui est basée sur l'entropie qui considère le premier plan et l'arrière-plan de l'image comme deux sources de signal différentes. A la différence de cette dernière, la méthode de Kapur prend en compte la distribution de probabilité moyenne de l'objet et du fond dans la détermination de l'entropie de division. Ainsi, le seuil optimal  $T$  est lorsque la somme des entropies des deux classes atteint son maximum.

$$S_{optimal} = \max_t (H_1(t) + H_2(t)) = \max_t \left( - \sum_{i=0}^t \left( \frac{p(i)}{P_1} \times \log \left( \frac{p(i)}{P_1} \right) \right) - \sum_{i=t+1}^{255} \left( \frac{p(i)}{P_2} \times \log \left( \frac{p(i)}{P_2} \right) \right) \right) \quad (2.20)$$

➤ **Méthode de Cheng et Chen** [11]

La méthode de Cheng et Chen est basée sur le principe la partition floue, le choix du seuil se fait par la maximisation de d'entropie. Considérons les deux ensembles flous objet et fond où les fonctions d'appartenance sont définis par :

$$\mu_{objet} = \begin{cases} 1 & x \leq a \\ \frac{x-c}{a-c} & a < x < c \\ 0 & x \geq c \end{cases} \quad \mu_{fond} = \begin{cases} 0 & x \leq a \\ \frac{x-a}{c-a} & a < x < c \\ 1 & x \geq c \end{cases} \quad (2.21)$$

Dans cette méthode, le seuil de binarisation optimal  $T$  est choisi comme le niveau de gris dont la fonction d'appartenance égale 0.5 et donc c'est le centre de l'intervalle  $[a_{opt}, c_{opt}]$ , tel que  $a_{opt}, c_{opt}$  sont les valeurs de  $a, c$  en maximisant l'entropie de la division donnée par :

$$H = -(P_1 \times \log P_1 + P_2 \times \log P_2) \quad (2.22)$$

$$P_1 = \sum_{i=0}^{L-1} \mu_1(i) \times p(i) \quad P_2 = \sum_{i=0}^{L-1} \mu_2(i) \times p(i) \quad (2.23)$$

➤ **Méthode de Li-Lee** [12]

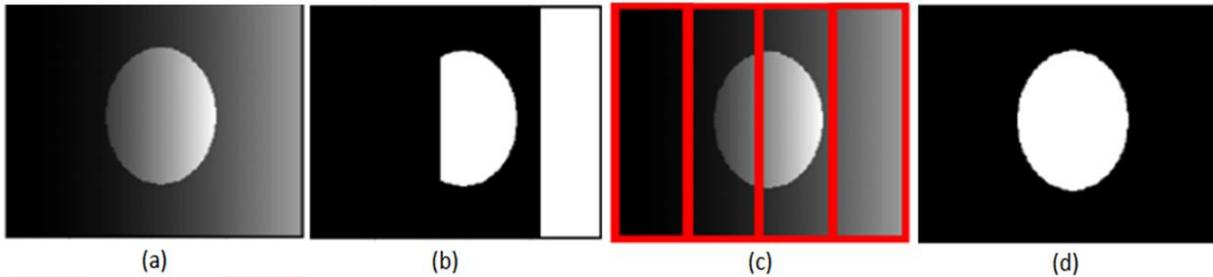
Li et Lee ont proposé une méthode de seuillage d'images où le regroupement des niveaux de gris en deux classes (objet et fond) est fondé sur la minimisation de l'entropie croisée. Le seuil optimal  $T$  est calculé de façon à minimiser la somme de l'entropie croisée de l'objet et du fond où:

$$\eta(t) = \eta_1(t) + \eta_2(t) = \sum_{i=0}^t i \times h_i \times \log \left( \frac{i}{\mu_1(t)} \right) + \sum_{i=t+1}^{255} i \times h_i \times \log \left( \frac{i}{\mu_2(t)} \right) \quad (2.24)$$

### 2.3.2. Méthodes de binarisation locale

Il existe des images pour lesquelles il est difficile d'effectuer une bonne binarisation au moyen d'un seul seuil, si celui-ci est le même pour chaque point de l'image. Ceci peut être dû à une non uniformité de l'éclairage de l'image ou au fait que les différents objets qui composent l'image ont des dynamiques de luminances différentes. Pour cela, les méthodes de binarisation locale sont mieux adaptées.

A l'inverse des méthodes globales qui ne considèrent que la valeur du pixel, les méthodes locales prennent en considération la valeur des pixels voisins pour le calcul des seuils. Deux stratégies sont souvent employées. La première consiste à diviser, en premier lieu, l'image en fenêtres non chevauchantes. Dans un deuxième temps, un seuil est déterminé dans chaque fenêtre selon la nature de l'histogramme local. Un seuil égal à zéro est affecté à la fenêtre possédant un histogramme uni-modal tandis qu'une procédure de seuillage est appliquée pour déterminer un seul seuil dans le cas d'un histogramme bimodal ou multimodal. La figure suivante illustre le principe du seuillage local.



**Figure 2.2:** Illustration du seuillage adaptatif. (a) Image originale avec un fond non uniforme, (b) Image segmentée avec un seuil global, (c) Image originale découpée en sous images, (d) Image segmentée avec un seuillage adaptatif.

#### ➤ Méthode de Bernsen [13]

Bernsen propose que le seuil local soit calculé comme la moyenne des intensités de niveau de gris minimal  $I_{\min}(i, j)$  et maximale  $I_{\max}(i, j)$  dans la fenêtre carré  $w$  centré sur le pixel  $I(i, j)$ . Ainsi pour chaque pixel de coordonnées  $(x, y)$ , le seuil local est donné par

$$T_w(i, j) = \frac{I_{\max}(i, j) + I_{\min}(i, j)}{2} \quad (2.25)$$

Avec un contraste calculé par :

$$C_w(i, j) = I_{\max}(i, j) - I_{\min}(i, j) \quad (2.26)$$

Cependant si la mesure du contraste  $C_w(i, j)$  est inférieure à un certain seuil  $T$ , alors le voisinage consiste en une seule classe: fond ou objet. La méthode de Bernsen est très sensible au bruit du fond. A cause de la prise en compte du maximum et du minimum uniquement, dans le cas où la fenêtre est uniquement sur du fond, le bruit sera interprété comme objet, car le seuil sera bas.

➤ **Méthode de Niblack** [14]

L'algorithme de Niblack est une amélioration de la méthode de Bernsen qui calcule un seuil local pour chaque pixel  $I(i, j)$ . Le seuil  $T$  est calculé en fonction de la moyenne  $m$  et de l'écart-type  $\sigma$  des tous les pixels de la fenêtre carré  $w$  centrée sur le pixel de coordonnées  $(i, j)$ . Ainsi le seuil  $T_w(i, j)$  est donné par :

$$T_w(i, j) = m + k \times \sigma_w(i, j) \quad (2.27)$$

$$T_w(i, j) = m + k \times \sqrt{\frac{\sum (I_w(i, j) - m)^2}{NP}} \quad (2.28)$$

Le paramètre  $k$  est utilisé pour déterminer le nombre des pixels des contours considérés comme des pixels d'objet et il prend des valeurs négatives,  $k$  est fixé à (-0.2) par les auteurs et  $NP$  représente le nombre des pixels de la fenêtre carré  $w$  et  $m$  est la valeur moyenne des pixels de la fenêtre carré  $w$ . Cette méthode donne de bons résultats car le seuil dépend du pixel et de l'information extraite à partir de son voisinage, mais n'est pas efficace quand le fond n'est pas uniforme.

➤ **Méthode de Sauvola** [15]

L'algorithme de Sauvola est une amélioration de celui de Niblack, où il insère un paramètre  $R$  qui est le rang dynamique de l'écart type de la valeur de gris de l'image, et cela pour donner plus de performance pour les documents avec un fond uniforme et aussi pour avoir augmenté la performance pour les documents avec un fond contenant des textures claires ou bien une illumination inégale [1]. Dans la modification de Sauvola le seuil de binarisation est donné par:

$$T_w(i, j) = m \times \left[ 1 + k \times \left( \frac{\sigma_w(i, j)}{R} - 1 \right) \right] \quad (2.29)$$

Le paramètre  $R$  est le rang dynamique de l'écart-type  $\sigma$  et le paramètre  $k$  prend des valeurs positives dans l'intervalle  $[0.2, 0.5]$ . Dans la littérature on prend  $R=128$  et  $k=0.5$ .

Dans [4], la méthode de Sauvola se montre plus efficace que celle de Niblack dans le cas où le niveau de gris de l'objet est proche du niveau de gris 0, et celui du fond est proche du niveau de gris 255. Cependant dans les images où les niveaux de gris des pixels du fond et de l'objet sont proches, les résultats sont peu satisfaisants.

➤ **Méthode de Wolf** [16]

Pour résoudre les problèmes rencontrés avec l'algorithme de Sauvola en présence de document de faible contraste, Wolf et al. Proposent de normaliser le contraste et la moyenne de niveaux de gris de l'image. Ainsi le seuil se calcule par :

$$T_w(i, j) = (1 - k) \times m + (k \times M) + k \times \frac{\sigma_w(i, j)}{R} \times (m - M) \quad (2.30)$$

Wolf propose de fixer  $k = 0.5$  et  $M$  c'est le niveau de gris minimum de toute l'image et  $R$  c'est l'écart-type maximum obtenu sur toutes les fenêtres. Cette méthode surpasse dans la plupart des cas les méthodes de Niblack et de Sauvola. Cependant, si les valeurs de gris d'arrière-plan sont très différentes alors une dégradation de ses performances est observée. Cela est dû au fait que les valeurs des paramètres  $M, R$  sont calculées à partir de l'image globale.

➤ **La méthode de Nick** [17]

L'avantage majeur de cette méthode est qu'elle améliore considérablement la binarisation des images de page « blanches » et claires, et dans le cas où l'image présente de faible contraste, en déplaçant vers le bas, le seuil de binarisation.

Le calcul du seuil est réalisé comme suit :

$$T_w = m + k \sqrt{\frac{(\sum I_w(i, j)^2 - m^2)}{NP}} \quad (2.31)$$

Tel que :

$k$  est le facteur de Niblack et il varie entre  $[-0.1, -0.2]$  selon les besoins de l'application et  $NP$  représente les nombre des pixels de la fenêtre carré  $w$ .

### 2.3.3. Méthodes de binarisation hybride

Les méthodes de binarisation hybride traitent les images des documents anciens en niveaux de gris, considèrent que les caractéristiques spécifiques d'un document tels que le texte ou les graphiques sont placés sur le premier plan et le fond comme deuxième plan. Les images d'entrée sont décrites par l'équation :

$$I(x, y) = r, r \in [0, 255] \quad (2.32)$$

Où  $x$  et  $y$  sont les coordonnées horizontale et verticale d'un pixel dans une image, et  $r$  peut prendre n'importe quelle valeur entre 0 et 255, la valeur 0 représentant un pixel noir et 255 la valeur d'un pixel blanc. L'objectif est la transformation de tous les intermédiaires de pixels évalués pour finalement trouver la valeur du fond, le pixel vaut 255 et pour le premier plan, la valeur de pixel vaut 0. Ainsi, une technique globale peut être utilisée. En plus de cela, une méthode qui conduit à l'approche hybride pour résoudre ces problèmes de documents anciens doit être proposée.

➤ **Algorithme de Gatos** [18]

Le principe de cette méthode est d'estimer le fond d'une image, pour ensuite faire un seuillage sur la différence entre le fond et l'image d'origine, en faisant une combinaison des différentes méthodes ayant des propriétés complémentaires pour réaliser une binarisation automatique capable de traiter des documents dégradés. Gatos décrit sa méthode comme une succession de cinq phases.

- La première améliore la qualité globale du document.
- L'étape suivante consiste à estimer la valeur du fond de l'image. Cette opération est réalisée en deux phases.
- Une première estimation de l'avant plan est effectuée en utilisant l'algorithme de Sauvola.
- Une fois les régions détectées, les pixels objet sont remplacés par une interpolation avec les pixels voisins.

$$B(x, y) = \begin{cases} I(x, y) & \text{si } S(x, y) = 0 \\ \frac{\sum_{x-dx}^{x+dx} \sum_{y-dy}^{y+dy} (I(i_x, i_y)(1 - S(i_x, i_y)))}{\sum_{x-dx}^{x+dx} \sum_{y-dy}^{y+dy} (1 - S(i_x, i_y))} & \text{si } S(x, y) = 1 \end{cases} \quad (2.33)$$

Avec :

$B(x, y)$  : Image du fond estimée.

$I(x, y)$  : Image après le filtrage de Wiener.

$S(x, y)$  : Image binaire obtenue à partir de la segmentation de Sauvola.

A la fin, en comparant l'image filtrée avec l'image du fond pour avoir un seuillage final, les zones de texte sont détectées si la distance entre l'image filtrée et l'image de fond est supérieure à un seuil  $T$ .

Ce seuil  $T$  doit varier en fonction de l'image du fond, pour que dans le cas d'un fond sombre, le seuil soit fixé plus bas pour pouvoir séparer le texte du fond. Cette opération est donnée par la formule suivante:

$$D(x, y) = \begin{cases} 1 & \text{Si } B(x, y) - I(x, y) > T(B(x, y)) \\ 0 & \text{Sinon} \end{cases} \quad (2.34)$$

Avec :

$D(x, y)$  : Image binaire finale.

$T$  : Fonction de seuillage.

Dans le cas d'images de document à éclairage uniforme, le seuil  $T$  est défini en fonction de la distance moyenne  $\delta$  entre l'image du fond et l'image filtrée est calculé par les équations suivantes :

$$T = q \times \delta \quad (2.35)$$

$$\delta = \frac{\sum_x \sum_y (B(x, y) - I(x, y))}{\sum_x \sum_y S(x, y)} \quad (2.36)$$

Cependant, dans le cas d'une image sombre, le seuil  $T$  (pour  $k$  fixé à 0.8) n'est plus bon. Pour que la méthode soit correcte dans le cas où l'illumination du document ne serait pas constante, cette constante est adaptée en fonction de l'intensité moyenne du fond. Pour cela une fonction sigmoïde permettant d'adapter en douceur la constante est utilisée. Finalement,  $T$  est calculé ainsi :

$$T(x, y) = q \times \delta \times \left( \frac{(1 - p_2)}{1 + \exp\left(\frac{-4 \times B(x, y)}{b \times (1 - p_1)} + \frac{2 \times (1 + p_1)}{1 - p_1}\right)} \right) \quad (2.37)$$

$$b = \frac{\sum_x \sum_y (B(x, y) \times (1 - S(x, y)))}{\sum_x \sum_y (1 - S(x, y))} \quad (2.38)$$

$T(x, y)$  : Les valeurs de Seuillage final.

$q, p_1$  et  $p_2$  : Paramètre fixé par Gatos  $q = 0.6, p_1 = 0.5$  et  $p_2 = 0.8$ .

$B(x, y)$  : Image du fond estimée.

➤ **Algorithme de Lu** [19]

La méthode de binarisation de Lu repose sur la combinaison de plusieurs techniques. Tout d'abord elle estime la surface du background du document par le biais d'une procédure de lissage polynomial itératif unidimensionnel horizontale et verticale [20]. Une fois avoir obtenu une bonne approximation de l'arrière-plan, l'étape suivante de l'algorithme, est la compensation du contraste en utilisant l'image résultante de l'estimation du background en utilisant l'équation suivante:

$$I' = \frac{C}{BG} \times I \quad (2.39)$$

Avec :

$I$  : L'image originale.

$C$  : Contraste de l'image généralement calculée par la médiane de toute l'image.

$BG$  : L'image du background estimée.

Différentes méthodes pour la détection de contour du texte ont été rapportées [21], [22]. Cependant, Lu estime que les pixels du contour peuvent être mieux détectés en combinant le gradient de (norme L1) dans le sens horizontal et vertical, comme suit:

$$V(x, y) = V_h(x, y) + V_v(x, y) \quad (2.40)$$

$$V_h(x, y) = |I'(x, y + 1) - I'(x, y - 1)| \quad (2.41)$$

$$V_v(x, y) = |I'(x + 1, y) - I'(x - 1, y)|$$

Une fois que les contours du texte sont détectés, l'image binarisée est extraite en fonction de ces contours, comme suit:

$$R(x, y) = \begin{cases} 0 & N_e \geq N_{\min} \ \&\& \ I'(x, y) \leq E_{mean} \\ 1 & \text{Otherwise} \end{cases} \quad (2.42)$$

$N_e$  : Nombre des pixels de contour du texte détectés dans une fenêtre  $w$ .

$N_{\min}$  : Valeur du seuil qui spécifie le nombre minimum de contour du texte détecté

$$E_{mean} = \frac{\sum I'(x, y) \times (1 - E(x, y))}{N_e} \quad (2.43)$$

$E$  : Fait référence à l'image du contour du texte déterminée par l'équation (2.40).

➤ **Méthode de Su** [23]

La méthode de Su détermine l'image hautement contrastée, en supprimant la variation de l'arrière-plan par l'application de l'algorithme d'Otsu sur l'image contrastée, calculée en fonction du maximum et du minimum locaux de l'image [24] comme suit :

$$D(x, y) = \frac{f_{\max}(x, y) - f_{\min}(x, y)}{f_{\max}(x, y) + f_{\min}(x, y) + \varepsilon} \quad (2.44)$$

$f_{\max}(x, y)$  et  $f_{\min}(x, y)$  font référence aux intensités d'image maximale et minimale dans une fenêtre carré  $w(3 \times 3)$ , et le terme  $\varepsilon$  est un nombre positif mais infiniment petit, il est ajouté si le maximum local est égal à 0.

La binarisation se fait sur la base des seuils locaux estimés en fonction de l'intensité du pixel par rapport à celle des pixels voisins à contraste élevé, comme suit:

$$R(x, y) = \begin{cases} 0 & N_e \geq N_{\min} \ \&\& \ I(x, y) \leq E_{mean} + \frac{E_{std}}{2} \\ 1 & \text{Otherwise} \end{cases} \quad (2.45)$$

$E_{mean}$  et  $E_{std}$  représentent respectivement la moyenne et l'écart type de l'intensité de l'image hautement contrasté, ils peuvent être calculé comme suit :

$$E_{mean} = \frac{\sum I(x, y) \times (1 - E(x, y))}{N_e} \quad (2.46)$$

$$E_{std} = \sqrt{\frac{\sum ((I(x, y) - E_{mean}) \times (1 - E(x, y)))^2}{2}} \quad (2.47)$$

$I$  : Image originale.

$E$  : Image binaire hautement contrastée.

$N_e$  : Nombre des pixels hautement contrasté dans une fenêtre  $w$ .

$N_{\min}$  : Valeur du seuil qui spécifie le nombre minimum des pixels hautement contrastés.

➤ **Algorithme de Hedjam** [25]

Le procédé proposé par cette méthode est constitué de deux principales étapes : La première étape consiste à binariser l'image du document en éliminant les informations de dégradation et de ne garder que la partie noire du texte. Dans ce cas, les motifs d'interférence, le bruit, et les parties avec l'intensité faibles seront effacées, en utilisant la méthode de binarisation Sauvola pour produire une image binarisée avec un seuil optimum  $T$ .

La deuxième étape récupère les parties de texte manquantes (perdues dans l'étape précédente) en utilisant une binarisation locale basée sur la classification des maximums de vraisemblances.

Cela nous permet de prendre deux modèles gaussiens, l'un pour le fond et l'autre pour le texte :

$$\begin{cases} P(u(x)|H_0) = P(u(x)|text) = \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(\frac{-(u(x)-u_t)^2}{2\sigma_t^2}\right) \\ \text{et} \\ P(u(x)|H_1) = P(u(x)|fond) = \frac{1}{\sqrt{2\pi\sigma_b^2}} \exp\left(\frac{-(u(x)-u_b)^2}{2\sigma_b^2}\right) \end{cases} \quad (2.48)$$

$u_w$  et  $\sigma_w$  représentent la moyenne et la variance respectivement avec  $w \in \{t, b\}$ ,  $u(x)$  représente le niveau de gris d'un pixel. A la fin il faut résoudre le problème de classification équivalent au problème de minimisation suivant :

$$u_{fond}(x) = \arg \min_{w \in \{t, b\}} \left\{ \frac{(u(x) - u_w)^2}{2\sigma_w^2} \right\} \quad (2.49)$$

## 2.4. Implémentation d'algorithmes de calcul du seuil de binarisation

Le Seuillage automatique est un domaine très ancien, mais toujours très actif dans la recherche. De nombreuses techniques ont été développées pour cette tâche, allant de solutions ad-hoc simples à des algorithmes complexes avec des fondations théoriques solides, tel que documenté dans plusieurs revues et études d'évaluation dans la littérature.

La binarisation d'images est également considérée comme une technique de «segmentation par seuillage», elle est donc souvent classée sous ce terme. Dans cette partie nous allons donner quelques implémentations des méthodes de base, les plus utilisées pour le calcul du seuil, la majorité de ces méthodes utilisent la notion de distribution de probabilité et histogramme. Pour cela, nous commençons par les méthodes simples, puis par les méthodes de seuillage globales et enfin les méthodes locales.

### 2.4.1. Implémentation d'une méthode simple de calcul du seuil

La valeur de seuil ne doit pas être fixée d'une certaine manière, mais sur la base du contenu de l'image. Dans le cas le plus simple, nous pourrions utiliser comme seuil de binarisation la moyenne ou la médiane de tous les pixels de l'image, ou même aussi la moyenne entre la valeur maximale et minimale des pixels.

Comme valeur du seuil la valeur de la moyenne :

$$T = \text{mean}(I) \quad (2.50)$$

Comme valeur du seuil la valeur de la médiane :

$$T = \text{median}(I) \quad (2.51)$$

Comme valeur du seuil la moyenne entre la valeur maximale et minimale :

$$T = \frac{\text{Max}(I) + \text{Min}(I)}{2} \quad (2.52)$$

```

1: Simple_Threshold (I,T)
   Input: I           -- Image originale.
   Return: T         -- La valeur du seuil optimal.
2:   $I_{\min} \leftarrow \text{Min}(I)$       -- Valeur minimale d'intensité du pixel
3:   $I_{\max} \leftarrow \text{Max}(I)$       -- Valeur maximale d'intensité du pixel
4:   $T \leftarrow \frac{I_{\min} + I_{\max}}{2}$     -- Calcul du seuil
5:  Return T .

```

**Algorithme 2.1** : Calcul simple d'un seuil de binarisation

### 2.4.2. Implémentation méthode globale de calcul du seuil

L'approche globale du calcul du seuil est calculée à partir d'une mesure globale sur toute l'image (par exemple, l'histogramme). Dans cette partie on va présenter deux implémentations d'algorithme de seuillage global.

L'algorithme d'Otsu suppose que l'image originale contient des pixels classés en deux, dont les distributions d'intensité sont inconnues. L'objectif est de trouver un seuil  $T$  tel que l'arrière-plan et l'avant plan soient au maximum séparés, ce qui signifie d'après l'histogramme qu'ils sont :

- Chacune soit aussi étroite que possible (avoir des écarts minimes)
- Leurs centres (moyens) sont les plus éloignés.

L'algorithme suivant donne une implémentation de l'algorithme d'Otsu pour le calcul du seuil par la maximisation de la variance interclasse, il suppose que l'image est en niveaux de gris  $K$  avec un total de  $N$  pixels.

```

1: Otsu_Threshold(h)
   Input : h           -- Histogramme.
   Return : T         -- La valeur du seuil optimal
2: k ← Size (h)           -- Nombre des niveaux d'intensité
3: N ← Sum (h)           -- Nombre de pixels
4: T ← 0,  $\sigma_{b\max}^2$  ← 0
5: for q ← 0 .. k do
6:   P1 ← 0, P2 ← 0,  $\mu_1$  ← 0,  $\mu_2$  ← 0
7:   for i ← 0 .. q do
8:      $P_1 \leftarrow P_1 + \frac{h(i)}{N}$            -- Calcul de la probabilité moyenne de la classe objet
9:   End for
10:  for i ← q + 1 .. k do
11:     $P_2 \leftarrow P_2 + \frac{h(i)}{N}$            -- Calcul de la probabilité moyenne de la classe fond
12:  End for
13:  for i ← 0 .. q do
14:     $\mu_1 \leftarrow \mu_1 + \left(\frac{h(i)}{N}\right) \times i$  -- Calcul de l'écart type de la classe objet
15:  End for
16:   $\mu_1 \leftarrow \frac{\mu_1}{P_1}$ 
17:  for i ← q + 1 .. k do
18:     $\mu_2 \leftarrow \mu_2 + \left(\frac{h(i)}{N}\right) \times i$  -- Calcul de l'écart type de la classe fond
19:  End for
20:   $\mu_2 \leftarrow \frac{\mu_2}{P_2}$ 
20:   $\sigma_b^2 \leftarrow P_1 \times P_2 \times (\mu_1 - \mu_2)^2$ 
21:  if  $\sigma_b^2 > \sigma_{b\max}^2$  do
22:     $\sigma_{b\max}^2 \leftarrow \sigma_b^2$            -- Maximisation de la variance pour trouver le seuil
23:    T ← q
24:  end if
25: End for
26: Return T

```

**Algorithme 2.2** : Calcul du seuil de binarisation par la méthode d'Otsu

Pour la méthode d'ISODATA elle détermine l'intervalle [min, max] des valeurs non nulles de l'histogramme puis elle fait une estimation initiale du seuil (en prenant la moyenne ou la médiane de l'image entière). Cela divise l'ensemble de pixels en deux classes (objet et fond). Ensuite, la moyenne de deux classes est calculée et le seuil est repositionné centré entre les deux moyennes jusqu'au moment où l'algorithme converge.

```

1 : ISODATA_Threshold (h)
   Input: h           -- Histogramme
   Return : T        -- La valeur du seuil optimal
2 :  $k \leftarrow \text{Size}(h)$       -- Nombre des niveaux d'intensité
3 :  $N \leftarrow \text{Sum}(h)$       -- Nombre de pixels
4 :  $T_0 \leftarrow 0, T_1 \leftarrow 0, m_1 \leftarrow 0, m_2 \leftarrow 0$ 
5 :  $T_0 \leftarrow \text{round}\left(\frac{\text{Sum}(h(1:k) \times (1:k))}{N}\right)$   -- Calculer l'intensité moyenne de l'image
6 :  $m_1 \leftarrow \frac{\text{Sum}(h(1:T_0) \times (1:T_0))}{\text{Sum}(h(1:T_0))}$   -- Calculer la moyenne en dessous du seuil
7 :  $m_2 \leftarrow \frac{\text{Sum}(h(T_0+1:256) \times (T_0+1:256))}{\text{Sum}(h(T_0+1:256))}$   -- Calculer la moyenne en dessus du seuil
8 :  $T_1 = \text{round}\left(\frac{m_1 + m_2}{2}\right)$   -- Calculer le nouveau seuil
9 : While ( $|T_1 - T_0| \geq 1$ ) do
10 :  $T_0 \leftarrow T_1$   -- Répéter la même procédure pour les
    autres itérations a une condition que
11 :  $m_1 \leftarrow \frac{\text{Sum}(h(1:T_0) \times (1:T_0))}{\text{Sum}(h(1:T_0))}$   la différence entre le nouveau et
    l'ancien seuil soit supérieure à 1
13 :  $T_1 = \text{round}\left(\frac{m_1 + m_2}{2}\right)$ 
14 : End While
15 :  $T \leftarrow T_1$ 
16 : Return T

```

**Algorithme 2.3:** Calcul du seuil de binarisation par la méthode D'ISODATA

### 2.4.3. Implémentation de la méthode locale de calcul du seuil

Dans beaucoup de situations, un seul seuil fixe pour toute l'image ne peut être approprié pour classifier les pixels de toute l'image, surtout lorsque cette dernière présente des zones avec des arrière-plans non uniformément éclairées. La solution est de recourir à chercher des seuils locaux adaptatifs. Pour cela nous trouvons dans la littérature plusieurs méthodes de seuillage local. La méthode de Bernsen c'est une méthode locale adaptative dont le seuil est calculé pour chaque pixel de l'image. Ainsi le seuil est le centre entre le niveau de gris le plus haut et le plus bas dans une fenêtre carré de taille  $w$  centré pour chaque pixel de coordonnées  $(x, y)$ .

1: *Bernsen\_Threshold* ( $I, w$ )

*Input:*  $I, w$  -- Image originale,  $w$  fenêtre carrée

*Return:*  $T$  -- Valeurs du seuil pour chaque pixel de l'image

2:  $(M, N) \leftarrow \text{Size}(I)$  -- Taille de l'image

3: Create map  $Q : M \times N \rightarrow w$

4:  $n \leftarrow 0$

5: for all pixel  $I(i, j)$  do

6:  $n \leftarrow n + 1$

7:  $I_w \leftarrow \text{map}(I, w)$  -- Les pixels de la fenêtre centrée sur le pixel  $I(i, j)$

8:  $I_{\min} = \min I_w$  -- Valeur minimale dans la fenêtre  $w$

7:  $I_{\max} = \max I_w$  -- Valeur maximale dans la fenêtre  $w$

9:  $T_n = \left( \frac{I_{\max} + I_{\min}}{2} \right)$

10: End for

11:  $T \leftarrow T_n$

12: return  $T$

**Algorithme 2.4:** Calcul du seuil de binarisation par la méthode Bernsen

La méthode de Niblack calcule le seuil localement pour chaque pixel en glissant la fenêtre rectangulaire de taille  $w$  sur toute l'image. Le seuil  $T$  est calculé en utilisant la moyenne et l'écart-type de tous les pixels dans la fenêtre (voisinage du pixel en question).

```

1: Niblack_Threshold ( $I, w, k$ )

Input:  $I, w, k$       -- Image originale,  $w$  taille fenêtre carrée,  $k$  paramètre fixé à -0,2

Return:  $T$           -- Valeurs du seuil pour chaque pixel de l'image

2:  $(M, N) \leftarrow \text{Size}(I)$   -- Taille de l'image

3: Create map  $Q : M \times N \rightarrow w$ 

4:  $n \leftarrow 0$ 

5: for all pixel  $I(i, j)$  do

6:    $n \leftarrow n + 1$ 

7:    $I_w \leftarrow \text{map}(I, w)$   -- Les pixels de la fenêtre centré sur le pixel  $I(i, j)$ 

8:    $T_n \leftarrow m + k \times \sigma$   --  $m$  : la moyenne des pixels,  $\sigma$  : l'écart type dans la fenêtre  $w$ 

9:    $T_n \leftarrow \text{mean}(I_w) + k \times \sqrt{\text{mean}((I_w - \text{mean}(I_w))^2)}$ 

10: End for

11:  $T \leftarrow T_n$ 

12: return  $T$ 

```

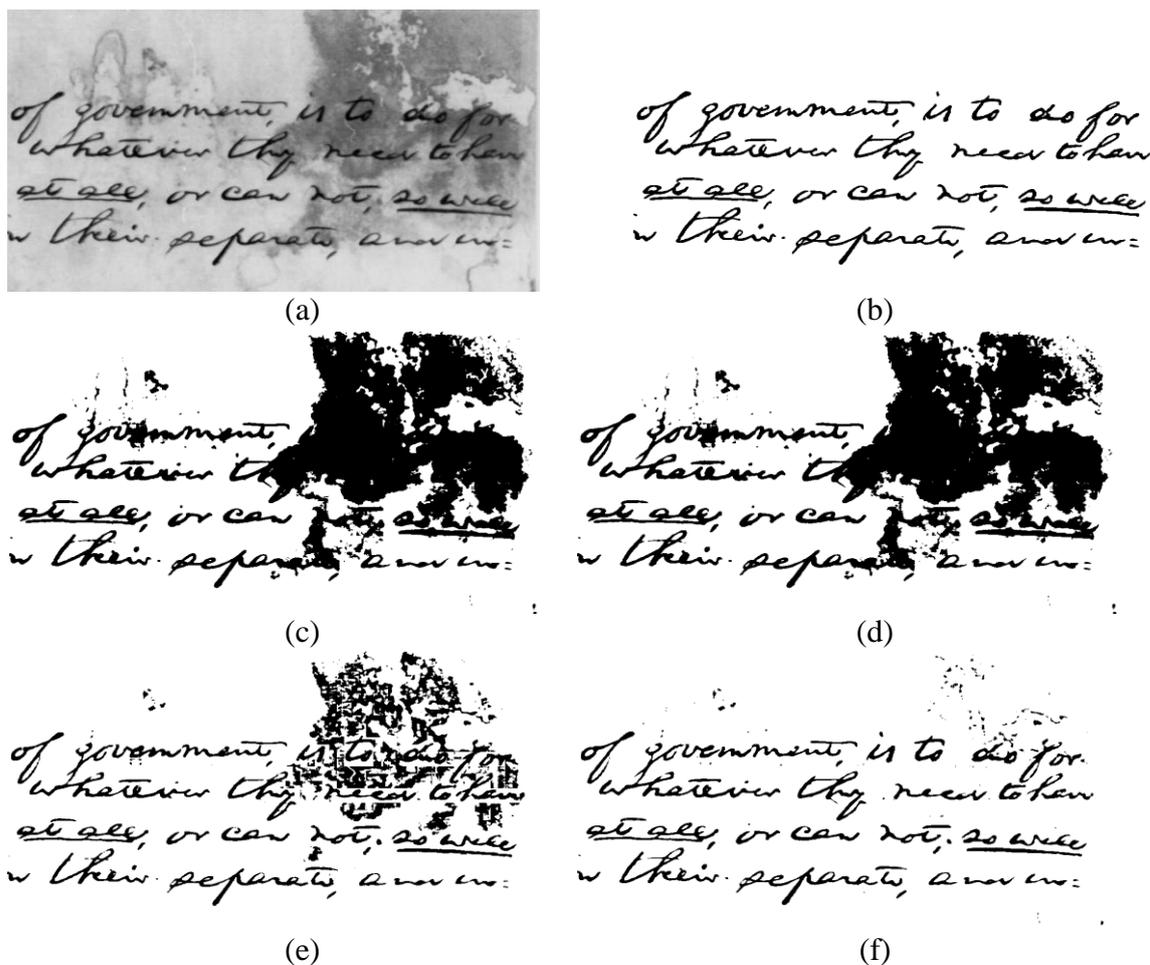
**Algorithme 2.5:** Calcul du seuil de binarisation par la méthode Niblack

## 2.5. Résultats de binarisation sur quelques images vieux documents

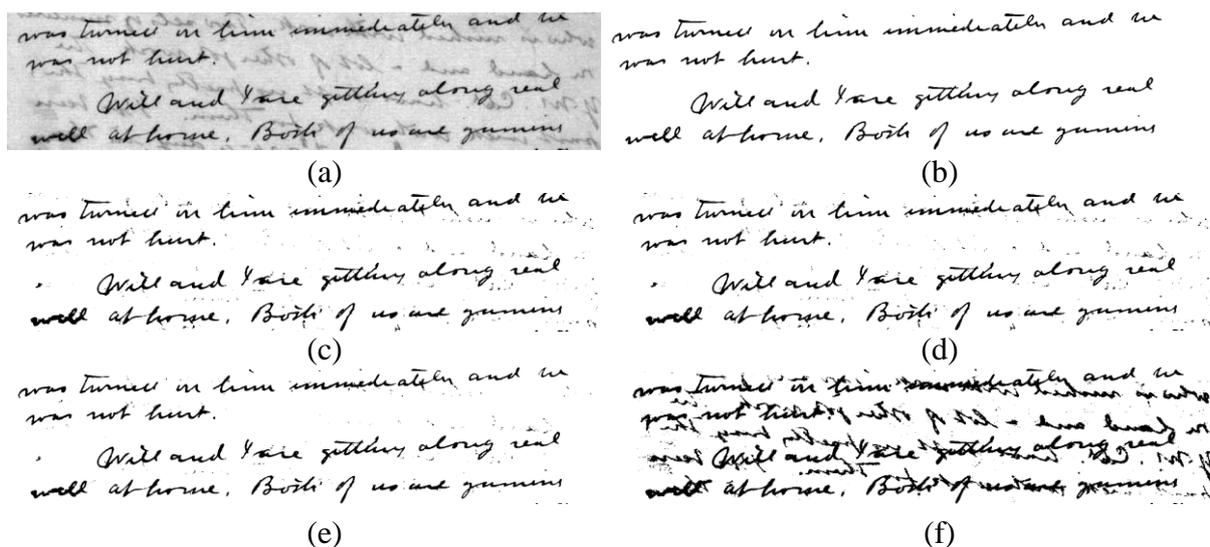
Dans cette partie, nous allons présenter quelques résultats en appliquant les différents algorithmes de binarisation (global, local) implémentés en Matlab.

Le choix s'est porté sur plusieurs images de la base des données DIBCO avec ses versions de 2009 à 2013. Nous nous contentons de présenter les tests surtout sur les types d'images qui présentent:

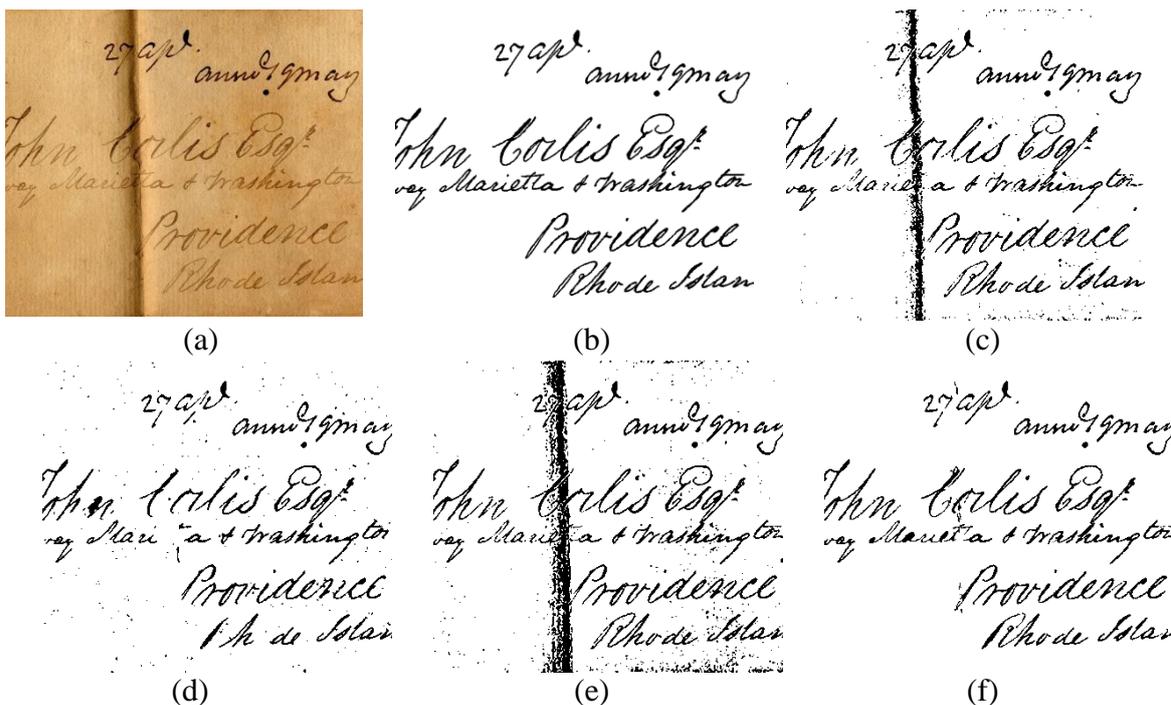
- Des dégradations dues à un éclairage non uniforme
- Des tâches dues à un vieillissement.
- Des transparences de l'arrière-plan



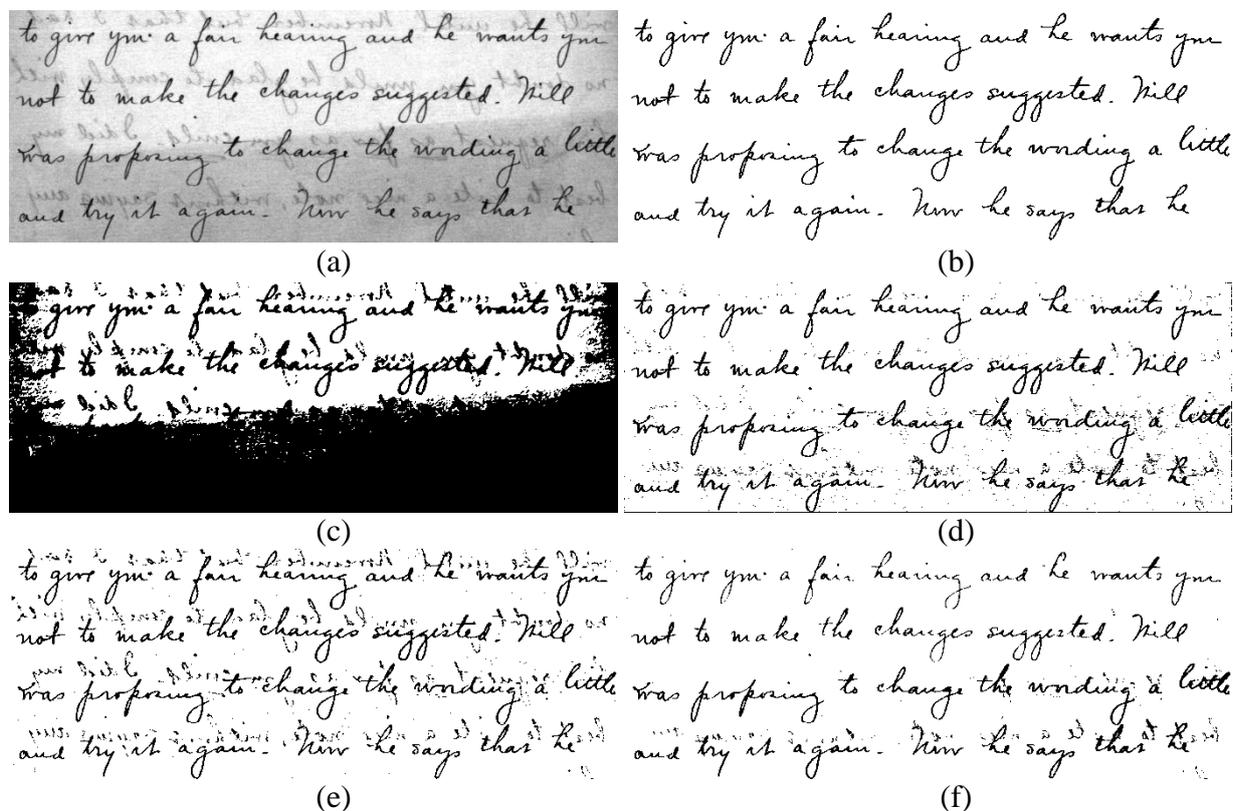
**Figure 2.3:** Résultat de binarisation sur la base de données DIBCO 2009. (a) Image originale, (b) Image de référence, (c) Méthode d’Otsu, (d) Méthode d’ISODATA, (e) Méthode de Bernsen, (f) Méthode de Niblack



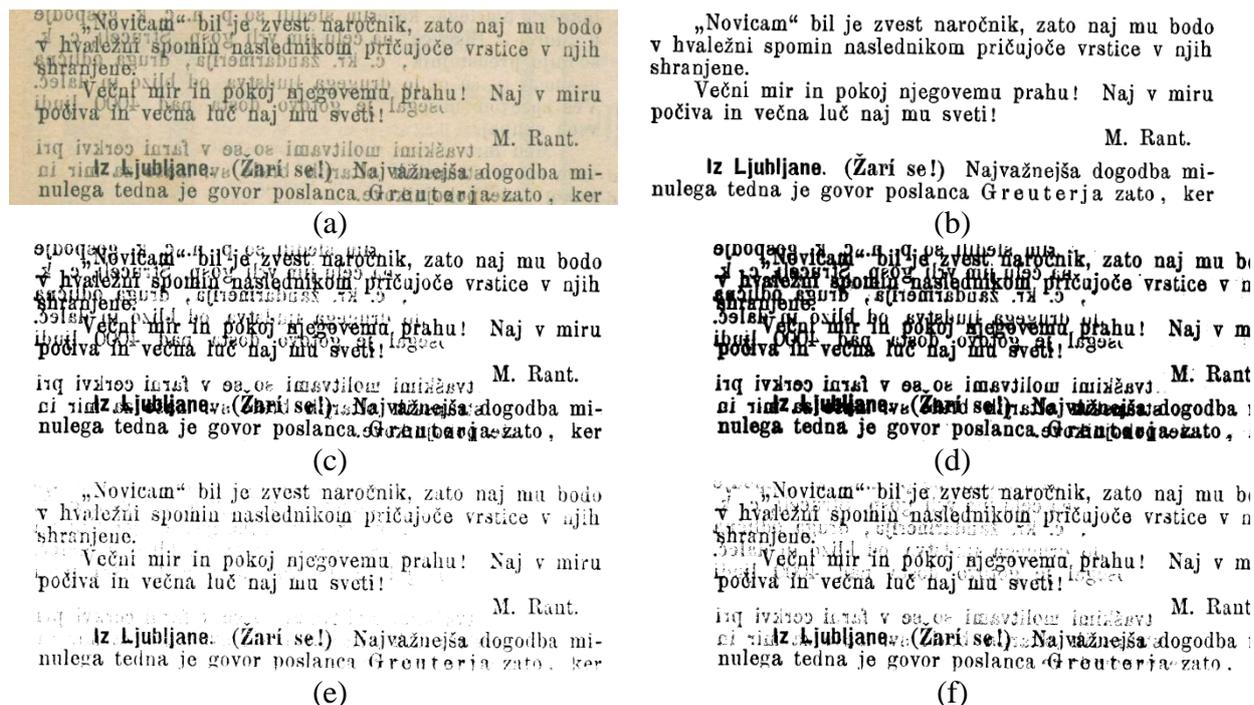
**Figure 2.4:** Résultat sur la base de données DIBCO 2010. (a) Image originale, (b) Image de référence, (c) Méthode d’Otsu, (d) Méthode d’ISODATA, (e) Méthode de Bernsen, (f) Méthode de Niblack



**Figure 2.5:** Résultat sur la base de données DIBCO 2011. (a) Image originale, (b) Image de référence, (c) Méthode d’Otsu, (d) Méthode d’ISODATA, (e) Méthode de Bernsen, (f) Méthode de Niblack



**Figure 2.6:** Résultat sur la base de données DIBCO 2012. (a) Image originale, (b) Image de référence, (c) Méthode d’Otsu, (d) Méthode d’ISODATA, (e) Méthode de Bernsen, (f) Méthode de Niblack



**Figure 2.7:** Résultat sur la base de données DIBCO 20013. (a) Image originale, (b) Image de référence, (c) Méthode d’Otsu, (d) Méthode d’ISODATA, (e) Méthode de Bernsen, (f) Méthode de Niblack

## 2.6. Conclusion

Dans ce chapitre, Nous avons présenté différentes méthodes de binarisation et donné quelques implémentations, en particulier les méthodes utilisant l’analyse par histogramme. Les méthodes de binarisation peuvent être regroupées en trois catégories

- Les méthodes globales sont plus ou moins rapides et faciles à mettre au point. Cependant elles présentent des limitations surtout lorsque l’image présente des variations de luminance dans le fond et l’objet, c.à.d. lorsque l’histogramme de l’image est multimodal, ainsi que la nécessité de définir les paramètres manuellement (taille de la fenêtre, facteur  $k$ ).
- Les méthodes locales sont plus efficaces, plus complexes et sont composées de filtres et d’opérations existantes [26] ainsi elles nécessitent des capacités de calcul plus en fonction du choix de la taille de la fenêtre.
- Les algorithmes hybrides combinent en même temps entre les techniques globales et locales, ainsi ils peuvent être rapides comme les méthodes globales, et fournissant des résultats de haute qualité comme les algorithmes de binarisation locale.

En fin il est rare qu'une image scannée d'un document historique, soit suffisamment de bonne qualité pour appliquer un algorithme de binarisation qui est encore un sujet de recherche très actif, malgré le nombre de publications dans ce domaine. En effet, il est bien souvent nécessaire d'adapter une technique en fonction de la problématique.

## *Bibliographie*

- 
- [1] B. Sankur, « Survey over image thresholding techniques and quantitative performance evaluation », *J. Electron. Imaging*, vol. 13, n° 1, pp. 146, janv. 2004.
- [2] B. Su, S. Lu, C. L. Tan, « Combination of Document Image Binarization Techniques », in *2011 International Conference on Document Analysis and Recognition*, Beijing, China, 2011, pp. 22-26.
- [3] N. Arica, F. T. Yarman-Vural, « An overview of character recognition focused on off-line handwriting », *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.*, vol. 31, n° 2, pp. 216-233, mai 2001.
- [4] A. Kefali, T. Sari, M. Sellami, « Evaluation de plusieurs techniques de seuillage d'images de documents Arabes anciens », in *5ème Symposium International Image, Multimédia, Application Graphique et Environnement (IMAGE'2009)*, Biskra, Algeria, 2009.
- [5] S. M. Ismail, S. N. H. Sheikh Abdullah, F. Fauzi, « Statistical Binarization Techniques for Document Image Analysis », *J. Comput. Sci.*, vol. 14, n° 1, pp. 23-36, janv. 2018.
- [6] N. Otsu, « A Threshold Selection Method from Gray-Level Histograms », *IEEE Trans. Syst. Man Cybern.*, vol. 9, n° 1, pp. 62-66, janv. 1979.
- [7] J. Kittler, M. Hatef, R. P. W. Duin, J. Matas, « On combining classifiers », *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, n° 3, pp. 226-239, mars 1998.
- [8] V. FLAVIO, « Thresholding Using the ISODATA Clustering Algorithm », *IEEE Trans. Syst. Man Cybern.*, vol. 10, n° 11, pp. 771-774, 1980.
- [9] T. Pun, « A new method for grey-level picture thresholding using the entropy of the histogram », *Signal Process.*, vol. 2, n° 3, pp. 223-237, juill. 1980.
- [10] J. N. Kapur, P. K. Sahoo, A. K. C. Wong, « A new method for gray-level picture thresholding using the entropy of the histogram », *Comput. Vis. Graph. Image Process.*, vol. 29, n° 3, pp. 273-285, mars 1985.
- [11] H. D. Cheng, J.-R. Chen, J. Li, « Threshold selection based on fuzzy c-partition entropy approach », *Pattern Recognit.*, vol. 31, n° 7, pp. 857-870, juill. 1998.
- [12] C. H. Li, C. K. Lee, « Minimum cross entropy thresholding », *Pattern Recognit.*, vol. 26, n° 4, pp. 617-625, avr. 1993.
- [13] J. Bernsen, « Dynamic thresholding of gray level images », in *Proceedings of the International Conference on Pattern Recognition (ICPR '86)*, pp. 1251-1255, 1986.
- [14] W. Niblack, *An introduction to digital image processing*. Englewood Cliffs, N.J: Prentice-Hall International, 1986.
- [15] J. Sauvola, M. Pietikäinen, « Adaptive document image binarization », *Pattern Recognit.*, vol. 33, n° 2, pp. 225-236, févr. 2000.
- [16] C. Wolf, J. M. Jolion, F. Chassaing, « Extraction de texte dans des vidéos : le cas de la binarisation », *Proceedings of 13ème Congrès Francophone de Reconnaissance des Formes et Intelligence Artificielle*, pp. 145-152, 2002.

- [17] K. Khurshid, I. Siddiqi, C. Faure, N. Vincent, « Comparison of Niblack inspired binarization methods for ancient documents », présenté à IS&T/SPIE Electronic Imaging, San Jose, CA, 2009, pp. 724.
- [18] B. Gatos, I. Pratikakis, S. J. Perantonis, « Adaptive degraded document image binarization », *Pattern Recognit.*, vol. 39, n° 3, pp. 317-327, mars 2006.
- [19] S. Lu, B. Su, C. L. Tan, « Document image binarization using background estimation and stroke edges », *Int. J. Doc. Anal. Recognit. IJDAR*, vol. 13, n° 4, pp. 303-314, déc. 2010.
- [20] S. J. Lu, C. L. Tan, « Binarization of Badly Illuminated Document Images through Shading Estimation and Compensation », in *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*, Curitiba, Parana, Brazil, 2007, pp. 312-316.
- [21] R. F. Moghaddam, M. Cheriet, « Application of Multi-Level Classifiers and Clustering for Automatic Word Spotting in Historical Document Images », in *2009 10th International Conference on Document Analysis and Recognition*, Barcelona, Spain, 2009, pp. 511-515.
- [22] Q. Chen, Q. Sun, P. Ann Heng, D. Xia, « A double-threshold image binarization method based on edge detector », *Pattern Recognit.*, vol. 41, n° 4, pp. 1254-1267, avr. 2008.
- [23] B. Su, S. Lu, C. L. Tan, « Binarization of historical document images using the local maximum and minimum », in *Proceedings of the 8th IAPR International Workshop on Document Analysis Systems - DAS '10*, Boston, Massachusetts, 2010, pp. 159-166.
- [24] M. van Herk, « A fast algorithm for local minimum and maximum filters on rectangular and octagonal kernels », *Pattern Recognit. Lett.*, vol. 13, n° 7, pp. 517-521, juill. 1992.
- [25] R. Hedjam, R. F. Moghaddam, M. Cheriet, « A spatially adaptive statistical method for the binarization of historical manuscripts and degraded document images », *Pattern Recognit.*, vol. 44, n° 9, pp. 2184-2196, sept. 2011.
- [26] B. Bataineh, S. N. H. S. Abdullah, K. Omar, « An adaptive local binarization method for document images based on a novel thresholding method and dynamic windows », *Pattern Recognit. Lett.*, vol. 32, n° 14, pp. 1805-1813, oct. 2011.

## CHAPITRE 3

BINARISATION PAR ESTIMATION DU BACKGROUND ET SELECTION  
AUTOMATIQUE DE FENETRE DE TRAITEMENT

### 3.1. Introduction

Comme nous l'avons déjà présenté, la binarisation est souvent utilisée dans le traitement et l'analyse des documents anciens souffrant de différentes dégradations. Elle est considérée comme une étape de base indispensable et cruciale dans ce domaine, qui doit être rapide et précise.

Dans la littérature, les techniques de binarisation, que ce soit nouvelles ou anciennes, donnent des résultats presque identiques dans le cas des images non dégradées. Le problème se pose dans le cas des images dégradées ou fortement dégradées, où les sources de bruits ne sont pas du même type. Dans ce cas l'approche hybride est la plus adoptée, qui combine en même temps les techniques globales et locales. Cette approche est basée sur le principe qu'un algorithme hybride hérite la rapidité des méthodes globales et la précision des algorithmes locaux.

Dans ce chapitre nous allons présenter la méthode (*Improved Degraded Document Image Binarization Using Median Filter for Background Estimation*) [1] concurrente à la méthode du polynôme smoothing proposée par Lu et al. [2]. En utilisant des outils d'évaluation performants, tels que la F-Measure, PSNR, DRD, NRM, MPM et P-Recal, nous allons démontrer que cette approche est plus efficace pour différentes images ayant différentes tailles et niveaux de dégradation.

### 3.2. Binarisation des documents ancien par utilisation du filtre médian pour l'estimation du background

Dans les méthodes de seuillage adaptatif rapportées dans [2-8], certains utilisent le contour pour détecter la limite du texte. Chen et al. proposent un algorithme [3] qui obtient une image de document binaire primaire. Il se base sur l'information déterminée à travers l'image résultante donnée par une étape de détection et fermeture des contours du texte de l'image. Moghaddam et al. [4] utilisent plutôt le profil de contour pour localiser la région de texte. Alors que Su et al. proposent dans [5] le calcul du maximum et minimum local du contraste pour déterminer le contour du texte. D'autres méthodes utilisent l'estimation du background pour calculer un seuil. La méthode proposée par Gatos et al. [6] estime le background du document par le calcul du seuil par la méthode de Sauvola [7]. Moghaddam et al. [8] montrent que l'estimation de l'arrière-plan du document se fait à partir de la moyenne itérative de l'image adaptative. Quant à Lu et al. [2] estiment tout d'abord le background du document par une procédure de lissage polynomiale itérative pour compenser le contraste,

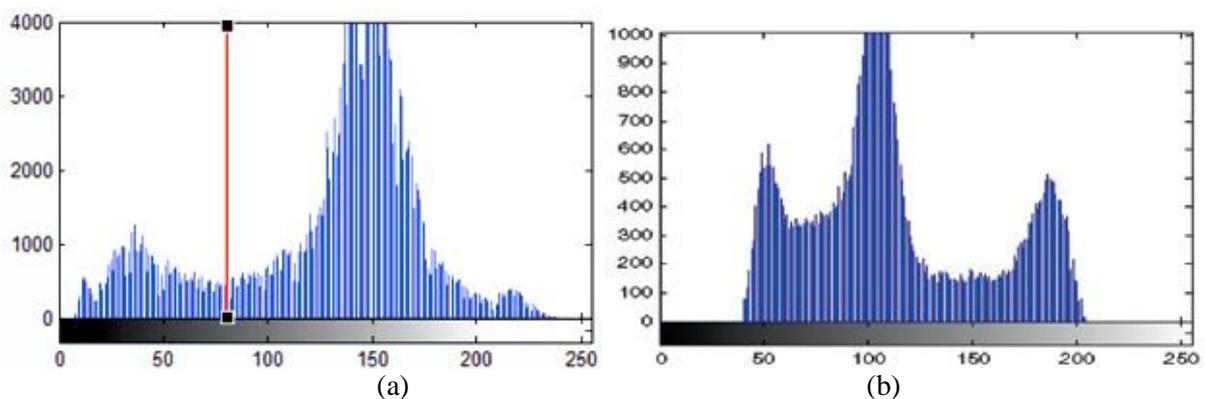
puis ils extraient le contour du texte en combinant le gradient dans les deux directions verticale.

La méthode proposée utilise le filtre médian à fenêtres variables, pour estimer l'arrière-plan et traiter l'image en général. Ce type de filtre a été couramment utilisé dans la littérature comme dans le cas des méthodes décrites dans [9-13] pour faire seulement un post-traitement afin d'améliorer les résultats de binarisation d'images de documents. L'originalité de la méthode proposée réside dans le fait que les tailles des fenêtres seront déterminées d'une manière intelligente en fonction du type de document ancien et l'état de dégradation. L'aspect nouveauté de la méthode concernant l'estimation du background se résume selon les étapes suivantes :

- Détection du type du document.
- Estimation de l'arrière-plan.
- Compensation du contraste de l'image.
- Estimation du seuil globale.
- Poste traitement pour enlever les pixels isolé.

### 3.2.1. Détection du type du document

Les images de document ancien se divisent en deux types selon le niveau de dégradation ; images non dégradées et images dégradées ou fortement dégradées. Nous allons montrer comment l'algorithme arrive d'une façon intelligente à obtenir les dimensions de fenêtre de traitement et par conséquent déterminer les paramètres du filtre (algorithme Sharpeak [14]) en analysant le degré de dégradation de l'image. L'allure de l'histogramme détermine selon le nombre de vallées et de pics aigus l'état de l'image à traiter. Les structures suivantes montrent quelques cas de figure.



**Figure 3.1:** Type d'histogramme : (a) Histogramme bimodale, (b) Histogramme multimodale

**1: Shar-Peak** ( $h$ )

**Input:**  $h$  -- Histogramme de l'image originale

**Return:**  $N\_peak$  -- Nombre de pique aiguë

2:  $freq \leftarrow h$

3: *for*  $k \leftarrow 0 .. 255$  *do*

4:   *if* ( $k = 0$ ) *do*

5:      $P_1(k) = 255, P_2(k) = 254, N_1(k) = 1, N_2(k) = 2$

6:   *else if* ( $k = 1$ ) *do* -- Les positions sont prises de manière circulaire où  $P_1, P_2$  et  $N_1, N_2$  représente respectivement les deux positions précédentes et suivantes pour chaque valeur de niveau de gris  $k$

7:      $P_1(k) = 0, P_2(k) = 255, N_1(k) = 2, N_2(k) = 3$

8:   *else if* ( $k = 255$ ) *do*

9:      $P_1(k) = 254, P_2(k) = 253, N_1(k) = 0, N_2(k) = 1$

10:   *else if* ( $k = 254$ ) *do*

11:      $P_1(k) = 253, P_2(k) = 252, N_1(k) = 255, N_2(k) = 0$

12:   *else*

13:      $P_1(k) \leftarrow (k - 1), P_2(k) \leftarrow (k - 2), N_1(k) \leftarrow (k + 1), N_2(k) \leftarrow (k + 2)$

14:   *end if*

15: *end for*

16:  $n \leftarrow 0$

17: *for*  $k \leftarrow 0..255$  *do*

18:   *if*  $freq(k) > \max(freq(P_1(k)), freq(P_2(k)), freq(N_1(k)), freq(N_2(k)))$  *do*

19:      $n \leftarrow n + 1$  -- Le nombre des piques

20:      $P(n) \leftarrow freq(k)$

21:   *end if*

22: *end for*

23:  $Avg\_peak \leftarrow \frac{\sum P}{n}$  -- Valeur moyenne des piques

23:  $N\_peak \leftarrow find (P > Avg\_peak)$  -- Valeur des piques aiguës

24: *return*  $N\_peak$

**Algorithme 3.1:** Algorithme Shar-Peak.

### 3.2.2. Estimation de l'arrière-plan

L'estimation de fond par la méthode de lissage polynomial a été utilisée dans plusieurs travaux pour différentes applications. Krzysztof et al. [15] pour le traitement des images d'empreintes digitales, Seeger et al. [16] pour la pré-détection de régions de texte. Dans la méthode proposée, nous estimons le fond du document en utilisant un lissage basé sur un calcul itératif de la médiane. Nous allons démontrer à travers les résultats que cette nouvelle méthode fournit de bonnes estimations du background dépendamment du choix de la taille de la fenêtre de traitement. Il est à noter qu'une méthode similaire mais non intelligente, utilisant les mêmes notions, a été utilisée pour les opérations de filtrage dans les travaux [9-13].

Il est à rappeler que dans le cas du traitement d'image, le filtre médian est une opération non linéaire qui parcourt tous les pixels de niveau de gris  $f(x, y)$  et leurs voisinages déterminés en faisant glisser une fenêtre  $w$  de dimension impaire  $r \times r$  sur l'image entière. Cette fenêtre comprend  $N$  pixels avec des coordonnées  $(x_0 \pm r, y_0 \pm r)$ . La médiane sera calculée en triant d'abord toutes les valeurs de niveau de gris des pixels de la fenêtre centrée sur le pixel  $f(x_0, y_0)$  en ordre numérique croissant. Le pixel central sera remplacé par la valeur médian du vecteur formé par le tri.

Mathématiquement parlant, le vecteur trié et ordonné, noté par  $f_i$  est caractérisé par :

$$f_1 < f_2 < \dots < f_{\frac{(N+1)}{2}} < f_{(N-1)} < f_N \quad (3.1)$$

Le filtrage consiste donc à remplacer la valeur du pixel  $f(x_0, y_0)$  par la médiane  $f_{median}$ , représentant la valeur médiane de son voisinage calculée par l'équation suivante:

$$f_{median}(i, j) = \begin{cases} f_{\left(\frac{N+1}{2}\right)} \\ \frac{1}{2} \left( f_{\left(\frac{N}{2}\right)} + f_{\left(\frac{N}{2}+1\right)} \right) \end{cases} \quad (3.2)$$

où:

$N$  : Nombre d'éléments de la fenêtre  $w$ .

$f_N$  : Intensité d'un pixel de la fenêtre  $w$ .

Pour justifier le choix des tailles de fenêtre de traitement, nous nous sommes référé aux travaux de [17-19] dans lesquels la taille de la fenêtre a été choisie entre 10 et 100 pixels et aux travaux de Gatos et al [6] qui suggèrent une taille de fenêtre de  $60 \times 60$  afin de couvrir un ou deux caractères. Généralement, la détection de la taille d'un caractère nécessite une segmentation d'image [20] et ceci étant difficile pour les documents dégradés, puisque la valeur de la taille de la fenêtre est un paramètre très important [21], qui influe sur l'obtention de bons résultats de binarisation. Moghaddam et Cheriet [22] ont proposé une méthode qui commence avec une grande taille de fenêtre et se réduit itérativement à une taille appropriée. Notre méthode utilise les informations fournies par l'algorithme Sharp-Peak décrit dans [14] pour décider automatiquement du type de l'image et donc, en fonction des observations expérimentales, nous fixons la taille de la fenêtre à  $40 \times 40$  ou  $100 \times 100$  selon le niveau de la dégradation. La figure 3.2 (d) montre l'effet résultant de l'estimation de l'arrière-plan par la méthode proposée appliquée à l'image de la figure 3.2 (a) qui contient la transparence du texte. Comparée à la méthode de lissage polynomiale de Lu et al. [2], nous remarquons que l'image obtenue sur la figure 3.2 (c) est uniforme, contrairement à la méthode proposée dans laquelle l'image résultante n'est pas uniforme dans la zone du côté du texte. Ainsi, la méthode proposée peut fournir directement un meilleur résultat d'estimation de ces zones. Cependant, la méthode de lissage polynomial est réalisée en plusieurs itérations doublées (ligne/ligne et colonne/colonne).



**Figure 3.2:** (a) Image originale. (b) Estimation du background par le lissage polynomial ligne par ligne [2]. (c) Estimation du background par le lissage polynomial colonne par colonne [2]. (d) Estimation du background par la méthode proposée.

### 3.2.3. Compensation du contraste de l'image

Si l'image originale contient des points d'ancre, des zones de texte avec un fond sombre figure 3.3 (a), ou même un éclairage inégal du fond, alors l'estimation du fond de la figure 3.3 (b) sera utilisée pour la compensation du contraste du document. Ceci se fait par l'application

de l'équation (3), les conséquences donnent un changement des zones sombres en zones claires et un texte plus lisible et séparé de l'arrière-plan figure 3.3 (c).

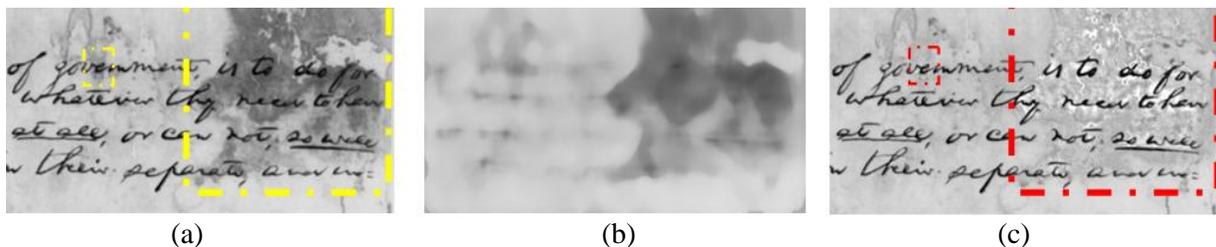
$$I_c = \frac{C}{I_{bG}} \times I = \alpha \times I \quad (3.3)$$

$C$  : Constante qui contrôle la luminosité du document d'arrière-plan compensé qui sera calculé par la médiane de l'image originale.

$I_{bG}$  : L'image de l'estimation de l'arrière-plan.

$I$  : L'image originale.

Ainsi, la variation de l'image de fond originale sera compensée par le facteur  $\alpha$ , qui sera plus élevé dans la zone la plus sombre du document (faible valeur de  $I_{bG}$  dans cette zone) et inversement dans la zone claire, elle sera faible (valeur élevée de  $I_{bG}$ ).



**Figure 3.3:** Compensation du contraste: (a) Image originale. (b) Background estimé par la méthode proposée. (c) Image compensée

### 3.2.4. Estimation du seuil global

Une fois la compensation du background est obtenue (étape précédente), un seuillage par une méthode globale (Otsu dans notre cas) est appliqué à l'image. Ce choix est lié à la non complexité et au paramètres temporels. Nous rappelons que le seuillage global consiste à calculer un seuil optimal à partir d'une mesure globale sur l'ensemble de l'image.

Nous adoptons ici la méthode d'Otsu [23]. Qui fournit un seuil optimal calculé à partir de la maximisation de la variance entre classes. Cette méthode a été utilisée pour séparer l'image en niveaux de gris en deux classes de premier plan et d'arrière-plan.

### 3.2.5. Étape de poste traitement

L'étape de post-traitement est indispensable. Lors de cette phase, tous les points noirs ou blancs séparés et isolés seront éliminés par les deux opérations suivantes :

- **Dilatation** : elle consiste à éliminer les points "noirs" isolés. La méthode consiste à faire balayer l'image par une fenêtre  $(2n + 1)^2$  qui est centré sur le pixel courant, et

appliquer une opération OU logique sur les deux fois  $(2n + 1)$  pixels de voisinage. Si le résultat est 1, le pixel est mis à 1 tandis que si le résultat est 0, le pixel est conservé.

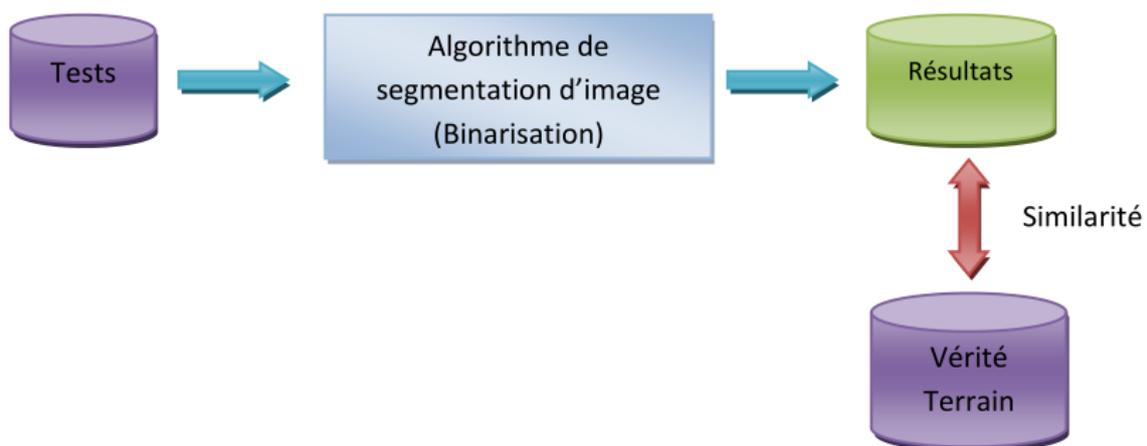
- **Erosion** : Il s'agit ici d'éliminer les points "blancs" isolés. La méthode est assez similaire à celle de la dilatation sauf qu'il faut utiliser une opération ET logique au lieu de OU logique. Si le résultat est 1, le pixel est conservé tandis que si le résultat est 0, le pixel est mis à 0.

Remarque : L'opération d'érosion précède toujours l'opération de dilatation.

### 3.3. Bibliothèque et outils d'évaluation des images

La diversité des méthodes de binarisation des images de documents anciens exige d'avoir des outils métriques d'évaluation fiable pour mesurer le degré de similarité entre les images résultantes des méthodes de binarisation et les images de références voir figure 3.4, Ces outils métriques, dites de performance nous permettent de faire la comparaison entre les différentes méthodes de binarisation d'images de document ancien. Néanmoins, il n'existe pas vraiment un outil de mesure absolu de la performance d'une méthode. Cependant un maximum d'outils combiné d'évaluation peut être mis à notre disposition [24] pour classer les méthodes de point de « vue performance ».

Dans cette partie nous allons définir les différents critères et outils d'évaluations utilisés et voir leurs impacts sur la classification des résultats expérimentaux entre les différentes méthodes et la méthode proposée.



**Figure 3.4** : Evaluation d'un algorithme de binarisation.

Le problème d'évaluation a intéressé plusieurs chercheurs dans le domaine du traitement d'images [25]. Ces chercheurs ont proposé plusieurs critères d'évaluation et les ont regroupés en deux classes principales selon l'existence ou non d'image de références ou dites

*vérité-terrain*. C'est-à-dire : critères d'évaluation non-supervisés et critères d'évaluation supervisés [26].

Dans notre cas, nous avons opté pour une évaluation hybride (supervisé et non-supervisé). Ce choix est utilisé dans la plupart des travaux de recherche.

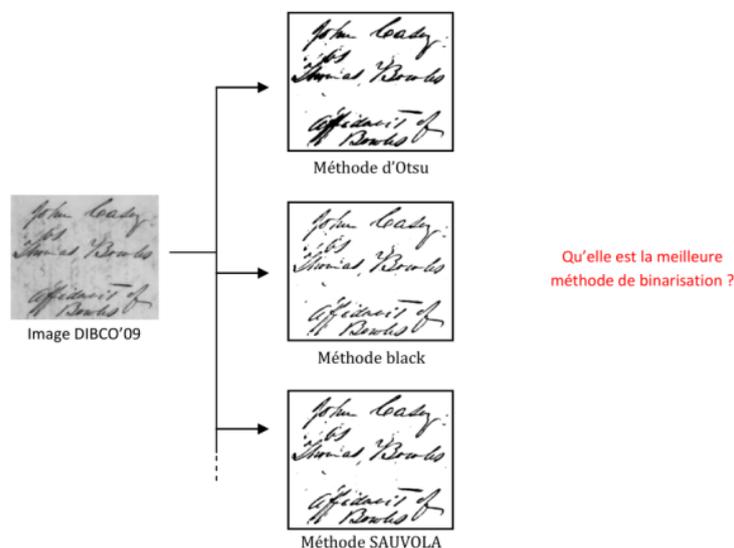
### 3.3.1. Bibliothèque des images

Dans le domaine de la binarisation, des bibliothèques ou base de données particulières qui contiennent un ensemble d'images à différentes dégradations sont utilisées pour tester et comparer des algorithmes de binarisation. Ces bibliothèques se divisent en deux groupes :

Le premier regroupe les images originales, ou images de test, Le second contient les images de référence, appelées *vérité-terrain*. Ces images sont construites d'une façon manuelle par des experts [27].

### 3.3.2. Evaluation des méthodes de binarisation

L'évaluation des méthodes de binarisation se fait grâce à des descripteurs reposant sur une comparaison entre l'image à mesurer et celle de référence, sans l'image de référence on ne peut pas déterminer quel algorithme donne le meilleur résultat étant donné que l'observateur n'est pas déterminant figure 3.5.



**Figure 3.5 :** Évaluation des méthodes de binarisation

Il existe trois approches d'évaluation de la qualité de l'image résultante des méthodes de binarisation. La première approche est purement mathématique, elle ignore l'observateur. La deuxième repose sur des critères complètement subjectifs et fait appel à une série de protocoles lourds et coûteux, la dernière combine entre les deux aspects.

Cependant, malgré l'effort considérable déployé dans la recherche de critères objectifs, ou hybrides de qualité de l'image, il n'existe pas, à l'heure actuelle une mesure de qualité objective qui est universelle. La métrique de qualité objective la plus utilisée est le PSNR (Peak Signal to Noise Ratio).

L'insuffisance de ce genre de mesures purement mathématiques a conduit au développement d'autres métriques inspirées ou basée sur quelques connaissances limitées sur des mécanismes de la vision humaine. En effet, la prise en compte de l'appréciation subjective de la qualité visuelle dans la mise au point de métriques de qualité est primordiale.

Dans chaque approche, il y a trois types d'évaluation de qualité d'image. La première, et la plus développée. Elle concerne les métriques de qualité avec référence. La seconde regroupe les mesures de qualité avec référence réduite. Dans ce cas on dispose d'une information a priori sur l'image originale (caractéristiques spatiales : les contours des objets ou des statistiques sur la distribution des paramètres extrait de l'image ou de sa transformée). Enfin, le troisième concerne les métriques de qualité où l'on ne dispose pas d'information sur l'image originale. Dans ce dernier cas, on parle de méthode aveugle où l'on évalue le niveau de dégradation de façon ad-hoc ou en s'appuyant sur des modèles de dégradation.

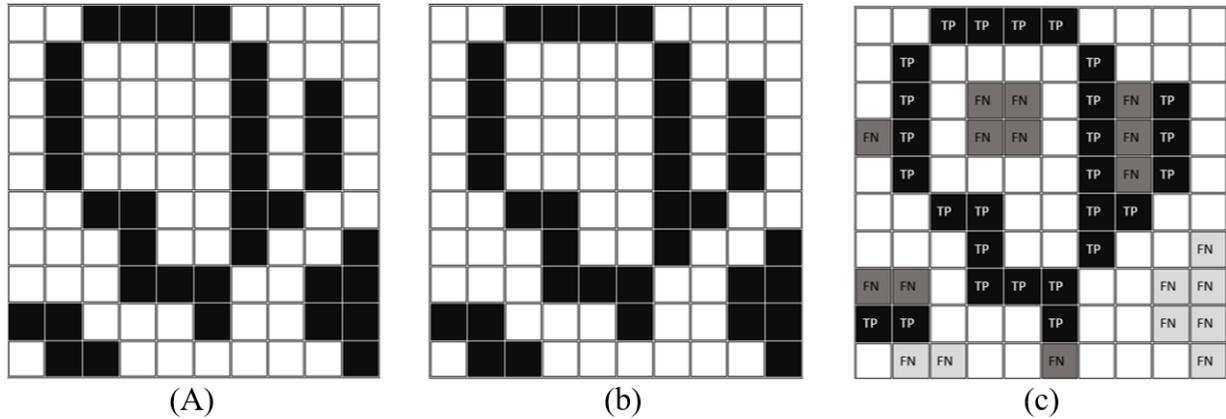
Les mesures d'évaluation utilisées dans la littérature ont été utilisées dans les compétitions DIBCO et H-DIBCO. L'évaluation de la similarité (ou la différence) de l'image binarisée et l'image de référence peut être réalisée sur la base de nombreuses mesures métriques. Dans le cadre du concours DIBCO 2009 [28], quatre mesures d'évaluation ont été introduites : la F-measure, PSNR, NRM et MPM. La compétition H-DIBCO 2010 [29] a vu l'introduction de la P-Recall pour produire une cinquième mesure qui est le pseudo F-Measure. Dans la DIBCO 2011 [30] une nouvelle mesure appelée DRD a été utilisée.

### ➤ **F-Measure**

La F-measure «FM» est une mesure d'évaluation utilisée dans plusieurs domaines, surtout pour la récupération de donnée ou d'information. Elle est calculée au niveau des pixels par la moyenne harmonique de la précision (p) et du rappel (r). Chaque pixel est classé comme suite :

- Vrai positive (VP) : on dit que le pixel est VP lorsqu'il est noir dans l'image de référence et aussi dans l'image binarisée.
- Faut positive (FP) : on dit que le pixel est FP lorsqu'il est blanc dans l'image de référence et noir dans l'image binarisée.

- Faut positive (FP) : on dit que le pixel est FP lorsqu'il est blanc dans l'image de référence et noir dans l'image binarisée.
- Faut négative (FN) : on dit que le pixel est FN lorsqu'il est noir dans l'image de référence et blanc dans l'image binarisée.



**Figure 3.6** : Exemple pour le calcul de la F-mesure.

La F-mesure se calcule par l'équation suivante :

$$FM = 2 \times \frac{recall \times precision}{recall + precision} \times 100\% \quad (3.4)$$

$$recall = \frac{VP}{VP + FN} \quad et \quad precision = \frac{VP}{VP + FP} \quad (3.5)$$

Le rappel est la proportion des pixels de l'avant plan correctement binarisée. Lorsque le rappel est à 100%, il n'y a pas de faux négatifs et donc aucun pixel de l'avant plan n'a été incorrectement classé comme arrière-plan. Pour la précision lorsqu'elle est de 100% il n'y a pas de faux positifs et donc aucun pixel de l'arrière-plan n'a été incorrectement classé comme avant plan.

Une mesure FM élevée signifie une meilleure correspondance entre l'image binarisée et l'image de référence.

### ➤ PSNR (Peak Signal to Noise Ratio)

C'est une mesure de distorsion utilisée en image numérique. Il s'agit de quantifier la performance de la binarisation en mesurant la qualité de reconstruction de l'image binarisée par rapport à l'image de référence. Le PSNR est calculé par :

$$PSNR = 10 \times \log_{10} \frac{d^2}{EQM} \quad (3.6)$$

Où  $d$  est la dynamique du signal dans notre cas nous travaillons avec des images aux niveaux de gris où chaque pixel est codé sur 8 bits alors  $d = 255$ .

$EQM$  est l'erreur quadratique moyenne, elle est définie par les deux images, binarisée et de référence de taille  $m \times n$

$$EQM = \frac{1}{m \times n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (I_{référence(i,j)} - I_{binarisée}(i, j))^2 \quad (3.7)$$

Le PSNR est inversement proportionnel à l'erreur quadratique moyenne entre l'image binarisée et l'image de référence alors pour une valeur basse du  $EQM$ , un PSNR élevé. Cela signifie une meilleure correspondance entre l'image binarisée et l'image de référence.

➤ **NRM (Negative Rate Metric)**

Le NRM est basé sur les discordances des pixels entre l'image binarisée et l'image de référence où il combine le taux de faux négatifs  $NR_{FN}$  et le taux de faux positifs  $NR_{FP}$ . Il est calculé comme suit :

$$EQM = \frac{1}{m \times n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (I_{référence(i,j)} - I_{binarisée}(i, j))^2 \quad (3.8)$$

$$NRM = \frac{NR_{FN} + NR_{FP}}{2} \quad (3.9)$$

Où

$$NR_{FN} = \frac{FN}{FN + VP} \quad NR_{FP} = \frac{FP}{FP + VN} \quad (3.10)$$

Contrairement à la F-measure et le PSNR, pour une valeur basse du NRM cela signifie une meilleure correspondance entre l'image binarisée et l'image de référence.

➤ **MPM (Misclassification Penalty Metric)**

La métrique MPM est utilisée pour évaluer la reconstitution du contour de l'objet de l'image binarisée par rapport à l'image de référence. Le calcul du MPM se fait par l'équation suivante :

$$MPM = \frac{MP_{FN} + MP_{FP}}{2} \quad (3.11)$$

Où

$$MP_{FN} = \frac{\sum_{i=1}^{FN} d_{FN}^i}{D} \quad MP_{FP} = \frac{\sum_{i=1}^{FP} d_{FP}^i}{D} \quad (3.12)$$

$d_{FN}^i, d_{FP}^i$  expriment la distance de  $i^{eme}$  faut négative et  $j^{eme}$  faut positive des pixels de la segmentation de contour de l'objet de l'image de référence.  $D$  est la somme des pixels formant le contour de l'objet dans l'image de référence.

Une valeur basse du MPM signifie une meilleure détection du contour de l'objet entre l'image binarisée et l'image de référence.

### ➤ DRD (Distance-Reciprocal Distortion metric)

La métrique de la distance réciproque de distorsion (DRD) a été utilisée pour mesurer la distorsion visuelle dans les images de documents binaires [25]. Le DRD se calcule comme suit :

$$DED = \frac{\sum_{k=1}^S DRD_K}{NUBN} \quad (3.13)$$

$$DRD_K = \sum_{i=-2}^2 \sum_{j=-2}^2 \left| GT_K(i, j) - B_K(i, j) \right| \times W_{NM}(i, j) \quad (3.14)$$

Où  $GT$  est l'image de référence et  $B_k$  est l'image binarisée et  $W_{nm}$  est la matrice normalisé définie en [31] en effet cette métrique s'est révélée être corrélée avec la perception humaine du niveau de dégradation de l'image binarisée.

Une valeur basse du DRD signifie une meilleure correspondance entre l'image binarisée et l'image de référence.

### 3.4. Expérimental et résultat

Pour valider l'efficacité de notre méthode surtout pour les images fortement dégradées en présence de taches d'encre et d'un fond non uniformément éclairé, une étude comparative a été accomplie.

Dans cette comparaison nous avons utilisé deux groupes d'images de vieux documents (Handwritten & Printed) appartenant à différentes bases de données de DIBCO dans [28], [30], [32] et la base de données BICKLEY DIARY [33] connue pour ses images très détériorées. Les différentes bases de données rassemblent pour chaque image originale une

image de référence représentant le résultat idéal de la binarisation car certaines mesures d'évaluation l'exigent.

Nous comparons les performances de la méthode proposée avec celles des méthodes classiques de binarisation, les méthodes récentes de l'état de l'art et les algorithmes les mieux classés dans chaque compétition de DIBCO. La comparaison et le test sont effectués en termes des mesures d'évaluation largement utilisés dans plusieurs travaux qui concernent la binarisation des documents anciens. Ces mesures consistent en:

- F-measure.
- PSNR "Peak Signal to Noise Ratio".
- NRM
- MPM
- DRD
- Précision

### 3.4.1. Test sur les différentes bases de données DIBCO

Le concours DIBCO a attiré 35 groupes de recherche qui sont actuellement actifs dans l'analyse d'image qui a prouvé l'importance des méthodes de binarisation comme une étape préliminaire vers la reconnaissance d'un document, d'où la nécessité de poursuivre une référence qui mènera à une évaluation significative et objective. Ces images sont issues du corpus DIBCO [28] qui propose des images des documents fortement dégradés accompagnés d'une annotation contenant une binarisation manuelle (*vérité-terrain*). Chaque pixel d'encre a été labellisé par un expert.

Le premier ensemble d'images utilisé dans les expérimentations est constitué des trois collections d'images proposées dans le cadre des compétitions internationales de binarisation DIBCO 2009 [28], DIBCO 2011 [30] et DIBCO 2013 [32]. DIBCO (Document Image Binarization Contest) et H-DIBCO (Handwritten Document Image Binarization Competition) sont des compétitions internationales de binarisation d'images de documents émergées en 2009 dans le contexte de la conférence ICDAR 2009 [28]. Le but de ces compétitions est de suivre les avancées dans le domaine de la binarisation d'images de documents et d'évaluer les performances en utilisant des mesures quantitatives sur la même base. Chaque collection contient un certain nombre d'images de documents dégradées, et d'images de *vérité-terrain* correspondantes créées semi automatiquement en utilisant la procédure proposée dans [34].

Les images incluses dans les cinq collections DIBCO/H- DIBCO ont été sélectionnées attentivement par les organisateurs pour qu'elles couvrent la plupart des dégradations liées aux documents anciens (intensité variable dans l'arrière-plan, faible contraste, ombre, taches d'encre et d'humidité, effet de transparence, plis et déchirures, etc.).

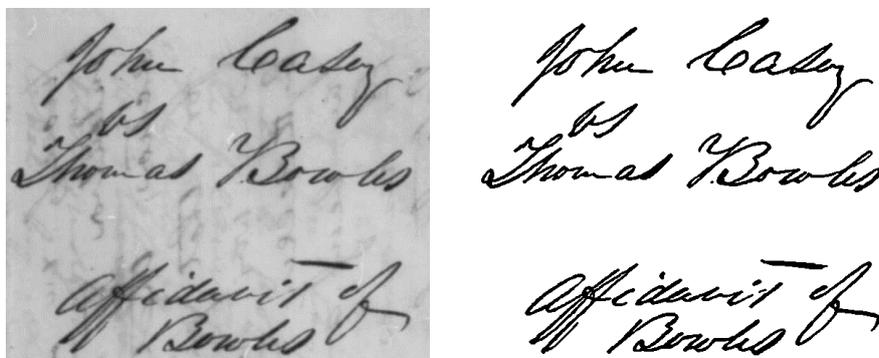
La collection de DIBCO 2009 contient 5 images de documents imprimés et 5 manuscrits et la collection de DIBCO 2011 comprend 8 images de documents imprimés et 8 manuscrits qui proviennent de plusieurs bibliothèques (bibliothèque de l'état et de l'université Göttingen, bibliothèque de l'état de Bavière, bibliothèque britannique et bibliothèque du Congrès), avec les images de vérité-terrain associées. L'ensemble d'images de DIBCO 2013 est composé de 8 images de documents imprimés et 8 manuscrits, et d'images de références associées.

Un total de 42 images est ainsi obtenu. Ces images sont réparties comme dans le tableau 3.1.

**Tableau 3.1:** Répartition d'images de documents dans les collections publiques de DIBCO

	Manuscrit	Imprimé	Somme
<b>DIBCO 2009</b>	5	5	10
<b>DIBCO 2011</b>	8	8	16
<b>DIBCO 2013</b>	8	8	16
<b>Somme</b>	21	21	42

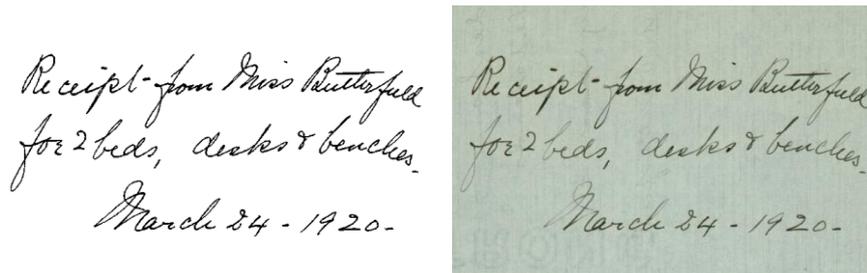
Les figures 3.7 à 3.12 présentent quelques exemples d'images extraites des collections DIBCO avec les images de vérité terrain correspondantes



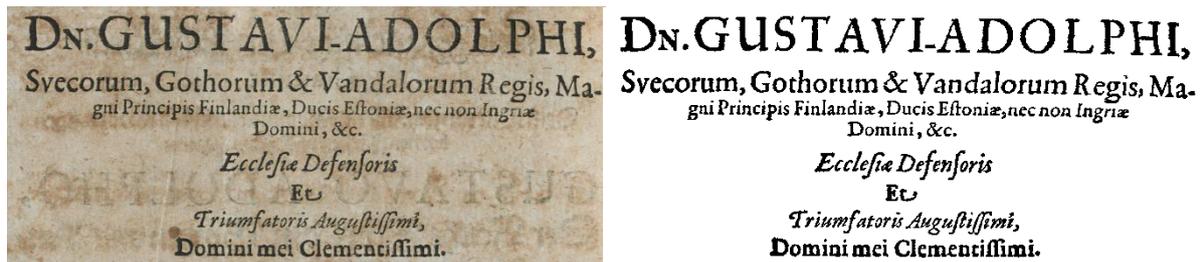
**Figure 3.7 :** Image originale de la collection DIBCO 2009 Handwritten avec l'image de vérité-terrain correspondante.



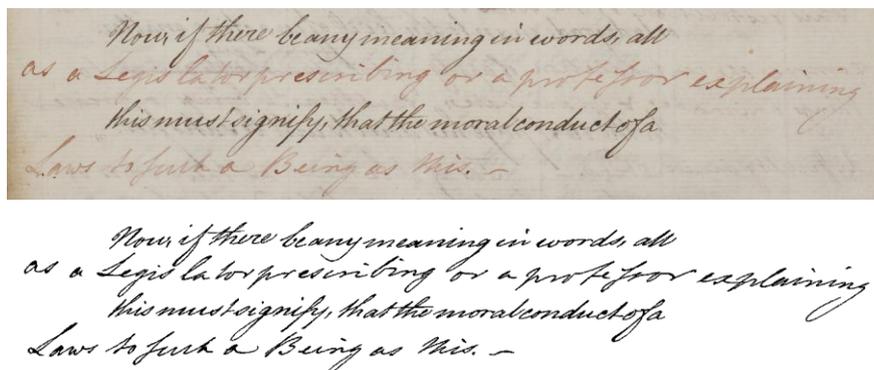
**Figure 3.8 :** Image originale de la collection DIBCO 2009 Printed avec l’image de vérité-terrain correspondante.



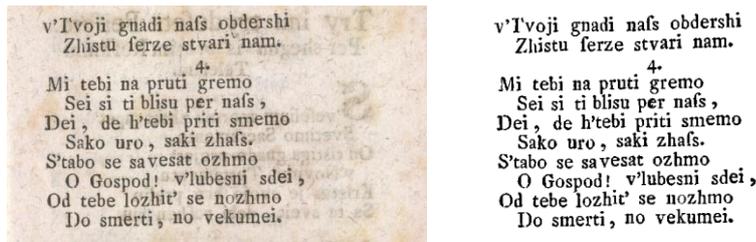
**Figure 3.9 :** Image originale de la collection DIBCO 2011 Handwritten avec l’image de vérité-terrain correspondante.



**Figure 3.10 :** Image originale de la collection DIBCO 2011 Printed avec l’image de vérité-terrain correspondante.



**Figure 3.11 :** Image originale de la collection DIBCO 2013 Handwritten avec l’image de vérité-terrain correspondante.

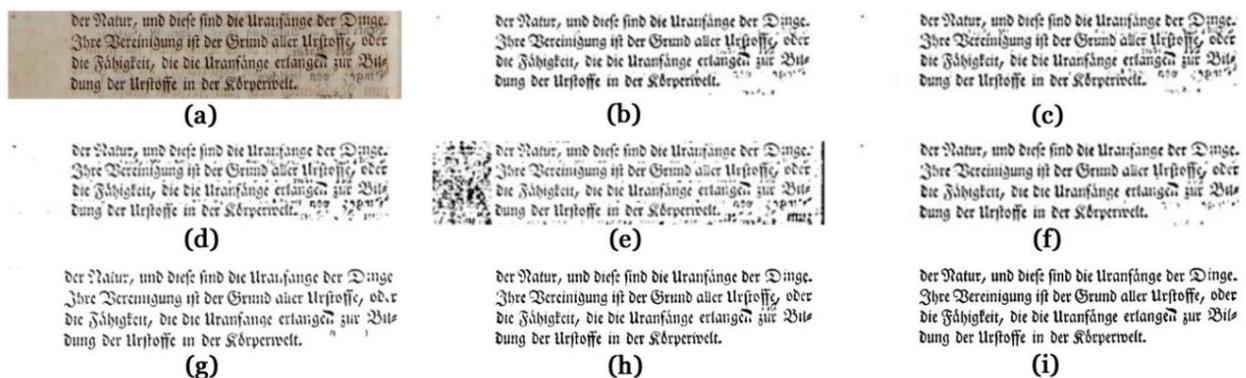


**Figure 3.12 :** Image originale de la collection DIBCO 2013 Printed avec l’image de *vérité-terrain* correspondante.

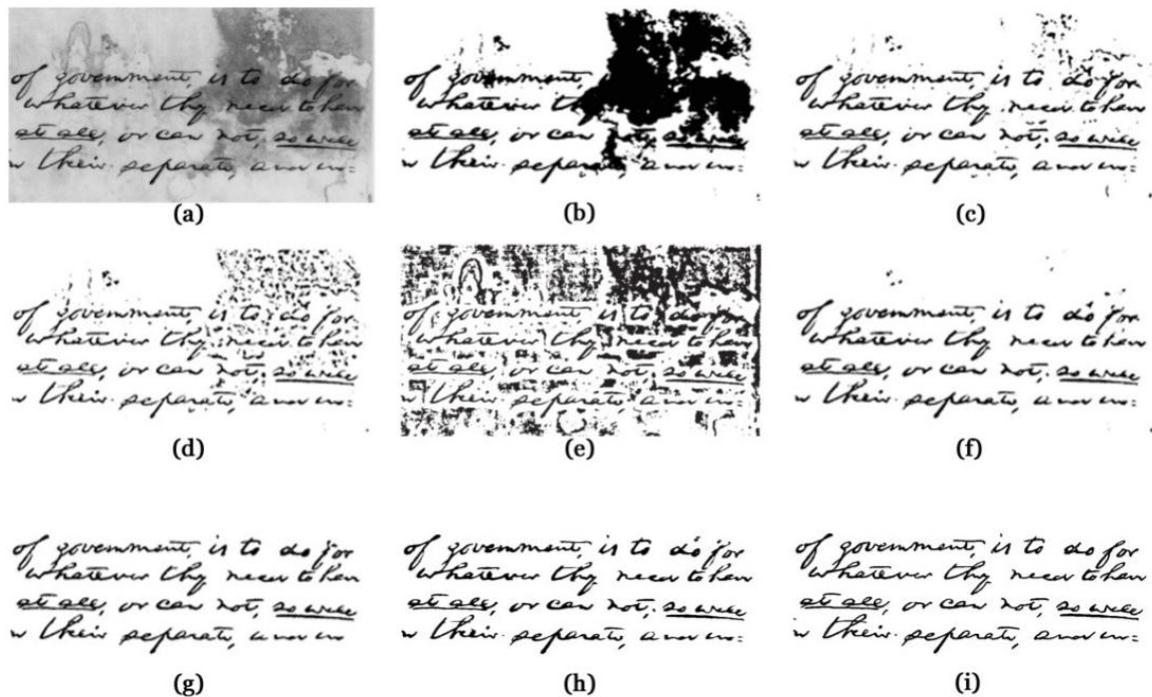
La méthode proposée permet d’obtenir des meilleurs résultats pour la FM et le PSNR ainsi que pour le NRP et MPM. Les figures 3.13 et 3.14 montrent deux exemples d’images de la bibliothèque DIBCO 2009 et leurs résultats de binarisation par différentes méthodes, sont présentés dans le tableau 3. 2.

**Tableau 3.2.** Résultat de mesure DIBCO 2009

	F-measure (%)	PSNR	NRM ( $\times 10^{-2}$ )	MPM( $\times 10^{-3}$ )
Otsu	78.72	15.34	5.77	13.30
Niblack	55.82	9.89	16.4	61.50
Sauvola	85.41	16.39	6.94	3.20
Bernsen	52.48	8.89	14.29	113.80
Gatos	85.25	16.5	10	0.70
BE	91.24	18.6	4.31	0.55
2 <sup>eme</sup> rank	90.06	18.23	4.75	0.89
Su	91.06	18.50	7	0.30
Méthode proposée	91.53	18.74	5.01	0.33



**Figure 3.13 :** Résultat de la binarisation par différentes méthodes (a) Image originale DIBCO 2009, (b) Otsu, (c) Sauvola,(d) Niblack,(e) Bernsen,(f) Gatos,(g) BE,(h) Méthode proposée, (i) Image de référence.



**Figure 3.14** : Résultat de la binarisation par différentes méthodes (a) Image originale DIBCO 2009, (b) Otsu,(c) Sauvola,(d) Niblack,(e) Bernsen,(f) Gatos,(g) BE, (h) Méthode proposée, (i) Image de référence.

Outre les méthodes de comparaison mentionnée auparavant, la méthode proposée a été également comparée avec trois algorithmes classés premier pour ICDAR 2011 [30], à savoir la méthode de Lore (LELO) [35], la méthode présentée par Su (SNUS) [36] et la méthode de Howe (HOWE) [37]. Les résultats des mesures sont présentés dans le tableau suivant.

**Tableau 3.3.** Résultat de mesure DIBCO 2011.

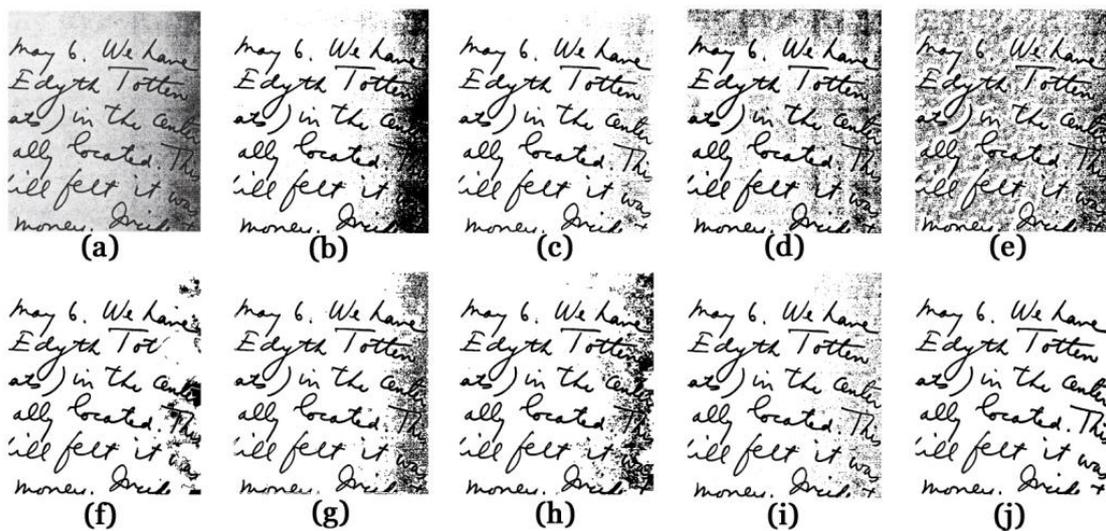
	F-measure (%)	PSNR	DRD	MPM
Otsu	82.22	15.77	8.72	15.64
Sauvola	82.54	15.78	8.09	9.20
Niblack	68.52	12.76	28.31	26.38
Bernsen	47.28	7.92	82.28	136.54
Gatos	82.11	16.04	5.42	7.13
BE	81.67	15.59	11.24	11.40
LELO 1 <sup>st</sup> rank	80.86	16.13	104.48	64.42
SNUS 2 <sup>nd</sup> rank	85.20	16.38	15.66	9.01
Howe 3 <sup>rd</sup> rank	88.74	17.84	5.37	8.64
Méthode proposée	86.30	16.89	4.88	5.66

La méthode proposée est la plus performante en termes de DRD et MPM, cela signifie que notre méthode engendre bien le contour du texte et fournit une meilleure qualité visuelle. Pour la F-measure et le PSNR, la méthode proposée donne des résultats qui effleurent les

meilleurs résultats des autres méthodes (tableau 3.3). Nous signalons que la méthode proposée donne de bons résultats sur toutes les images de test, ce qui reflète sur la valeur élevée de la F-measure et PSNR. Les figures 3.15 et 3.16 montre un exemple de binarisation pour la DIBCO 2011.

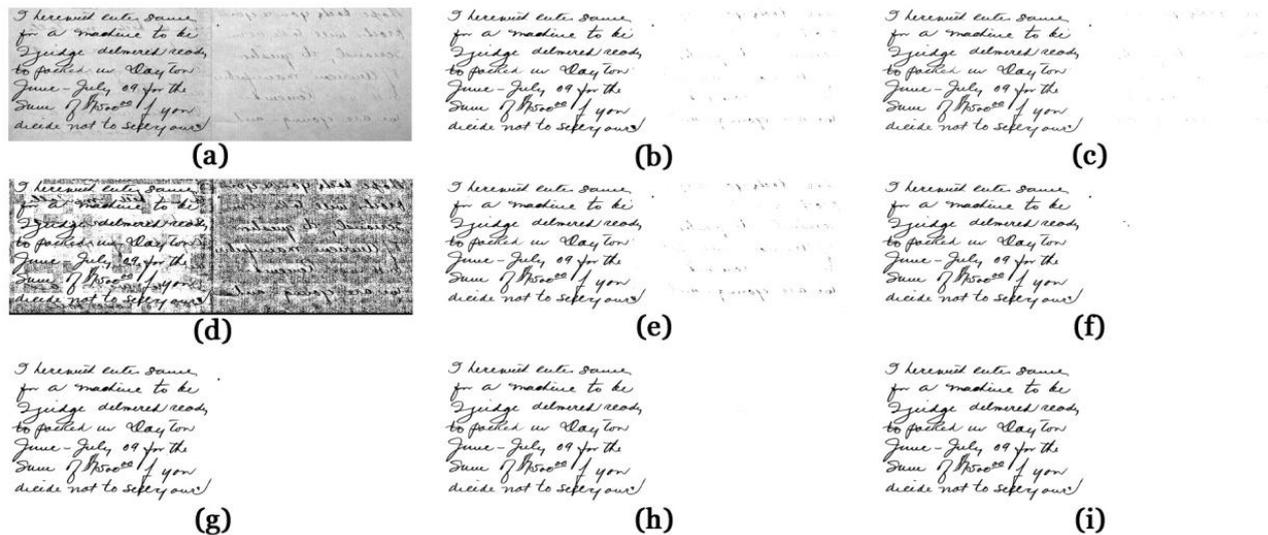


**Figure 3.15 :** Résultat de la binarisation par différentes méthodes (a) Image originale DIBCO 2011, (b) Otsu,(c) Sauvola,(d) Niblack,(e) Bernsen,(f) Gatos,(g) BE,(h) LELO,(i) Su SNUS , (j) Howe,(k) Méthode proposée,(l) Image de référence.



**Figure 3.16 :** Résultat de la binarisation par différentes méthodes (a)image originale DIBCO 2011, (b) Otsu,(c) Sauvola,(d) Niblack,(e) Bernsen,(f) LELO,(g) Su SNUS,(h) Howe, (i) Méthode proposée,(j) Image de référence.

Les figures 3.17 à 3.19 montrent d'autres exemples de trois images de la bibliothèque DIBCO 2013 et leurs résultats binaires obtenus par différentes méthodes (tableau 3.4). Tel qu'il est observé sur la Figure. 3.17, les méthodes Otsu, Sauvola, Bernsen introduisent un bruit de fond par contre la méthode Niblack récupère mais avec trop de bruit, pour les deux méthodes classées premières dans DIBCO 2013 et notre méthode proposée donne les même résultats.



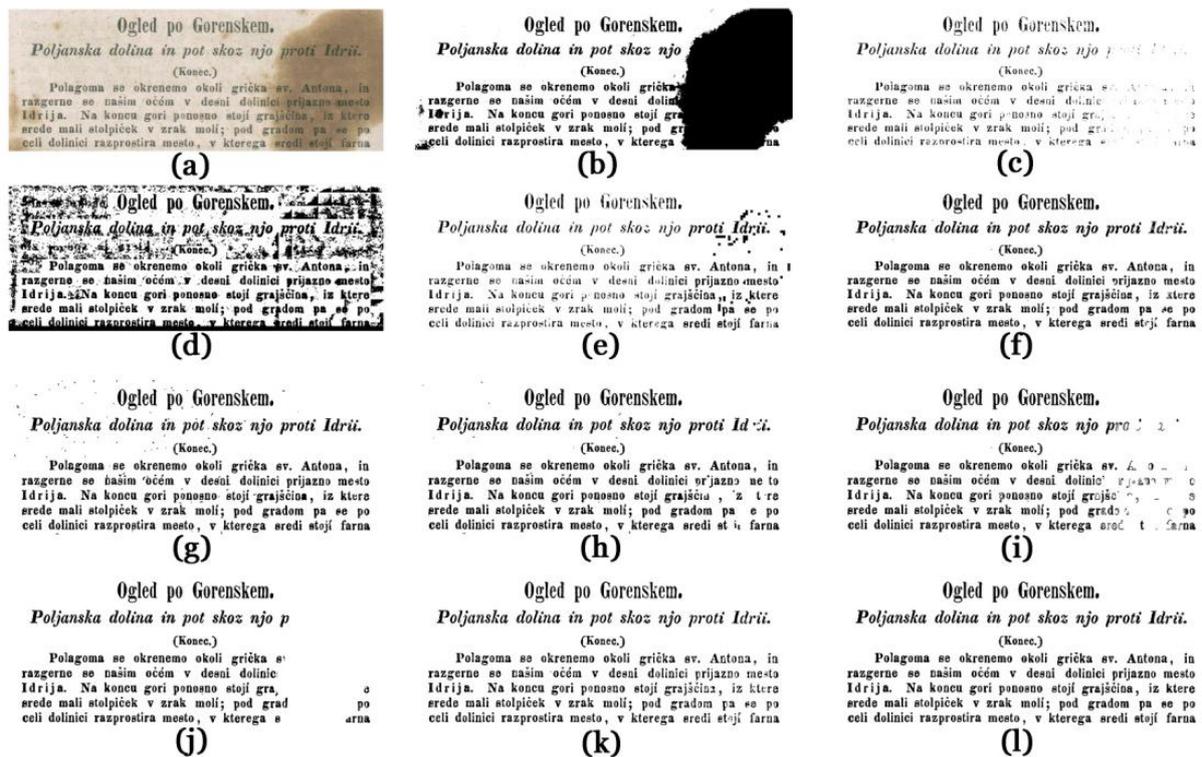
**Figure 3.17 :** Résultat de la binarisation par différentes méthodes (a) Image originale DIBCO13, (b) Otsu,(c) Sauvola,(d) Niblack,(e) Bernsen,(f) 1<sup>st</sup> Rank DIBCO 2013,(g) 2<sup>nd</sup> Rank DIBCO 2013,(h) Méthode proposée,(i) Image de référence.

**Tableau 3.4.** Résultat de mesure DIBCO2013.

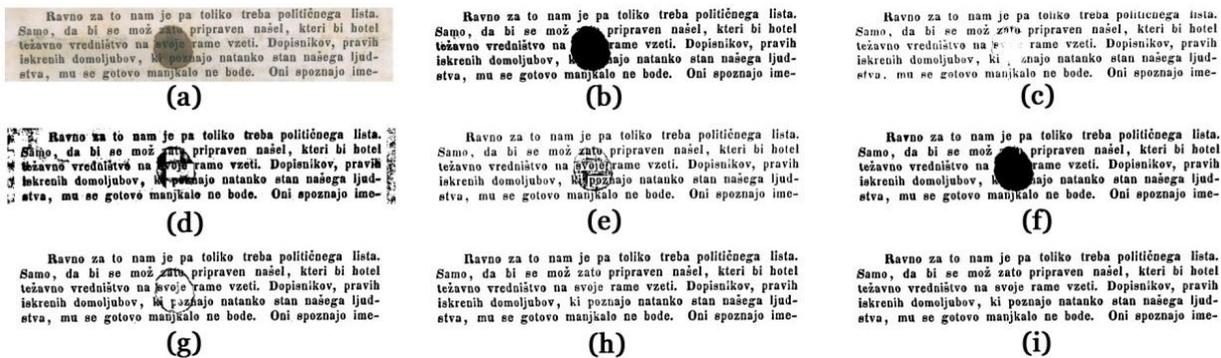
	F-measure (%)	PSNR	NRM (×10 <sup>-2</sup> )	Precision
Otsu	81.86	14.7271	6.94	77.58
Sauvola	82.84	14.7016	9.27	83.96
Niblack	46.52	8.02	10.3	33.20
Bernsen	80.63	14.4270	8.95	78.75
ALLT	77.45	14.5029	14.94	91.74
1 <sup>st</sup> rank	92,12	20,68	5.67	95.01
2 <sup>nd</sup> rank	92,70	21,29	4.47	92.21
Méthode proposée	86.44	18.76	9.15	94.64

Les méthodes Otsu, Sauvola, Niblack et Bernsen ne parviennent pas à produire des résultats raisonnables. En outre, la plupart des méthodes, y compris les méthodes 1<sup>ère</sup> et 2<sup>ème</sup> du classement DIBC 2013 induisent un bruit de fond dans les résultats finaux pour la figure 3.18. D'autre part on prend une autre image bruitée (figure 3.19) et on la compare aux méthodes classiques OUTSU, Sauvola, Niblack et Bernsen et aux cinq méthodes classées

premières pour DIBCO 2013. Notre méthode proposée, produit un résultat binaire avec une meilleure qualité visuelle et contient la plupart des informations de texte.



**Figure 3.18** : Résultat de la binarisation par différentes méthodes (a) Image originale DIBCO13, (b) Otsu,(c) Sauvola,(d) Niblack,(e) Bernsen,(f) 1<sup>st</sup> Rank DIBCO 2013,(g) 2<sup>nd</sup> Rank DIBCO 2013,(h) 3<sup>rd</sup> Rank DIBCO 2013,(i) 4<sup>th</sup> Rank DIBCO2013,(j) 5<sup>th</sup> Rank DIBCO 2013,(k) Méthode proposée,(l) Image de référence.



**Figure 3.19** : Résultat de la binarisation par différentes méthodes (a) Image originale DIBCO13, (b) Otsu,(c) Sauvola,(d) Niblack,(e) Bernsen,(f) 1<sup>st</sup> Rank DIBCO 2013,(g) 2<sup>nd</sup> Rank DIBCO 2013,(h) Méthode proposée,(i) Image de référence.

La méthode que nous avons proposé est concurrente aux nouvelles méthodes telles que la méthode de Su [36], la méthode de Nicholas Howe [37] et la méthode de Reza Farrahi Moghaddam [38]. Comme décrit dans les sections précédentes, notre méthode proposée comporte quelque paramètres qui peuvent être estimés automatiquement à partir des statistiques de l'image. Cela rend notre méthode proposée plus facile à utiliser pour des images dégradées.

La performance supérieure de notre méthode proposée peut être expliquée par plusieurs facteurs. Tout d'abord, la méthode proposée combine l'estimation de l'arrière-plan puis l'utilisé pour la compensation du contraste de l'image et l'aide à supprimer la variation de fond et éviter la récupération des taches sur l'image comme on le voir sur les résultats des tests des autres méthodes.

En deuxième lieu, la méthode proposée permet d'extraire le texte du premier plan à partir du fond du document avec précision et surtout pour les images très dégradées. Afin de valoriser notre méthode et son efficacité pour les images très dégradées on a fait deux autres testes le premier sur un ensemble d'image avec différentes dégradations en comparons notre méthode avec les dix meilleurs méthodes classé première pour DIBCO 2011 et DIBCO 2013 et le 2<sup>ème</sup> teste sur la bibliothèque [33] Diary qui contient un ensemble d'image très dégradé photocopie d'un journal avec les méthodes récentes comme la méthode de Su [36] et la méthode de Prashali Chaudhary [39].

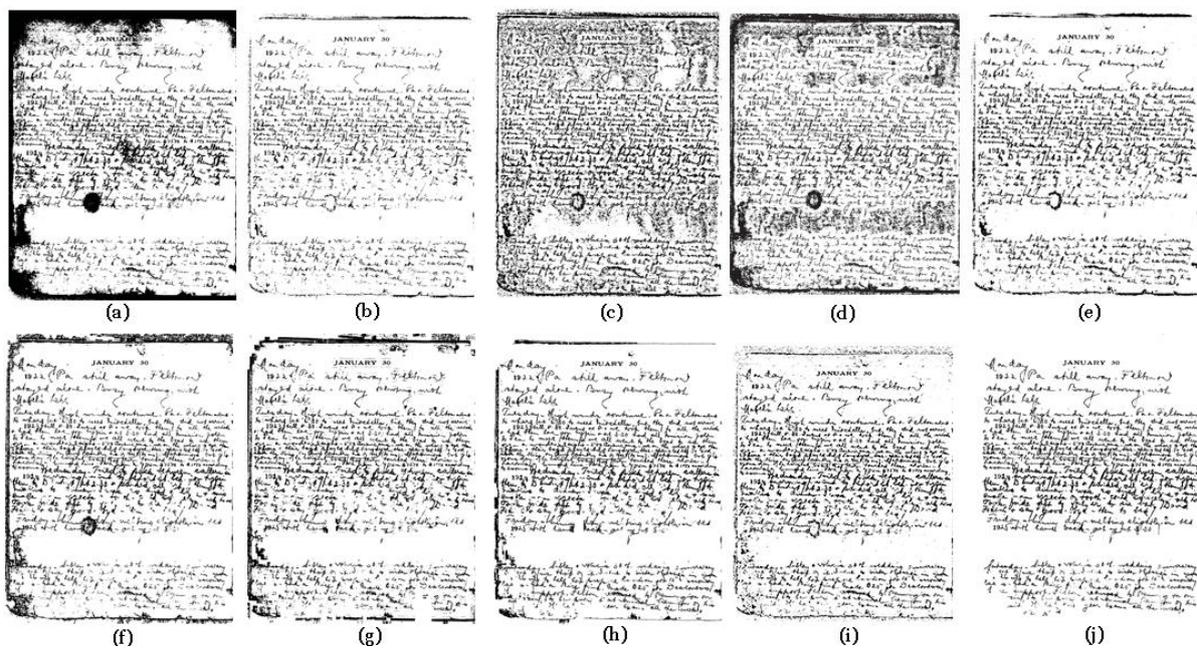
### 3.4.2. Teste sur la bibliothèque « Bickley Diary »

Pour donner plus d'impacte de la méthode, nous avons expérimenté notre algorithme sur une autre base de données (la bibliothèque Bickley Diary) qui est un ensemble d'images photocopées d'un journal écrit il y'a environ 100 ans. Ces images souffrent de différents types de dégradation, tels que les taches d'eau, encre transparent et contient d'autres cas de figures qui complètent les bases DIBCO's de point de vue degré de dégradation.

Pour cela, nous avons pris un ensemble de sept images originales et sept images de référence (figure 3.20) qui sont annotées manuellement à l'aide de Pix Labeler [40]. Notre méthode proposée permet d'obtenir une moyenne de 79.11% en termes de F-mesure et de 13.24 en termes de PSNR qui sont nettement supérieures aux autres méthodes et en particulier la méthode Su. Les résultats des évaluations détaillées sont illustrés dans le tableau 3.5.

**Tableau 3.5.** Résultat de mesure Bickley Diary.

	F-measure (%)	PSNR	NRM ( $\times 10^{-2}$ )	MPM ( $\times 10^{-3}$ )
Otsu	50.42	7.58	21.41	196.98
Sauvola	64.60	11.62	23.26	28.97
Niblack	67.71	9.79	9.52	105.17
Bernsen	52.97	7.71	18.86	193.35
Gatos	69.13	11.44	21.89	36.57
LMM [30]	66.44	10.76	17.50	72.08
BE	34.65	3.54	40.78	370.15
Su	78.54	13.15	12.92	16.71
Prashali	78.78	13.50	12.07	12.10
Méthode proposée	79.11	13.24	12.85	23.99



**Figure 3.20 :** Résultats de la binarisation d'une image dégradée de la bibliothèque Bickley diary obtenue par différentes méthodes de binarisation. (a) Otsu. (b) Sauvola. (c) Niblack. (d) Bernsen. (e) Gatos. (f) LMM. (g) BE. (h) Su. (i) Méthode proposée. (j) Image de référence.

### 3.5. Conclusion

Dans ce chapitre, nous avons décrit la méthode de binarisation proposée qui est basée sur quatre étapes consécutives:

- Proposition d'une méthode rapide pour déterminer le type d'ancienne image de document étudiée.

- Estimation et compensation du fond par une nouvelle approche, qui utilise le filtre médian.
- Segmentation de l'image de document résultante par une binarisation à seuil global.
- Un post-traitement pour supprimer les pixels isolés afin d'améliorer la binarisation de l'image finale.

Nous avons testé sur différentes bases de données couvrant de nombreux types de dégradation, à savoir DIBCO 2009, DIBCO 2011, DIBCO 2013 et BICKLEY DIARY. Les résultats expérimentaux confirment que la performance de la méthode proposée est généralement compétitive par rapport à celle des méthodes existantes, surtout pour les images dégradés.

# *Bibliographie*

- 
- [1] M. Khitas, L. Ziet, S. Bouguezel, « Improved Degraded Document Image Binarization Using Median Filter for Background Estimation », *Elektron. Ir Elektrotechnika*, vol. 24, n° 3, juin 2018.
- [2] S. Lu, B. Su, C. L. Tan, « Document image binarization using background estimation and stroke edges », *Int. J. Doc. Anal. Recognit. IJDAR*, vol. 13, n° 4, pp. 303-314, déc. 2010.
- [3] Q. Chen, Q. Sun, P. Ann Heng, D. Xia, « A double-threshold image binarization method based on edge detector », *Pattern Recognit.*, vol. 41, n° 4, pp. 1254-1267, avr. 2008.
- [4] R. F. Moghaddam, M. Cheriet, « Application of Multi-Level Classifiers and Clustering for Automatic Word Spotting in Historical Document Images », in *2009 10th International Conference on Document Analysis and Recognition*, Barcelona, Spain, 2009, pp. 511-515.
- [5] B. Su, S. Lu, C. L. Tan, « Binarization of historical document images using the local maximum and minimum », in *Proceedings of the 8th IAPR International Workshop on Document Analysis Systems - DAS '10*, Boston, Massachusetts, 2010, pp. 159-166.
- [6] B. Gatos, I. Pratikakis, S. J. Perantonis, « Adaptive degraded document image binarization », *Pattern Recognit.*, vol. 39, n° 3, pp. 317-327, mars 2006.
- [7] J. Sauvola, M. Pietikäinen, « Adaptive document image binarization », *Pattern Recognit.*, vol. 33, n° 2, pp. 225-236, févr. 2000.
- [8] R. Farrahi Moghaddam, M. Cheriet, « RSLDI: Restoration of single-sided low-quality document images », *Pattern Recognit.*, vol. 42, n° 12, pp. 3355-3364, déc. 2009.
- [9] Jyotsna, S. Chauhan, E. Sharma, A. Doegar, « Binarization techniques for degraded document images — A review », in *2016 5th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, Noida, India, 2016, pp. 163-166.
- [10] L. T. K. Van et G. Lee, « Stroke Width-Based Contrast Feature for Document Image Binarization », *J. Inf. Process. Syst.*, vol. 10, n° 1, pp. 55-68, mars 2014.
- [11] H. Z. Nafchi, R. F. Moghaddam, M. Cheriet, « Phase-Based Binarization of Ancient Document Images: Model and Applications », *IEEE Trans. Image Process.*, vol. 23, n° 7, pp. 2916-2930, juill. 2014.
- [12] T. I. Shobhika, S. P. Bhosale, « Highly Optimized and Robust Binarization Technique for Degraded Document Image », *Int. J. Sci. Res. IJSR*, vol. 3, n° 7, pp. 2016-2018, juill. 2014.
- [13] S. Ajith, P. Srivani, et B. N. Ravi Kumar, « Height Map Construction for Ancient Palm Leaf », *Int. J. Comput. Sci. Mob. Comput.*, vol. 4, n° 6, pp. 354-360, juin 2015.
- [14] N. Chaki, S. H. Shaikh, K. Saeed, « A New Image Binarization Technique Using Iterative Partitioning », in *Exploring Image Binarization Techniques*, vol. 560, New Delhi: Springer India, 2014, pp. 17-44.
- [15] K. Mieloch, P. Mihailescu, A. Munk, « Dynamic threshold using polynomial surface regression with application to the binarization of fingerprints », présenté à Defense and Security, Orlando, Florida, USA, 2005, pp. 94.
- [16] M. Seeger, C. Dance, « Binarising camera images for OCR », in *Proceedings of Sixth International Conference on Document Analysis and Recognition*, Seattle, WA, USA, 2001, pp. 54-58.
-

- 
- [17] R. Saini, « Document Image Binarization Techniques, Developments and Related Issues: A Review », *Int. J. Comput. Appl.*, vol. 116, n° 7, pp. 41-44, avr. 2015.
- [18] E. Kavallieratou, E. Stamatatos, « Improving the Quality of Degraded Document Images », in *Second International Conference on Document Image Analysis for Libraries (DIAL'06)*, Lyon, France, 2006, pp. 340-349.
- [19] T. Romen Singh, S. Roy, O. Imocha Singh, T. Sinam, kh Manglem Singh, « A New Local Adaptive Thresholding Technique in Binarization », *Int. J. Comput. Sci. IJCSI*, vol. 8, n° 6, pp. 271-277, nov. 2011.
- [20] Y.-H. Chiu, K.-L. Chung, W.-N. Yang, Y.-H. Huang, C.-H. Liao, « Parameter-free based two-stage method for binarizing degraded document images », *Pattern Recognit.*, vol. 45, n° 12, pp. 4250-4262, déc. 2012.
- [21] G. Lazzara, T. Géraud, « Efficient multiscale Sauvola's binarization », *Int. J. Doc. Anal. Recognit. IJDAR*, vol. 17, n° 2, pp. 105-123, juin 2014.
- [22] R. Farrahi Moghaddam, M. Cheriet, « A multi-scale framework for adaptive binarization of degraded document images », *Pattern Recognit.*, vol. 43, n° 6, pp. 2186-2198, juin 2010.
- [23] N. Otsu, « A Threshold Selection Method from Gray-Level Histograms », *IEEE Trans. Syst. Man Cybern.*, vol. 9, n° 1, pp. 62-66, janv. 1979.
- [24] B. McCane, « On the evaluation of image segmentation algorithms », *Digit. Image Comput. Tech. Appl. DICTA97 IVCNZ97*, pp. 455-459, oct. 1997.
- [25] Y. J. Zhang, « A survey on evaluation methods for image segmentation », *Pattern Recognit.*, vol. 29, n° 8, 1996.
- [26] S. Philip-Floguet, « Evaluation de la segmentation d'images: état de l'art, nouveau indices et comparaison », Technical report, ETIS, mars 2001.
- [27] H. Salam, « Détection automatique de fissures dans des images de chaussée par modélisation markovienne », *Institut de Recherche en Communications et Cybernétique de Nantes*, pp. 47-53, 2010.
- [28] B. Gatos, K. Ntirogiannis, I. Pratikakis, « ICDAR 2009 Document Image Binarization Contest (DIBCO 2009) », (:unav), 2009.
- [29] I. Pratikakis, B. Gatos, K. Ntirogiannis, « H-DIBCO 2010 - Handwritten Document Image Binarization Competition », (:unav), nov. 2010.
- [30] I. Pratikakis, B. Gatos, K. Ntirogiannis, « ICDAR 2011 Document Image Binarization Contest (DIBCO 2011) », sept. 2011.
- [31] H. Lu, A. C. Kot, Y. Q. Shi, « Distance-Reciprocal Distortion Measure for Binary Document Images », *IEEE Signal Process. Lett.*, vol. 11, n° 2, pp. 228-231, 2004.
- [32] I. Pratikakis, B. Gatos, K. Ntirogiannis, « ICDAR 2013 Document Image Binarization Contest (DIBCO 2013) », in *2013 12th International Conference on Document Analysis and Recognition*, Washington, DC, USA, 2013, pp. 1471-1476.
- [33] F. Deng, Z. Wu, Z. Lu, M. S. Brown, « BinarizationShop: a user-assisted software suite for converting old documents to black-and-white », in *Proceedings of the 10th annual joint conference on Digital libraries - JCDL '10*, Gold Coast, Queensland, Australia, 2010, pp. 255.
- [34] K. Ntirogiannis, B. Gatos, I. Pratikakis, « An Objective Evaluation Methodology for Handwritten Image Document Binarization Techniques », présenté à Proceedings of
-

- the 11 th International Conference on Frontiers in Handwriting Recognition (ICFHR), Montreal-Canada, 2008, pp. 586-591.
- [35] T. Lelore, F. Bouchara, « Super-Resolved Binarization of Text Based on the FAIR Algorithm », in *2011 International Conference on Document Analysis and Recognition*, Beijing, China, 2011, pp. 839-843.
- [36] Bolan Su, Shijian Lu, Chew Lim Tan, « Robust Document Image Binarization Technique for Degraded Document Images », *IEEE Trans. Image Process.*, vol. 22, n° 4, pp. 1408-1417, avr. 2013.
- [37] N. R. Howe, « A Laplacian Energy for Document Binarization », in *2011 International Conference on Document Analysis and Recognition*, Beijing, China, 2011, pp. 6-10.
- [38] R. F. Moghaddam, F. F. Moghaddam, et M. Cheriet, « Unsupervised Ensemble of Experts (EoE) Framework for Automatic Binarization of Document Images », in *2013 12th International Conference on Document Analysis and Recognition*, Washington, DC, USA, 2013, pp. 703-707.
- [39] C. Prashali, « AN EFFECTIVE AND ROBUST TECHNIQUE FOR THE BINARIZATION OF DEGRADED DOCUMENT IMAGES », *Int. J. Res. Eng. Technol.*, vol. 03, n° 06, pp. 140-145, juin 2014.
- [40] E. Saund, J. Lin, P. Sarkar, « PixLabeler: User Interface for Pixel-Level Labeling of Elements in Document Images », in *2009 10th International Conference on Document Analysis and Recognition*, Barcelona, Spain, 2009, pp. 646-650.

## CHAPITRE 4

IMPLEMENTATION HARDWARE D'ALGORITHMES DE BINARISATION  
D'IMAGES DE DOCUMENTS ANCIENS SUR FPGA

## 4.1. Introduction

Le traitement des images en temps réel est long et difficile à réaliser sur un processeur classique. Cela est dû au grand nombre de données représenté par l'image et les opérations complexes nécessaires effectuées sur l'image.

Les FPGA's (Field Programmable Gate Array) sont souvent utilisés comme plateformes de mise en œuvre pour les applications de traitement d'images en temps réel, vu leurs structures capables d'exploiter le parallélisme spatial et temporel. Fondamentalement, ils sont composés d'un certain nombre de modules programmables et peuvent être configurés sur place, en émulant le comportement de n'importe quel circuit numérique. Ces dispositifs électroniques sont souvent utilisés dans les systèmes embarqués pour implémenter des algorithmes et applications logiciels dans le matériel afin de les exécuter en considérant un niveau de parallélisme donné.

L'objectif de cette partie de thèse est de présenter la conception et la mise en œuvre d'un système de traitement d'images « binarisation d'images de vieux documents » basé sur la méthode d'Otsu, au tour du processeur intégré à noyau souple NIOS II d'*Intel-Altera*.

Le système proposé est implémenté sur un Système sur une puce programmable SOPC (System On Programmable Chip) avec la suite de logiciels intégrés dans Quartus II (SoPC-builder, Qsys, NIOSII software builds for eclipse). Ce système utilise également en plus d'un ensemble de périphériques externes comprenant une carte SD (Secure Data), probablement une caméra et une interface transmission internet, un ensemble de composants softcores IP's (Intellectual property) délivrés par *Altera-terasic-university program* et d'autres IP's que nous avons spécialement conçus pour cette application (*core-histogram* et *core-Otsu*).

## 4.2. Plateforme et outils de conception d'un système SoPC

La plus part des systèmes embarqués font appel à des circuits numériques de type FPGA sur lesquels peuvent être implantés des processeurs pour la conception de fonctionnalités numériques complexes.

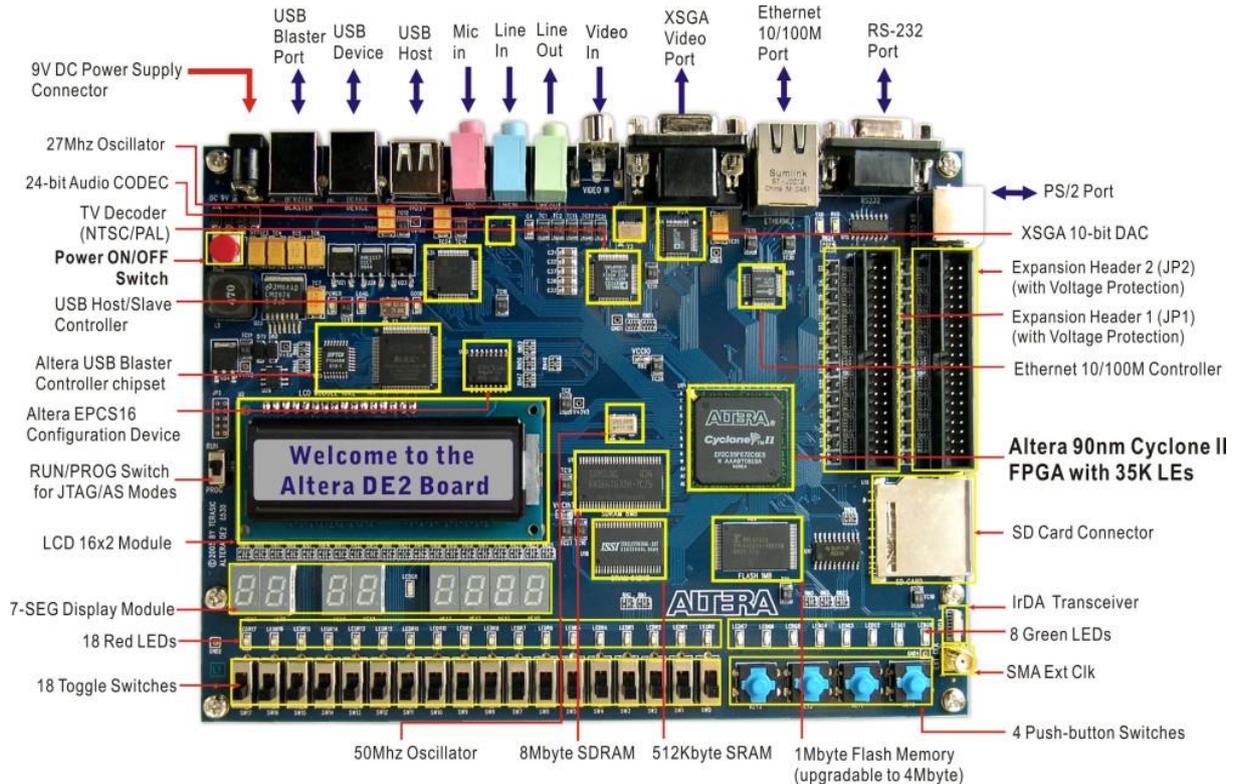
L'approche de conception est alors dans ce cas une approche de conception conjointe matérielle/logicielle appelée « *codesign* » où une fonctionnalité complexe allie à la fois une logique programmée pour sa flexibilité et une logique câblée pour ses performances.

Pour la conception de tels systèmes nous avons besoin de suites d'outils délivrés par de grandes sociétés dans ce domaine.

La société *Intel-Altera* [1] propose une série de cartes de développement et un ensemble de logiciels Quartus II, (SOPC-BUILDER, Qsys, SignalTrap-analyser JTAG Chain debugger, NIOS\_II Software Build for Eclipse...) qui permettent la gestion complète d'un flot de conception FPGA.

#### 4.2.1. Carte de développement

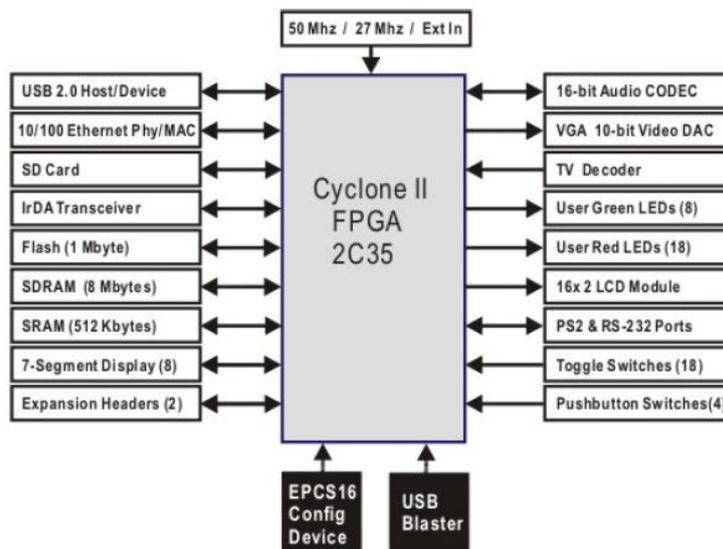
Pour faire le prototypage d'un système SoPC, nous avons besoin d'un matériel à base d'FPGA. La cyclone Altera DE2 est l'une des cartes la plus réputée proposée par terasic dans le cadre de ses programmes universitaires. Ces cartes intègrent généralement toutes une puce FPGA associée à des périphériques externes (figure 4.1).



**Figure 4.1 :** Carte DE2

La carte utilise la puce FPGA Cyclone II EP2C35F672C6FPGA comme cœur et dispose de mémoire (SRAM, SDRAM et Flash), de convertisseurs audio, vidéo et TV, ainsi que d'interfaces IrDA, Ethernet et USB.

Le schéma block interne est représenté par :



**Figure 4.2:** Périphérique de Cyclone FPGA II 2C35

La description de ces éléments est comment suit :

- Cœur :
  - FPGA : (Cyclone II EP2C35F672C6) : 35.000 cellules (LE) utilisables dans un boîtier de 672 broches (BGA) ;
  - Flash de configuration : EPCS16 : cette mémoire permet de stocker de manière non-volatile la configuration du FPGA.
  - Horloges : Oscillateurs 27 et 50 MHz et entrée externe (SMA).
- Memoirs :
  - FLASH (S29AL032D) : 4 M\*8 (10ns).
  - SDRAM (IS42S16400-8) : 4 M\*16 (100 MHz <2-2-2>).
  - SRAM (61LV25616) : 256 k\*16 (10ns).
- Périphériques :
  - Codec Audio (WM8731) : 24 bits mono (8 - 96 kHz).
  - Décodeur Vidéo/TV (ADV7181B) : NTSC/PAL 50/60 Hz
  - Convertisseur VGA (ADV7123) : jusqu'à 1600\*1200@100 Hz
- Entrées/Sorties :
  - Connecteur RS232 : type DB9
  - Connecteur USB 2.0 : types A (host) et B (device), contrôleur : ISP 1362
  - Connecteur Ethernet 10/100 : type RJ45, contrôleur : DMA9000A, full duplex
  - Connecteur Clavier/Souris : type PS/2
  - Connecteurs d'extension : 2 connecteurs HE10-40b (72 signaux utilisables).
  - Module IrDA : Emission/Réception infra-rouge (jusqu'à 115,2 kbauds)

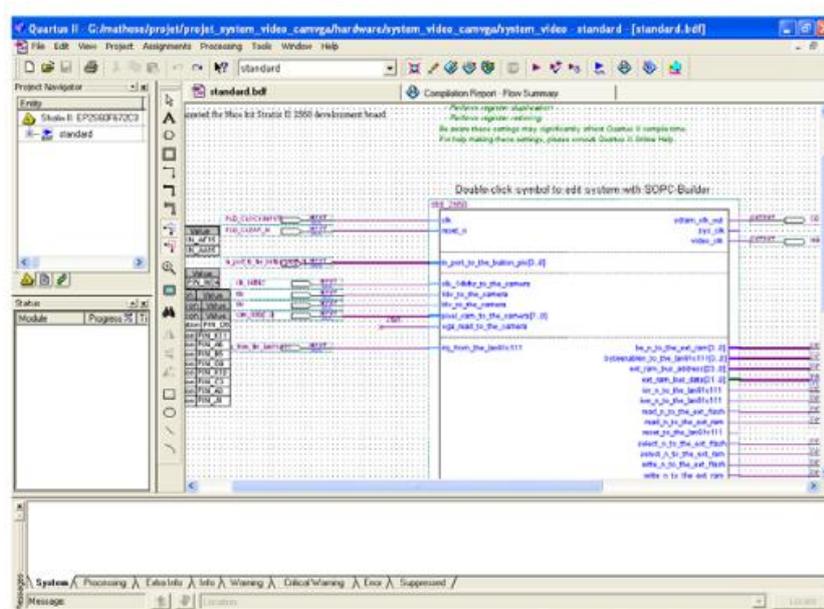
- Interfaces :
  - Boutons : il y a 18 switches et 4 boutons poussoirs.
  - Voyants LED : il y a 18 LED rouge et 9 LED vertes.
  - Afficheurs : 8 digits 7-Segments et 1 module LCD (2 lignes de 16 caractères).
- Modules :
  - LCM : Afficheur LCD 320\*240 couleur 3,6".
  - DC2 : 2 caméras CMOS.

Pour pouvoir bien utiliser cette carte la maison mère met à notre disposition une suite de logiciels chapoté par Quartus.

La figure 4.3 présente l'interface graphique de Quartus II.

#### 4.2.2. Logiciels utilisés

##### a. Quartus



**Figure 4.3:** IDE Quartus II

Ce logiciel permet de faire une saisie graphique ou une description textuelle matérielle de haut niveau (VHDL (Very high speed integrated circuit Hardware Description Language), Verilog, SystemVerilog...ect) d'architectures numériques, d'en réaliser une simulation en utilisant un simulateur « *Model Sim de Mentor Graphiques* » par exemple, une synthèse et une implantation sur FPGA. Il comprend une suite de fonctions de conception au niveau système permettant d'accéder à une large bibliothèque d'IP's d'*Intel-Altera* et un moteur de placement/routage intégrant la technologie d'optimisation et des solutions de vérification [2]. D'une manière générale, un flot de conception ayant pour but la configuration de composants programmables se déroulent de la manière suivante (figure 4.4)

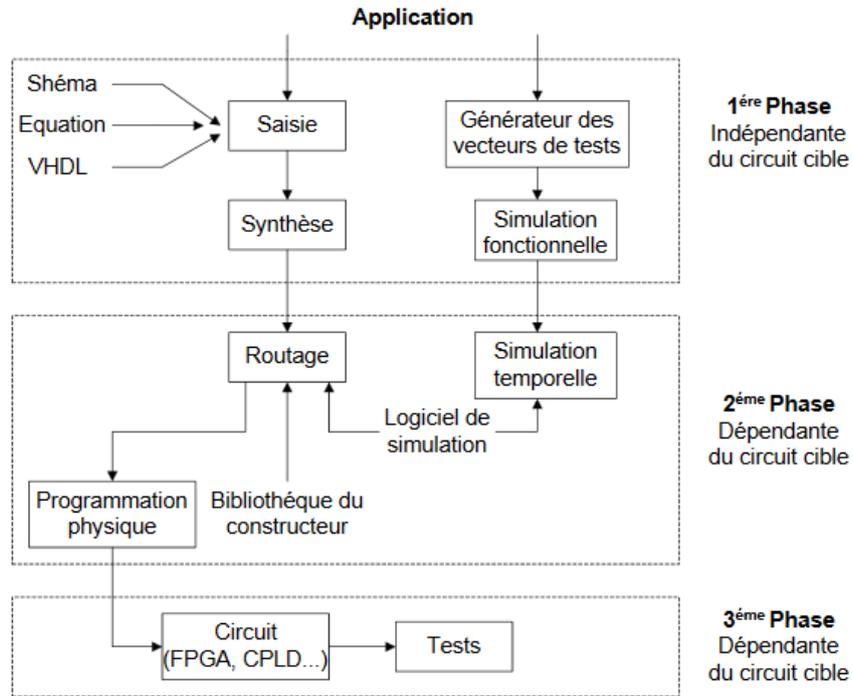


Figure 4.4: Flot de conception

Ce flot est composé en trois phases essentielles : une phase comportant la saisie, la présynthèse et la simulation fonctionnelle, une autre phase dans laquelle se fait le routage, la simulation temporelle, la programmation physique et la dernière phase où nous obtenons le produit prototype pour le tester réellement.

Puisque nous avons utilisé une approche de conception mixte, La phase de saisie est très importante surtout lorsqu'il faudra bien gérer l'aspect partitionnement HW/SW. Pour cela le flot suivant sera adopté.

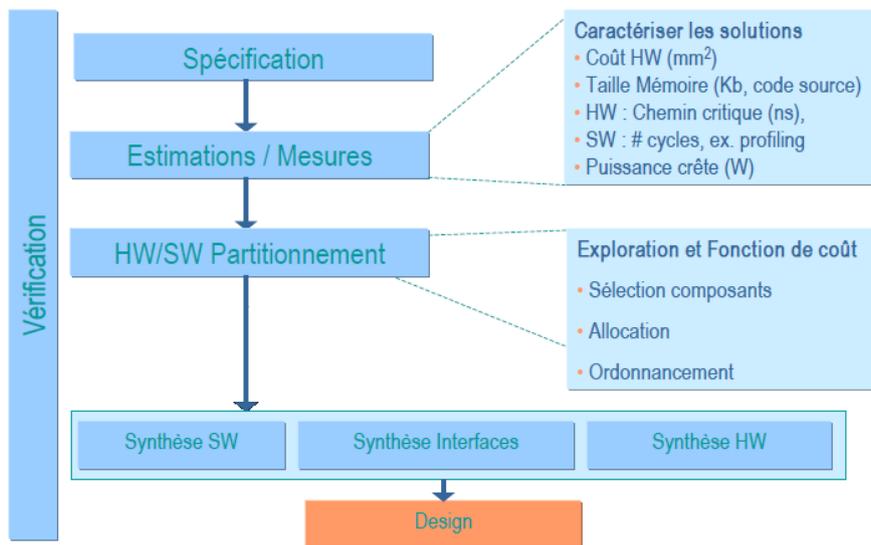


Figure 4.5: Flot de conception mixte

Deux outils intégrés dans Quartus sont en plus nécessaires.

### b. SOPC-BUILDER ou QSYS

Le SOPC-Builder ou Qsys permet de construire un système sur puce programmable SoPC intégrant divers périphériques d'E/S associés au processeur NIOS II tels que : les contrôleurs de SRAM et de SDRAM, un contrôleur DMA (Direct Memory Access), des interfaces parallèles, des interfaces séries, des interface UART JTAG...ect. De même, on peut intégrer son propre composant dans le design sous forme d'un bloc IP externe (Interface caméra, VGA...).

On peut ainsi intégrer autant de périphériques que l'on veut, n'étant limité que par le nombre de broches et de cellules logiques du circuit FPGA. Le mapping mémoire et le niveau des interruptions du design sont fixés durant cette phase. La figure 4.6 montre la mise en œuvre de l'outil Qsys . C'est en fait la première passerelle avec le logiciel embarqué.

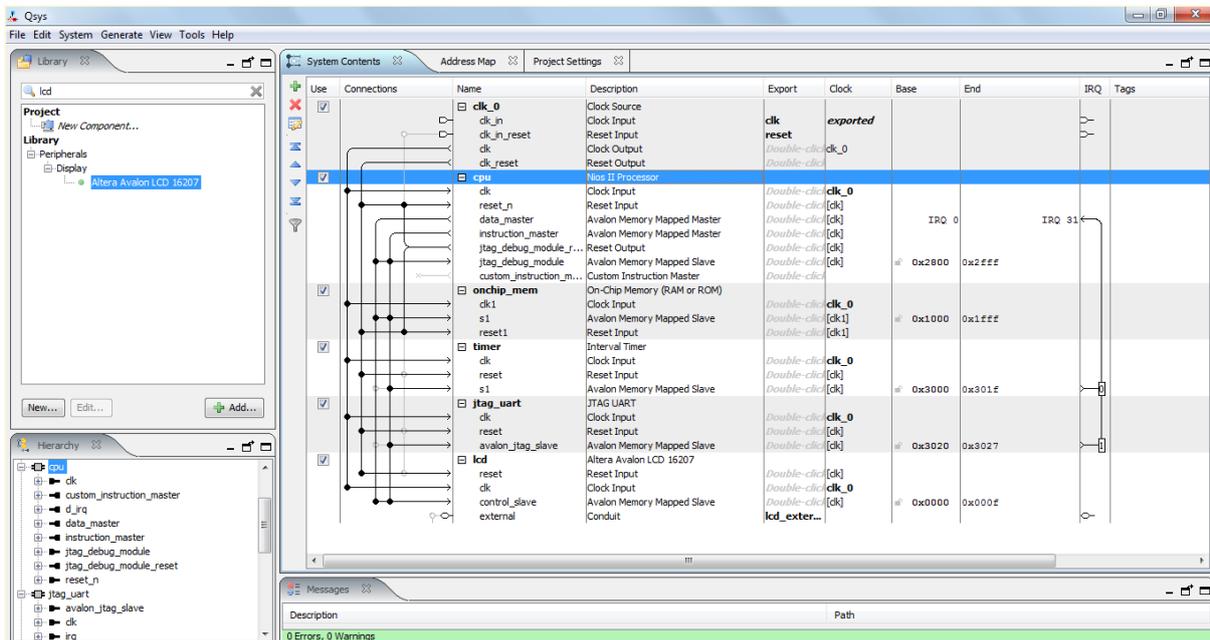


Figure 4.6: Interface Qsys et mapping mémoire

### c. Nios II Software Build Tools for Eclipse

Le développement d'une application embarquée dans le système SoPC nécessite l'utilisation de d'un programme software capable de gérer les différentes unités qui la composent, tout en tenant compte des critères de performance [3].

Lorsque nous concevons avec le processeur intégré Nios® II, nous avons accès à un ensemble d'outils de développement logiciels et de composants robustes disponibles. Nios® II

Embedded Design Suite (EDS) est un package de développement complet pour la conception logicielle Nios® II. Nios® II EDS contient non seulement des outils de développement, mais également des logiciels, des pilotes de périphérique, une bibliothèque HAL (Hardware Abstraction Layer), un logiciel de pile réseau de niveau commercial et une version d'évaluation d'un système d'exploitation en temps réel. La figure suivante présente la fenêtre d'interface EDS.

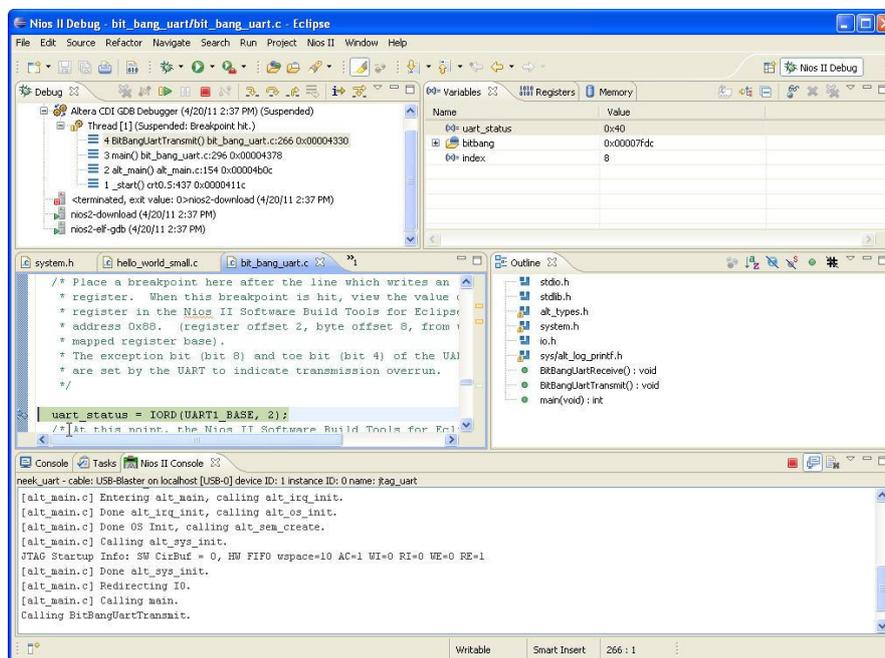


Figure 4.7: Nios II Software Build Tools pour Eclipse IDE

### 4.3. Système proposé

Nous avons réalisé un système de binarisation des documents à base d'une conception mixte logicielle/matérielle implémenté sur une carte de prototypage Intel-Altera DE2 FPGA.

La figure suivante représente ce système.

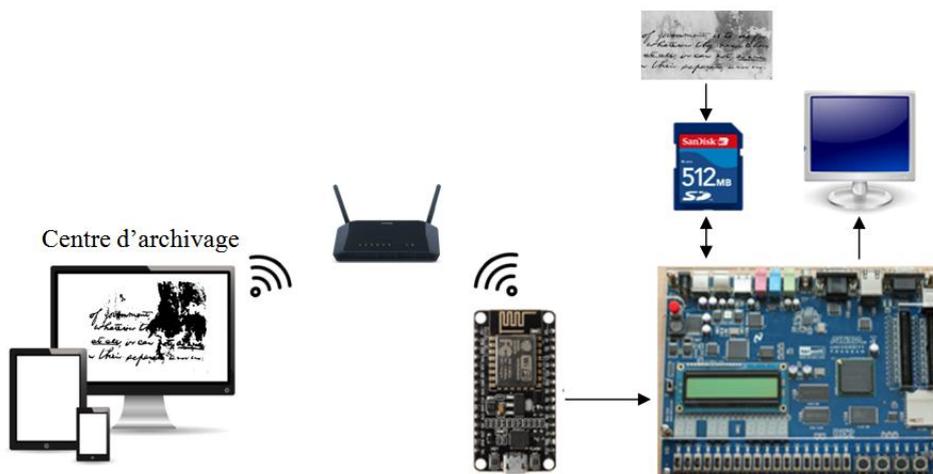


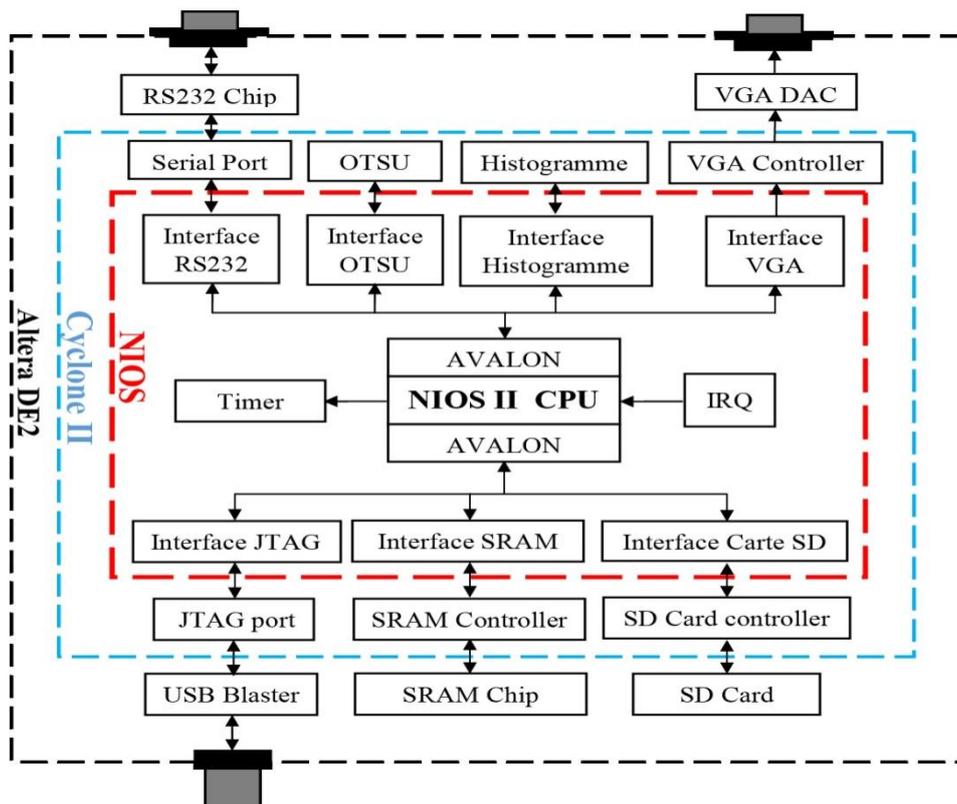
Figure 4.8: Système de binarisation et de transfert

Le principe est que l'image a binarisée sera capturée par une camera, dans notre cas nous avons utilisé des images originales sauvegardées déjà dans la carte SD (Secure Digital), ces images seront traitées par les différents modules du système, puis seront transmises vers un système de stockage et d'archivage via internet.

En plus de tous les composants associés au processeur NIOS, deux modules essentiels ont été conçus. Ces modules représentent des accélérateurs matériels implémentant les algorithmes de calcul d'histogramme d'image et le module de calcul du seuil de binarisation.

Le schéma block de tout le système est présenté comme suit :

Nous rappelons que les composants IP's utilisés se trouvent dans des bibliothèques délivrés par Intel-Altera sous une compilation Altera-Terasic université programme [4].



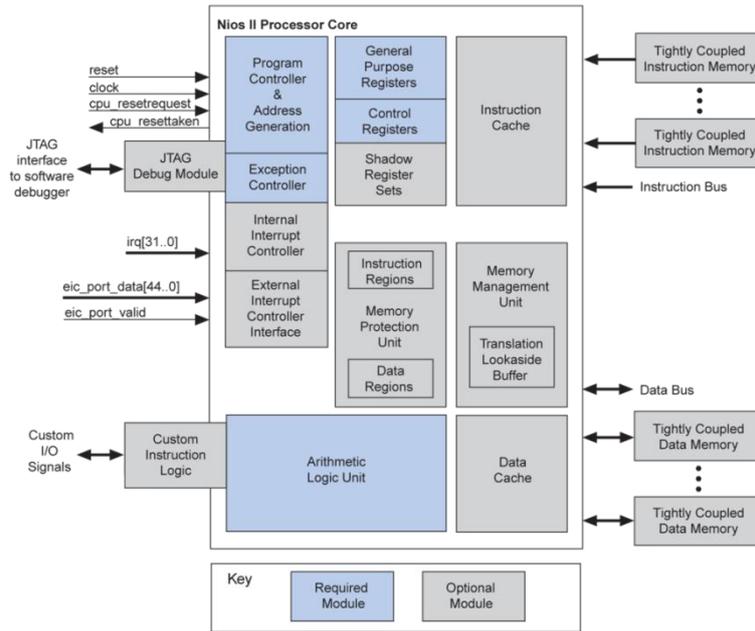
**Figure 4.9:** Conception du système embarqué proposé sur Altera DE2

Nous allons maintenant décrire d'une manière très brève ces composants :

#### 4.3.1. Le processeur Nios II

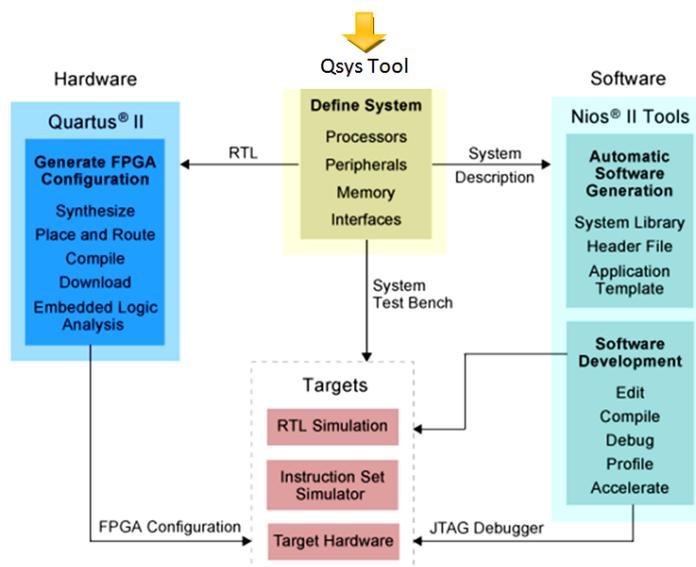
Le processeur Nios II est utilisé pour mettre en œuvre toutes les fonctions de notre système. Il est basé sur une architecture RISC (Reduced Instruction Set Computer). Les opérandes du registre à usage général sont utilisés pour effectuer des opérations arithmétiques et logiques [5]. Dans l'architecture Nios II, nous avons deux bus distincts pour l'envoi

d'instructions et le transfert de données (architecture de Harvard). Il possède également des registres 32 bits et une longueur de mot de 32 bits. Il existe trois configurations disponibles dans lesquelles le processeur Nios II peut être implémenté (Economic, Standard, Fast). La figure 4.9 décrit la structure matérielle du processeur Nios II, y compris toutes les unités fonctionnelles de l'architecture et les principes fondamentaux de la mise en œuvre matérielle du processeur Nios II.



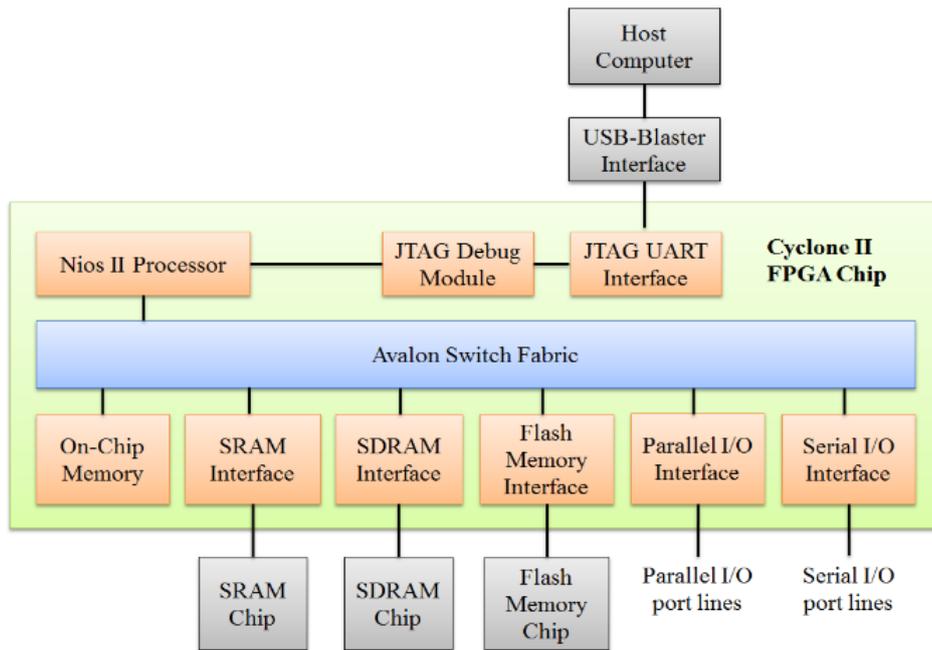
**Figure 4.10:** Architecture interne du processeur Nios II

Le système complet NIOS se fait par deux processus distincts: Processus de génération de matériel et Processus de création du logiciel.



**Figure 4.7:** Etape de conception Nios II

Le système NIOS communique avec les différents composants à travers un bus particulier appelé le bus Avalon [6] décrit dans la figure suivante.



**Figure 4.12:** Interface Avalon

#### 4.3.1.1. IP contrôleur VGA

Les corps vidéo IP de Altera UP (University Program) facilitent le décodage, le traitement et l'affichage des données image ou vidéo [7]. Ils sont conçus pour être utilisés sur les cartes de la série Altera DE. Le décodeur vidéo convertit l'entrée vidéo en paquets pouvant être traités par les cœurs de traitement vidéo. Le contrôleur VGA central affiche les images en créant les signaux de synchronisation requis par les moniteurs compatibles VGA connectés au port VGA de la carte série DE.

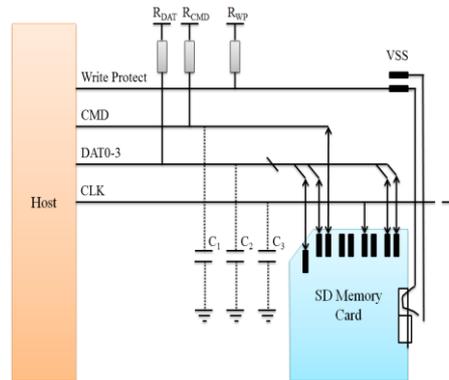
#### 4.3.1.2. Contrôleur SD-Card

La carte SD c'est un périphérique de stockage de données. Il est souvent utilisé dans les appareils numériques tels que les appareils photo, les caméscopes, les lecteurs MP3. Il est très pratique et permet de transférer les données stockées sur plusieurs appareils.

La carte de développement d'Intel-Altera DE2 contient un port de carte SD (figure 4.12) ce qui permet d'enregistrer une grande quantité de données. Dans notre projet, la carte SD utilisée comme une source des fichiers d'image originale. Dans cette conception, le processeur NIOS II utilise le protocole de communication SPI «Serial Peripheral Interface» [8] pour accéder à la carte SD.

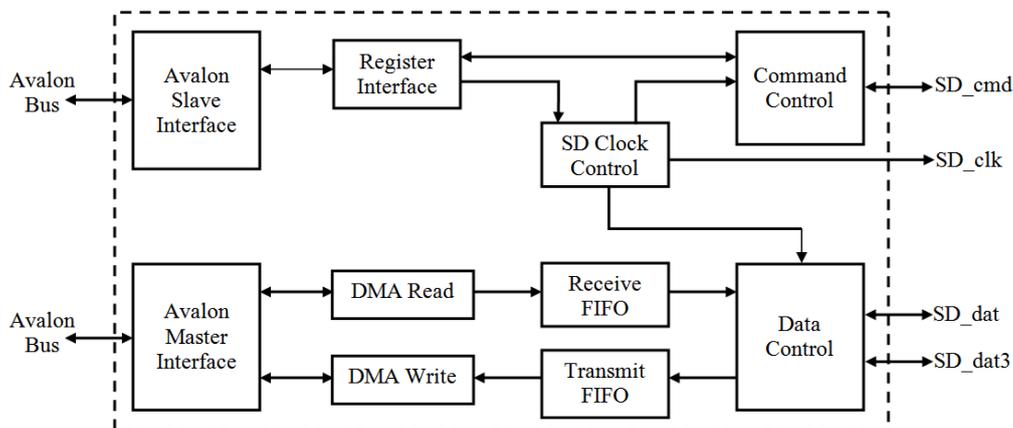
La liaison SPI est un bus de données série synchrone baptisé ainsi par Motorola qui opère en mode full-duplex. Les circuits communiquent selon un schéma maître-esclaves, où le maître contrôle la communication. Plusieurs esclaves peuvent coexister sur un même bus, dans ce cas, la sélection du destinataire se fait par une ligne dédiée entre le maître et l'esclave appelée SS (Slave Select). Ainsi, Les broches utilisées dans SPI sont:

- SD\_CLK: utilisé pour contrôler la vitesse d'horloge de la carte SD.
- SD\_DATA: utilisé pour envoyer des commandes à la carte SD.
- SD\_CMD: utilisé pour le transfert des données.



**Figure 4.8:** Brochage de la carte SD

Le contrôleur SD-Card d'Altera Université Programme [9] est un circuit matériel qui permet l'utilisation d'une carte SD. Lorsqu'il est inclus dans une conception et connecté au port de carte SD, le cœur détecte la carte SD connectée au port et en autorise l'accès. Le noyau a été conçu pour être utilisé dans un système implémenté Qsys. Lorsque ce système comprend également un processeur softcore Nios II, il est possible d'accéder aux données de la carte SD à l'aide de programmes exécutés sur le processeur Nios II. La figure suivante décrit le schéma bloc de l'IP de la carte SD.



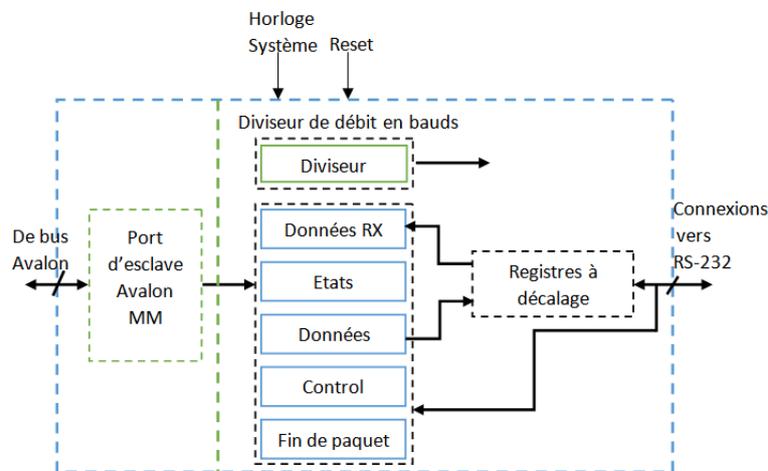
**Figure 4.9:** Schéma bloc d'IP SD carte

### 4.3.1.3. IP UART

Le noyau émetteur / récepteur asynchrone universel avec l'interface Avalon (nommé UART) met en œuvre une méthode de communication de flux de caractères en série entre un système embarqué sur un FPGA *Intel-Altera* [10] et un dispositif externe.

Le noyau implémente la synchronisation de protocole RS-232, et offre des taux de transmission ajustables, parité, bits de données et d'arrêt et des signaux de contrôle de flux optionnels RTS / CTS.

Le jeu de fonctionnalités est configurable, permettant aux concepteurs de mettre en œuvre tous les fonctionnalités nécessaires pour un système donné. Il fournit une interface esclave Avalon-MM qui permet des périphériques maîtres Avalon-MM (comme un processeur Nios II) de communiquer avec le noyau par une simple lecture et écriture dans les registres de control et de données, l'image suivante présente le schéma bloc de cet IP.



**Figure 4.10:** Schéma bloc d'IP UART

Le noyau comporte deux parties visibles à l'utilisateur :

- Le fichier de registre, qui est accessible via le port esclave Avalon-MM.
- Les signaux RS-232, RXD, TXD, CTS et RTS.

Le noyau UART fournit une interface esclave Avalon-MM au fichier de registre interne. L'interface utilisateur au noyau UART se compose de six registres de 16 bits: Contrôle, Etats, Données RX, Données TX, Diviseur, et Fin de paquet. Un périphérique maître, comme un processeur Nios II, accède aux registres pour contrôler le noyau et transférer les données sur la liaison série.

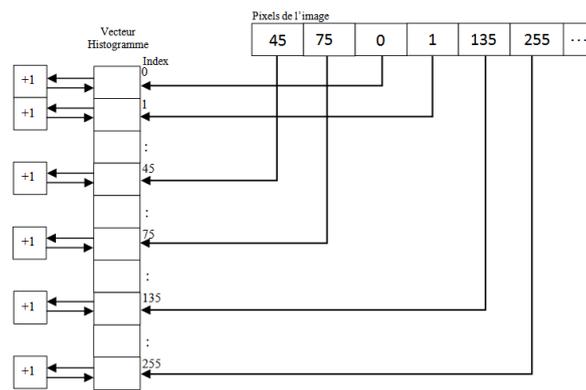
Le noyau de l'UART fournit une demande d'interruption actif-haut (IRQ) qui peut demander une interruption lorsque de nouvelle donnée a été reçue, ou lorsque le noyau est

prêt à transmettre un autre caractère. Le noyau peut être utilisé en conjonction avec un contrôleur DMA pour l'envoi et / ou réception de flux de données exemple : « Mémoire FIFO ».

### 4.3.2. Module histogramme

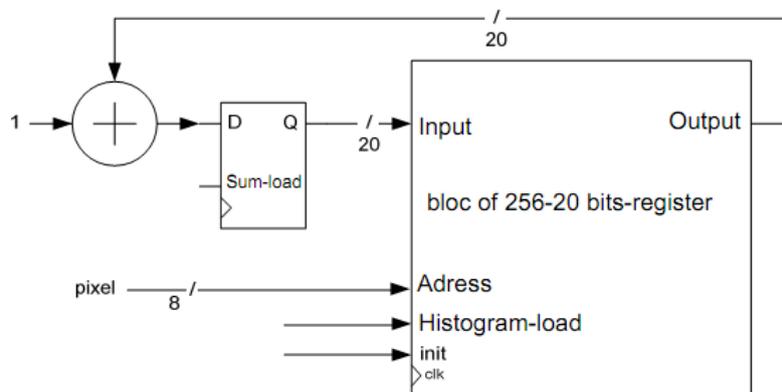
L'histogramme est utilisé dans de nombreuses applications de traitement d'images et de vidéos [11]. Son implémentation matérielle va énormément accélérer le processus de calcul.

Le pixel dans une image au niveau de gris est codé sur 8 bits, alors 256 valeurs différentes qui représentent l'intensité sont possibles pour chaque pixel. Pour une conception de base [12] il serait donc nécessaire de disposer de 256 compteurs de 20 bits pour implémenter l'histogramme voir figure 14.



**Figure 4.11 :** Exemple de calcul de l'histogramme

Pour modéliser le calcul de l'histogramme, nous avons utilisé un modèle basé sur le fonctionnement d'une machine à état. La figure suivante donne l'allure du système ou son entrée est une image et la sortie donne les états des 256 registres [13].



**Figure 4.12:** Schéma block du module histogramme

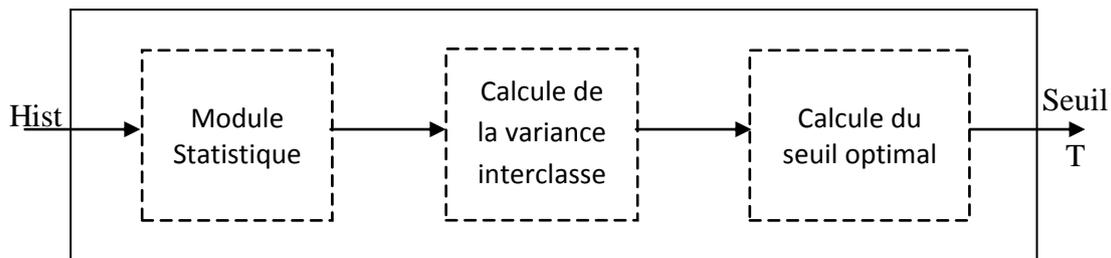
### 4.3.3. Module Otsu

L'implémentation matérielle de l'algorithme Otsu [14] tire parti des capacités de calcul parallèle offertes par la technologie FPGA et utilise la bibliothèque *Mégacore d'Intel-Altera* afin d'éliminer les divisions et multiplications complexes de la méthode. Pour simplifier le calcul de la variance interclasse indiqué dans équation 2,  $N$  est une constante pour chaque image et peut être ignoré, ainsi équation 2 peut être davantage réduit à l'équation 3.

$$\begin{aligned} \sigma_B^2 &= P_1 P_2 (\mu_1 - \mu_2)^2 \\ &= \frac{1}{N^2} \left( \sum_{i=0}^t h_i \right) \times \left( \sum_{i=t+1}^{L-1} h_i \right) \times \left( \frac{\sum_{i=0}^t i \times h_i - \sum_{i=t+1}^{L-1} i \times h_i}{\sum_{i=0}^t h_i - \sum_{i=t+1}^{L-1} h_i} \right)^2 \end{aligned} \quad (1)$$

$$\sigma_B^2 = \left( \sum_{i=0}^t h_i \right) \times \left( \sum_{i=t+1}^{L-1} h_i \right) \times \left( \frac{\sum_{i=0}^t i \times h_i - \sum_{i=t+1}^{L-1} i \times h_i}{\sum_{i=0}^t h_i - \sum_{i=t+1}^{L-1} h_i} \right)^2 \quad (2)$$

Selon l'analyse ci-dessus, l'architecture de l'algorithme d'Otsu illustré dans la figure suivante.



**Figure 4.13:** Schéma block du module calcul de seuil

- **Le module de statistique**

Après avoir obtenu l'histogramme complet, nous pouvons calculer les valeurs de l'histogramme cumulatif (probabilités moyennes) et les valeurs d'intensité de l'histogramme cumulée (l'écart type) en contrôlant la machine à états, puis nous envoyons les données stockées dans le module de statistiques dans le module de calcul de la variance conformément à l'équation 3.

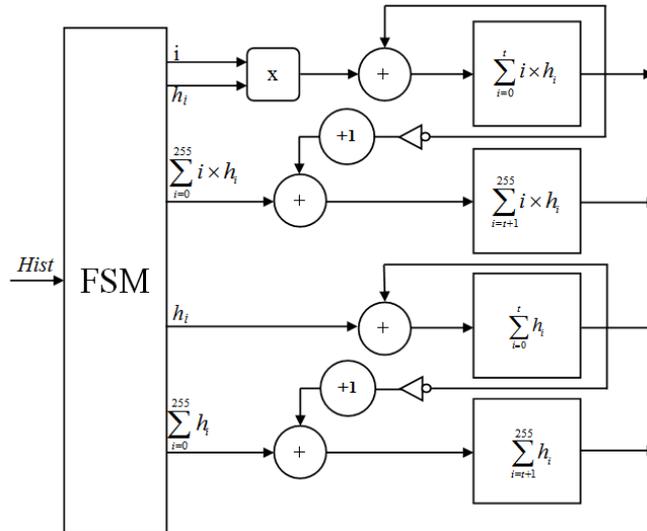


Figure 4.14 : Schéma block du calculé probabilités moyennes et l'écart type

- **Le module du calculé de la variance interclasse**

Pour réaliser la fonction décrite par l'équation 3, du module de calcul de la variance interclasse illustrée à la figure 4.17, on utilise le noyau IP du multiplicateur et du diviseur d'Intel-Altera Pour obtenir les meilleures performances. La figure suivante représente l'architecture du module de calculé de la variance interclasse avec les entrées 1 jusqu'à 4 qui

représente respectivement  $\sum_{i=0}^t h_i$ ,  $\sum_{i=t+1}^{L-1} h_i$ ,  $\sum_{i=0}^t i \times h_i$  et  $\sum_{i=t+1}^{L-1} i \times h_i$

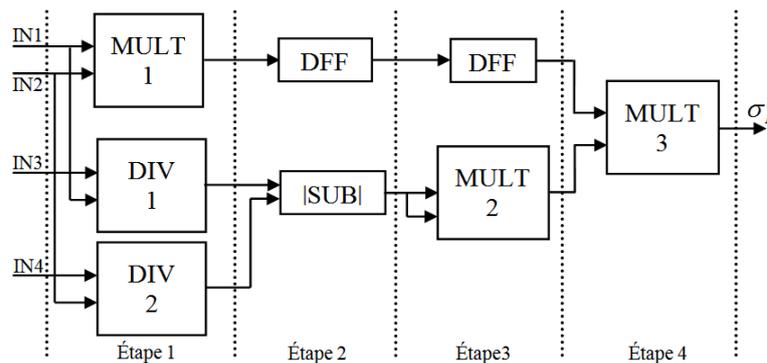
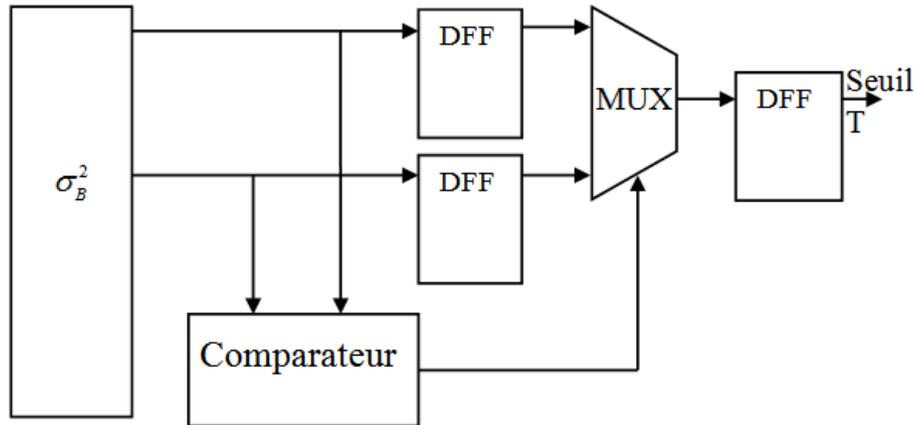


Figure 4.15: Schéma block du calculé de la variance interclasse

- **Le module du calculé du seuil optimal**

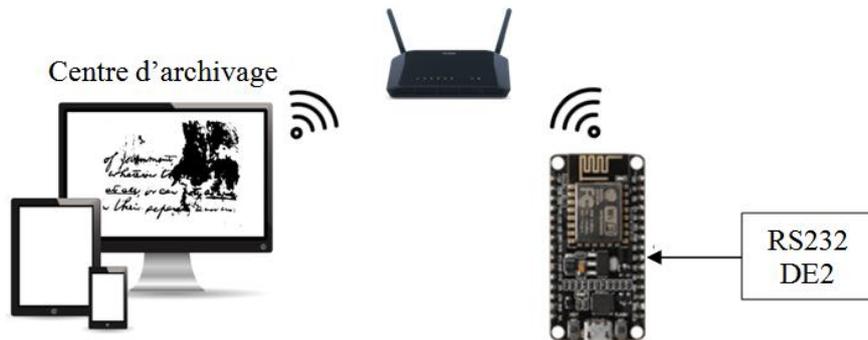
Le calculé du seuil optimal se fera par la maximisation de la variance interclasse. La figure suivante représente l'architecture du module de maximisation pour déduire le seuil optimale pour la binarisation



**Figure 4.16:** Schéma block de la maximisation de la variance interclasse

#### 4.3.4. Transmission internet de l'image

Pour la partie transmission nous proposons le système suivant :



**Figure 4.17:** Schéma du mécanisme de transmission

Pour la transmission de l'image binarisée par notre système vers le centre d'archivage on a choisi d'implanter un serveur web sur NodeMCU [15] relié au processeur NIOS par communication RS232 qui est une plate-forme open source, elle contient le firmware qui fonctionne sur le module ESP8266 WiFi (Wireless Fidelity) SoC (System On a Chip) (figure 4.22) qui est configuré et génère une page web accessible à l'utilisateur par liaison WiFi pour accéder à l'image binarisée, en bref les caractéristiques du module sont :

- Supporte les protocoles 802.11 b/g/n.
- Fonctionne comme serveur web avec sa propre adresse IP.
- Peut fonctionner en communiquant avec d'autres cartes comme dans notre cas la DE2 d'Intel-Altera ou être programmé pour fonctionner en toute autonomie.

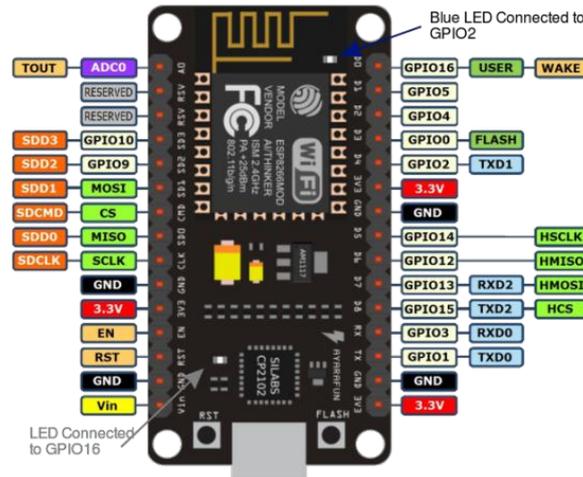


Figure 4.18: Brochage NodeMCU 1.0 ESP8266

➤ La liaison série UART

Une liaison série consiste à envoyer une information bit après bit avec un délai entre chacun. La transmission UART c'est une transmission série qui s'effectue sans signal d'horloge. Pour que le récepteur reçoive sans erreur la donnée série il doit :

- Connaître le format des données utilisées.
- Connaître la fréquence de transmission de chaque bit.

Ainsi, les UART's d'émission et de réception doivent être configurés de manière identique et conformément aux options choisies concernant le format des trames (figure9) et la vitesse de transmission.

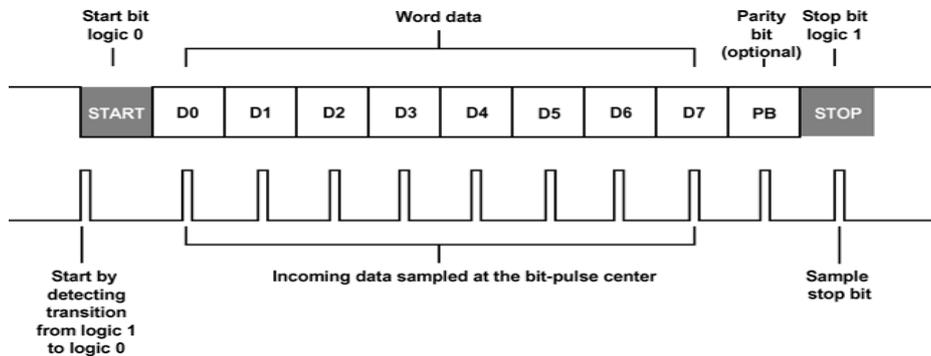


Figure 4.19: Exemple d'une trame UART

➤ La technologie WiFi

Le WiFi apparu en 1999, est un ensemble de protocoles de communication sans fil régis par les normes du groupe IEEE 802.11 (ISO/CEI 8802-11) et il permet à nos appareils informatiques de communiquer entre eux par ondes radio. Au départ limité aux ordinateurs de

bureau, le WiFi est aujourd'hui intégré à tout type d'équipement et en particulier aux équipements mobiles tels que les Smartphones, les tablettes et les ordinateurs portables.

Le standard WiFi met en œuvre essentiellement trois types de trames :

- Les trames de données, utilisées pour la transmission des données.
- Les trames de contrôle, utilisées pour contrôler l'accès au support (RTS, CTS, ACK).
- Les trames de gestion, pour l'échange d'informations de gestion.

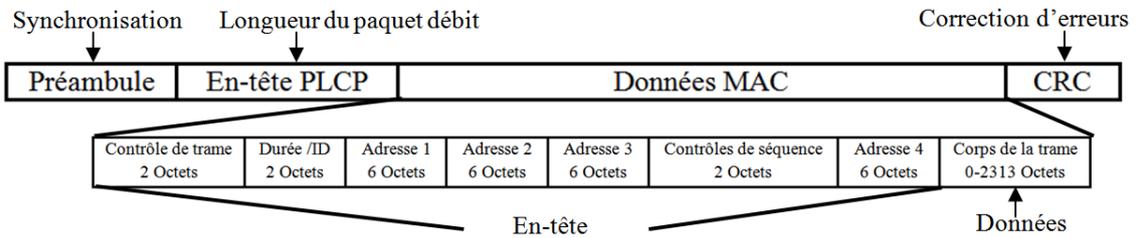


Figure 4.20: Format d'une trame WiFi

#### 4.4. Résultats

Les résultats que nous allons présenter seront divisés en deux parties

##### a- Résultats de point de vue matériel

Le processeur Nios II et d'autres composants seront utilisés comme partie de l'ensemble du système à développer et intégré dans un la carte Intel-Altera. Ils sont instanciés et connectés à l'aide de l'outil Qsys. Les figures 4.25, 4.26 et 4.27 donnent la structure de mappage et l'instanciation des différents composants par rapport au processeur.

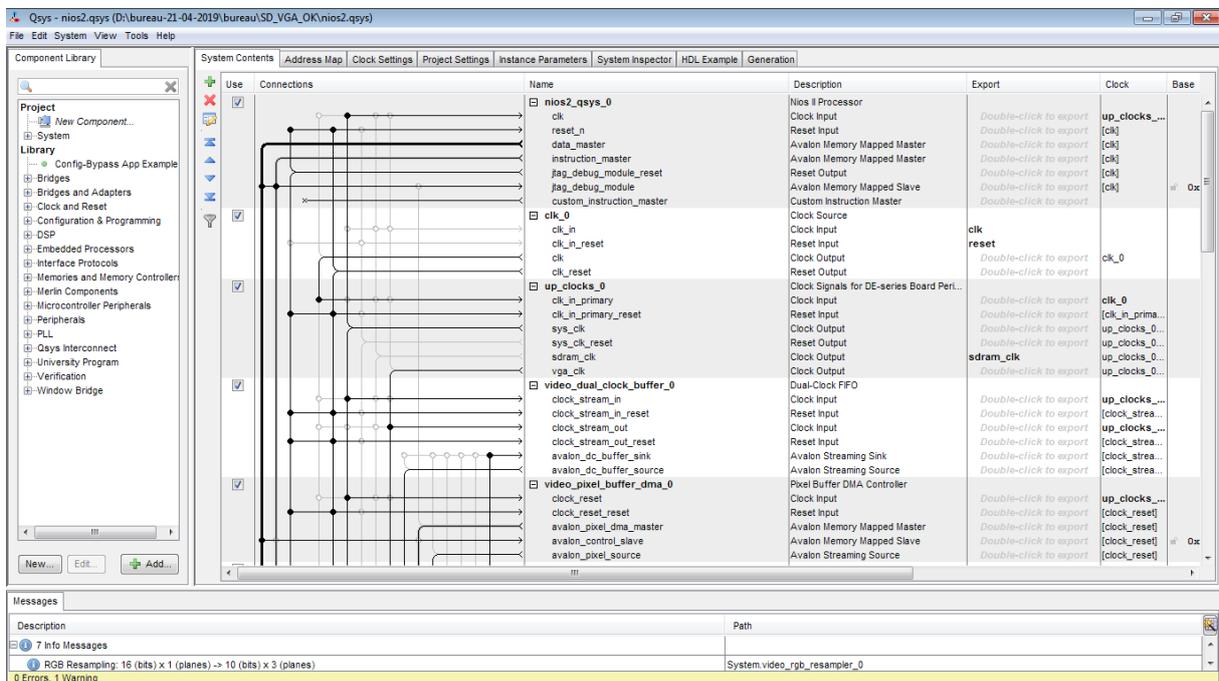


Figure 4.21: Système complet sur Qsys partie-1

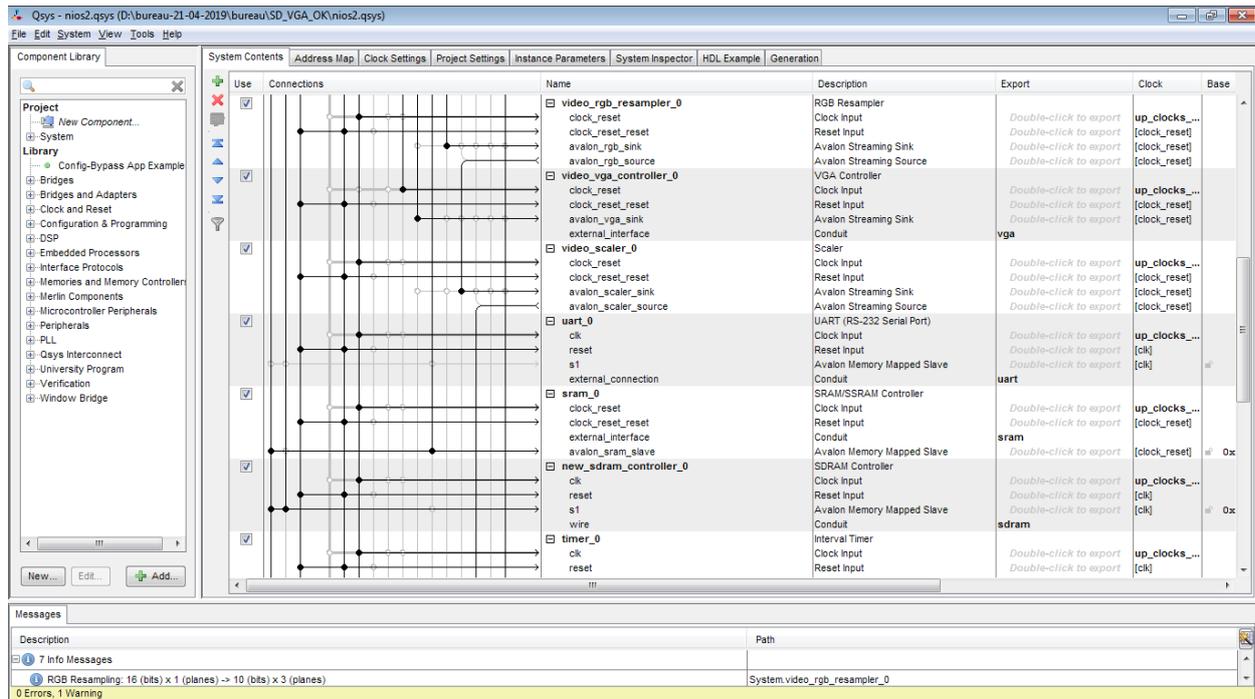


Figure 4.22: Système complet sur Qsys partie-2

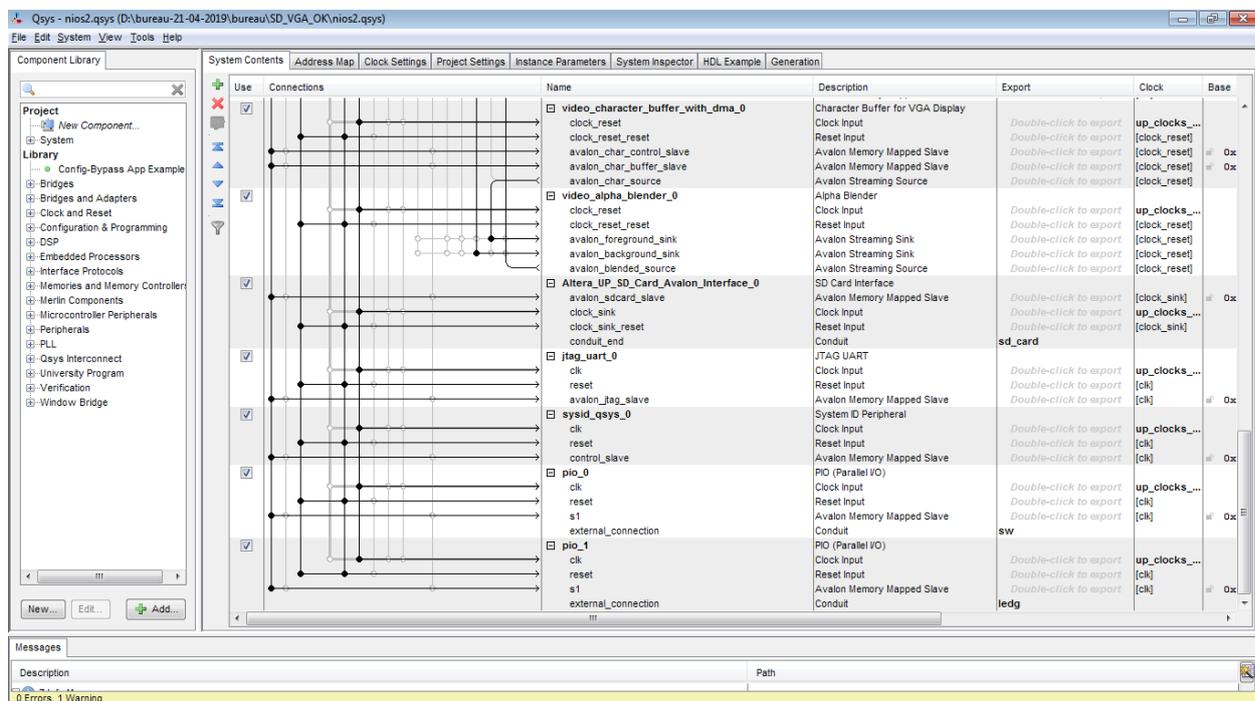


Figure 4.23: Système complet sur Qsys partie-3

Une fois la génération du système NIOS après compilation sous Qsys et assignement des pins, Figure 4.28. Ce système sera instancié dans un projet Quartus recompilé pour obtenir le fichier Bit-Stream à télécharger vers le FPGA. Nous obtenons ainsi la partie matérielle de la conception.

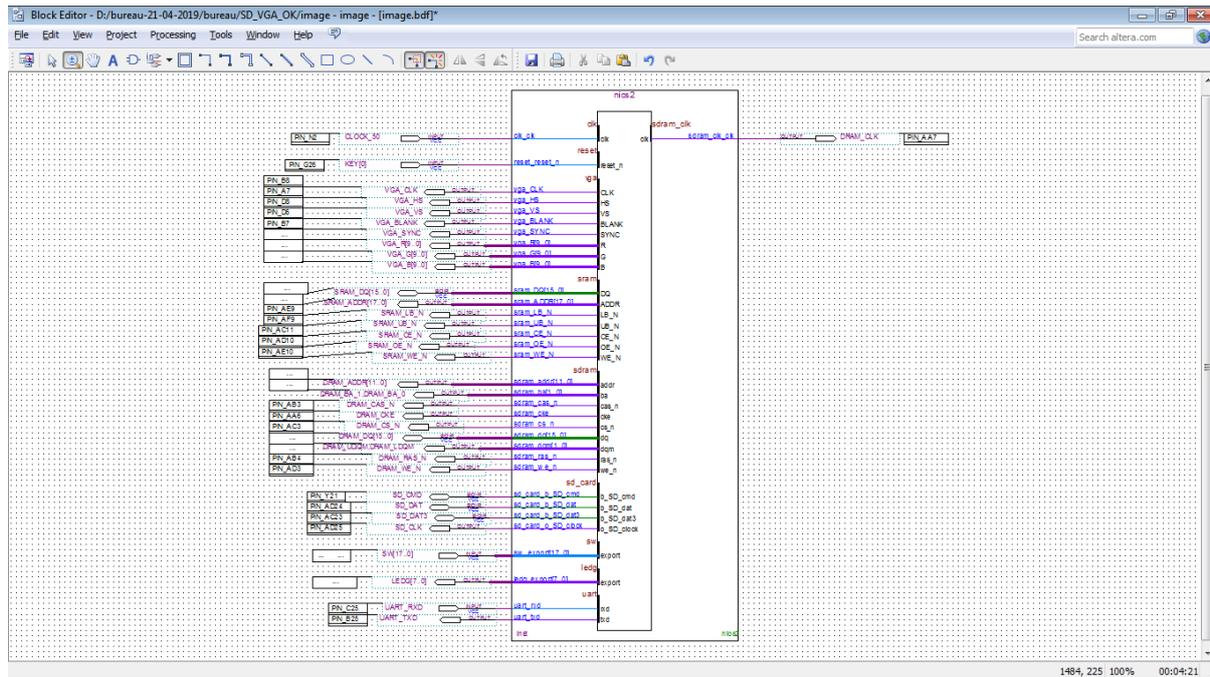


Figure 4.24: block diagramme du système

La figure suivante donne le rapport de compilation et les ressources matérielles consommées.

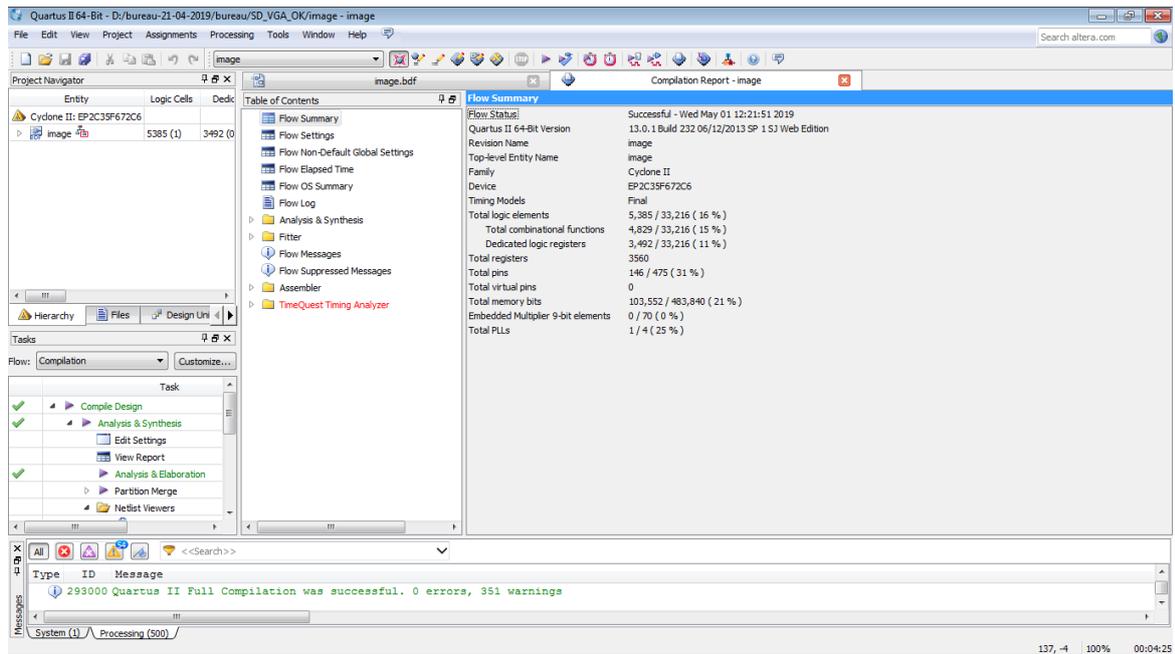
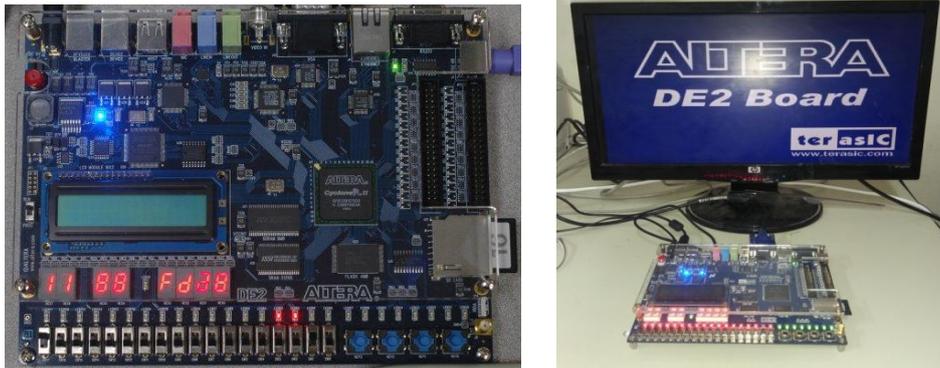


Figure 4.25 : Rapport de compilation du système

### b- Résultats expérimentaux

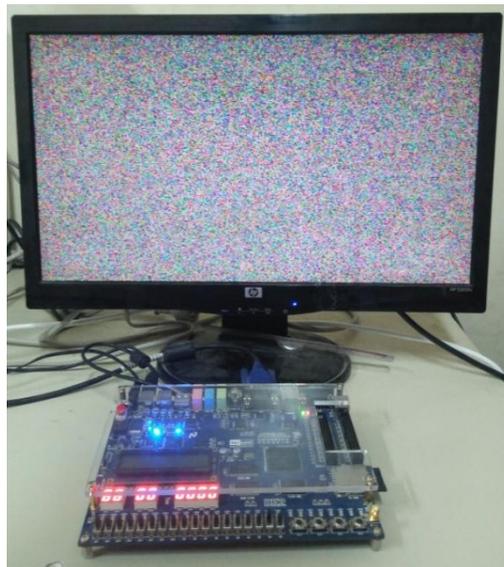
Le prototype développé donne les fonctionnalités suivantes :

➤ **Test de la carte sans l'implémentation hardware**



**Figure 4.26 :** État de système avant d'implémentation

➤ **Etat du système après implémentation matérielle**



**Figure 4.27:** Affichage VGA le cas d'une RAM vide avant transfert du programme

➤ **Etat du système après exécution de la partie software pour différentes images de tests**



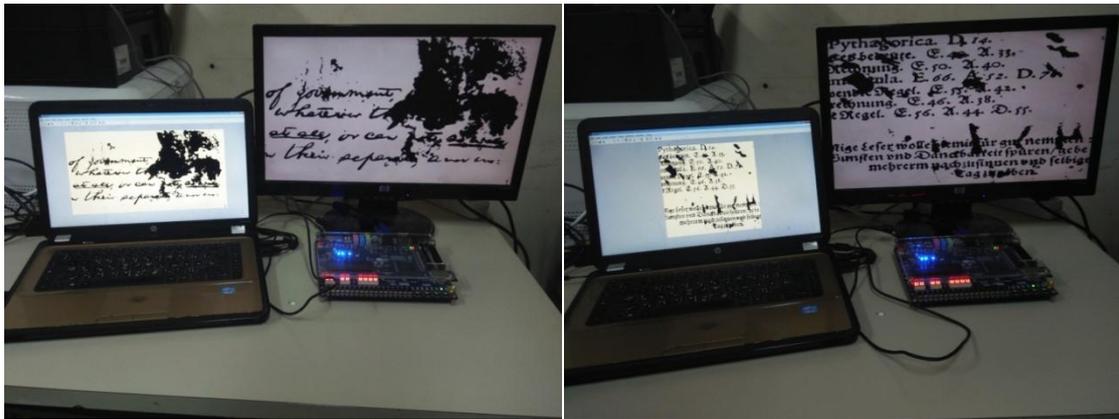
**Figure 4.28** : Résultat de binarisation : (a,c,e) Images originales, (b,d,f) Images binarisées

➤ **Comparaison avec le système classique**

Les figures qui suivent montrent une comparaison entre une binarisation d'images utilisant un système à usage général avec l'outil de programmation Matlab et une application implémentée dans un FPGA.

Les outils de mesures ont prouvé que les mêmes résultats dans les deux cas.

En note que les deux approches donne les même résultats ce qui concerne l'image binarisée.



**Figure 4.29** : Résultats de comparaison

#### 4.5. Conclusion

Dans ce chapitre, nous avons exploité les avancées actuelles dans la technologie des semi-conducteurs et des méthodologies de conception permettant le développement de systèmes numériques complexes sur puce.

Nous avons développé un système de binarisation en utilisant une approche de conception mixte logicielle/matérielle. La partie matériel comporte le système physique constitué de circuits numériques modélisés par un langage de description de matérielle VHDL, la partie programme est réalisée en utilisant le langage C sous Eclipse.

Nous pensons que le prototype du système décrit dans le chapitre sera finalisé en ajoutant les modules de l'estimation du background et le prétraitement prochainement.

# *Bibliographie*

- [1] Intel-Altera Corporation Disponible sur: <http://www.altera.com/products/software/products/quartus2/qts-index.html>.
- [2] P. Kadionik, P. Nouel, A. Atitallah, P. Dondon, « L'Enseignement des Systèmes Numériques », présenté à 5ième Colloque sur l'Enseignement des Technologies et des Sciences de l'Information et des Systèmes (CETIS '05), Nancy, France, 2005.
- [3] P. Kadionik, P. Nouel, A. Ben Atitallah, « Linux et le System on Chip SoC », présenté à LSM '05, Dijon, France, 2005.
- [4] S. Malhotra, T. P. Borer, D. P. Singh, S. D. Brown, « The Quartus University Interface Program: enabling advanced FPGA research », in *Proceedings. 2004 IEEE International Conference on Field- Programmable Technology (IEEE Cat. No.04EX921)*, Brisbane, NSW, Australia, 2004, pp. 225-230.
- [5] L. Abdelhai, I. Zohir, S. Samir, N. Abdelkrim, et H. Mustapha, « Contribution to the implementation of image mosaicing algorithm on FPGA using NIOS II softcore », in *2015 3rd International Conference on Control, Engineering & Information Technology (CEIT)*, Tlemcen, Algeria, 2015, pp. 1-5.
- [6] Institute of Electrical and Electronics Engineers, Éd., *2011 International Conference on Advanced Computer Science and Information Systems (ICACSIS 2011): Jakarta, Indonesia, 17 - 18 December 2011 ; [proceedings]*. Piscataway, NJ: IEEE, 2011.
- [7] Institute of Electrical and Electronics Engineers, Éd., *2009 International Conference on New Trends in Information and Service Science: (NISS 2009) ; Beijing, China, 30 June - 2 July 2009 ; [including 7 workshops]*. Piscataway, NJ: IEEE, 2009.
- [8] S. Saha, Md. A. Rahman, A. Thakur, « Design and implementation of SPI bus protocol with Built-in-self-test capability over FPGA », in *2014 International Conference on Electrical Engineering and Information & Communication Technology*, Dhaka, Bangladesh, 2014, pp. 1-6.
- [9] L. Pyrgas, A. Kalantzopoulos, E. Zigouris, « Design and Implementation of an Open Image Processing System based on NIOS II and Altera DE2 - 70 Board », *J. Eng. Sci. Technol. Rev.*, vol. 9, n° 5, pp. 51-55, oct. 2016.
- [10] Y. Fang et X. Chen, « Design and Simulation of UART Serial Communication Module Based on VHDL », in *2011 3rd International Workshop on Intelligent Systems and Applications*, Wuhan, China, 2011, pp. 1-4.
- [11] A. Ramakrishna.M, K. Kumar.M, et A. Purna Ramesh, « Hardware Acceleration of Histogram Equalization and Image Sharpening Filter on NIOSII Processor Based SOC on FPGA », *Int. J. Comput. Appl.*, vol. 31, n° 1, pp. 48-54, oct. 2011.
- [12] Institute of Electrical and Electronics Engineers, IEEE Industrial Electronics Society, et Boğaziçi Üniversitesi, Éd., *2014 IEEE 23rd International Symposium on Industrial Electronics (ISIE 2014): Istanbul, Turkey, 1 - 4 June 2014 ; [proceedings]*. Piscataway, NJ: IEEE, 2014.

- [13] M. khitas, L. Ziet, et F. RADJAH, « FPGA design and Implementation of Histogram Algorithm for image processing », International Workshop on Cryptography and its Applications (IWCA'16), Oran, Algerie, 2016, pp. 23.
- [14] L. Ziet, M. Khitas, et F. RADJAH, « SOPC implementation of image histogram for image binarization processing », International Conference on Advances in Sciences (ICAS'16), Istanbul, Turkey, 2016.
- [15] C.-H. Yang, H.-C. Wu, et S.-F. Su, « Implementation of Encryption Algorithm and Wireless Image Transmission System on FPGA », *IEEE Access*, vol. 7, p. 50513-50523, 2019.

## CONCLUSION GENERALE

## Conclusion générale

Les objectifs principaux visés par ce thème de recherche ont été atteints. En effet, il s'agit d'un premier point où il fallait développer un Framework concurrentiel qui doit traiter et binariser les images. Le deuxième est d'utiliser ce Framework dans une application embarquée dans un système électronique reconfigurable.

Nous avons montré que la méthode utilisée par le Framework est efficace pour différentes images présentant différentes tailles et niveaux de dégradation. Cette méthode est concurrente vis-à-vis des méthodes les plus récentes rencontrées dans la littérature. La méthode proposée a été testée et validée sur différentes bases de données couvrant de nombreux types de dégradation, à savoir DIBCO 2009, DIBCO 2011, DIBCO 2013 et BICKLEY DIARY. Les résultats expérimentaux confirment que la performance de la méthode proposée est compétitive par rapport à celle des méthodes existantes. Pour les images de documents fortement dégradés, notamment lorsque l'image de document contient des points d'ancrage ou des zones de texte avec un arrière-plan sombre et surtout pour la bibliothèque BICKLEY DIARY connue pour ses images fortement dégradées

D'autre part, nous avons pu implémenter un système embarqué qui réalise l'opération de binarisation d'images par le calcul du seuil d'Otsu. Le système général se présente sous forme d'une conception mixte autour d'un processeur NIOS accompagné d'un accélérateur hardware pour faire le traitement, la transmission et l'archivage des images.

Comme perspectives nous envisageons :

Afin d'améliorer d'avantage la performance de la méthode de binarisation proposée :

- Régler minutieusement la taille de la fenêtre de traitement qui joue un rôle important pour les résultats de binarisation :
- Appliquer un algorithme de binarisation qui utilise les caractéristiques du contour du texte pour déterminer le seuil de binarisation.

Diversifier les modèles de binarisation en ajoutant d'autres implémentations hardware des techniques de binarisation (globales et locales).

Ajouter les modules de l'estimation du background et du prétraitement au système pour réaliser une implémentation hardware la méthode proposée.

## ملخص

عملت رقمنة المستندات القديمة على فتح مجالات بحثية جديدة تتمثل في استعادة المستندات المتدهورة والتعرف البصري على الحروف اين تكون عملية ترميز الصورة مطلوبة كخطوة ما قبل المعالجة . نتناول هذه الأطروحة ترميز المستندات المتدهورة عبر وسائل الاتصال وفقاً لطريقتين . الطريقة الأولى تقدم تقنية جديدة لترميز المستندات المتدهورة . حيث قدمنا تقنية جديدة تعتمد على عامل التصفية المتوسطة لتقدير الخلفية . بالنسبة للطريقة الثانية ، نقدم تنفيذ خوارزمية ترميز الصورة باستخدام طريقة العتبة Otsu على بنية قابلة لإعادة التشكيل (FPGA)

## Résumé

La numérisation des documents anciens ouvre de nouvelles perspectives de recherche telles que la restauration des documents dégradés et la reconnaissance optique des caractères où le processus de binarisation d'images est requis en tant qu'étape de prétraitement. Cette thèse traite la binarisation des documents dégradés via des moyens de communication selon deux approches. La première approche, présente une nouvelle technique pour la binarisation des documents dégradés. Dans laquelle nous avons introduit une nouvelle technique basée sur le filtre médian pour l'estimation du background. Pour la deuxième approche, nous présentons l'implémentation de l'algorithme de binarisation d'image en utilisant la méthode de seuillage Otsu sur architecture reconfigurable (FPGA).

## Abstract

Scanning of old documents opens new research perspectives such as the restoration of degraded documents and optical character recognition where the image binarization process is required as a pretreatment step. This thesis deals with the binarization of degraded documents via means of communication according to two approaches. The first approach presents a new technique for binarisation of degraded documents. In which we introduced a new technique based on the median filter for the estimation of the background. For the second approach, we present the implementation of the image binarization algorithm using the Otsu method on reconfigurable architecture (FPGA)