

UNIVERSITE SAAD DAHLAB DE BLIDA 1

Faculté de Technologie
Département d'Electronique

THESE DE DOCTORAT EN SCIENCES

Spécialité : Electronique

**STRATEGIES DE TRAITEMENT ET DE REHAUSSEMENT
AUTOMATIQUES DES IMAGES DE MANUSCRITS ANCIENS**

Par

HADJADJ Zineb

Devant le jury composé de :

N. BENBLIDIA	Professeur, U. de Blida 1	Présidente
Y. CHIBANI	Professeur, USTHB	Examineur
Y. SMARA	Professeur, USTHB	Examineur
N. BELKHAMSA	Maître de Conférences A, U. de Blida 1	Examineur
M. CHERIET	Professeur, U. de Québec, Montréal, Canada	Directeur de thèse
Y. CHERFA	Professeur, U. de Blida 1	Co-Directeur de thèse

Blida, Février 2019

RÉSUMÉ

Les documents anciens représentent un patrimoine culturel très précieux. Ces documents encourent une dégradation progressive, et sont menacés d'un réel danger de disparition. Les détériorations les plus communes que l'on retrouve dans les images de documents anciens sont : l'effet de transparence, la présence de taches d'humidité absorbée par le papier, les variations de la couleur du papier, la présence de plis et de déchirures et les déformations dues à la courbure naturelle des pages. Le traitement automatique de ces documents en vue de leur restauration offre un avantage certain, mais il est confronté à de nombreuses difficultés dues au mauvais état de ces documents.

Nous nous intéressons dans le présent travail aux prétraitements d'images de documents anciens, afin de les rendre lisibles et accessibles, et plus précisément à la binarisation, considérée comme l'une des opérations de prétraitement les plus importantes. La binarisation est l'une des techniques de rehaussement qui produit des informations binaires, faciles à manipuler par des méthodes d'analyse de haut niveau (OCR, par exemple), et moins coûteuse en termes de calcul, par rapport aux images couleurs ou en niveaux de gris. Deux contributions majeures sont présentées dans cette thèse pour la binarisation d'images de documents anciens.

Notre première contribution est une nouvelle méthode de binarisation destinée aux images de manuscrits anciens en niveaux de gris ou en couleurs. L'approche proposée utilise l'image de contraste définie par le maximum et le minimum locaux, pour détecter en premier lieu un sous-ensemble de pixels de texte les plus probables, qui est utilisé pour générer une carte d'initialisation. Elle effectue ensuite une segmentation par un nouveau modèle de contour actif que nous avons proposé, et utilise en dernier un seuillage moyen pour obtenir l'image binarisée finale.

Notre deuxième contribution est une amélioration de la méthode de binarisation de Sauvola. La méthode proposée coopère parallèlement la méthode de Sauvola et la méthode de contraste, définie par le maximum et le minimum locaux. Les deux méthodes de binarisation sont intégrées dans le but d'atteindre une meilleure détection des caractères, aussi bien ceux qui sont bien contrastés, que ceux qui sont faiblement contrastés, présents dans le même document, sans ajuster manuellement les valeurs des paramètres utilisateur. La méthode est proposée dans le but de parvenir à une technique de binarisation rapide et efficace en même temps.

Les deux méthodes de contribution ont été validées sur les bases d'images des DIBCOs sur lesquelles elles ont été quantifiées, et comparées aux autres algorithmes existants. La validité des deux méthodes proposées est aussi montrée sur des images de documents anciens arabes, issues de la Bibliothèque Nationale d'Algérie et du Centre National des Manuscrits d'Adrar. Différentes expérimentations ont été effectuées, et les résultats obtenus par les deux méthodes de contribution sont encourageants.

Mot-clés : Binarisation d'images de documents anciens, restauration d'images de documents anciens, contours actifs, méthode de Sauvola.

ABSTRACT

Old documents represent a very valuable cultural heritage. These documents are subject to progressive degradation and are threatened with a real danger of disappearance. The most common deterioration found in old document images is the transparency effect, the presence of moisture stains absorbed by the paper, variations in the color of the paper, the presence of folds and tears, and deformations due to the natural curvature of the pages. The automatic processing of these documents for restoration offers a definite advantage, but it faces many difficulties due to the poor state of these documents.

In the present work, we are interested in the pre-processing of old documents images, in order to make them readable and accessible, and more precisely to binarization, which is considered one of the most important pre-processing operations. Binarization is one of the enhancement techniques that produces bi-level information which is easy to handle using methods of analysis methods (OCR, for instance) and is less computationally costly to process than 256 levels of grey or color images. Two major contributions are presented in this thesis for the binarization of old document images.

Our first contribution is a new binarization method for images of old manuscripts in gray or color levels. The proposed approach uses the contrast image defined by the local image maximum and minimum to first obtain a rough binarization that retains the main features of the input image, the result is used to generate a stroke map initialization, and then performs a segmentation using a propose active contour model, at the end it uses an average thresholding to produce the final binarization.

Our second contribution is an improvement of Sauvola's binarization method. In the proposed method we use the image contrast that is defined by the local image minimum and maximum in combination with the computed Sauvola's binarization step to guarantee good quality binarization for both low and correctly contrasted objects inside a single document, without adjusting manually the user-defined parameters to the document content. The method is proposed in order to achieve a fast and efficient binarization technique at the same time.

The two contribution methods were validated on the image bases of the DIBCOs on which they were quantified and compared with the other existing algorithms. The validity of the two proposed methods is also shown on images of old arabic documents from the National Library of Algeria and the National Center of Manuscripts of Adrar. Various experiments have been carried out and the results obtained by the two contribution methods are encouraging.

Keywords: Old manuscript images binarization, old manuscript images restoration, active contours, Sauvola's method.

ملخص

تمثل الوثائق القديمة تراثا ثقافيا قيما للغاية، هذه الوثائق معرضة لتدهور تدريجي وخطر حقيقي للاختفاء. إن التدهور الأكثر شيوعا و الموجود في صور الوثائق القديمة هو ظهور النص المكتوب خلف الصفحة بفعل شفافية الورق، وجود بقع رطوبة ممتصة من طرف الورق، تغيرات في لون الورق، وجود طيات و تمزقات، اعوجاج الصفحات وانحناء وتقوس الكتابة. المعالجة الآلية لهذه الوثائق من أجل ترميمها أصبح ضرورة ملحة لبقائها بصفة دائمة في حالة حسنة للأجيال القادمة، ولكنها تواجه صعوبات كثيرة بسبب سوء حالة هذه الوثائق.

العمل المقدم في هذه الأطروحة يهتم بالمعالجة الأولية لصور الوثائق القديمة من أجل بقائها بصفة دائمة في حالة حسنة، وبدقة أكثر يهتم بالتثنية التي تعتبر واحدة من أهم عمليات المعالجة الأولية. تعتبر تثنية صور الوثائق القديمة واحدة من تقنيات التحسين التي تنتج معلومات ثنائية من السهل التعامل معها من طرف أساليب التحليل عالية المستوى (OCR، على سبيل المثال)، وأقل تكلفة مقارنة بالصور الملونة او الرمادية. مشاركتين رئيسيتين سيتم عرضهما في هذه الأطروحة من أجل استخراج النصوص من صور الوثائق القديمة.

المشاركة الاولى عبارة عن طريقة جديدة لتثنية صور الوثائق القديمة الرمادية او الملونة. تستخدم الطريقة المقترحة صورة التباين التي يحددها الحد الأقصى والحد الأدنى المحليين للكشف أولا عن مجموعة فرعية من بكسالات النص الأكثر احتمالا، والتي تستخدم لإنشاء بطاقة البدء، ثم تجزئة الصورة بنموذج جديد مقترح من المحيطات المتحركة، في النهاية يستخدم متوسط العتبة للحصول على الصورة الثنائية النهائية.

أما المشاركة الثانية فتهدف الى تحسين طريقة التثنية ل Sauvola. الطريقة المقترحة تدمج بالتوازي طريقة Sauvola وطريقة التباين التي يحددها الحد الأقصى والحد الأدنى المحليين، يتم دمج الطريقتين من أجل استخراج أفضل للأحرف المتباينة جيدا وحتى الأحرف المنخفضة التباين مع الخلفية الموجودة في نفس الصورة، وذلك دون ضبط قيم العوامل يدويا. الطريقة المقترحة تهدف الى تحقيق تقنية تثنية سريعة وفعالة في نفس الوقت.

تم تجربة طريقتي المشاركة على مجموعات صور DIBCOs حيث تم مقارنتهما مع الطرق الأخرى. وتظهر فعالية الطريقتين المقترحتين أيضا على صور الوثائق العربية القديمة من مكتبة الجزائر الوطنية والمركز الوطني للمخطوطات في أدرار. قمنا بعدة تجارب و النتائج المتحصل عليها بطريقتي المشاركة مشجعة.

كلمات المفتاح: تثنية صور الوثائق القديمة، ترميم صور الوثائق القديمة، المحيطات المتحركة، طريقة 'Sauvola'.

REMERCIEMENTS

Les travaux présentés dans cette thèse ont été réalisés au Centre de Recherche sur L'Information Scientifique et Technique (CERIST).

Je tiens tout d'abord à remercier ALLAH de m'avoir donné la volonté et le courage qui m'ont permis de réaliser ce travail. Veuillez t-il me guider dans le droit chemin.

Je tiens à remercier mon directeur de thèse, professeur Mohamed Cheriet, qui est à l'origine de cette thèse. Je le remercie de m'avoir aidé à bien mener ces recherches par sa grande rigueur scientifique et ses conseils précieux.

Je remercie le professeur Yazid Cherfa, mon co-directeur de thèse pour la troisième fois, pour son aide précieuse et ses conseils, et pour m'avoir fait profiter de son expérience.

Je remercie aussi tous les membres du jury pour m'avoir fait l'honneur d'accepter de juger cette thèse:

N. BENBLIDIA : Professeur, U. de Blida 1

Y. CHIBANI : Professeur, USTHB

Y. SMARA : Professeur, USTHB

N. BELKHAMSA : Maître de Conférences A, U. de Blida 1

Je tiens à remercier mon chef d'équipe au CERIST, Monsieur Abdelkrim Meziane, pour son soutien scientifique, mais également pour sa bonne humeur, sa disponibilité et ses conseils. Simplicité, gentillesse, compréhension et patience ne sont que quelques aperçus de ses qualités humaines qui font qu'il a tout mon respect.

Mes remerciements spéciaux vont bien sûr à ma famille, qui m'a aidé, soutenu, encouragé et qui m'a offert la chance de faire toutes mes études.

Mes derniers remerciements, mais non des moindres, vont à toute personne ayant participé de près ou de loin à la réalisation de ce travail de recherche.

Zineb HADJADJ

DEDICACES

Je dédie cette thèse aux deux personnes que je chérisse le plus au monde :

Ma mère et mon père.

A toute ma famille.

A ceux que j'aime...et à ceux qui m'aiment

Zineb HADJADJ

TABLE DES MATIERES

RÉSUMÉ	
REMERCIEMENTS	
TABLE DES MATIERES	
LISTE DES FIGURES	
LISTE DES TABLEAUX	
INTRODUCTION GÉNÉRALE	13
Contexte et problématique.....	13
Objectifs généraux, motivations et contributions.....	14
Plan général de la thèse.....	15
PARTIE I : ÉTAT DE L'ART	
1. CHAPITRE 1 : DEGRADATIONS ET PRETRAITEMENT DES IMAGES DE DOCUMENTS ANCIENS	17
1.1. Introduction.....	17
1.2. Définition d'une dégradation.....	17
1.3. Types de dégradations.....	19
1.3.1. Dégradations intrinsèques au document.....	19
1.3.2. Dégradations dues à la numérisation.....	20
1.4. Sources de dégradations.....	23
1.4.1. Sources de dégradations physiques.....	23
1.4.1.1. Dégradations chimiques.....	23
1.4.1.2. Dégradations biologiques.....	24
1.4.1.3. Dégradations physiques d'origine humaine.....	25
1.4.2. Sources de dégradations externes.....	25
1.5. Etude faite sur les manuscrits du centre d'Adrar.....	26
1.6. Prétraitement des images de documents anciens.....	27
1.7. Conclusion.....	30
2. CHAPITRE 2 : BINARISATION D'IMAGES DE DOCUMENTS: ÉTAT DE L'ART	31
2.1. Introduction.....	31
2.2. Définition de la binarisation.....	31
2.3. Catégorisations des méthodes de binarisation.....	31
2.3.1. Binarisation basée sur le seuillage.....	33
2.3.1.1. Méthodes de seuillage global.....	33
2.3.1.2. Méthodes de seuillage local 'adaptatif'.....	35
2.3.2. Binarisation basée sur la classification.....	42
2.3.2.1. Classification basée sur 'k-means'.....	42
2.3.2.2. Classification basée sur les champs de Markov 'Markov random fields'.....	43
2.3.2.3. Classification basée sur 'SVM'.....	44

2.3.2.4. Classification basée sur les réseaux de neurones artificiels (ANN pour <i>Artificial Neural Network</i>).....	45
2.3.3. Binarisation basée sur la segmentation.....	47
2.3.3.1. Binarisation basée sur les contours actifs.....	47
2.3.4. Combinaison de méthodes de binarisation.....	47
2.4. Conclusion.....	48

3. CHAPITRE 3 : CONTOURS DEFORMABLES (ACTIFS): DEFINITION ET ETAT DE L'ART **50**

3.1. Introduction.....	50
3.2. Définition des modèles des contours déformables (actifs).....	51
3.3. Modèles de contours déformables.....	51
3.3.1. Contour actif classique, paramétrique ou snake.....	51
3.3.2. Modèle déformable géométrique.....	54
3.3.3. Méthode des ensembles de niveaux d'Osher –Sethian.....	56
3.3.4. Modèle déformable géodésique.....	60
3.3.5. Modèle de Mumford-Shah.....	61
3.4. Conclusion.....	62

PARTIE II : CONTRIBUTIONS

4. CHAPITRE 4 : CONTRIBUTIONS A LA BINARISATION D'IMAGES DE DOCUMENTS DEGRADEES PAR CONTOURS ACTIFS **63**

4.1. Introduction.....	63
4.2. Motivations et objectifs.....	63
4.3. Description de la technique proposée.....	64
4.3.1. Etape 1 : Initialisation du contour actif.....	65
4.3.1.1. Construction de l'image de contraste.....	65
4.3.1.2. Détection des pixels de haut contraste.....	68
4.3.2. Etape 2 : Evolution du contour actif.....	69
4.3.2.1. Force de régularisation de la longueur 'Length regularization force' (F_R).....	70
4.3.2.2. Force d'expansion ou de contraction de la zone 'Area expansion or contraction force' (F_A).....	71
4.3.2.3. Terme d'ajustement de données 'Data fitting term' (F_{DF}).....	72
4.3.2.4. Formulation avec les Level Set.....	73
4.3.3. Etape 3: Optimisation de la binarisation.....	75
4.4. Expérimentations, résultats et discussion.....	76
4.4.1. Bases de test.....	76
4.4.2. Sélection des paramètres.....	77
4.4.3. Implémentation de la méthode proposée.....	78
4.4.3.1. Validation de notre méthode sur les bases de données des DIBCOs, et comparaison avec les méthodes de l'état de l'art.....	78
4.4.3.2. Validation de la méthode proposée sur des images issues de la Bibliothèque Nationale d'Algérie et du Centre National des Manuscrits d'Adrar.....	84
4.4.4. Mesures de performance.....	88

4.4.4.1. $F - \text{mesure '}F - \text{measure'}$ (FM).....	90
4.4.4.2. Rapport signal maximum sur bruit (<i>PSNR</i>).....	90
4.4.4.3. Métrique de taux négatif (<i>NRM</i>).....	90
4.4.4.4. Métrique de pénalité de fausse-classification (<i>MPM</i>).....	91
4.4.5. Influence de la valeur du paramètre d'échelle σ	92
4.4.6. Matériel utilisé, coûts de calculs et complexité de la méthode proposée.....	93
4.5. Conclusion.....	94

5. CHAPITRE 5 : CONTRIBUTIONS A LA BINARISATION D'IMAGES DE DOCUMENTS DEGRADEES PAR L'ALGORITHME ISAUVOLA **95**

5.1. Introduction.....	95
5.2. Motivations et objectifs.....	95
5.3. Algorithme de Sauvola et ses limites.....	96
5.3.1. Algorithme de Sauvola.....	96
5.3.2. Limites de la méthode de Sauvola.....	97
5.3.2.1. Mauvaise binarisation des objets faiblement contrastés.....	97
5.3.2.2. Mauvaise binarisation des documents, dont la taille des objets est très variable.....	98
5.3.2.3. Mauvaise binarisation du texte texturé.....	99
5.3.2.4. Interférence d'objet spatial.....	99
5.4. Description de la technique proposée.....	100
5.4.1. Etape 1 : Initialisation par la méthode du contraste.....	101
5.4.2. Etape 2 : Binarisation de Sauvola.....	101
5.4.3. Etape 3 : Coopération parallèle.....	102
5.5. Expérimentations, résultats et discussion.....	102
5.5.1. Validation de la méthode proposée sur les bases de données des DIBCOs, et comparaison avec les méthodes de l'état de l'art.....	102
5.5.2. Validation de la méthode proposée sur des images issues de la Bibliothèque Nationale d'Algérie et du Centre National des Manuscrits d'Adrar.....	108
5.5.3. Comparaisons avec la première méthode de contribution.....	112
5.6. Conclusion.....	113

CONCLUSION ET PERSPECTIVES	114
Bilan.....	114
Perspectives.....	116

ANNEXES

Annexe A : Interface graphique développée

Annexe B : Liste des publications de l'auteur

Bibliographie

LISTE DES FIGURES

Figure 1. Exemples d'images de documents dégradés.	13
Figure 1.1. Exemples d'images de documents.	18
Figure 1.2. Exemples d'images naturelles.	18
Figure 1.3. Exemples de dégradations intrinsèques au document.	20
Figure 1.4. Exemples de dégradations engendrées par la numérisation.	23
Figure 1.5. Exemples de dégradations chimiques.	24
Figure 1.6. Exemples de dégradations biologiques.	24
Figure 1.7. Exemples de dégradations d'origine humaine.	25
Figure 1.8. Etudes faites sur les manuscrits du centre d'Adrar.	26
Figure 1.9. Position du prétraitement dans le processus d'analyse et de reconnaissance de documents.	27
Figure 1.10. Les opérations de prétraitement les plus courantes.	30
Figure 2.1. Schéma représentant quelques approches de binarisation.	33
Figure 2.2. Résultats de la binarization par la méthode d'Otsu.	34
Figure 2.3. Résultats de la binarization par la méthode de Niblack.	36
Figure 2.4. Résultats de la binarization par la méthode de Sauvola.	37
Figure 2.5. Résultats de la binarisation de l'image HW4 de la base de données de DIBCO'11.	40
Figure 3.1. Énergie externe (indiquée en flèches) pour un objet qui contient des concavités.	53
Figure 3.2. Méthode des Snakes appliquée à la détection de plusieurs objets.	53
Figure 3.3. Sens de propagation d'une courbe sous l'action de la force de courbure, dans la direction de la normale en chaque point.	54
Figure 3.4. Représentation de la fonction des ensembles de niveaux $\Phi(x, y)$ à un instant donné t .	56
Figure 3.5. Fonction distance Φ signée au contour C .	57
Figure 3.6. Exemple de la binarisation d'une image de document.	59
Figure 4.1. Organigramme des étapes de la méthode proposée.	65
Figure 4.2. Image de gradient et de contraste.	67
Figure 4.3. Détection des pixels de haut contraste.	69
Figure 4.4. Effet de la force de régularisation de la longueur de (4.4) sur l'évolution du contour actif. De gauche à droite : initialisation, à $t = 0.4$, $t = 4$ et $t = 40$.	71
Figure 4.5. Effet de la force d'expansion ou de contraction de la zone de (4.5) avec $D = -1$ sur l'évolution du contour actif. De gauche à droite : initialisation, à $t = 0.4$, $t = 1.4$ et $t = 2$.	72

Figure 4.6. Effet du terme d'ajustement de données 'Data fitting term' sur l'évolution du contour actif.	73
Figure 4.7. Résultat de la deuxième étape.	74
Figure 4.8. Les caractères du texte faiblement contrastés (entourés par des cercles bleus).	75
Figure 4.9. Le contour actif final du modèle proposé (lorsque le contour actif stagne).	76
Figure 4.10. Résultat de la binarisation finale par la méthode proposée.	76
Figure 4.11. Résultats de la binarisation de l'image HW2 de la base de données de DIBCO'11.	80
Figure 4.12. Résultats de la binarisation d'une image dégradée.	81
Figure 4.13. Résultats de la binarisation de l'image H04 de la base de données de DIBCO'09.	83
Figure 4.14. Résultats de la binarisation de l'image P04 de la base de données de DIBCO'09.	83
Figure 4.15. Etiquette de la première page du livre (أشرف الوسائل الى فهم الشمائل).	84
Figure 4.16. Résultats de la restauration de quelques images du livre (أشرف الوسائل الى فهم الشمائل) par notre méthode.	87
Figure 4.17. Etiquette de la première page du livre (الأصول).	87
Figure 4.18. Résultats de la restauration de quelques images du livre (الأصول) par notre méthode.	88
Figure 4.19. Exemple d'évaluation d'un résultat.	89
Figure 4.20. Résultats de la binarisation de notre méthode avec le paramètre d'échelle $\sigma = 0.5, 1, 3$ et 5 (de gauche à droite et de haut en bas).	93

Figure 5.1. Principe de la méthode de Sauvola.	96
Figure 5.2. Influence du paramètre k sur la binarisation des objets faiblement contrastés.	98
Figure 5.3. Influence du paramètre ω sur le résultat de la binarisation de Sauvola.	99
Figure 5.4. Influence d'une grande fenêtre ω (interférence d'objet).	100
Figure 5.5. Organigramme de la méthode proposée.	101
Figure 5.6. Résultats de la binarisation de l'image HW2 de la base de données de DIBCO'11.	103
Figure 5.7. Résultats de la binarisation de l'image H12 de la base de données de HDIBCO'12.	105
Figure 5.8. Résultats de la binarisation de l'image HW4 de la base de données de DIBCO'13.	105
Figure 5.9. Résultats de la binarisation de l'image H04 de la base de données de HDIBCO'10.	106
Figure 5.10. Résultats de la binarisation de l'image PR3 de la base de données de DIBCO'13.	107
Figure 5.11. Résultats de la binarisation de l'image P03 de la base de données de DIBCO'09.	108
Figure 5.12. Résultats de la binarisation de quelques images de documents par la méthode proposée.	111
Figure 5.13. Comparaison entre les résultats des deux méthodes de contribution.	112

Figure A.1. Notre système de traitement d'image et de binarisation des images de documents “ (ADIP) Arabic Document Images Processing ”.	118
Figure A.2. Menu de fonctions permettant d'appliquer des traitements de base à l'image.	119

LISTE DES TABLEAUX

Tableau 1.1. Les dégradations de 43 manuscrits du centre d'Adrar.	26
Tableau 2.1. Comparaison des méthodes globales et locales.	38
Tableau 3.1. Les principes du modèle de contour actif : paramétrique et géométrique.	56
Tableau 4.1. Résultats d'évaluation sur la base de données de DIBCO 2009.	91
Tableau 4.2. Résultats d'évaluation sur la base de données de H-DIBCO 2012.	92
Tableau 4.3. Comparaison des temps de calcul de l'implémentation.	94

INTRODUCTION GÉNÉRALE

1. Contexte et problématique

Les documents anciens conservés dans la bibliothèque nationale d'Algérie et le centre national des manuscrits d'Adrar forment une bonne partie du patrimoine culturel et scientifique algérien (Figure 1). Ces manuscrits encourrent une dégradation progressive et sont menacés d'un réel danger de disparition. La conservation de ces documents constitue aujourd'hui un nécessité incontournable. Une méthode de conservation consiste à les numériser par caméra, scanner ou autres équipements, et à les stocker ainsi sous format image. Cependant, seule, la numérisation n'est pas suffisante, elle doit être accompagnée d'outils informatiques, permettant leur analyse et leur reconnaissance.

La littérature dans le domaine, et l'étude d'un nombre important de documents anciens, montrent que différents types de problèmes peuvent affecter les documents anciens. Ces problèmes sont principalement de deux types :

- Des problèmes dus à la numérisation : variation d'éclairage, position du document sur le scanner qui peut créer des inclinaisons, des bombages et des pliures du papier ;
- Des problèmes inhérents au document : le vieillissement des documents et le lieu de stockage font apparaître des taches d'humidité sur le papier, ainsi que la transparence de l'encre sur le recto. En outre, l'utilisation massive de ces documents fait apparaître des déchirures du papier et des traces de pliage des documents.



Figure 1. Exemples d'images de documents dégradés.

Pour remédier à ces problèmes, plusieurs projets de numérisation de documents anciens ont vu le jour, et de multiples techniques ont été développées par les chercheurs pour l'analyse des documents anciens. Que l'on fasse de l'identification du contenu (écriture, symboles ou objets, scripteur), de l'indexation et de la recherche, ou plus simplement de la classification, de l'organisation et du catalogage, on passe inévitablement par une étape de prétraitement. Cette étape regroupe un ensemble de traitements, appliqués sur les images de documents numérisés, afin de les préparer aux étapes suivantes dans le processus d'analyse.

De multiples techniques ont été développées pour le prétraitement des documents anciens, mais aucune d'elles n'est « universelle ». Elles sont généralement spécifiques à un seul défaut, et font souvent apparaître des effets secondaires sur d'autres éléments de l'image considérée. A cet effet, le centre national des manuscrits d'Adrar a mené une étude sur un grand nombre d'images, issues de ses documents anciens, qui a montré qu'il est plutôt rare qu'un problème existe seul dans un document. Très souvent, dans le même document, coexistent deux ou plusieurs défauts à la fois, ce qui nécessite des techniques de prétraitement plus élaborées.

2. Objectifs généraux, motivations et contributions

Le travail présenté dans cette thèse s'intègre dans la démarche de sauvegarde et de valorisation de documents patrimoniaux. Nous nous intéressons dans ce travail à une des opérations de prétraitement les plus usuelles et les plus importantes, à savoir la binarisation, pour l'amélioration de la qualité des images de documents historiques numérisés. La binarisation consiste à segmenter le document en deux classes, le fond, d'une part, et l'objet, c'est-à-dire les autres composantes (texte, graphique...), d'autre part. Elle permet de réduire de manière appréciable la quantité d'informations à manipuler. La qualité du résultat obtenu est cruciale, puisqu'elle conditionne fortement celle des étapes postérieures du processus d'analyse et de reconnaissance de documents, ce qui fait que l'opération de binarisation doit être effectuée le plus fidèlement possible. Elle doit, d'une part, conserver le maximum d'informations et détails présents dans l'image en entrée, et d'autre part, elle doit éliminer le bruit présent dans l'image.

Deux contributions majeures sont présentées dans cette thèse pour la binarisation d'images de documents anciens :

La première est une technique basée sur les contours actifs. D'après la recherche bibliographique que nous avons fait dans le domaine du traitement d'images de documents anciens, nous avons remarqué que les contours actifs qui sont très utilisés et très performants pour le traitement d'autres types d'images, comme les images médicales et satellitaires, sont très peu utilisés pour le traitement des images de documents anciens. Nos recherches se sont donc tournées vers ces outils.

Au lieu de concevoir une nouvelle méthode de binarisation, nous essayons, au cours de notre deuxième contribution, d'apporter des améliorations sur la méthode de binarisation existante de Sauvola, dont le but est d'améliorer, non seulement la performance de la méthode originale de Sauvola, mais aussi la robustesse sur différents types d'images de

documents dégradés. La méthode est proposée dans le but de parvenir à une technique de binarisation rapide et efficace en même temps.

Deux ensembles d'images de test ont été utilisés pour la validation des deux méthodes de contribution. Le premier est un ensemble d'images de documents arabes anciens, issues de la Bibliothèque Nationale d'Algérie et du Centre National des Manuscrits d'Adrar, alors que le second est une collection d'images, issues des bases d'images des DIBCOs. Différentes expérimentations ont été effectuées, et les résultats obtenus par les deux méthodes de contribution sont encourageants.

3. Plan général de la thèse

Cette thèse est structurée en cinq chapitres, répartis en deux parties. La première partie regroupe les chapitres 1, 2 et 3, et est consacrée à la présentation des principaux concepts, techniques, et travaux relatifs au sujet de thèse. La deuxième partie regroupe les chapitres 4 et 5, où nous présentons de manière détaillée les différentes contributions, la mise en œuvre, ainsi que les résultats obtenus à chaque niveau. Pour faciliter la lecture de ce document, les objectifs des différents chapitres de cette thèse sont présentés aux paragraphes suivants:

Chapitre 1 : Dégradations et prétraitement des images de documents anciens

Ce chapitre est consacré à la présentation des principales notions relatives à ce travail, à savoir les images de documents dégradés. Il présente les notions et les sources de bruit, de même que la dégradation des images de documents anciens, pour bien montrer leurs effets sur les images, et leurs influences sur les étapes d'analyse. Puis, il aborde la tâche des prétraitements, leurs objectifs, et leurs résultats sur les images de documents. Il aborde par la suite les opérations usuelles, utilisées pour le rehaussement de la qualité des images de documents numérisés.

Chapitre 2 : Binarisation d'images de documents : État de l'art

Le deuxième chapitre propose un état de l'art sur la binarisation des images de documents, sans se limiter à un type particulier de documents. Nous nous focalisons dans ce chapitre sur la binarisation d'images en niveaux de gris, car dans la plupart du temps, les images de documents en couleurs peuvent être transformées en niveaux de gris, sans perte significative d'information, exceptions faites des documents publicitaires et magazines. Nous exposons les différentes approches de binarisation existantes, et les diverses taxonomies proposées. Le reste du chapitre est réservé à la présentation des différentes techniques de binarisation les plus fréquemment citées dans la littérature.

Chapitre 3 : Contours déformables (actifs) : Définition et état de l'art

Ce chapitre présente le contexte des contours actifs, des premiers travaux de Kass et al. [Kass et al., 1988] aux méthodes, plus récentes, basées sur des critères de régions. Le chapitre présente le principe des contours actifs, et en décrit quelques modèles. Nous ressortons pour chaque modèle ses avantages et ses limites, ainsi que les différentes

contributions introduites pour pallier à certaines de ses limitations. En outre, une description de quelques travaux de traitement, en particulier de binarisation, d'images de documents, basés sur les contours actifs, proposés dans la littérature, sera aussi présentée.

Chapitre 4 : Contributions à la binarisation d'images de documents dégradés par contours actifs

Ce chapitre se détache des aspects théoriques abordés dans les trois premiers chapitres, et s'oriente vers la présentation de notre première contribution, qui s'intéresse à la binarisation d'images de documents dégradés. Cette contribution s'articule autour d'une binarisation, basée sur les contours actifs, et non pas sur le seuillage, comme la majorité des méthodes de la littérature. La technique proposée binarise des images de documents dégradés, en trois étapes afin d'avoir un résultat plus sûr. Au cours de ce chapitre, nous mentionnons tout d'abord les motivations et les objectifs de la méthode proposée, nous présentons la démarche suivie, nous décrivons les bases d'images utilisées, et nous exposons les expérimentations effectuées et les résultats obtenus.

Chapitre 5 : Contributions à la binarisation d'images de documents dégradés par l'algorithme ISauvola 'Improved Sauvola'

Le cinquième chapitre décrit notre deuxième contribution, à savoir la méthode de binarisation d'images de documents dégradés, proposée ISauvola pour « Improved Sauvola », et qui est une amélioration de la méthode de binarisation de Sauvola, trop sensible au bruit de fond de l'image. La méthode ISauvola a été choisie dans le but de parvenir à une binarisation rapide et efficace en même temps. Dans ce chapitre, nous mentionnons tout d'abord les motivations et les objectifs de la méthode proposée, nous présentons la démarche suivie, nous décrivons les bases d'images utilisées, et nous exposons les expérimentations effectuées et les résultats obtenus.

Enfin, nous terminons par une conclusion générale et quelques perspectives, pour améliorer ce travail et poursuivre éventuellement d'autres recherches dans ce domaine.

CHAPITRE 1

Dégradations et prétraitement des images de documents anciens

1.1. Introduction

La numérisation des documents anciens est une opération qui contribue à la préservation, à la valorisation et à la diffusion du patrimoine culturel, scientifique, technique, juridique... En effet, la consultation des livres anciens dans les bibliothèques est le plus souvent réservée aux spécialistes du domaine, afin de garantir la conservation du précieux patrimoine, souvent fragile, abîmé et exposé au risque de vol du document entier ou de quelques pages. A cet effet, plusieurs projets de numérisation de documents anciens ont été lancés tels que DEBORA¹, Philectre² et BAMBI³ [Kricha et al., 2006], afin de préserver et de mettre à la disposition des utilisateurs des documents originaux. Cependant la tâche n'est pas triviale. Les images de documents numérisés laissent apparaître souvent de nombreux défauts et dégradations. Certains dépendent de l'état de conservation du document (la qualité du papier, l'acidité de l'encre, l'humidité du lieu de stockage, vieillissement du support...) ou des modifications apportées par l'homme (annotations, soulignements, mauvaise restauration physique...). D'autres reviennent au processus même de numérisation (résolution insuffisante, courbure apparente, restitution non fidèle des couleurs...).

Dans ce chapitre, nous mettons le point sur les différents défauts et dégradations qui apparaissent souvent dans les images de documents anciens, pour bien comprendre leurs effets nuisibles sur la qualité des images, et afin de mieux sentir la difficulté de traitement de ces images de documents anciens. Les différents défauts et dégradations présentés dans ce chapitre sont classifiés en fonction de leurs causes.

1.2. Définition d'une dégradation

Drira [Drira, 2007] appelle dégradation tous les effets indésirables cumulés, nuisant à la lisibilité, au traitement ou à la conservation des images. Les dégradations ont plusieurs origines, et l'accumulation des défauts rend difficile leur séparation. Les dégradations ont des effets plus destructeurs sur les images de documents (Figure 1.1) que sur autres images, dites «naturelles» (Figure 1.2).

¹ Digital accEss to BOoks of the RenAissance : Projet européen ayant pour but de numériser et rendre plus accessibles des livres du XVIème siècle.

² Philologie électronique : Projet de numérisation.

³ Better Access to Manuscripts and Browsing of Images : Projet européen de numérisation de manuscrits médiévaux.



Figure 1.1. Exemples d'images de documents.



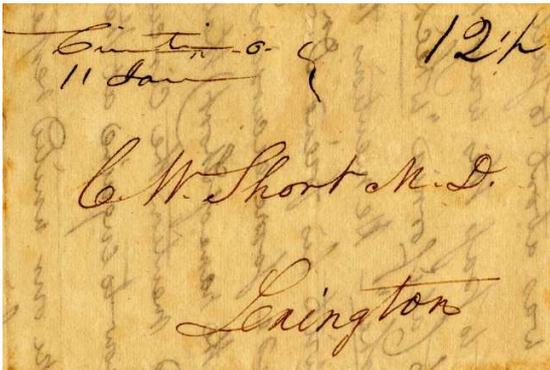
Figure 1.2. Exemples d'images naturelles.

1.3. Types de dégradations

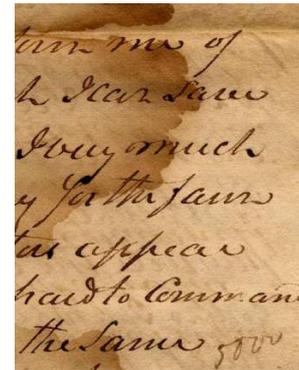
1.3.1. Dégradations intrinsèques au document

Cette catégorie de dégradations regroupe tous les défauts qui proviennent essentiellement de l'état du document original [Rabeux, 2013]. Elle constitue ainsi la majeure partie des défauts affectant les images de documents anciens numérisés. Parmi les dégradations les plus courantes, nous trouvons :

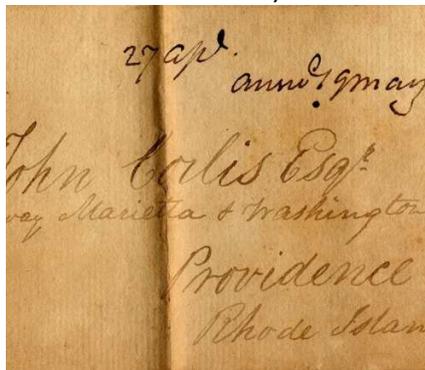
- **L'effet de transparence du verso 'Bleed-through'** : certains papiers, trop fins, laissent apparaître l'encre présente au verso du document à travers le recto. Selon la quantité d'encre qui est diffusée, la transparence peut être plus ou moins visible. Un exemple est présent sur la figure 1.3.a.
- **Les taches** : Elles sont souvent dues à l'acidité de l'encre, à l'humidité du lieu de stockage, mais peuvent avoir d'autres causes. La figure 1.3.b montre un exemple de tache pouvant avoir plusieurs origines.
- **Les pliures et ondulations** : peuvent avoir plusieurs origines. Voir les figures 1.3.c et 1.3.e.
- **Les défauts touchant les caractères** : L'encrage faible ou de mauvaise qualité peut dégrader le contraste entre l'écriture et le fond, et créer la rupture des traits, et l'effacement de parties de caractères. Dans d'autres cas, l'encre peut se diffuser de façon trop importante dans le papier, et créer des caractères épais avec des formes non uniformes, des trous bouchés et la fusion de plusieurs caractères. Un exemple de dégradation touchant les caractères est montré sur la figure 1.3.d.
- **Les déchirures et trous** : sont des dégradations plus rares qui peuvent avoir des origines diverses (usure, mauvaise utilisation, insectes, etc.). Voir la figure 1.3.e.
- **Les annotations et les ajouts indésirables** : tels que les tampons apposés par les bibliothécaires, les explications, et les corrections ajoutées par les chercheurs et les conservateurs (Figure 1.3.e).



(a) Effet de transparence 'Bleed-through'
(image de la base de données de
DIBCO'11)



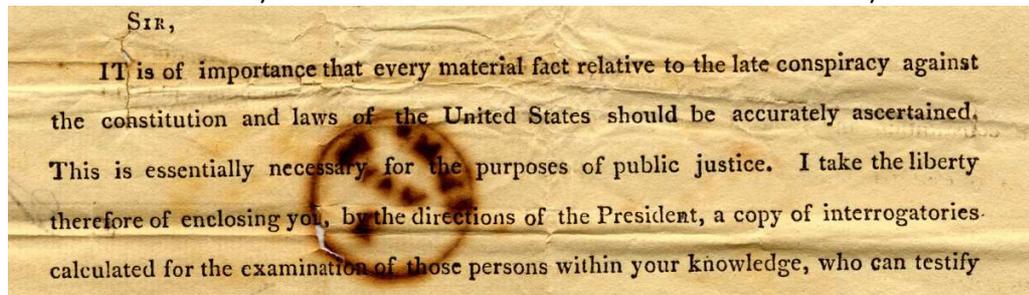
(b) Taches d'humidité sur le document (image
de la base de données de DIBCO'11)



(c) Document plié
(image de la base de données de
DIBCO'11)



(d) Caractères faiblement contrastés 'faint
characters' (image de la base de données de
DIBCO'13)



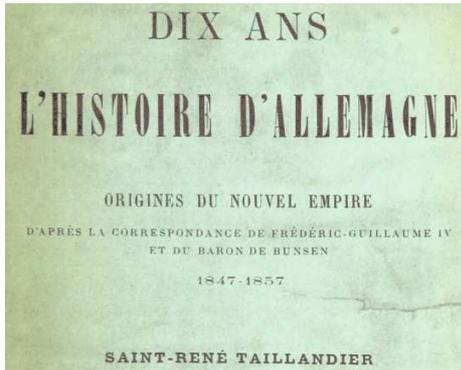
(e) Pliures, déchirures, trous, tampon et acidité de l'encre du tampon
(image de la base de données de DIBCO'13)

Figure 1.3. Exemples de dégradations intrinsèques au document.

1.3.2. Dégradations dues à la numérisation

La numérisation, même de qualité, n'est qu'une représentation imparfaite, à un moment donné, d'un objet physique [Drira, 2007], et, de ce fait, elle peut introduire du bruit aux images reçues du capteur. Les défauts causés par la numérisation peuvent avoir plusieurs formes, citons [Rabeux, 2013]:

- **L'illumination non uniforme 'Uneven illumination'**: due à une trajectoire de lumière non uniforme entre la caméra et le document à acquérir, ce qui entraîne un contraste non uniforme le long de la surface de l'image du document (voir les figures 1.4.a et 1.4.b), et qui rend difficile la différenciation du premier plan de l'arrière-plan.
- **Les défauts géométriques**: tels que l'inclinaison résultante d'un mauvais positionnement du document sur la vitre du scanneur (Figure 1.4.c), et la courbure du texte sur les bords, due à l'épaisseur du livre (Figure 1.4.d).
- **L'effet du flou**: peut être le résultat de plusieurs mauvaises opérations ou réglages. Par exemple, c'est l'opérateur de numérisation qui réalise le réglage du focus sur les scanners manuels. Or, certains livres anciens sont très volumineux, et le scanner doit être réajusté au bout d'un certain nombre de pages numérisées. Le défaut de flou est présent sur toute la page. Dans d'autres cas, le flou est local: reliure trop profonde, mouvements de l'ouvrage pendant la numérisation (Figure 1.4.e), . . .
- **Apparition des éléments inutiles**: comme les doigts tenant le document, le marque page, etc. et qui couvrent une partie du document, entraînant la perte d'informations importantes. Voir la figure 1.4.f.
- **Le bruit marginal**: formé de l'ensemble d'ombres qui apparaissent dans les marges verticales ou horizontales d'une image [Farahmand et al., 2013]. Ce type de bruit est le résultat de la numérisation de documents épais ou des bords de pages dans les livres; il peut être textuel ou non textuel. La figure 1.4.g montre ce genre de bruit.
- **La mauvaise colorimétrie d'un document**: provient d'un mauvais réglage du capteur.
- **Des défauts chromatiques**: qui se créent souvent autour des pixels d'encre, et sont le résultat d'une mauvaise synchronisation des capteurs [Rabeux, 2013].
- **Le bruit des images numériques**: causé par le capteur du scanner. Le bruit le plus courant est le bruit impulsif, connu sous le nom de « Bruit sel et poivre ». Il apparaît sous forme de pixels noirs et blancs, répartis aléatoirement dans l'image numérique.
- **Le bruit de fouillis (clutter noise)**: se réfère au contenu non désiré du premier plan, dont la taille est plus grande que le texte dans les images binaires [Farahmand et al., 2013]. Il est causé par les perforations, l'inclinaison du document, ou la connexion d'énormes quantités de bruit poivre (Figure 1.4.c).



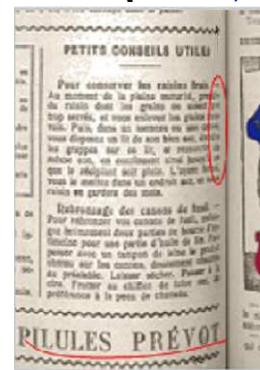
(a) Illumination non uniforme 'Uneven illumination' (image de la base de données de DIBCO'13)



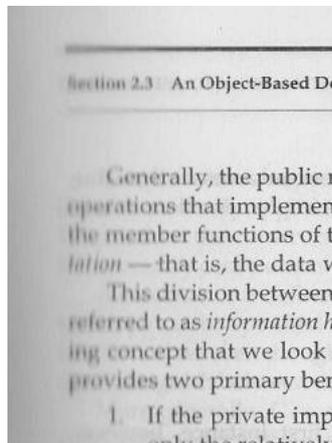
(b) Illumination non uniforme 'Uneven illumination' [Lu et al., 2010]



(c) Inclinaison du document et bruit de fouillis [Agrawal et Doermann, 2009]



(d) Courbure et écrasement des caractères au niveau de la reliure [Drira, 2007]



(e) Effet du flou [Kefali, 2015]



(f) Apparition des doigts tenant le document [Rabeux, 2013]



(g) Bruit marginal [Kefali, 2015]

Figure 1.4. Exemples de dégradations engendrées par la numérisation.

1.4. Sources de dégradations

La dégradation dans les documents historiques est acquise avec le temps, et peut rendre ces documents difficiles, voire impossibles à lire. Les sources de dégradations sont trop nombreuses à compter. Certaines d'entre elles sont causées par un phénomène physique (chimique, biologique ou humain), et certaines sont causées par des outils d'acquisition inadéquats [Hedjam, 2013, Drira, 2007].

1.4.1. Sources de dégradations physiques

1.4.1.1. Dégradations chimiques

Ce type de dégradation du papier est en grande partie causé par les processus de fabrication. Une feuille de papier est fabriquée à partir de fibres de cellulose, d'origine végétale ou animale.

Les liaisons glucosidiques de la cellulose sont stables en milieu neutre, et faiblement alcalins. Par contre, elles sont rapidement hydrolysées, en présence d'un acide fort ou d'une base forte, ce qui se traduit par la diminution plus ou moins importante du degré de polymérisation et, par conséquent, de la masse moléculaire. L'hydrolyse de la cellulose est également favorisée par la présence de groupements oxydés (aldéhydes, carboxyles). D'autre part, la lignine et la colophane peuvent participer à la formation de peroxydes, même à température ambiante : il s'agit d'agents d'oxydation très puissants qui réagissent directement sur les différents groupements chimiques de la cellulose. Le dioxyde de soufre, les oxydes d'azote et l'ozone favorisent aussi bien l'hydrolyse que l'oxydation de la cellulose [Hedjam, 2013]. Des exemples de dégradations chimiques de manuscrits sont montrés à la figure 1.5.



Figure 1.5. Exemples de dégradations chimiques.

1.4.1.2. Dégradations biologiques

Des agents, comme les moisissures, les insectes et les rongeurs, contribuent principalement à la dégradation des documents et des manuscrits, car ils se nourrissent de tous substrats organiques qu'ils trouvent. Leur prolifération est favorisée par un manque de ventilation et de lumière, ainsi que par des températures élevées et des niveaux élevés d'humidité. Ces agents peuvent détruire le papier, la reliure et l'encre. Des exemples de dégradations biologiques sur des manuscrits sont montrés à la figure 1.6.



Figure 1.6. Exemples de dégradations biologiques.

1.4.1.3. Dégradations physiques d'origine humaine

Un stockage inadéquat et une manipulation négligente des documents peuvent causer des déchirures, des plis dans le papier, etc., alors que certaines interventions humaines, telles que la restauration chimique, l'ajout de notes pour l'archivage, etc., peuvent cacher des parties de l'écriture. La consultation fréquente des lecteurs accélère également la destruction du papier. Des exemples de dégradations d'origine humaine sur des manuscrits sont montrés à la figure 1.7.



Figure 1.7. Exemples de dégradations d'origine humaine.

1.4.2. Sources de dégradations externes

Différents défauts peuvent être générés lors du processus de numérisation d'un document. La qualité des images numérisées dépend des conditions de la prise d'image (placement, éclairage, réglages...), et de la qualité du numériseur (capteur linéaire ou matriciel, bruit du capteur, précision de l'optique et de la mécanique, codage interne des couleurs...). Les choix techniques effectués (choix des paramètres du matériel, des prétraitements et des formats de fichiers) par les acteurs de la numérisation ont plus d'incidence sur la qualité des images que les problèmes précédemment cités.

Drira [Drira, 2007] propose une autre classification, orientée traitement d'image, qui regroupe les dégradations décrites précédemment, en fonction des méthodes de restauration possibles. La classification proposée regroupe toutes les dégradations, et les subdivise en *dégradations uniformes* et *dégradations non-uniformes*.

1.5. Etude faite sur les manuscrits du centre d'Adrar

Le centre national des manuscrits d'Adrar nous a fourni les éléments, présentés au tableau 1.1 et à la figure 1.8, concernant la nature des dégradations les plus apparentes, sur 43 des manuscrits en leur possession.

Tableau 1.1. Les dégradations de 43 manuscrits du centre d'Adrar.

Types d'altérations	Nombre d'altérations
Salissures générales	34
Humidité	31
Auréoles	9
Micro-organismes	2
Particules solides	3
Gras	1
Acidité	4
Décoloration	2
Friabilité	23
Oxydation	3
Jaunissement	2
Insectes	13
Rongeurs	17
Autres...	8
Cassures	12
Déchirures	38
Lacunes	30
Pliures	38
Coupures	3

Nombre d'altérations par type

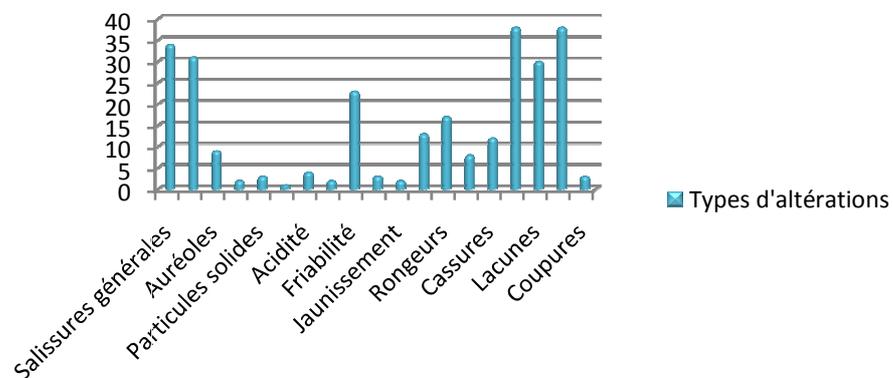


Figure 1.8. Etudes faites sur les manuscrits du centre d'Adrar.

Cette étude, menée par le centre national des manuscrits d'Adrar sur 34 images issues de ses documents anciens, a montré qu'il est plutôt rare qu'un problème existe seul dans un document. Très souvent, dans le même document coexistent deux ou plusieurs défauts à la fois, ce qui nécessite des techniques de prétraitement plus élaborées.

1.6. Prétraitement des images de documents anciens

Comme nous l'avons déjà mentionné, les images de documents numérisées contiennent divers types de bruit et de dégradations, rendant difficile l'exploitation directe de ces images. C'est pourquoi la plupart des systèmes de traitement et d'analyse d'images de documents disposent d'un module de prétraitement.

La figure 1.9 illustre la position du prétraitement dans le processus d'analyse et de reconnaissance de documents.

Le prétraitement consiste à améliorer la qualité des images, en éliminant les défauts liés à l'image numérisée, afin de faciliter l'étape de reconnaissance. Ces défauts peuvent être de deux types : ceux qui sont liés à la chaîne de numérisation (inclinaison, luminosité, bruit, ...), ou ceux qui sont liés à la qualité intrinsèque du document (les tâches d'humidité, l'apparition du verso, des trous, ...).

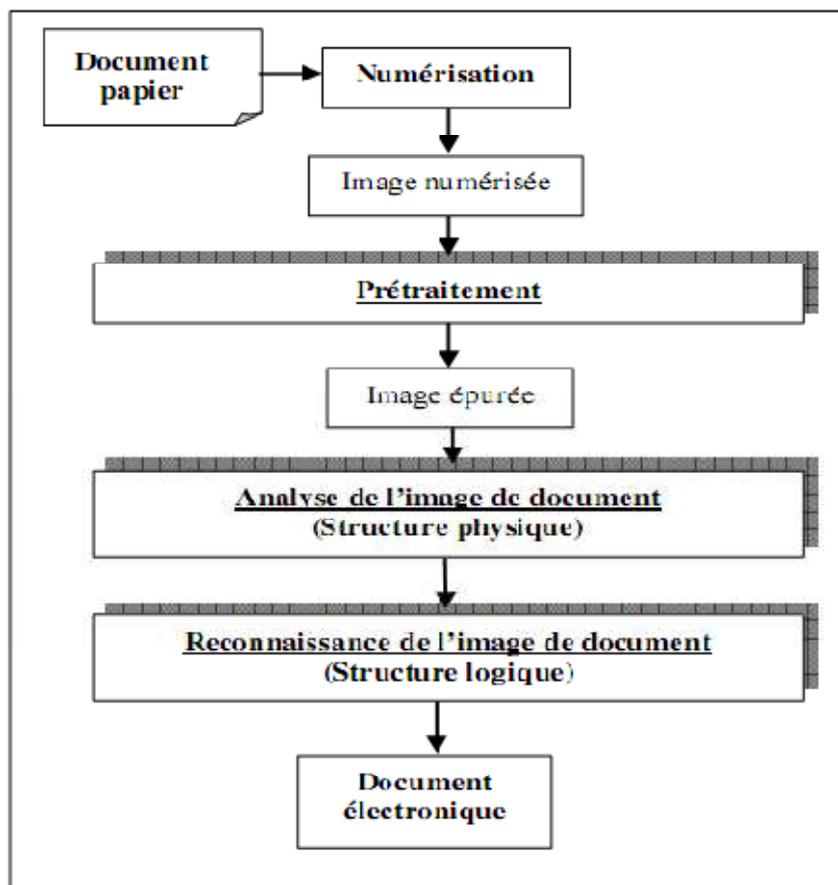
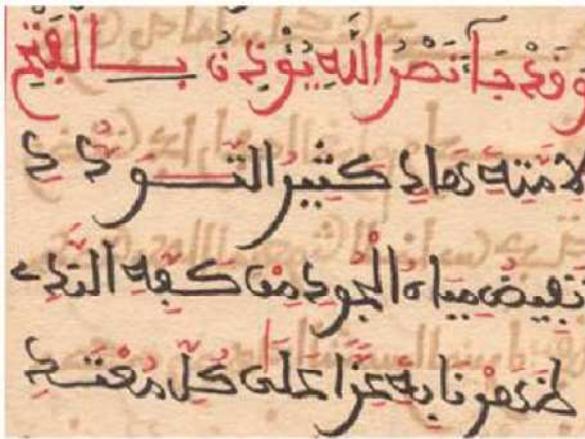


Figure 1.9. Position du prétraitement dans le processus d'analyse et de reconnaissance de documents selon [Ketata et Khemakhem, 2010].

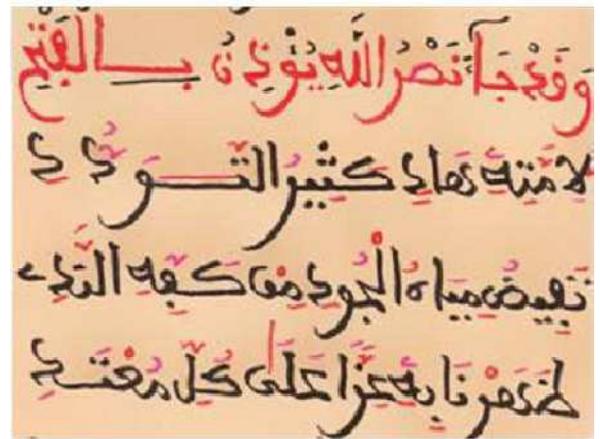
Les opérations du prétraitement d'images les plus couramment utilisées sont :

- la suppression du bruit et des points parasites,
- la séparation avant-plan/ arrière-plan (binarisation),
- la séparation recto/verso,
- le rehaussement du contraste.
- la correction de l'inclinaison et de la courbure,

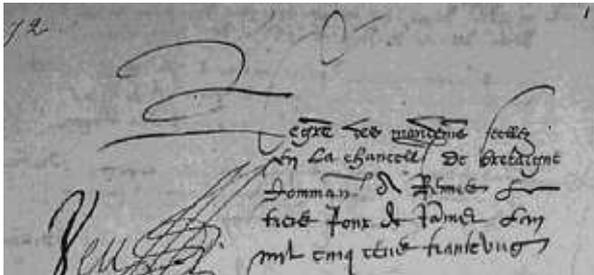
La figure 1.10 présente les résultats des prétraitements les plus courants de quelques images de documents dégradés.



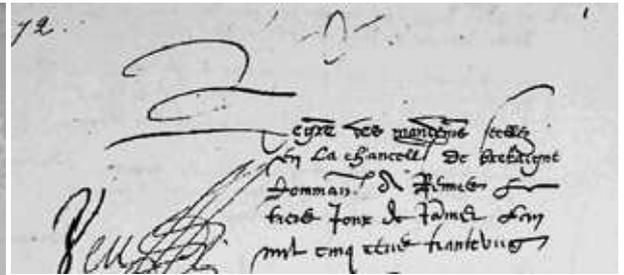
(a) Image dégradée



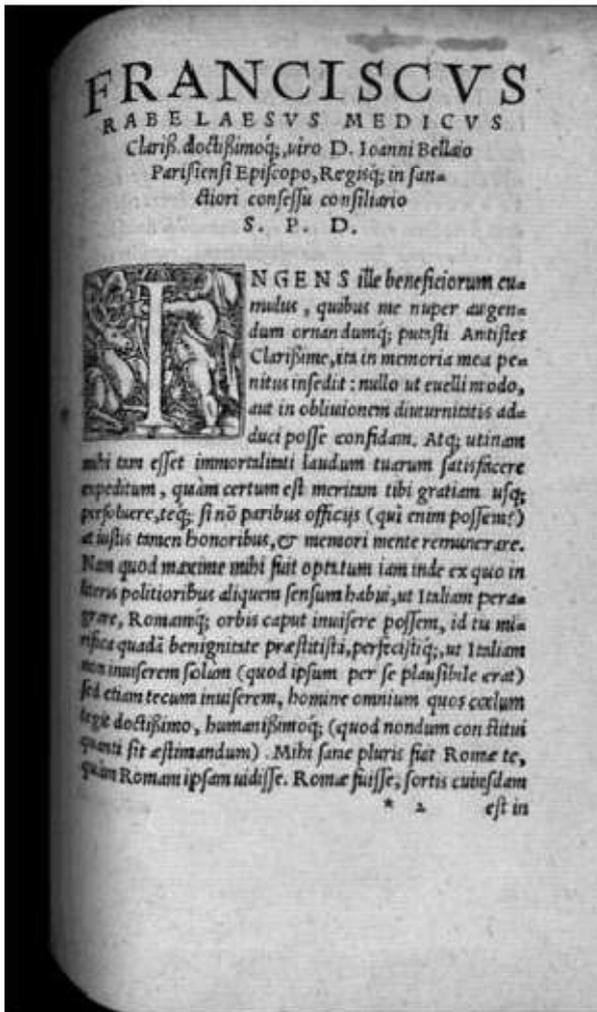
(b) Image restaurée par une séparation recto/verso [Kefali, 2015]



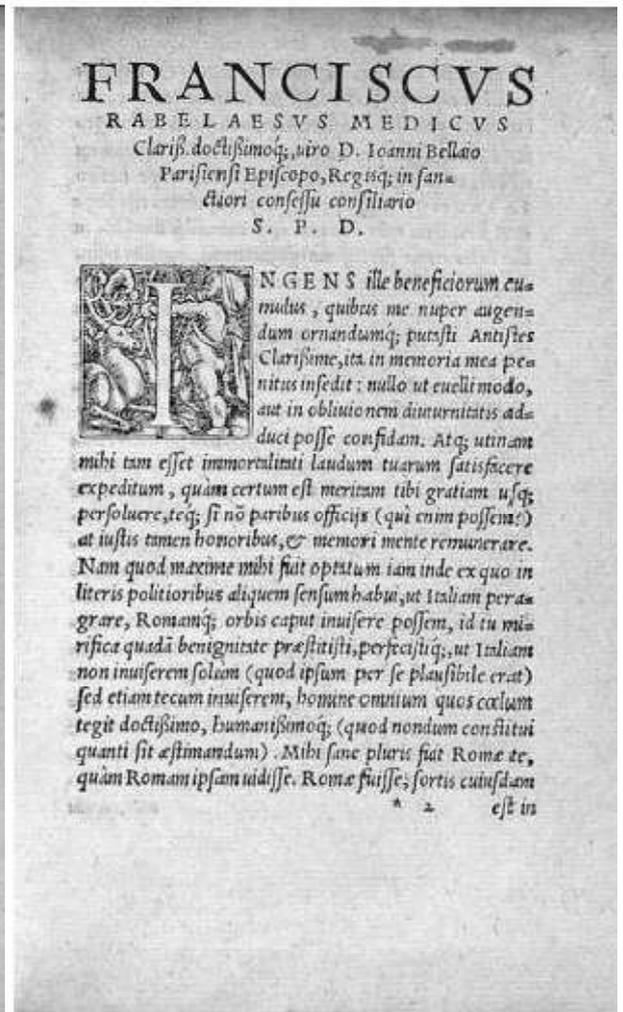
(c) Image dégradée



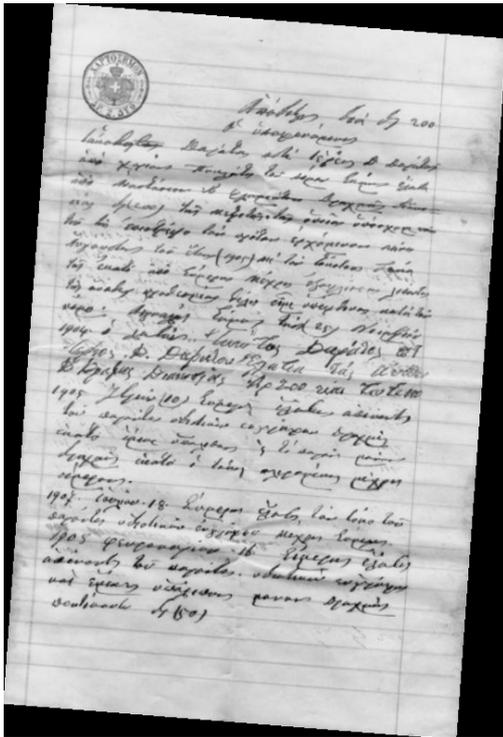
(d) Image restaurée par un rehaussement du contraste [Likforman-Sulem, 2003]



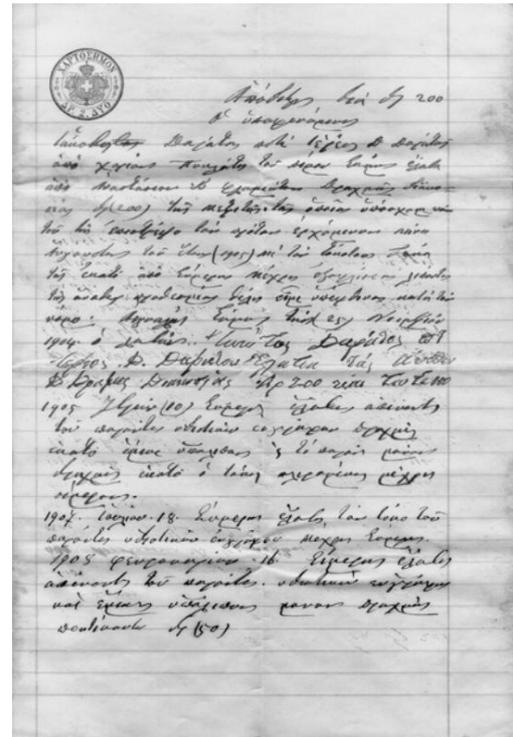
(e) Image dégradée



(f) Image restaurée par une correction de la courbure et un rehaussement de contraste [Rabeux, 2013]



(g) Image dégradée



(h) Image restaurée par une correction de l'inclinaison [Kefali, 2015]

Figure 1.10. Les opérations de prétraitement les plus courantes.

Likforman-Sulem [Likforman-Sulem, 2003] indique les prétraitements les plus couramment utilisés pour améliorer la qualité des images de documents anciens, ainsi que quelques références, présentant des méthodes de prétraitements pour chaque défaut.

1.7. Conclusion

Nous venons de voir que les dégradations peuvent provenir de causes multiples. Nous avons montré qu'elles affectent la qualité visuelle des images, et peuvent engendrer une perte significative de l'information. Le chapitre a montré que les images de documents présentent de nombreux défauts physiques ou numériques que l'on doit corriger par des opérations de traitements pour améliorer la qualité de ces images.

Nous nous intéressons dans le présent travail aux prétraitements d'images de documents anciens, afin de les rendre lisibles et accessibles, et plus précisément à la binarisation, considérée comme l'une des opérations de prétraitement les plus importantes. Dans le chapitre suivant, nous dressons un bilan des méthodes de binarisation des images de documents existantes dans la littérature.

CHAPITRE 2

Binarisation d'images de documents : État de l'art

2.1. Introduction

L'une des premières étapes d'un système traitant un document graphique, voire la première, est la binarisation, qui consiste à segmenter le document en deux classes, le fond, d'une part, et l'objet, c'est-à-dire les autres composantes (texte, graphique...), d'autre part. Elle permet de réduire de manière appréciable la quantité d'informations à manipuler. La qualité du résultat obtenu est cruciale, puisqu'elle conditionne fortement celle des étapes postérieures de la chaîne de traitements appliqués au document.

La binarisation est devenue plus difficile dans le cas d'images de documents anciens, caractérisés par différents types de dégradations, provenant du processus de numérisation lui-même, des effets du vieillissement, de l'humidité, etc. L'importance de l'étape de binarisation justifie le grand nombre de méthodes proposées, dont le but est d'arriver à une séparation parfaite entre le fond et les autres composantes (texte, graphique...).

Dans le présent chapitre, nous nous focalisons sur la binarisation d'images de documents en niveaux de gris, car dans la plupart du temps, les images de documents couleurs peuvent être transformées en niveaux de gris, sans perte significative d'information, sauf quelques exceptions, comme les documents publicitaires et certains styles de magazines. Nous présentons ainsi un état de l'art étendu des techniques de binarisation, les plus connues et les plus fréquemment citées dans la littérature, classées selon leurs catégories.

2.2. Définition de la binarisation

En traitement d'image, la binarisation est l'opération qui permet de transformer une image en niveaux de gris ou en couleurs, en une image à deux classes de pixels, en général, représentées par des pixels noirs et des pixels blancs. La différence entre la binarisation et la segmentation est que la binarisation produit toujours deux classes, alors que la segmentation peut en produire plusieurs. Cependant, ces deux termes sont souvent confondus, par abus de langage.

Dans la section suivante, nous allons présenter les techniques de binarisation les plus connues, en les organisant selon la catégorie qui les régit.

2.3. Catégorisations des méthodes de binarisation

Un grand nombre de techniques ont été proposées dans la littérature pour la binarisation d'images, et ce nombre augmente continuellement chaque année, la majeure partie de ces techniques ont été destinées à la binarisation d'images en niveaux de gris. Plusieurs catégorisations ont été décrites :

- ✓ La catégorisation la plus connue est celle qui considère uniquement les méthodes de binarisation par seuillage, et qui sépare ces méthodes en deux classes principales : méthodes par seuillage global et méthodes par seuillage local [Trier et Taxt, 1995] [Khurshid et al., 2009].
- ✓ Sezgin et Sankur [Sezgin et Sankur, 2004] ont établi une classification des méthodes de binarisation, selon l'information qu'elles exploitent, en six catégories : les méthodes basées sur la forme de l'histogramme, celles basées sur le regroupement (clustering), celles basées sur l'entropie, celles basées sur les attributs d'objets, les méthodes de binarisation spatiales, et enfin les méthodes locales.
- ✓ D'après [Stathis et al., 2008], cinq groupes de méthodes globales peuvent être distingués, selon la manière de trouver le seuil optimal de binarisation ; les méthodes basées sur des procédures de classification, sur l'histogramme, sur le regroupement (*clustering*), sur l'entropie, et sur la distribution des Gaussiennes.
- ✓ Cinq groupes de méthodes locales existent aussi selon [Stathis et al., 2008], des méthodes basées sur des procédures de regroupement, sur la variation locale, sur l'entropie, sur les informations du voisinage, et sur la méthode d'Otsu.
- ✓ Une troisième classe de *méthodes hybrides* peut être ajoutée, et les méthodes de cette classe combinent des informations globales et locales pour la binarisation [Stathis et al., 2008, Singh et al., 2012].
- ✓ Une autre catégorisation a été décrite dans [Fernando et al., 2010]. Dans cette dernière, les méthodologies de binarisation d'images sont divisées en deux catégories principales : méthodes basées sur le seuillage, et méthodes basées sur le regroupement.
- ✓ Kumar et al. [Kumar et al., 2013], ont considéré que les techniques de binarisation d'images peuvent être réparties en *méthodes supervisées* et *méthodes non-supervisées*.
- ✓ Hedjam [Hedjam, 2013], Moghaddam et Cheriet [Moghaddam et Cheriet, 2010], et Rivest-Hénault et al. [Rivest-Hénault et al., 2012] ont considéré que les techniques de binarisation d'images peuvent être réparties, en méthodes basées sur le seuillage et méthodes basées sur la classification.
- ✓ ...

Dans le reste de ce chapitre, nous dressons un bilan des méthodes de binarisation existantes dans la littérature. Nous adoptons la catégorisation décrite dans [Hedjam, 2013], [Moghaddam et Cheriet, 2010] et [Rivest-Hénault et al., 2012] en **méthodes basées sur le seuillage** et **méthodes basées sur la classification**, et la classification des méthodes basées sur le seuillage en **méthodes globales** et **méthodes adaptatives (locales)**. Une troisième catégorie, décrite dans [Fernando et al., 2010], pour **méthodes basées sur la segmentation** sera aussi considérée. Nous suivons aussi la répartition décrite par Lu et al. [Lu et al., 2010] pour la classification des méthodes adaptatives, en **méthodes adaptatives basées sur la**

fenêtre 'window-based' et méthodes adaptatives utilisant des connaissances spécifiques au document, tels que l'estimation ou la soustraction du fond 'Background subtraction', et les contours de l'image 'image edges'.

Le schéma de la figure 2.1 illustre les approches présentées dans ce chapitre.

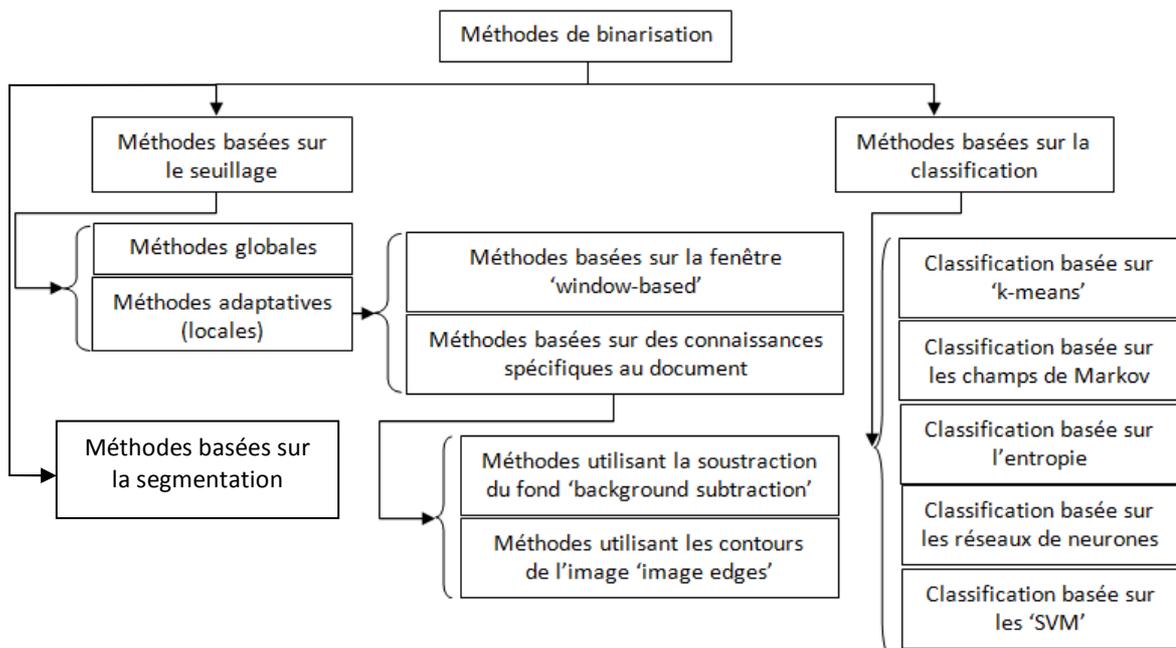


Figure 2.1. Schéma représentant quelques approches de binarisation.

2.3.1. Binarisation basée sur le seuillage

Le seuillage est la technique la plus répandue en binarisation, où la transformation s'effectue suite à un (des) seuil (s) à définir.

2.3.1.1. Méthodes de seuillage global

Les méthodes de seuillage global calculent pour l'image entière un seuil unique T , divisant l'image en deux classes *Objet* et *Fond*, en utilisant des informations globales sur l'image. Le seuillage s'effectue comme suit :

$$I_b(i, j) = \begin{cases} 255, & I(i, j) > T \\ 0, & I(i, j) \leq T \end{cases}$$

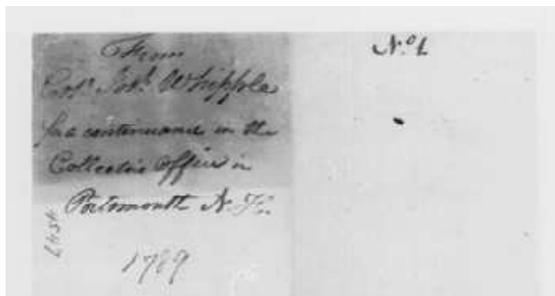
- ✓ L'algorithme d'Otsu [Otsu, 1979] est connu comme l'un des meilleurs dans cette catégorie. Il vise à trouver un seuil optimal pour l'ensemble de l'image, en maximisant la variance interclasses. L'algorithme est rapide, et donne de bons résultats lorsque l'image à binariser est bimodale (l'histogramme de l'image comporte deux pics) ; cependant, il

n'est pas bien adapté au cas d'une illumination non uniforme ou d'une présence de bruit aléatoire (Figure 2.2). Malgré les défauts de cette méthode, elle reste une référence pour comparer différentes méthodes de seuillage et de binarisation en général.

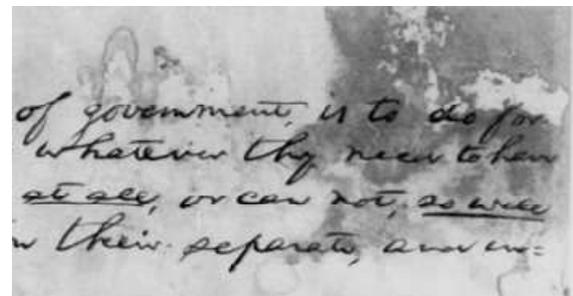
- ✓ [Abutaleb, 1989] Cette méthode exploite l'entropie de distribution des niveaux de gris pour l'estimation du seuil global de binarisation.

Il existe également d'autres méthodes de seuillage global :

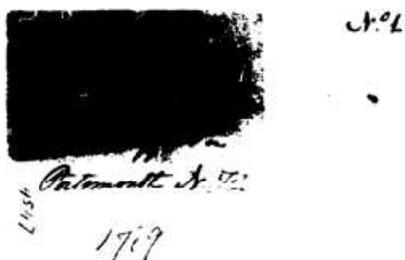
- ✓ [Kapur et al., 1985]
- ✓ [Kittler et Illingworth, 1986]
- ✓ [Brink, 1992]



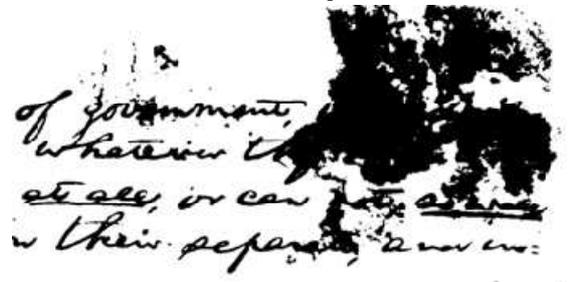
(a) Image de documents dégradée, extraites de la collection DIBCO 2009 [Gatos et al., 2009]



(b) Image de documents dégradée, extraites de la collection DIBCO 2009 [Gatos et al., 2009]



(c) Image binarisée de (a)



(d) Image binarisée de (b)

Figure 2.2. Résultats de la binarisation par la méthode d'Otsu [Otsu, 1979].

En effet, les méthodes globales ont été largement utilisées dans plusieurs applications, à cause de leur simplicité et rapidité. Cependant, ces méthodes sont performantes, uniquement lorsque les images de documents originales sont de bonne qualité, ayant un histogramme bimodal, séparant l'avant-plan de l'arrière-plan. Lorsque les documents sont de mauvaise qualité, ou contenant différents types de bruits et de dégradations (taches d'encre et d'humidité, effet de transparence, fond texturé, illumination non uniforme, etc.), ou bien lorsque les niveaux de gris des pixels de l'avant-plan et de l'arrière-plan se chevauchent, il n'est pas possible de trouver un seuil unique qui sépare complètement l'avant-plan de l'arrière-plan de l'image. Dans ce cas, les méthodes locales ou adaptatives, qui estiment un

seuil local pour chaque pixel de l'image, sont généralement capables de produire de meilleurs résultats de binarisation.

2.3.1.2. Méthodes de seuillage local 'adaptatif'

2.3.1.2.1. Seuillage adaptatif basée sur la fenêtre 'window-based'

Les méthodes locales sont dites aussi adaptatives, car elles calculent un seuil différent pour chaque pixel de l'image. Une approche de seuillage adaptatif typique est basée sur la fenêtre 'window-based' ([Niblack, 1986, Bernsen, 1986, Sauvola et Pietikainen, 2000]...); elle calcule un seuil $T(i,j)$ pour chaque pixel (i,j) de l'image, en se basant sur les informations de son voisinage. Ainsi, le seuillage s'effectue comme suit :

$$I_b(i,j) = \begin{cases} 255, & I(i,j) > T(i,j) \\ 0, & I(i,j) \leq T(i,j) \end{cases}$$

- ✓ La méthode proposée par Niblack [Niblack, 1986] peut être considérée comme la première méthode de seuillage local. Elle calcule le seuil $T(i,j)$ de chaque pixel (i,j) , en glissant une fenêtre sur l'image. Le seuil $T(i,j)$ est calculé en utilisant la moyenne $m(i,j)$ et l'écart type $s(i,j)$, de tous les pixels de la fenêtre par la formule suivante :

$$T(i,j) = m(i,j) + k * s(i,j)$$

Où k est une constante utilisée pour contrôler et ajuster l'effet de l'écart type s , elle a une valeur entre 0 et 1. La valeur de k et la taille de la fenêtre ω définissent la qualité de la binarisation. La méthode fournit de bons résultats pour des images de documents nettes et claires, cependant, elle détecte beaucoup de bruit de fond pour des images de documents relativement dégradées (Figure 2.3).

- ✓ Bernsen, dans [Bernsen, 1986], calcule le seuil $T(i,j)$ du pixel (i,j) en faisant la moyenne de la plus haute $max(i,j)$ et la plus basse $min(i,j)$, valeur de niveau de gris dans la fenêtre centrée en (i,j) . Mathématiquement, le calcul du seuil peut s'écrire ainsi :

$$T(i,j) = \frac{(max(i,j) + min(i,j))}{2}$$

Cependant, la méthode est très sensible au bruit du fond, à cause de la prise en compte du maximum et du minimum uniquement. Dans le cas où la fenêtre est située uniquement sur le fond, le bruit sera interprété comme objet, car le seuil sera bas.

- ✓ Sauvola et Pietikainen ont proposé dans [Sauvola et Pietikainen, 2000] une amélioration de la méthode de Niblack pour augmenter la robustesse de la binarisation des images de documents dégradées. La méthode a diminué le bruit de fond par rapport à celle de Niblack, mais le taux de détection du texte a également diminué, alors que le bruit de l'effet de transparence reste dans la plupart des cas (Figure 2.4). La méthode fournit de

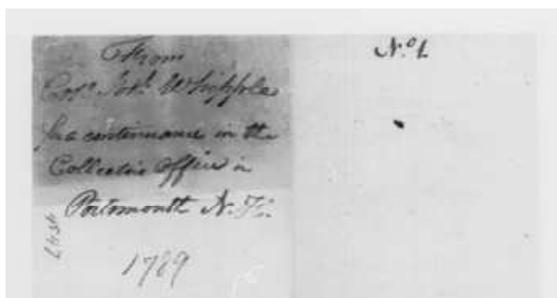
bons résultats pour des images de documents nettes et claires, cependant, elle présente quelques inconvénients comme sa sensibilité à la variation du contraste entre le texte et le fond de l'image.

Plusieurs travaux ont été proposés récemment pour améliorer la performance de la méthode de Sauvola, comme le travail de :

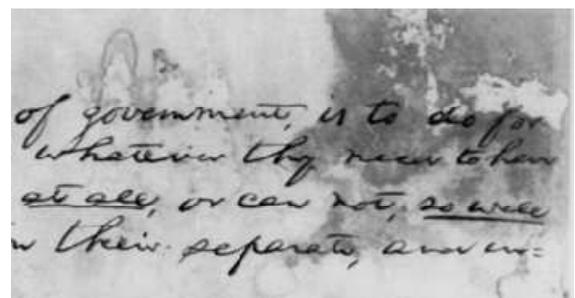
- ✓ Lazzara et Géraud dans [Lazzara et Géraud, 2014]
- ✓ Hadjadj et al. dans [Hadjadj et al., 2016] ; la description de cette méthode fait l'objet du cinquième chapitre.
- ✓ Hadjadj et Meziane dans [Hadjadj et Meziane, 2017]

Il existe également d'autres méthodes de seuillage local :

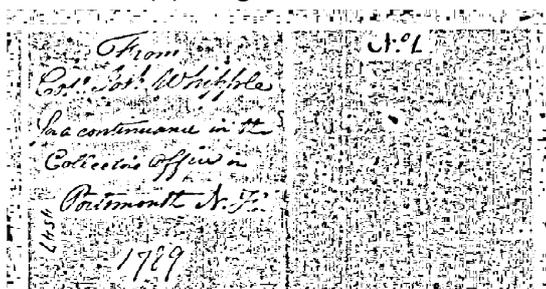
- ✓ [Eikvil et al., 1991]
- ✓ [Parker et al., 1993]
- ✓ [Yang et al., 1994]



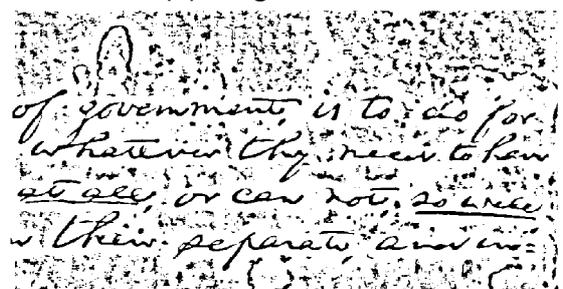
(a) Image à binariser



(b) Image à binariser



(c) Image binarisée de (a)



(d) Image binarisée de (b)

Figure 2.3. Résultats de la binarization par la méthode de Niblack [Niblack, 1986].

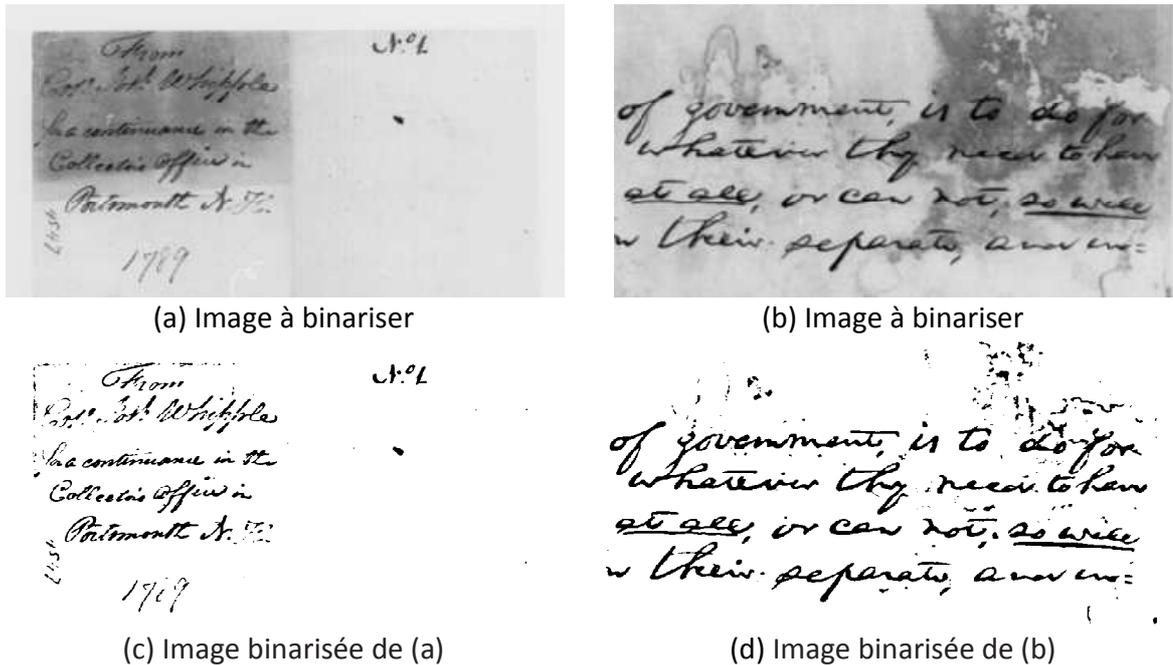


Figure 2.4. Résultats de la binarisation par la méthode de Sauvola [Sauvola et Pietikainen, 2000].

Les méthodes locales, basées sur la fenêtre sont plus performantes que les globales, et sont souvent plus robustes contre l'illumination inégale, le faible contraste et la variation de couleurs que les méthodes globales.

Cependant,

- La performance des méthodes locales basées sur la fenêtre dépend fortement de la taille de la fenêtre qui ne peut pas être déterminée correctement sans connaissance préalable de la localisation du texte.
- Ces méthodes sont plus complexes en termes de calcul, car un seuil distinct est calculé pour chaque pixel de l'image, en se basant sur les informations de ses voisinages. Ainsi, le calcul des seuils devient de plus en plus complexe, chaque fois que l'on augmente la taille du voisinage considéré.
- Elles sur-segmentent souvent le bruit et la texture de l'arrière plan de l'image.
- Le traitement devient plus difficile, pour les documents, dont la taille des caractères est très variable, la fenêtre d'analyse étant fixe durant tout le traitement (elles ont tendance à ne pas segmenter les parties intérieures des caractères épais).

Le tableau 2.1 résume les avantages et les inconvénients des méthodes globales et locales, basées sur la fenêtre.

Tableau 2.1. Comparaison des méthodes globales et locales.

Méthodes	Avantages	Faiblesses
Méthodes de seuillage global	Simple, Rapides Produisent de bons résultats sur des images de documents nettes	Produisent de mauvais résultats sur des images dégradées
Méthodes de seuillage local, basées sur la fenêtre	Produisent de meilleurs résultats sur des images dégradées	Sensibles à la taille de la fenêtre

D'après ce tableau, on peut conclure qu'aucune des méthodes classiques ne remplit toutes les conditions imposées (efficacité et rapidité).

La méthode de seuillage global d'Otsu, et les méthodes locales de Niblack et Sauvola sont largement incorporées dans d'autres méthodes de binarisation d'images de documents, comme :

- ✓ Kim et al. [[Kim et al., 2002](#)] considèrent l'image originale comme un terrain 3D, avec des vallées et des montagnes, correspondant aux régions de texte et du fond respectivement, sur lesquelles l'eau a été versée pour remplir les vallées qui représentaient les composantes textuelles. Le résultat de la binarisation finale a été produit, en appliquant la méthode d'Otsu à l'image compensée (la différence entre l'image originale et l'image remplie d'eau).
- ✓ Moghaddam et Cheriet, dans [[Moghaddam et Cheriet, 2010](#)] ont proposé une méthode de binarisation multi-échelle, dans laquelle l'image du document d'entrée a été binarisée à plusieurs reprises, en utilisant différentes échelles. Ensuite, ces images de sortie ont été combinées pour former l'image de sortie finale. Cette méthode a été étendue à la méthode d'Otsu [[Otsu, 1979](#)] avec de meilleurs résultats, est a été nommée AdOtsu [[Moghaddam et Cheriet, 2012](#)].

Il existe également d'autres méthodes :

- ✓ [[Gatos et al., 2006](#)]
- ✓ [[Lu et al., 2010](#)]
- ✓ [[Ntirogiannis et al., 2012](#)]
- ✓ [[Hadjadj et al., 2014](#)]
- ✓ [[Hadjadj et al., 2016](#)]
- ✓ ...

Certaines méthodes adaptatives, rapportées dans la littérature, utilisent des connaissances spécifiques au document, tels que l'estimation ou la soustraction du fond 'Background subtraction', et les contours de l'image 'image edges'.

2.3.1.2.2. Binarisation utilisant la soustraction du fond 'Background Subtraction'

Ces méthodes essaient de soustraire ou d'estimer le fond 'background' de l'image dégradée, et de l'utiliser pour binariser la même image de document ; mais il est difficile de modéliser le fond du document, et de le séparer du texte du premier plan.

- ✓ Gatos et al. ont utilisé dans [Gatos et al., 2006] le filtre de Wiener pour le prétraitement, et ont estimé le fond, en tenant compte du résultat de binarisation de Sauvola. Le seuil final est obtenu par la différence entre le fond estimé et l'image prétraitée, alors que le post-traitement a amélioré le résultat final. Bien que cette méthode réalise de meilleures performances OCR que Sauvola, elle hérite de certains inconvénients de Sauvola. Par exemple, les caractères faiblement contrastés ne peuvent pas être restaurés avec succès, alors que le bruit de l'effet de transparence reste.
- ✓ Moghaddam et al. dans [Moghaddam et Cheriet, 2009a] estiment la surface de l'arrière-plan du document, grâce à une procédure adaptative et itérative de calcul de la moyenne d'image.
- ✓ Dans le travail de Lu et al. [Lu et al., 2010], une procédure de lissage polynomial itératif est d'abord appliquée pour estimer efficacement le fond. Les contours du texte sont ensuite détectés, en fonction de la variation de l'image locale, à l'intérieur de l'image du document compensée, en utilisant la surface du fond de document estimé. Enfin, le seuil local est estimé, en fonction des pixels de contours du texte, détectés dans une fenêtre de voisinage locale. Bien que cette méthode donne de bons résultats (en obtenant la première place dans le concours DIBCO'09), elle a une certaine limitation, comme l'ont expliqués les auteurs. Elle utilise le contraste local pour le seuillage final et, par conséquent, le bruit de fond, dû au contraste élevé, de même que le bruit de l'effet de transparence, restent.
- ✓ Messaoud et al. [Messaoud et al., 2011] ont proposé une combinaison entre une étape de prétraitement et une étape de localisation. Dans l'étape de prétraitement, le filtrage de Wiener a été appliqué, après avoir estimé l'arrière-plan, en utilisant un filtrage médian. Dans l'étape de localisation, les contours de Canny [Canny, 1986] ont été extraits avec leurs boîtes de délimitation 'bounding boxes'. La binarisation finale a été effectuée dans les 'bounding boxes', en utilisant le seuil d'Otsu, de Sauvola ou de Lu et al. [Lu et al., 2010]. Bien que le détecteur de contours de Canny, puisse manquer certaines informations ou détecter de bruit, cette méthode fournit des résultats relativement bons, et la méthode a été classée à la 4^{ème} position à DIBCO'11. Sur la figure 2.5.b, on peut voir le bruit introduit par l'utilisation des 'bounding boxes'.

Les méthodes d'estimation de l'arrière-plan sont performantes, mais elles peuvent ne pas fonctionner correctement sur certaines images de documents, avec un arrière-plan complexe, ou avec différents objets de texte, tels que les cartes et les dessins.

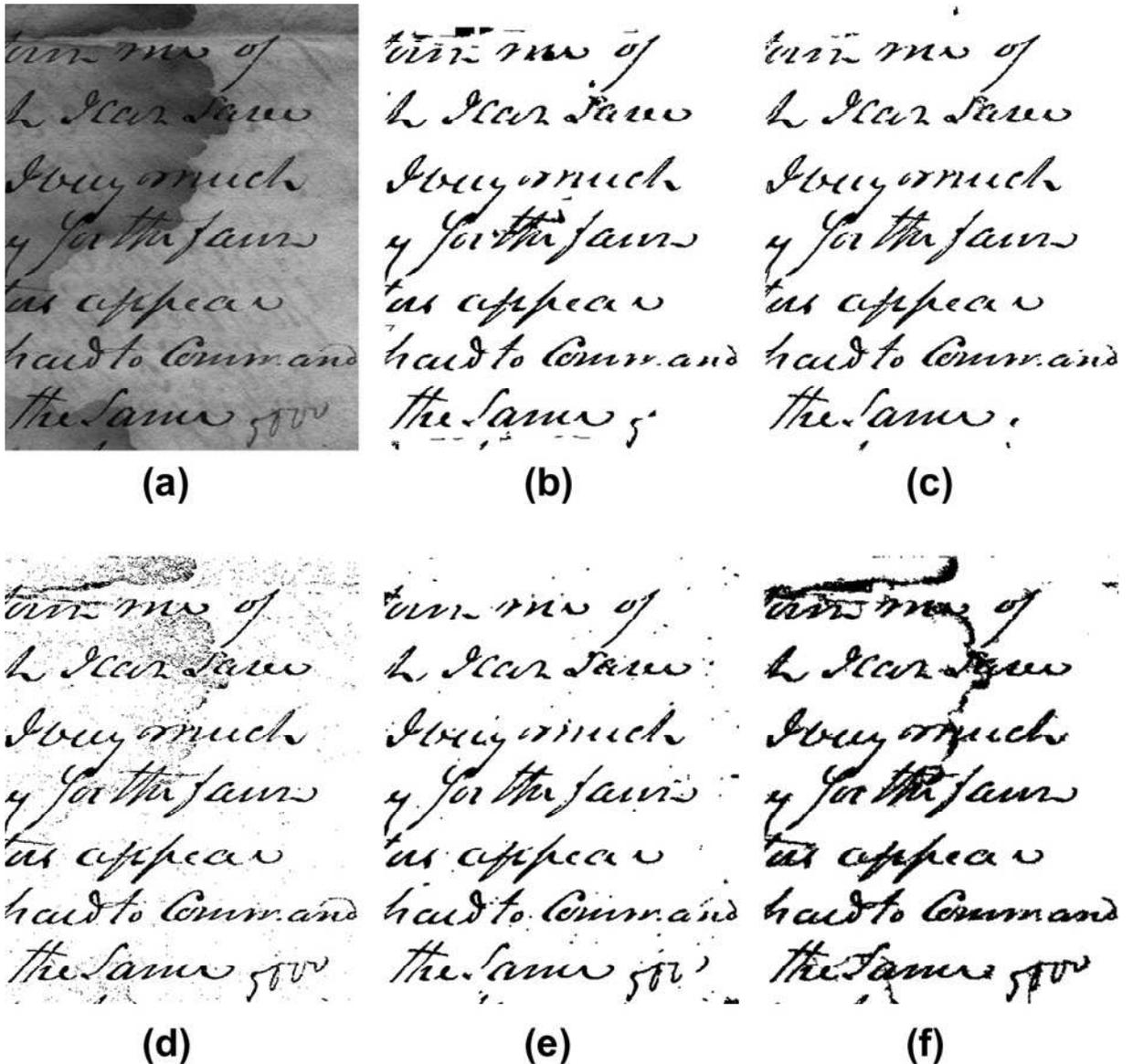


Figure 2.5. Résultats de la binarisation de l'image HW4 de la base de données de DIBCO'11 :

- (a) Image à binariser ;
- (b) Résultat de la méthode de Messaoud et al. [Messaoud et al., 2011] ;
- (c) Méthode de Su et al. [Su et al., 2010] ;
- (d) Méthode de Ntirogiannis et al. [Ntirogiannis et al., 2009] ;
- (e) Méthode de Howe [Howe, 2011] ;
- (f) Méthode de Gatos et al. [Gatos et al., 2008] .

2.3.1.2.3. Binarisation utilisant l'image de contraste et de l'information du contour ('Image contrast and edge information')

Le contraste et l'information du contour, qui sont de bons indicateurs du texte, sont utilisés par certaines méthodes de binarisation, pour supprimer le fond non uniforme, bien qu'il soit difficile d'identifier la différence entre les bords du texte et le bruit de fond du document. Les méthodes basées sur les contours, utilisent habituellement une mesure de changement, à travers un contour, tels que le gradient ou d'autres dérivées partielles, ou utilisent un détecteur de contour, tel que le détecteur de contour de Canny directement.

Parmi les méthodes de binarisation, basées sur les contours de l'image, on peut citer :

- ✓ Chen et al. proposent dans [Chen et al., 2008] de détecter et de fermer d'abord les contours de l'image, puis d'obtenir une image de document binaire primaire, en se basant sur les informations de contour déterminées.
- ✓ Moghaddam et al. utilisent dans [Moghaddam et Cheriet, 2009] le profil de contour, pour localiser la région du texte et, en conséquence, estimer le seuil local de l'image.
- ✓ Ntirogiannis et al. dans [Ntirogiannis et al., 2009], ont modifié la technique de niveau logique de Yang et Yan [Yang et Yan, 2000], qui considère à la fois le niveau de gris et le contraste local, calculé à partir de points antidiagonaux. Les auteurs ont utilisé le résultat de la binarisation de [Gatos et al., 2006] pour définir de manière adaptative la largeur du trait local, et pour modifier le critère de seuil local, et augmenter le contraste en faveur d'une détection de caractère faible. Cette méthode a une performance similaire à [Gatos et al., 2006], et ne peut pas éliminer efficacement le bruit de l'effet de transparence 'bleed through', et le bruit dans les régions de changements de contraste locaux (Figure 2.5.d).
- ✓ Dans [Su et al., 2010], les auteurs calculent le contraste de l'image (en fonction des intensités maximale et minimale locales [van Herk, 1992]) et binarisent l'image de contraste, en utilisant la méthode d'Otsu, pour détecter les pixels du contour du texte. Ils utilisent une formule de seuillage local, semblable à celle utilisée par Lu et al. dans [Lu et al., 2010], et présentent de meilleurs résultats que Lu et al. [Lu et al., 2010]. Les auteurs ont également participé avec une version modifiée de cette méthode (les contours de Canny sont également considérés) dans les concours de binarisation H-DIBCO'10 [Pratikakis et al., 2010] et DIBCO'11 [Pratikakis et al., 2011], où ils ont gagné la 1ère et la 2ème place, respectivement. Cette méthode est capable de supprimer la majorité du bruit de fond et de l'effet de transparence, mais elle ne détecte pas les faibles caractères, de manière efficace (Figure 2.5.c, [Pratikakis, 2011]).
- ✓ Hedjam et al. [Hedjam et al., 2011] ont utilisé deux modèles gaussiens, l'un pour le fond et l'autre pour le texte, et ont réalisé un inpainting 'inpainting' [Bertalmio et al., 2000] en utilisant le résultat de binarisation de Sauvola comme masque d'inpainting. Le masque de

l'inpainting et la version inverse de celui-ci ont été déterminés, en utilisant des statistiques locales (e.g. moyenne, écart type) du fond et du texte, respectivement.

- ✓ Howe propose dans [Howe, 2011] une méthode basée sur le Laplacien de l'intensité de l'image, dans laquelle une fonction d'énergie a été minimisée par un calcul de coupe de graphe 'graph cut'. Elle incorpore les informations de contours de Canny dans la construction du 'graph', pour identifier les discontinuités probables. C'est une méthode efficace, qui utilise plusieurs paramètres. L'auteur a réglé les paramètres sur la base de données de DIBCO '09, et a obtenu la 3ème position à H-DIBCO'10, et la même position à DIBCO'11. Cependant, cette méthode ne détecte pas les parties des caractères faiblement contrastées, et présente un bruit de fond (Figure 2.5.e).

2.3.2. Binarisation basée sur la classification

Les méthodes de classification utilisent divers critères d'homogénéité pour classer les pixels de l'image en deux classes ou plus, et localiser les pixels du texte. Plusieurs méthodes de classification, comme l'algorithme 'K-means', les champs de Markov (MRF), les réseaux de neurones artificiels (ANN pour *Artificial Neural Network*), la machine à vecteurs de support (SVM pour *Support Vector Machine*), ..., sont utilisées pour la binarisation d'images de documents.

2.3.2.1. Classification basée sur 'k-means'

K-moyennes (K-means)¹ [Diday, 1971] est une méthode de classification non supervisée ; elle nécessite cependant de choisir, *à priori*, le nombre de classes de l'ensemble d'arrivée. A l'initialisation, on choisit de façon aléatoire « n » centres de gravité 'germes' (« n » étant le nombre de classes souhaitées). On affecte alors à chaque pixel l'étiquette de la classe la plus proche (en prenant par exemple la distance comme critère d'affectation), puis les centres de gravité sont recalculés. On réitère le processus jusqu'à stabilisation des centres de gravité (pas de variation entre deux itérations successives).

L'algorithme *K-means* a été utilisé pour l'extraction de l'avant-plan, dans des images de documents, mais le résultat dépend fortement des germes choisis, et du nombre de classes (choisi *à priori*). Par ailleurs, l'application de cet algorithme sur plusieurs millions de pixels nécessite un traitement considérable.

- ✓ Bottou et al. [Bottou et al., 1998] ont implémenté une séparation entre l'avant-plan et l'arrière-plan, dans le cadre d'une nouvelle méthode de compression d'images de documents, baptisée « DjVu ». L'approche de séparation avant-plan/ arrière-plan proposée, est basée sur l'algorithme *k-means*, en considérant plusieurs grilles de résolutions croissantes. Dans cette méthode, chaque grille délimite des blocs carrés de pixels, et la taille des blocs d'une grille est une fraction de la taille de blocs de la grille

¹ L'algorithme fut nommé nuées dynamiques par son auteur mais il est connu de la plupart sous son appellation anglophone.

précédente. L'algorithme *k-means* est appliqué sur les blocs de la première grille, pour trouver les deux couleurs dominantes (une de l'avant-plan et l'autre de l'arrière-plan) dans chaque bloc. Les blocs de la grille suivante sont ensuite traités par un algorithme de regroupement de couleurs légèrement modifié, et ce traitement continue jusqu'à la convergence. Finalement, plusieurs filtres ont été appliqués, en se basant sur divers critères, dans le but de nettoyer les composantes de l'avant-plan, et d'éliminer les erreurs de classification.

- ✓ Leydier et al. [Leydier et al., 2004] propose une méthode de séparation avant-plan/arrière-plan, basée sur une sérialisation de l'algorithme *k-means*. Il s'agit d'une classification automatique des couleurs sur une fenêtre glissante parcourant toute l'image. Cependant, l'approche implique une intervention manuelle, pour décider du nombre de classes et de nuées, ainsi que de leurs centres initiaux. Lors du déplacement de la fenêtre, l'algorithme réutilise des informations provenant des nuées calculées dans la fenêtre précédente, et les ajuste automatiquement pour adapter le classifieur à toute variation locale des couleurs. Afin d'améliorer les résultats, les auteurs proposent de définir plusieurs nuées dans l'espace des caractéristiques, pour chaque classe logique. Trois paramètres ont été utilisés dans le but de prévenir le croisement des centres des nuées, qui se produit fréquemment lors de l'utilisation des classifieurs non supervisés.
- ✓ [Garain et al., 2006] décrivent une méthode, basée sur l'algorithme *K-means*, pour la séparation avant-plan/ arrière-plan, appliquée sur des images en niveaux de gris et en couleurs. La méthode proposée commence par l'étiquetage des composantes connexes de l'image, pour capturer les pixels de couleurs similaires spatialement. Les composantes du fond dominant sont ensuite déterminées, en se basant sur leur taille, et l'image entière est divisée en blocs rectangulaires de différentes tailles. Ces blocs représentent l'uniformité locale d'illumination, le fond, etc., et les parties d'avant-plan respectives sont traitées contre ces uniformités locales. Puis, les blocs sont arrangés sous forme d'arbre, et connectés hiérarchiquement. Des relations père-fils sont établies entre les blocs, en se basant sur leur taille et sur leur disposition géométrique. Les blocs dans l'arbre sont par la suite traités d'une manière descendante. Les composantes connexes dans chaque bloc subissent une classification bi-couleurs par l'algorithme *k-means*, dont l'initialisation est faite, en utilisant les résultats de classification obtenus par le traitement de son parent. La combinaison des résultats du traitement de tous les blocs fournit l'image binarisée finale.

2.3.2.2. Classification basée sur les champs de Markov 'Markov random fields'

Les méthodes basées sur les champs de Markov sont des méthodes de classification non supervisée. Les champs de Markov offrent un cadre mathématique puissant, qui permette

une corrélation spatiale des étiquettes entre pixels voisins [Geman et Geman, 1984]. Il peut être modélisé par deux champs Y et E , tels que : Y est le champ des observations (niveau de gris, couleur...), représentant l'ensemble des N pixels de l'image. Nous cherchons à mettre en correspondance le champ Y avec un champ des étiquettes E , prenant ses valeurs dans un ensemble fini $\mathcal{F} = \{1 \dots L\}$, tel que L est le nombre total des classes présentes dans l'image. Le problème de la segmentation est vu comme l'estimation de la réalisation de E , connaissant la réalisation observée Y . On cherche, en tout pixel de l'image, une bonne estimation de la configuration du champ des étiquettes E , connaissant les données Y . En se plaçant dans le contexte de la maximisation d'une probabilité à posteriori, on cherche à maximiser la probabilité $P(E/Y)$.

Les champs de Markov ont également été utilisés pour le traitement et la binarisation des images de documents :

- ✓ Cao et Govindaraju, dans [Cao et Govindaraju, 2009] proposent aussi une méthode, basée sur les champs de Markov, pour la binarisation d'images de documents manuscrits dégradés.
- ✓ Dans [Wolf, 2010], une méthode, basée sur MRF cachés, est présentée, et nécessite des images de documents recto-verso, qui ne sont pas toujours disponibles. La méthode a été couronnée de succès sur des documents du 18^{ème} siècle, bien que sa performance diminue pour les grandes images de documents. En outre, le petit voisinage utilisé dans la méthode, limite l'étendue de son applicabilité.
- ✓ Hedjam et al., dans [Hedjam et al., 2010] proposent une méthode de segmentation robuste, pour l'extraction de texte, à partir des images de documents historiques. La méthode est basée sur la classification bayésienne markovienne. Elle traite les graphes locaux sur les deux échelles, pixel et régionale.

2.3.2.3. Classification basée sur 'SVM'

La technique des machines à vecteurs de support (en anglais « Support Vector Machines » ou SVM) est une méthode de classification supervisée, qui consiste à déterminer un classificateur *linéaire*, appelé *hyperplan*, pour séparer deux classes de données. Les points de données, les plus proches de celui-ci, définissent des points appelés, *vecteurs de support*. Pour obtenir une segmentation robuste, il faut maximiser la *marge*, i.e. la distance entre l'hyperplan et les vecteurs de support. Si le problème de classification est non linéairement séparable, la méthode projette les données dans un autre espace, de dimension supérieure à celle de l'espace d'entrée, de façon que les données transférées deviennent linéairement séparables, cela, en utilisant des fonctions de transformation non linéaire ϕ .

Récemment, les SVM ont été utilisés pour la binarisation d'images de documents, dans les travaux suivants :

- ✓ Kita et Wakahara ont combiné dans [Kita et Wakahara, 2010] l'algorithme *k-means* avec l'algorithme SVM, pour la binarisation de caractères multi-couleurs, dans les images de scènes contenant diverses dégradations et fonds complexes. La méthode proposée est composée de trois étapes. La première étape est la génération provisoire des images binarisées, via chaque dichotomisation des K classes obtenues de la classification, en utilisant l'algorithme *K-means* des pixels, dans l'espace de couleurs HSI (Teinte, Saturation, Intensité). Le nombre de classes K est déterminé à l'avance. La deuxième étape consiste à utiliser un SVM, pour déterminer si chaque image binarisée provisoire représente ou non un caractère, et à quel degré. Deux caractéristiques extraites de chaque image binarisée provisoire ont été utilisées, pour entraîner le SVM, qui estime le degré d' « apparence du caractère ». Dans la dernière étape, on sélectionne l'image binarisée, ayant le maximum de degré d' « apparence de caractère », comme le résultat de binarisation optimale.
- ✓ Une autre méthode, basée sur les SVMs, pour la binarisation d'images de documents, a été décrite par Chou et al., dans [Chou et al., 2010]. La méthode proposée divise l'image en plusieurs régions, et décide comment binariser chaque région séparément. Les règles de décision sont déterminées par un SVM entraîné, en utilisant des caractéristiques extraites de chaque région. Ainsi, trois caractéristiques sont utilisées : le seuil d'Otsu, la moyenne, et l'écart-type.
- ✓ Dans [Wakahara et Kita, 2011], les mêmes auteurs ont étendu la méthode précédente, pour qu'elle soit applicable sur des images de chaînes de caractères. Dans la nouvelle version de la méthode, chaque image binarisée provisoire, obtenue dans la première étape, est divisée en une séquence d'images, d'un caractère chacune.

2.3.2.4. Classification basée sur les réseaux de neurones artificiels (ANN pour *Artificial Neural Network*)

Les réseaux de neurones artificiels sont des techniques de classification supervisée très connues. Il existe dans la littérature plusieurs travaux, utilisant ces réseaux de neurones artificiels, pour la binarisation d'images de documents [Kefali, 2015] :

- ✓ N. Papamarkos et al. dans [Papamarkos et al., 2000] ont proposé une approche de multi-seuillage, utilisant l'analyse en composantes principales (ACP), et un ANN de type carte de Kohonen (SOFM pour *Self-Organized Feature Map*). La couche d'entrée du SOFM code les 255 éléments de l'histogramme de niveaux de gris, et la couche de sortie est une carte 1D, dans laquelle un seul neurone gagnant est activé, pour chaque vecteur d'entrée.
- ✓ Dans [Papamarkos et al., 2000], l'auteur a proposé une technique neuronale floue, basée sur la combinaison d'une carte de Kohonen (SOFM) et de l'algorithme *C-moyens flous* (FCM pour *fuzzy Cmeans*) pour la binarisation d'images de documents en niveaux de gris

et en couleurs. La technique proposée utilise les couleurs des pixels, en plus d'autres caractéristiques spatiales locales, extraites à partir du voisinage de chaque pixel pour entraîner une SOFM. Après l'apprentissage, deux classes de neurones de sortie sont obtenues. En utilisant le contenu de ces classes, deux fonctions d'appartenance floues sont définies et utilisées par la suite par l'algorithme FCM, afin de déterminer la valeur binaire finale de chaque pixel.

- ✓ [Hamza et al., 2005] a introduit trois approches de binarisation neuronales. Les trois approches utilisent une carte de Kohonen (SOFM), entraînée sur une partie de l'image, afin d'extraire les niveaux de gris ou les couleurs les plus représentatives, suivie d'un autre classifieur. Les entrées de la SOFM sont les pixels de la sous-image. Dans la première approche, les neurones de la SOFM sont regroupés en 2 régions, en utilisant l'algorithme *Kmeans*. Une fois les pixels regroupés, la SOFM est utilisée pour classifier les pixels de l'image entière, et chaque pixel est donné comme entrée à la SOFM, en vue d'être classifié. La deuxième approche utilise un MLP (Perceptron multi couches), conjointement à la SOFM, pour la binarisation d'images couleurs. La SOFM est d'abord entraîné sur une partie de l'image, et ses neurones sont regroupés, en utilisant l'algorithme *Kmeans* comme auparavant. Le MLP est ensuite entraîné, en se basant sur les neurones étiquetés. Les entrées du MLP correspondent aux neurones de la SOFM, alors que les sorties sont les classes des mêmes neurones. Après l'apprentissage, le MLP est appliqué sur l'image entière, en vue de classifier tous les pixels. Dans la troisième approche, les seuils de Sauvola ou de Niblack sont appliqués sur les neurones de la SOFM entraînée, et un seuil global est extrait. Ce dernier est appliqué pour binariser l'image entière.
- ✓ Dans [Khashman et Sekeroglu, 2007], les auteurs ont combiné une méthode de seuillage global, notamment le seuillage MD (pour *Mass-Difference thresholding*), et un ANN supervisé, dans le but de sélectionner la valeur optimale du seuil global. Chaque image d'apprentissage est d'abord subdivisée en plusieurs sous-images, de taille 32×32 pixels. Ensuite, l'algorithme MD est appliqué sur chacune des sous-images, produisant plusieurs valeurs de seuils locaux, et sur l'image entière, résultant d'un seul seuil global. Après, l'ANN supervisé est entraîné, en utilisant les valeurs des seuils locaux comme entrées, et la valeur du seuil global comme sortie. L'ANN entraîné est utilisé par la suite pour trouver le seuil global le plus optimal pour d'autres images.
- ✓ Saidane et Garcia [Saidane et Garcia, 2007] ont utilisé un réseau de neurones de convolution [Garcia et Delakis, 2004], pour la binarisation d'images de texte couleurs. Le réseau de neurones est entraîné sur un ensemble d'images couleurs, fournies avec leur image binaire correspondante, en utilisant l'algorithme de rétro-propagation.
- ✓ Badekas et Papamarkos [Badekas et Papamarkos, 2007] ont proposé un système pour la binarisation d'images de documents imprimés, normales et dégradés, basé sur la

combinaison des résultats de plusieurs algorithmes de binarisation, et utilisant une carte de Kohonen (SOFM). Premièrement, les valeurs les plus optimales des paramètres de techniques de binarisation sont estimées. Ensuite, la SOFM prend en entrée les résultats de binarisation obtenus par les différentes techniques, en utilisant les valeurs de paramètres estimés. Après l'apprentissage, la SOFM est utilisée pour classifier chaque pixel de l'image, en *Objet* ou en *Fond*.

- ✓ Lázaro et al. [Lázaro et al., 2010] ont proposé d'obtenir un seuil optimal pour chaque image, en utilisant un ANN à régression générale, et une description sémantique de son histogramme. L'histogramme de l'image d'entrée est d'abord lissé, afin d'éliminer les faux minima et maxima, et sa dérivée discrète est obtenue. En utilisant une version polygonale de la dérivée, de même que l'histogramme lissé, une nouvelle description de l'histogramme est calculée, à partir de laquelle une description sémantique est inférée. Les descriptions sémantiques des images d'apprentissage constituent les entrées de l'ANN, et les valeurs de seuils optimaux correspondants représentent les sorties désirées.

2.3.3. Binarisation basée sur la segmentation

2.3.3.1. Binarisation basée sur les contours actifs

Les contours actifs sont aussi utilisés pour le traitement d'images de documents, mais très peu de travaux les utilisent pour la binarisation. Dans le chapitre suivant, nous allons citer quelques méthodes, utilisant les contours actifs, pour le traitement des images de documents en général, et pour la binarisation en particulier.

2.3.4. Combinaison de méthodes de binarisation

D'autres méthodes, visant à une performance de binarisation améliorée, proposent des méthodologies, basées sur la combinaison de deux ou plusieurs méthodes de binarisation existantes, supposées complémentaires. Il existe deux catégories de combinaison connues. La première combine plusieurs méthodes de binarisation, avec un algorithme de vote, comme le travail proposé dans [Gatos et al., 2008]. La deuxième combine deux méthodes ou plus, en se basant sur certaines caractéristiques, comme le travail proposé dans [Su et al., 2011].

- ✓ Un autre cadre pour la combinaison des méthodes de binarisation a été proposé par Gatos et al. dans [Gatos et al., 2008], dans lequel les pixels du foreground, provenant de différentes méthodes de binarisation, ont été sélectionnés à l'aide d'un vote majoritaire. Les contours de Canny [Canny, 1986] ont également été incorporés, pour améliorer le résultat de la combinaison, en remplissant les pistes blanches entre les contours de Canny, qui correspondent au foreground. Cette méthode améliore les résultats des travaux antérieurs des auteurs [Gatos et al., 2006], mais cela dépend des contours de Canny, par

lesquelles, des caractères faiblement contrastés peuvent être non détectés, alors que le bruit de fond, y compris l'effet de transparence, peut être détecté (Figure 2.5.f).

- ✓ Su et al. dans [Su et al., 2011] ont proposé un cadre pour la combinaison de méthodes de binarisation, dans lequel les pixels sont classés en 3 catégories, avant-plan 'foreground', arrière-plan 'background' et incertain. Plus précisément, un pixel est considéré comme un foreground s'il s'agit d'un pixel de foreground dans toutes les sorties de binarisation combinées ; il en est de même pour les pixels background. Les pixels incertains sont classés (itérativement jusqu'à ce que tous les pixels incertains soient classés), en fonction de leur différence (en termes d'intensité et de contraste local), à partir des pixels classés en background et foreground, dans une fenêtre de voisinage. Le contraste local est calculé en fonction de l'intensité de l'image et de l'intensité maximale. Dans [Su et al., 2011], les auteurs ont présenté de bons résultats sur l'ensemble des données (imprimées et manuscrites) du DIBCO'09. Par ailleurs, le calcul de contraste proposé augmente le contraste dans les caractères faibles, mais il augmente également le contraste du bruit de fond et de l'effet de transparence.
- ✓ Ntirogiannis et al. dans [Ntirogiannis et al., 2014] ont développé une nouvelle méthode de binarisation, basée sur la combinaison. Ils ont proposés une procédure d'inpainting pour l'estimation du fond et la normalisation de l'image qui aide à combiner efficacement les résultats de binarisation globale et locale d'Otsu et de Niblack.

2.4. Conclusion

Dans ce chapitre, nous avons abordé le problème de la binarisation, considérée comme une des étapes clés dans le processus de traitement et d'analyse d'images de documents.

Nous avons mené une étude globale d'un grand nombre d'algorithmes de binarisation. Nous avons considéré la répartition des différentes techniques de binarisation, en techniques basées sur le seuillage, et techniques basées sur la classification. Dans les techniques de seuillage, la binarisation est effectuée en fonction d'un ou de plusieurs seuils, calculés à partir des informations statistiques de l'image, et nous distinguons deux groupes : seuillage global et seuillage local.

Les méthodes du seuillage global sont largement utilisées dans de nombreuses applications d'analyse d'images, grâce à leur simplicité, rapidité, et efficacité. Ces techniques tendent à calculer un seuil unique pour l'image entière, afin d'éliminer tous les pixels du fond, en préservant tous les pixels de l'avant-plan. Cependant, les méthodes globales sont performantes, uniquement lorsque les documents originaux sont de bonne qualité, bien contrastée, et ayant un histogramme bimodal clair, qui sépare le premier plan de l'arrière-plan.

Pour les images de documents historiques, qui sont généralement de mauvaise qualité, et contenant différents types de détériorations et de dégradations, avec un fond texturé et une

illumination inégale, ou lorsque les niveaux de gris des pixels de l'avant-plan et de l'arrière-plan se chevauchent, les méthodes de seuillage global deviennent inappropriées. Cela est dû principalement au fait qu'aucun unique seuil n'est capable de séparer parfaitement l'avant-plan de l'arrière-plan de l'image, puisqu'il n'y a pas de distinction suffisante entre la plage de gris des pixels du premier plan et de l'arrière-plan. Ce type de documents nécessite une analyse plus fine, et, dans ce cas, on a recours aux méthodes locales. Ces dernières sont souvent plus robustes contre l'illumination inégale, le faible contraste et la variation de couleurs, que les méthodes globales, mais elles sont plus coûteuses en termes de calcul, car un seuil distinct est calculé pour chaque pixel de l'image en se basant sur les informations de ses voisinages. Ainsi, le calcul des seuils devient de plus en plus complexe, chaque fois que l'on augmente la taille du voisinage considéré.

Les techniques basées sur la classification, citées dans ce chapitre, utilisent la méthode K-moyennes 'k-means', les champs de Markov 'Markov random fields', les 'SVM' et les réseaux de neurones artificiels.

Au cours de notre recherche bibliographique, nous avons remarqué que les contours actifs sont très peu utilisés pour le traitement des images de documents anciens, malgré que ces outils sont très utilisés et très performants pour le traitement d'autres types d'images, comme les images médicales et satellitaires. Nos recherches se sont donc tournées vers cette approche.

Dans le chapitre suivant, nous présentons quelques modèles de contours actifs, en mettant en avant, pour chaque modèle, leurs avantages et leurs limites, ainsi que les différentes contributions introduites, pour pallier certaines de ses limitations. Nous citons ensuite les travaux de la littérature, qui utilisent les modèles de contours actifs pour le traitement d'images de documents, et en particulier pour la binarisation d'images de documents. Le but de cette présentation est de trouver le modèle de contours actifs le plus adéquat pour la binarisation d'images de documents, afin de l'intégrer à l'approche que nous proposons au chapitre 4.

CHAPITRE 3

Contours déformables (actifs) : Définition et état de l'art

3.1. Introduction

Comme nous l'avons vu dans les chapitres 1 et 2, notre objectif est de binariser des images de documents. La binarisation d'images consiste à extraire une région spécifique de l'image que nous appellerons objet d'intérêt. Pour les images de documents, l'objet d'intérêt est le texte et les éléments de l'avant-plan. Nous cherchons donc à regrouper l'ensemble des portions de l'image de documents, possédant les mêmes caractéristiques. Nous nous attachons principalement à définir une fonctionnelle, permettant de caractériser au mieux l'objet d'intérêt, celui-ci étant représenté par la zone de l'image qui minimise cette fonctionnelle. Nous nous intéressons donc aux méthodes variationnelles, c'est-à-dire celles qui minimisent un critère. Une grande difficulté repose dans la définition de ce critère. Comment définir les caractéristiques qui nous intéressent, et qui vont représenter au mieux l'objet d'intérêt ? De plus, comment minimiser ce critère ?

Les contours actifs apparaissent comme un outil pertinent pour parvenir à implémenter ce problème de minimisation, puisqu'ils permettent de faire évoluer un contour vers l'objet d'intérêt. De plus, nous pouvons ajouter un a priori géométrique sur l'objet à segmenter, en imposant des contraintes de régularisation sur le contour. Pour faire évoluer le contour actif vers le minimum de la fonctionnelle, nous devons dériver la fonctionnelle, puis en déduire l'équation aux dérivées partielles (EDP) qui fera évoluer le contour vers la région qui correspond, idéalement, au minimum de la fonctionnelle. Notons que toutes les méthodes de contours actifs ne cherchent pas à minimiser une fonctionnelle, certaines introduisent directement l'EDP pour faire évoluer le critère. Nous préciserons cela dans la section 3.3.

Ce chapitre d'état de l'art se décompose de la façon suivante : dans la section 3.2, nous présentons le principe des contours actifs. La section 3.3 décrit quelques modèles de contours actifs. Nous ressortons pour chaque modèle les avantages et les limites, ainsi que les différentes contributions introduites, pour pallier certaines de ses limitations. Nous citons ensuite les travaux de la littérature, qui utilisent ces modèles de contours actifs pour le traitement d'images de documents, et en particulier pour la binarisation d'images de documents. Le but de cette présentation est de trouver le modèle de contours actifs le plus adéquat pour la binarisation d'images de documents, afin de l'intégrer à l'approche que nous proposons au chapitre 4.

3.2. Définition des modèles des contours déformables (actifs)

Les modèles déformables, contours actifs ou snakes ont été initialement décrits par Kass, Witkins et Terzopoulos en 1987 dans leur article: "Snakes : Active Contour Models" [Kass et al., 1988] pour l'extraction d'objets dans une image.

Un contour actif, ou modèle déformable, est une courbe (une surface pour une image 3D), qui est positionnée initialement par l'utilisateur, et qui se déplace successivement, sous certaines contraintes issues de la courbe elle-même et des données de l'image, jusqu'à ce qu'elle coïncide avec la frontière de l'objet à détecter.

3.3. Modèles de contours déformables

Plusieurs modèles de contours déformables existent dans la littérature, qui peuvent être classés en deux catégories principales [Meziou, 2013] :

- Modèles basés sur des critères «contour» :
 - Le modèle déformable paramétrique (snake) [Kass et al., 1988].
 - La méthode de Level set [Osher et Sethian, 1988].
 - Le modèle déformable géométrique [Caselles et al., 1993]
 - Le modèle déformable géodésique [Caselles et al., 1997].
 - ...
- Modèles basés sur des critères «région» :
 - Le modèle de Mumford-Shah [Mumford et Shah, 1989].
 - ...

L'utilisation de chaque modèle dépend de l'application étudiée, et du type d'images traitées. Dans les méthodes basées «contour», les énergies définies sont liées directement à l'information extraite des images de type gradient par exemple, mettant en avant les contours alors que pour les méthodes basées «région», l'énergie externe est associée aux régions mêmes, c'est-à-dire à l'ensemble des pixels les constituant et pas seulement aux informations frontières.

La suite du chapitre est consacrée à la présentation des méthodes existantes dans la littérature, notamment celles qui sont les plus utilisées, dans le cadre de la segmentation par contours actifs, que ce soit par des approches «contour» ou des approches «région». On commence par le modèle initial des « Snakes », on fera ressortir ses avantages et ses limites, ainsi que les différentes contributions introduites, pour pallier certaines de ses limitations.

3.3.1. Contour actif classique, paramétrique ou snake

Les contours actifs classiques, paramétriques ou plus communément appelés snakes, sont les premiers modèles de contours actifs. Ils ont été introduits par Michael Kass, Andrew Witkins

et Demetri Terzopoulos en 1987 [Kass et al., 1988]. L'origine du mot anglais « Snakes » provient du fait que les contours déformables ressemblent à des serpents (snakes en anglais) lorsqu'ils se déplacent. Leur principe de base consiste à positionner, à proximité de l'objet à détecter, une courbe (ou une surface pour une image 3D), notée \mathcal{C} , qui sera l'initialisation du contour actif, puis de la déformer successivement, selon un processus de déformation itératif, contrôlé par un test de convergence, jusqu'à ce qu'elle coïncide avec la frontière de l'objet à détecter.

Le snake est un ensemble de points d'un contour $\mathcal{C}(s) = (x(s), y(s))$, où $x(s)$, $y(s)$ sont les coordonnées de ces points, et $s \in [0, 1]$ est une approximation linéaire des coordonnées des points. L'approximation s représente l'abscisse curviligne, permettant de paramétrer la longueur de l'arc dans un intervalle très petit. Le contour final peut être obtenu, en minimisant la fonctionnelle d'énergie de l'équation (3.1) [Kass et al., 1988]:

$$\mathbf{E} = \mathbf{E}_{int} + \mathbf{E}_{image} \quad (3.1)$$

Cette fonctionnelle d'énergie est construite de telle sorte qu'un minimum local se trouve à la frontière de l'objet à détecter. L'énergie du snake se décompose en deux types :

– \mathbf{E}_{int} (énergie interne), qui permet au snake de se rétrécir ou de s'allonger ; en général, elle contrôle les caractéristiques géométriques (régularité, lissage) de la courbe \mathcal{C} recherchée.

$$\mathbf{E}_{int}(\mathcal{C}(s)) = \int_0^1 \alpha(s) \left| \frac{d\mathcal{C}}{ds} \right|^2 ds + \int_0^1 \beta(s) \left| \frac{d^2\mathcal{C}}{ds^2} \right|^2 ds \quad (3.2)$$

où α est le coefficient d'élasticité, et β le coefficient de rigidité [Bechar, 2009, Chenoune, 2008]. Cette énergie se décompose en deux parties. La première correspond à l'énergie d'élasticité, qui permet à la courbe de se rétrécir ou de s'allonger. La deuxième correspond à l'énergie de lissage. Notons que $\mathcal{C}''(s) = \left| \frac{d^2\mathcal{C}}{ds^2} \right|$ correspond à la courbure du snake au point s .

– \mathbf{E}_{image} (énergie de l'image ou énergie extérieure), qui attire la courbe \mathcal{C} vers l'objet, dont on cherche les frontières.

$$\mathbf{E}_{image}(\mathcal{C}(s)) = - \int_0^1 \lambda |\nabla I(\mathcal{C}(s))| ds \quad (3.3)$$

∇I est le gradient de l'intensité de l'image, et λ une constante réelle, dont le signe permet de créer localement un mouvement de contraction ou de dilatation de la courbe \mathcal{C} . L'énergie externe est minimale sur les contours de l'objet. Cela oblige à initialiser le snake à proximité du contour cible. Par ailleurs, les valeurs du gradient ont des sens inverses sur les deux côtés du même contour, ce qui empêche le snake d'entrer dans les concavités (voir figure 3.1.a). Pour cette raison, Xu et al. [Xu et Prince, 1997] ont proposé un nouveau type d'énergie externe, qui permet d'initialiser le snake loin du contour cible, et

d'aller vers les concavités (voir figure 3.1.b). Cette énergie est nommée flux du vecteur du gradient (GVF : Gradient Vector Flow).

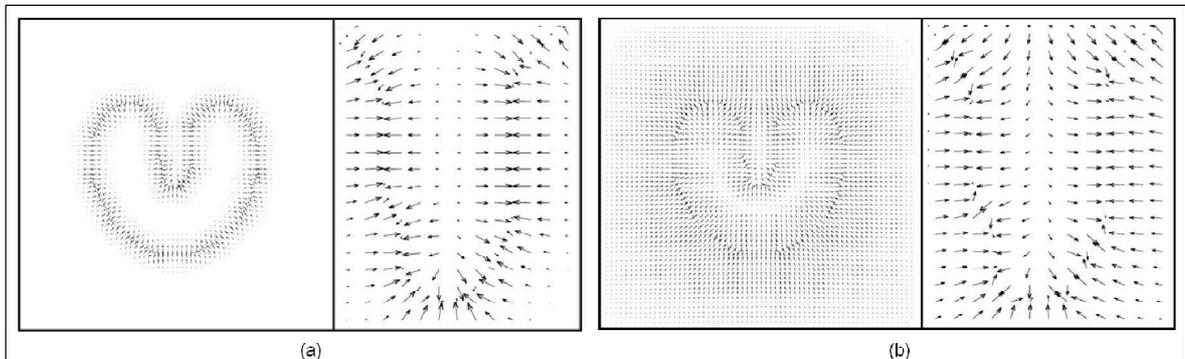


Figure 3.1. Énergie externe (indiquée en flèches) pour un objet qui contient des concavités :
 (a) Gradient ;
 (b) Flux du vecteur du gradient (GVF) [Xu et Prince, 1997].

Le principal avantage, offert par le modèle des Snakes classiques, est d'obtenir une convergence de la courbe vers les frontières recherchées, par une simple définition d'une fonctionnelle d'énergie. Il offre, en outre, la possibilité d'introduire de l'information, là où elle est manquante ou difficile à prendre en compte (performance que l'on ne peut pas avoir par un détecteur de contours classique) [Elomary, 2009].

En revanche, il souffre de plusieurs problèmes, les plus importants étant [Elomary, 2009] :

- Le modèle des snakes classiques fournit une solution qui n'est pas unique et qui dépend des conditions initiales (forme et position du contour initial,...).
- Il ne gère pas les ruptures topologiques et ne peut donc pas assurer la répartition d'une courbe initiale en plusieurs « courbes filles », afin de détecter plusieurs objets dans une image, comme l'exemple illustré en figure 3.2.
- De plus, aucune indication n'est donnée sur le choix des paramètres α et β , ce soin étant souvent laissé à l'utilisateur, en fonction des caractéristiques du contour de l'objet.
- Le contour initial doit être suffisamment proche de l'objet à détecter.

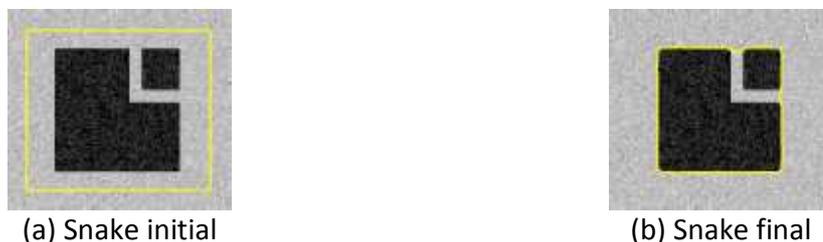


Figure 3.2. Méthode des Snakes appliquée à la détection de plusieurs objets [Bresson et al., 2007].

Le snake paramétrique a été beaucoup utilisé dans le domaine du traitement d'images de documents, surtout pour l'extraction des lignes de documents manuscrits :

- ✓ Un modèle de contour actif pour la restauration de caractères en niveau de gris dégradés a été développé par Allier et al. [Allier et al., 2006]. Les auteurs ont incorporé un modèle de forme unique à un contour actif paramétrique, basé sur une représentation de flux du vecteur du gradient (GVF) [Xu et Prince, 1998] de l'image et de la forme antérieure.
- ✓ Bukhari et al. ont proposé dans [Bukhari et al., 2009b] une méthode de snake paramétré, pour l'extraction des lignes de documents manuscrits. Le contour ou snake de départ est la ligne centrale des lignes du texte. Cette ligne est estimée, en se basant sur les valeurs de l'intensité dans une image de document en niveaux gris. Ensuite, le mécanisme de minimisation d'énergie de snake est appliqué pour trouver les lignes du texte.

Un autre modèle de contours actifs, « *le modèle géométrique* », permet de remédier à certains inconvénients des modèles paramétriques. Il a été développé par Caselles et ses collaborateurs [Caselles et al., 1993]. Le modèle sera présenté à la prochaine sous section.

3.3.2. Modèle déformable géométrique

Caselles et al. [Caselles et al., 1993] ont introduit un modèle géométrique des contours actifs, formulé par des équations aux dérivées partielles d'évolution de courbes, où chaque point de la courbe C se déplace dans la direction de son vecteur normal \vec{n} à une vitesse de propagation proportionnelle à la courbure k (équation (3.4)) (Figure 3.3). Ces modèles évoluent en utilisant uniquement des mesures géométriques, ce qui donne une évolution indépendante de la paramétrisation.

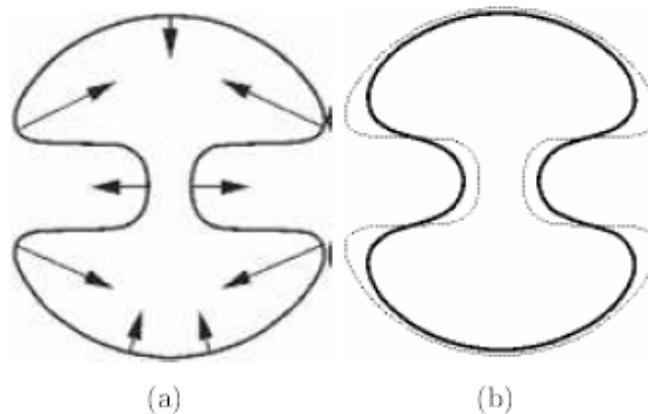


Figure 3.3. Sens de propagation d'une courbe sous l'action de la force de courbure, dans la direction de la normale en chaque point :

- (a) Mouvement initial ;
- (b) Forme finale (en noir).

L'équation d'évolution s'écrit alors sous la forme :

$$\frac{\partial C}{\partial t} = F(k)\vec{n} \quad (3.4)$$

Afin que la courbe converge vers les frontières de l'objet et s'arrête, la vitesse de propagation dépend aussi d'une fonction d'arrêt g (dépendant des données contenues dans l'image I), qui est monotone, positive et décroissante strictement. Les valeurs de cette fonction sont proche de 0 dans les régions où le gradient est élevé, et proche de 1 dans les régions homogènes, d'intensité constante (équation (3.5)).

$$g(I) = \frac{1}{1 + |\nabla G_\sigma * I|^p} \quad (3.5)$$

$G_\sigma * I$ est la convolution de l'image I avec une gaussienne, pour atténuer l'influence du bruit ; p prend pour valeurs 1 ou 2 .

Caselles et al. [Caselles et al., 1993] proposent le modèle géométrique représenté par l'équation (3.6) pour la segmentation :

$$\begin{cases} \frac{\partial C(s, t)}{\partial t} = g(I)(v + k)\vec{n} \\ C(s, 0) = C_0(s) \end{cases} \quad (3.6) \quad \text{La courbe initiale}$$

Où l'objet à détecter est donné par la solution de $\frac{\partial C(s, t)}{\partial t} = 0$.

Le modèle défini par cette EDP se déforme avec une vitesse associant deux termes [Elomary, 2009]:

- $k\vec{n}$ un terme de régularité qui lisse la courbe C .
- Le terme $v\vec{n}$ est introduit pour générer une force dans la direction normale à la courbe, permettant de faire évoluer cette dernière (courbe) vers les contours de l'objet, v est une constante que l'on peut prendre positive (resp négative) pour contracter (resp dilater) le contour actif.

L'équation (3.6) constitue le modèle géométrique, et le problème de segmentation est donc résolu en définissant de manière directe une équation d'évolution portant sur le contour. Cette EDP n'est plus associée à une minimisation d'énergie.

Contrairement au modèle des "Snakes classiques", les modèles géométriques permettent de gérer automatiquement, et sans aucune difficulté, les changements de topologie, comme la division ou la fusion de plusieurs contours, et offrent toujours une courbe / surface fermée.

Le tableau 3.1 explique la différence entre le snake paramétrique et le snake géométrique, selon cinq éléments : les paramètres, la détection (un ou plusieurs objets), la topologie, l'initialisation et l'utilisation.

Tableau 3.1. Les principes du modèle de contour actif : paramétrique et géométrique.

Contour actif paramétrique	Contour actif géométrique
Résultat dépendant de la paramétrisation (e.g. paramètres d'élasticité et de rigidité)	Résultat dépendant des informations géométriques de la courbe (e.g. vitesse de la courbure)
Détection d'un seul objet	Détection simultanée de plusieurs objets
Changement impossible de la topologie	Changement possible de la topologie
Initialisation proche du contour désiré	Initialisation non nécessairement proche du contour désiré
Utilisation pour des contours simples	Utilisation pour des contours complexes (e.g. images médicales)

Le snake géométrique a été beaucoup utilisé pour la segmentation d'images médicales complexes.

3.3.3. Méthode des ensembles de niveaux d'Osher –Sethian (Level set)

La méthode des ensembles de niveaux (Level Sets en anglais), introduite par Osher et Sethian [Osher et Sethian, 1988], est une technique établie pour représenter des courbes en mouvement, en deux dimensions ou plus. Elle consiste à représenter une courbe C implicitement, comme le niveau zéro d'une certaine fonction de niveaux, de dimension supérieure. Le mouvement de la courbe se traduit alors en une évolution de la fonction de niveaux, selon une EDP (équation aux dérivées partielles).

Dans le plan de l'image, la courbe C est donc considérée comme le niveau zéro d'une fonction Φ , de dimension supérieure (C est donc l'intersection de $\Phi(x, y)$ avec le plan défini par $xy(z = 0)$ $C: \{(x, y) / \Phi(x, y) = 0\}$), telles que les axes horizontaux représentent le plan de l'image, et l'axe vertical les différents niveaux de la fonction Φ (Figure 3.4).

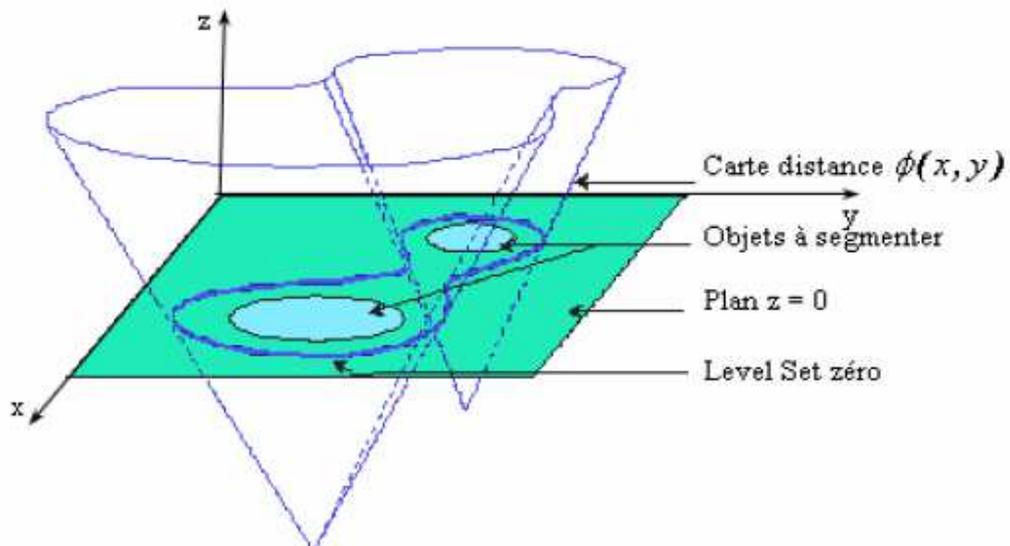


Figure 3.4. Représentation de la fonction des ensembles de niveaux $\Phi(x, y)$ à un instant donné t .

La courbe \mathcal{C} évolue dans le plan $xy(z = 0)$ vers les frontières des objets à segmenter (Figure 3.4), en utilisant l'image entière comme une carte de distance à la courbe \mathcal{C} ; en effet chaque pixel a pour valeur sa distance à la courbe (Figure 3.5).

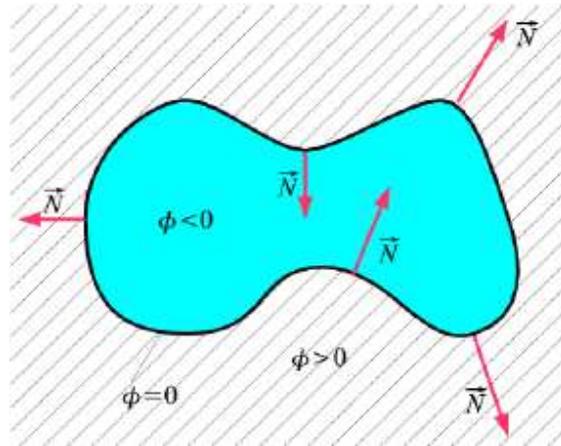


Figure 3.5. Fonction distance Φ signée au contour \mathcal{C} .

La fonction $\Phi(x, y)$ représente la distance signée de tous les points de l'image à la courbe. Elle est définie comme ceci :

- $\Phi(s, t) < 0$ si s est à l'intérieur de la courbe.
- $\Phi(s, t) > 0$ si s est à l'extérieur de la courbe.
- $\Phi(s, t) = 0$ si s est sur la courbe.

En transposant de ce fait l'étude de l'évolution de la courbe \mathcal{C} à celle de la fonction Φ , le problème de l'équation (3.6) devient [Elomary, 2009] :

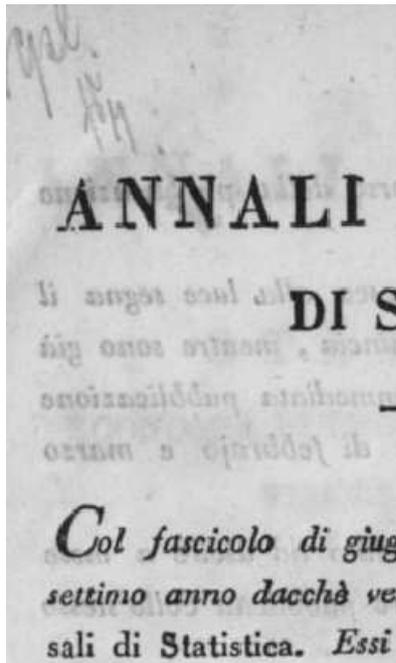
$$\begin{cases} \frac{\partial \Phi}{\partial t} = g(I) \left(v + \operatorname{div} \left(\frac{\nabla \Phi}{|\nabla \Phi|} \right) \right) |\nabla \Phi| \\ \Phi(s, 0) = \Phi_0(s) \end{cases} \quad \text{La courbe initiale} \quad (3.7)$$

Le terme $\operatorname{div} \left(\frac{\nabla \Phi}{|\nabla \Phi|} \right)$, désigne la courbure k .

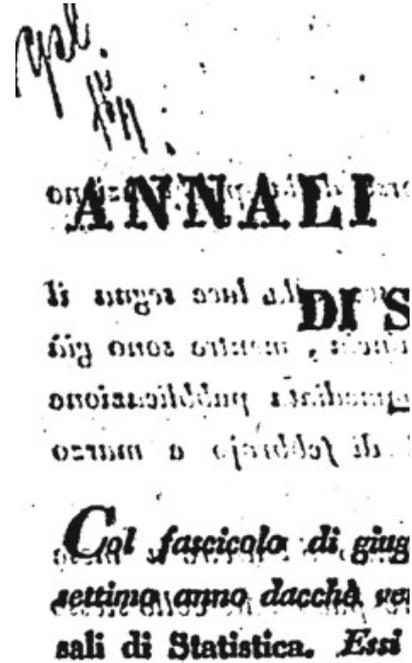
D'après l'équation (3.7), on remarque que la courbe s'arrête lorsque $g = 0$; ceci n'arrive que dans le cas d'un contour idéal, car pour des images réelles, g prend des valeurs proches de zéro le long des frontières de l'objet à segmenter. Il est donc nécessaire de considérer un intervalle dans lequel g puisse prendre des valeurs qui garantissent que la courbe s'arrête [Elomary, 2009].

La méthode de Level set a été beaucoup utilisée dans le domaine du traitement d'images de documents :

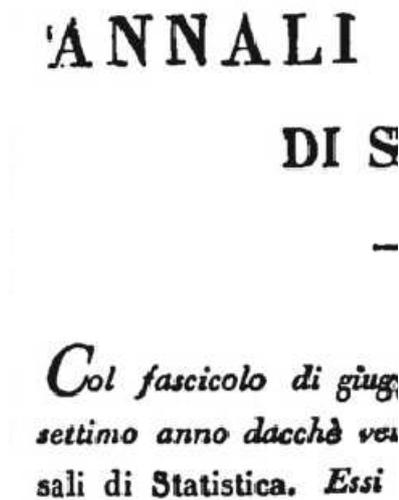
- ✓ Li et al. ont proposé dans [Li et al., 2008b] une telle technique, pour l'extraction des lignes de documents manuscrits complexes. La méthode démarre par l'application d'un filtre gaussien avec un noyau anisotrope, pour estimer la densité des pixels. Ensuite, une carte de probabilité est construite, donnant la probabilité pour qu'un pixel appartienne à une ligne du texte. Ensuite, le modèle de Level set est appliqué pour déformer le contour, afin de trouver les lignes. Enfin, deux étapes de post-traitement sont appliquées pour regrouper les lignes de texte voisines, qui ont une distance faible entre eux, et pour couper les lignes qui se connectent.
- ✓ Moghaddam et al. utilisent dans [Moghaddam et al., 2009] la méthode de Level set pour la restauration et la segmentation de caractères de texte fortement dégradés. Ils proposent de combiner deux approches complémentaires. Tout d'abord, les classificateurs multi-niveaux, qui profitent de l'information à priori de la largeur du trait du texte, permettent de localiser les pixels des caractères candidats. Deuxièmement, la méthode de Level set est utilisée pour identifier les contours d'un caractère.
- ✓ Yosef et al. utilisent dans [Yosef et al., 2009] la méthode de level set pour la restauration et la segmentation des caractères de texte fortement dégradés, ils prennent les bénéfices d'une forme préalable adaptative.
- ✓ Rivest-Hénault et al. ont proposé dans [Rivest-Hénault et al., 2012] une technique, basée sur la méthode de Level set, pour la binarisation d'images de documents historiques. La technique proposée a été testée sur la base de documents de la compétition DIBCO2009, et a été classée troisième dans cette compétition. Cette méthode est considérée comme la première qui utilise la méthode de Level set pour la binarisation d'images de documents. La méthode comprend trois étapes:
 - i) initialisation, en utilisant une carte de contour (SM) [Moghaddam et al., 2009];
 - ii) correction du SM, en utilisant des modèles linéaires locaux et la méthode level set en mode érosion;
 - iii) une deuxième série d'opérations de level set, cette fois avec une force de niveau de gris du texte, qui fournit les régions de texte finales, comme les régions intérieures de la fonction level set (Figure 3.6).



(a)



(b)



(c)

Figure 3.6. Exemple de binarisation d'une image de document tirée de [Rivest-Hénault et al., 2012] :

(a) Image à binariser ;

(b) Résultat de la binarisation par la méthode de Sauvola [Sauvola et Pietikainen, 2000] ;

(c) Résultat de la binarisation par la méthode de Rivest-Hénault [Rivest-Hénault et al., 2012].

La sous section suivante décrit l'approche géodésique pour la segmentation des objets, qui permette de relier les snakes classiques, basés sur la minimisation d'une énergie, et les contours actifs géométriques, basés sur la théorie d'évolution des courbes.

3.3.4. Modèle déformable géodésique

Les contours actifs géodésiques (CAG) sont introduits par Caselles et al. [Caselles et al., 1997] pour la détection des frontières d'objets. Ils sont fondés sur la relation entre les contours actifs classiques, et le calcul de géodésiques ou de courbes de distance minimale. En effet, les CAG découlent d'une formulation simplifiée du modèle des Snakes classiques (équation 3.1), mais ce dernier a été modifié, en éliminant le terme de rigidité ($\beta = 0$). L'élimination de ce terme permet de limiter les termes dérivés à un ordre élevé.

Dans le modèle proposé par Caselles et al. [Caselles et al., 1997], la formulation de l'énergie totale du contour, obtenue après modification des Snakes classiques, en supprimant le terme de rigidité, est de la forme :

$$E = \int_0^1 \alpha |C'(s)|^2 ds - \int_0^1 \lambda |\nabla I(C(s))| ds \quad (3.8)$$

Il ne reste ainsi que deux termes dans l'énergie totale : le premier assure le lissage du contour actif géodésique, et le second agit comme un détecteur de contours améliorable via l'introduction d'une fonction monotone $g : [0, +\infty[\rightarrow \mathcal{R}^+$ strictement décroissante telle que $\lim_{r \rightarrow +\infty} g(r) = 0$, appelée **fonction d'arrêt**. L'équation (3.8) devient :

$$E(C) = \underbrace{\int_0^1 \alpha |C'(s)|^2 ds}_{E_{int}} + \underbrace{\int_0^1 \lambda g(|\nabla I(C(s))|)^2 ds}_{E_{ext}} \quad (3.9)$$

Le rôle de la fonction g à l'équation (3.9) est de ralentir l'évolution des courbes dans les zones de fort gradient.

Les modèles géodésiques actifs se rapprochent de la détection de contours par snakes, mais ils ont pris le pas, car ils sont capables de se dédoubler et de fusionner, pour permettre la détection simultanée de plusieurs objets. Ils aboutissent à une détection de contours plus précise et indépendante de l'initialisation. Ils permettent également de résoudre certains problèmes liés à la paramétrisation.

Le modèle de contours actifs géodésiques a été utilisé dans le domaine du traitement d'images de documents :

- ✓ Allier et Emptoz utilisent dans [Allier et Emptoz, 2004] un modèle de contours actifs géodésiques pour la segmentation et l'extraction de la typographie dans les images de documents. Ils proposent l'utilisation de méthodes de traitement d'images texturées robustes (filtrage de Gabor et modèle de contours actifs géodésiques), au lieu des techniques classiques de traitement d'images de documents, qui fonctionnent au niveau des pixels, et qui sont très sensibles à de nombreux paramètres, tels que le bruit et l'inclinaison.
- ✓ Cohen et al. utilisent dans [Cohen et al., 2012] un modèle de contours actifs géodésiques modifié, pour la restauration des caractères brisés, qui ont été partiellement occultés par un texte.

Les méthodes basées «contour» mènent à de bonnes segmentations d'objets, dont les bords correspondent aux forts gradients de l'image. Mais, dans le cas où la frontière entre les régions n'est pas bien marquée, ce qui est le cas par exemple des images fortement bruitées, les informations locales type «gradient» ne sont plus suffisantes, et les résultats de segmentation deviennent très dépendants de l'initialisation de la courbe active.

Pour pallier les limites des méthodes basées uniquement sur des informations «contour», comme décrit précédemment, des approches de segmentation par contours actifs basés «région», ont été introduites initialement dans [Mumford et Shah, 1989].

3.3.5. Modèle de Mumford-Shah

Le modèle de Mumford-Shah [Mumford et Shah, 1989] déforme les contours, en minimisant la fonction d'énergie de l'équation (3.10) :

$$E(I, C) = \int_{\Omega} |I - I_0|^2 dx dy + \mu \int_{\Omega/C} |\nabla I|^2 dx dy + \nu |C| \quad (3.10)$$

où I_0 est l'image initiale, I l'image lissée après l'application d'un filtre gaussien, C l'ensemble des courbes (ou snakes) de départ, $|C|$ l'ensemble des longueurs des courbes, Ω le domaine de l'image, Ω/C le domaine de l'image sans l'ensemble des courbes, ∇I le gradient de l'image I , et μ et ν les paramètres qui équilibrent l'effet des autres termes dans cette équation.

- ✓ Du et al. ont proposé une telle méthode dans [Du et al., 2009], pour l'extraction des lignes de documents manuscrits latins. L'extraction des contours de départ est similaire à celle utilisée dans [Li et al., 2008b]. L'extraction des lignes se fait par l'application du modèle de Mumford-Shah.

D'après une étude comparative faite par Meziou dans [Meziou, 2013], les approches basées régions restent à ce jour celles permettant l'obtention des résultats de segmentation les plus

intéressants, en particulier dans le contexte d'images présentant des objets à segmenter que les descripteurs fondés sur les contours de l'image n'arrivent pas à discriminer; nos recherches ont donc été dirigées vers ces approches.

3.4. Conclusion

Les modèles déformables sont des modèles mathématiques ou géométriques qui, se déforment de façon à séparer, d'un côté l'objet qui nous intéresse, et de l'autre le reste de l'image.

Dans ce chapitre, nous avons présenté les principaux modèles de contours actifs, basés sur des critères «contour»: les modèles paramétriques classiques (Snakes), leur principe, leurs avantages, ainsi que leurs inconvénients qui ont été palliés par la suite par l'introduction des modèles déformables géométriques, qui sont implémentés par la méthode des ensembles de niveaux. On a vu aussi la combinaison entre l'approche variationnelle et l'approche géométrique, qui a donné la naissance aux modèles déformables géodésiques.

Malheureusement, les méthodes basées contours mènent à de bonnes segmentations d'objets, dont les bords correspondent aux forts gradients de l'image. Mais dans le cas où la frontière entre les régions n'est pas bien marquée, ce qui est le cas, par exemple, des images fortement bruitées, les informations locales, type «gradient», ne sont plus suffisantes et les résultats de segmentation deviennent très dépendants de l'initialisation de la courbe active.

Pour pallier les limites des méthodes, basées uniquement sur des informations «contour», des approches de segmentation par contours actifs, basés régions, ont été introduites. Ces approches restent à ce jour celles permettant l'obtention des résultats de segmentation les plus intéressants [Meziou, 2013], ce qui nous a poussé à diriger nos recherches vers ces approches.

Pour les images de documents anciens, qui sont généralement de mauvaise qualité, et contenant différents types de détériorations et de dégradations, avec un fond texturé et une illumination inégale, aucun modèle déformable basé contour n'est capable de séparer parfaitement l'avant-plan de l'arrière-plan de l'image, puisque la frontière entre les régions du premier plan et de l'arrière-plan n'est pas bien marquée. Ce type de documents nécessite une analyse plus fine, et, dans ce cas, on a recours à un nouveau modèle déformable basé région, présenté au chapitre suivant, avec des forces motrices spéciales, pour pouvoir gérer les inhomogénéités d'intensité qui se produisent souvent dans les images de documents réelles.

CHAPITRE 4

Contributions à la binarisation d'images de documents dégradés par contours actifs

4.1. Introduction

Comme nous avons déjà mentionné dans les chapitres précédents, la binarisation est une tâche critique, dans le sens où la mauvaise séparation entre l'avant-plan et l'arrière-plan entraîne la perte d'informations pertinentes, et/ou l'ajout d'informations incorrectes (bruit). Cette séparation est devenue plus difficile dans le cas d'images de documents anciens, caractérisés par différents types de dégradations, provenant du processus de numérisation, des effets du vieillissement, de l'humidité, etc. L'importance de l'étape de binarisation justifie le grand nombre de méthodes proposées, qui ont pour but d'arriver à une séparation parfaite entre l'avant-plan et l'arrière-plan, objectif qui reste loin d'être atteint. La séparation avant-plan/ arrière-plan dans les images de documents est devenue aujourd'hui un champ de recherche actif en analyse et reconnaissance de documents, qui attire une attention croissante chaque année.

Dans ce chapitre, une nouvelle technique de binarisation, pour les images de documents dégradés, est présentée. La technique proposée est basée sur l'évolution d'un contour actif, selon les mesures géométriques intrinsèques de l'image de document à binariser. Le contraste de l'image, qui est définie par le maximum et le minimum locaux, est utilisé pour générer automatiquement la carte d'initialisation de notre modèle de contour actif; un seuillage adaptatif moyen est également utilisé pour produire la binarisation finale. La suite de ce chapitre est consacrée à la présentation de cette méthode proposée, publiée au journal **Signal, Image and Video Processing (SIVP)** en 2017 [[Hadjadj et al., 2017](#)].

4.2. Motivations et objectifs

Dans ce chapitre, nous présentons une nouvelle méthode de binarisation, d'images de documents anciens, utilisant un modèle de contour actif. Les modèles de contour actif ont été largement appliqués à la segmentation d'images. Ils ont plusieurs avantages par rapport aux méthodes classiques de segmentation d'images. Premièrement, les modèles de contour actif peuvent atteindre la précision du sous-pixel 'sub-pixel accuracy' des limites de l'objet [[Caselles et al., 1997](#)]. Deuxièmement, les modèles de contour actifs peuvent être facilement formulés, dans un cadre de minimisation d'énergie, fondée sur des principes, et permettant l'incorporation de diverses connaissances antérieures 'prior knowledge', telles que la distribution de formes et d'intensité 'shape and intensity', pour une segmentation d'image robuste [[Chen et al., 2002](#)] [[Leventon et al., 2000](#)]. Troisièmement, ils peuvent fournir des

contours lisses et fermés, comme résultats de segmentation, qui sont nécessaires et qui peuvent être facilement utilisés pour d'autres applications, telles que l'analyse de forme et la reconnaissance [Li et al., 2008a].

Au cours de notre recherche bibliographique, nous avons remarqué que les contours actifs sont très utilisés et très performants dans le traitement d'images médicales, satellitaires, etc., mais très peu utilisés pour le traitement d'images de documents anciens. Nos recherches ont donc été orientées vers cette approche, et dans la méthode que nous proposons, nous adoptons les contours actifs pour la binarisation d'images de documents.

La méthode proposée est moins sensible au fort bruit de fond des images de documents, et produit de meilleurs résultats de binarisation que les approches antérieures, comme indiqué dans la section d'évaluation de la performance (Section 4.4.4). De plus, contrairement à d'autres approches, notre méthode garantit une bonne binarisation, des caractères de texte faiblement et correctement contrastés, présents dans la même image de document, ce qui conduit également à des résultats plus cohérents avec des contours plus continus.

Le reste de ce chapitre est organisé comme suit. La méthode de binarisation proposée, basée sur un modèle de contour actif, est présentée à la section 4.3. Les résultats expérimentaux de la méthode, appliquée sur des images de documents réels, de même que la comparaison des performances avec d'autres méthodes de binarisation sont également présentés à la section 4.4. Enfin, une conclusion et des perspectives pour de futures recherches sont prévues à la section 4.5.

4.3. Description de la technique proposée

Une nouvelle méthode de binarisation est introduite dans cette section. Celle-ci bénéficie de la flexibilité des modèles de contours actifs, pour intégrer diverses informations régionales en un seul modèle. Étant donné que les modèles de contours actifs sont des modèles d'optimisation locale, avec une tendance à tomber sur le minimum local le plus proche, donc une nouvelle méthode d'initialisation automatique adéquate est proposée.

La méthode proposée se compose d'un processus à trois étapes (Figure 4.1):

- (1) Une carte identifiant les pixels de texte les plus probables, est extraite de l'image du document, pour servir d'initialisation pour la fonction level set;
- (2) Le modèle de contour actif est évolué en mode de déplacement libre, afin d'enlever les pixels de texte les moins probables de la carte d'initialisation;
- (3) Le modèle de contour actif est évolué en mode de déplacement guidé, dans l'image du seuillage adaptatif moyen correspondante, pour produire la binarisation finale.

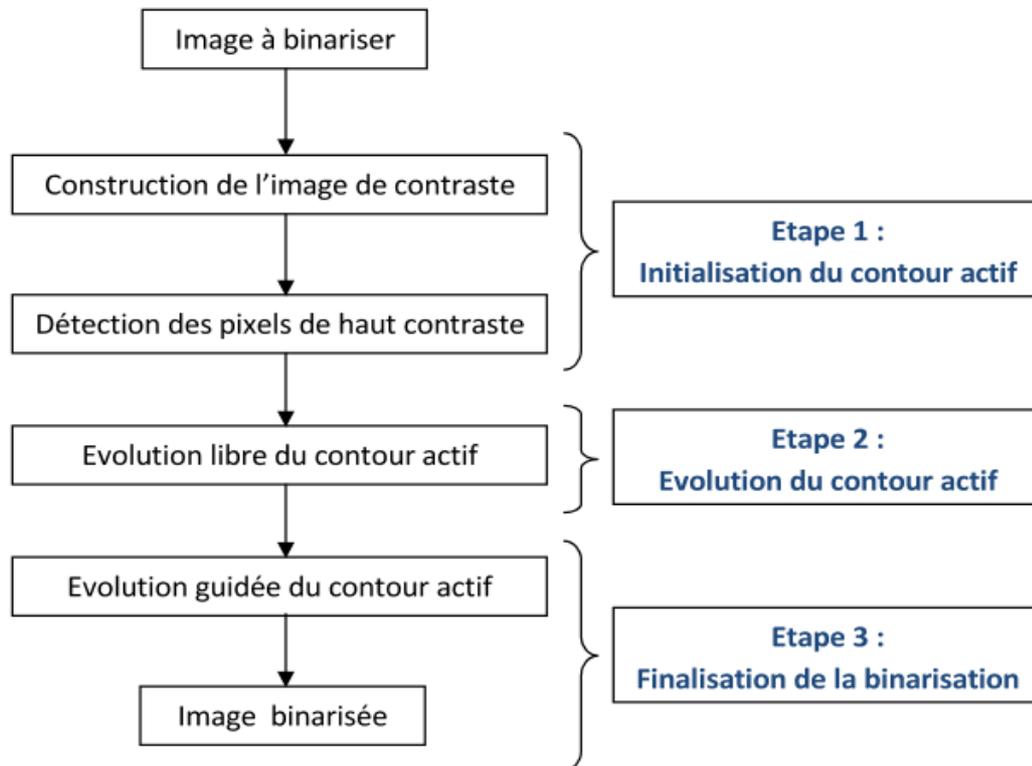


Figure 4.1. Organigramme des étapes de la méthode proposée.

Dans la sous-section suivante, nous présentons une estimation de la binarisation initiale, à l'aide du contraste de l'image. À ce stade, le but est d'obtenir une binarisation approximative, qui conserve le maximum des principales caractéristiques de l'image d'entrée.

4.3.1. Etape 1: Initialisation du contour actif

Une étape cruciale qui affecte la performance du modèle de contour actif ultime est l'initialisation. Le contour initial doit être assez proche de l'objet à détecter, sinon il convergera vers un mauvais résultat. Pour cela, nous proposons une nouvelle approche d'initialisation automatique, basée sur le contraste de l'image, pour identifier les pixels de l'avant-plan les plus probables. L'initialisation utilisée construit d'abord une image de contraste, évaluée par le maximum et le minimum locaux [van Herk, 1992, Su et al., 2010], puis détecte les pixels de haut contraste, qui se situent habituellement autour des contours du texte.

4.3.1.1. Construction de l'image de contraste

Le gradient de l'image a été largement utilisé dans la littérature, pour la détection de contours [Canny, 1986]. Cependant, le gradient de l'image est souvent obtenu par la différence absolue d'image, dans une fenêtre de voisinage locale, ce qui n'intègre pas l'intensité de l'image elle-même, et il est trop sensible à la variation du contraste/luminosité de l'image. Prenez une image de document historique, inégalement illuminée à titre

d'exemple, le gradient d'un pixel d'image (autour du texte), dans des régions lumineuses du document, peut être beaucoup plus élevé que dans les régions sombres du document. Pour détecter correctement les pixels à contraste élevé autour du texte, le gradient d'image doit être normalisé, pour compenser l'effet de la variation du contraste/luminosité de l'image. Au même temps, la normalisation supprime également la variation, dans l'arrière-plan du document.

Dans la technique proposée, nous supprimons la variation dans le fond, en utilisant un contraste d'image, calculé en fonction de l'image locale maximale et minimale [van Herk, 1992, Su et al., 2010], comme suit:

$$D(x, y) = \frac{f_{max}(x, y) - f_{min}(x, y)}{f_{max}(x, y) + f_{min}(x, y) + \varepsilon} \quad (4.1)$$

Où, $f_{max}(x, y)$ et $f_{min}(x, y)$ expriment l'intensité maximale et minimale, respectivement, dans une fenêtre de voisinage locale de taille 3×3 , centrée sur le pixel (x, y) . Le terme ε est un nombre positif très petit, qui est ajouté pour éviter la forme indéfinie, dans le cas où le maximum local est égal à 0.

Dans l'équation (4.1), le numérateur (c'est-à-dire la différence entre le maximum local et le minimum local) calcule la différence d'image locale, qui est similaire au gradient d'image traditionnel [Canny, 1986]. Le dénominateur agit comme un facteur de normalisation, qui diminue l'effet de la variation du contraste et de la luminosité. Pour les pixels d'image, dans les régions lumineuses autour du texte, le dénominateur est grand, ce qui neutralise le grand numérateur, et entraîne donc un contraste d'image relativement diminué. Mais pour les pixels d'image, dans les régions sombres autour du texte, le dénominateur est petit, ce qui compense le petit numérateur, et entraîne donc un contraste d'image relativement augmenté.

En conséquence, les contrastes des pixels d'image (situés autour du texte), dans les régions lumineuses et sombres des documents, convergent les uns sur les autres, ce qui facilite la détection des pixels à contraste élevé, situés autour du texte.

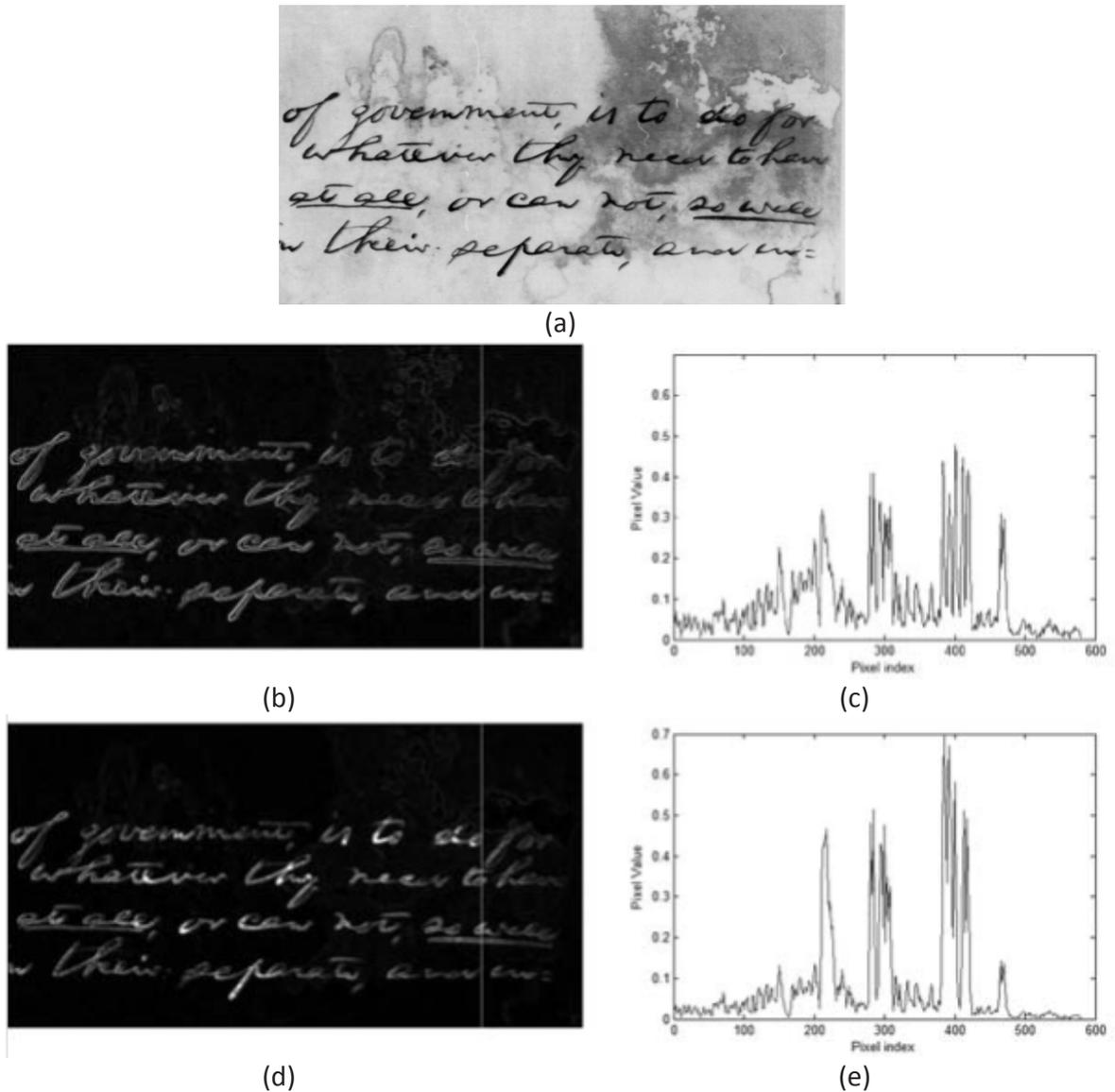


Figure 4.2. Image de gradient et de contraste :

(a) Image d'entrée ;

(b) Image de gradient, obtenue par le détecteur de contours de Canny [Canny, 1986] ;

(d) Image de contraste, obtenue par les maximum et minimum locaux [van Herk, 1992, Su et al., 2010] ;

(c) Une colonne de l'image de gradient de la figure 4.2.b (montrée comme une ligne blanche verticale) ;

(e) La même colonne de l'image de contraste de la figure 4.2.d.

La figure 4.2 illustre la différence, entre le gradient d'image et le contraste d'image défini dans l'équation (4.1). En particulier, les figures 4.2.b et 4.2.d montrent respectivement, l'image du gradient, et l'image du contraste. La figure 4.2.b montre que le gradient autour du texte varie visiblement, des régions lumineuses vers les régions sombres du document.

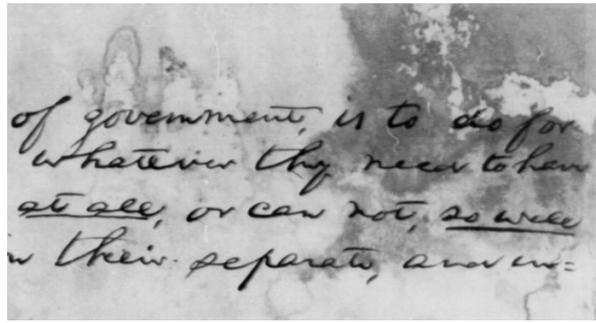
Cependant, comme le montre la figure 4.2.d, le contraste de l'image autour du texte varie peu, des régions lumineuses vers les régions sombres du document.

En outre, la discrimination entre le contraste de l'image autour du texte, et celui autour du fond du document, est beaucoup plus forte, par rapport à la discrimination du gradient d'image autour du texte, et celui autour du fond du document. Ces deux points peuvent être illustrés, plus en détail, sur les figures 4.2.c et 4.2.e, où la même colonne, de l'image du gradient de la figure 4.2.b et celle de l'image du contraste de la figure 4.2.d, sont tracées, respectivement.

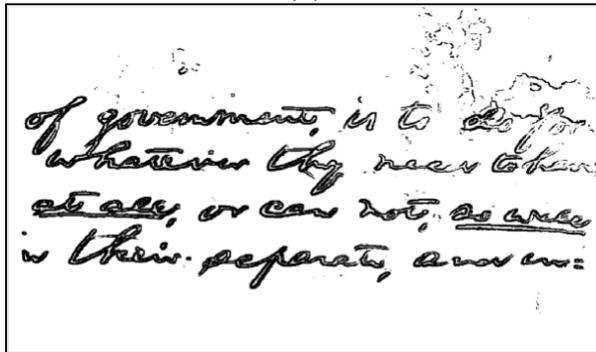
4.3.1.2. Détection des pixels de haut contraste

Le but de la construction de l'image de contraste D , est de détecter les pixels d'image, de contraste élevé désirés, situés autour des contours du texte. Comme décrit dans la sous-section 4.3.1.1, le contraste de l'image, autour des contours du texte est évidemment beaucoup plus grand, par rapport au contraste de l'image, dans l'arrière-plan du document, ce qui facilite donc la détection des pixels d'image, de contraste élevé désirés (situés autour des contours du texte), en utilisant la méthode de seuillage global d'Otsu [Otsu, 1979] (Figure 4.3.b).

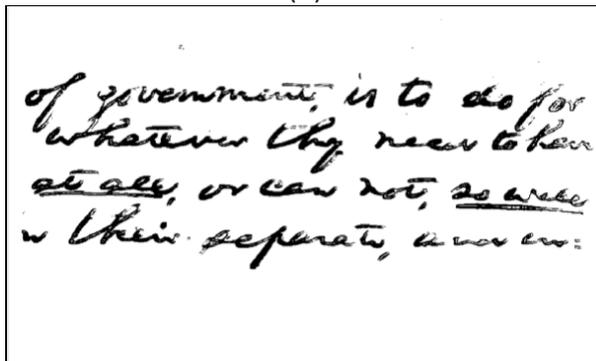
Les figures 4.3.a et 4.3.b montrent les résultats de la binarisation de l'image du gradient de la figure 4.2.b, et l'image du contraste de la figure 4.2.d, respectivement, en utilisant la méthode du seuillage global d'Otsu. Comme le montre la figure 4.3.b, la plupart des pixels, à contraste élevé, détectés par la binarisation de l'image de contraste, correspondent exactement aux pixels texte souhaités. D'autre part, la binarisation de l'image du gradient, dans la figure 4.3.a, introduit une certaine quantité de pixels indésirables, qui se situent habituellement dans l'arrière-plan dégradé du document.



(a)



(b)



(c)

Figure 4.3. Détection des pixels de haut contraste :

(a) Image originale ;

(b) Seuillage global de l'image de gradient de la figure 4.2.b par la méthode d'Otsu [Otsu, 1979] ;

(c) Seuillage global de l'image de contraste de la figure 4.2.d par la méthode d'Otsu [Otsu, 1979].

Une fois que la carte d'initialisation a été calculée à la fin de cette étape (Figure 4.3.b), l'évolution itérative de notre modèle de contour actif commence.

4.3.2. Etape 2: Evolution du contour actif

Le but de cette étape est de supprimer les pixels du bruit de fond, marqués comme des pixels texte, dans la carte du texte précédente, pour calculer, à l'étape suivante, la meilleure carte du texte possible. Cette étape est basée sur le modèle de contour actif proposé, dont les caractéristiques principales sont expliquées dans les paragraphes suivants.

Dans notre modèle de contour actif, la fonction level set ϕ est d'abord initialisée, comme une fonction de distance, signée à partir de la carte du texte extraite à l'étape précédente, à l'aide d'un algorithme de transformation de distance¹. La fonction level set est construite, de sorte que la région des pixels texte corresponde à $\phi > 0$. Pour déformer le contour actif, la carte du texte est érodée, en mettant à jour la fonction level set ϕ avec:

$$\frac{\partial \phi}{\partial t} = \delta_\varepsilon(\phi)(F_{DF} + \nu F_R + \beta F_A) \quad (4.2)$$

Où β et ν sont des nombres réels positifs, utilisés pour équilibrer les effets des trois forces, F_R est la force de régularisation de la longueur 'length regularization force', F_A est la Force d'expansion ou de contraction de la zone 'area expansion or contraction force', F_{DF} est la force d'ajustement de données 'data fitting force', δ_ε est la fonction de lissage de Dirac donnée par:

$$\delta_\varepsilon(x) = \frac{1}{\pi} \frac{\varepsilon}{\varepsilon^2 + x^2} \quad (4.3)$$

Les trois forces motrices de notre modèle de contour actif sont présentées et discutées dans les sous-sections suivantes.

4.3.2.1. Force de régularisation de la longueur 'Length regularization force' (F_R):

Avec la méthode des levels set, il faut ajouter une force motrice de courbure, à l'équation d'évolution. Cette force tend à minimiser la longueur du contour, augmentant ainsi sa régularité. La force de régularisation de la longueur 'length regularization force' est définie comme suit [Osher et Sethian, 1988]:

$$F_R(\phi) = \nabla \cdot \left(\frac{\nabla \phi}{|\nabla \phi|} \right) \quad (4.4)$$

Selon la figure 4.4 qui illustre l'effet de la force F_R sur l'évolution du contour actif, F_R a un effet de raccourcissement de longueur ou de lissage sur le contour de niveau zéro, ce qui est nécessaire pour maintenir la régularité du contour. Ce terme énergétique est également appelé le terme de longueur de l'arc 'arc length term'.

¹ Numériquement, cela se fait en utilisant la fonction `bwdist` de Matlab.

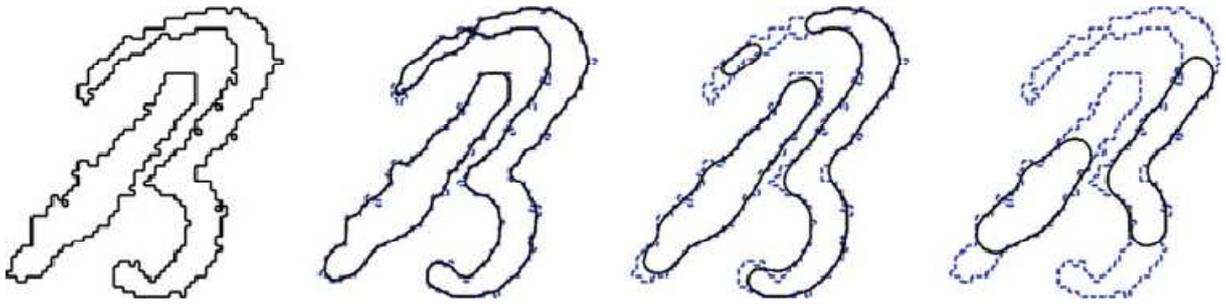


Figure 4.4. Effet de la force de régularisation de la longueur de (4.4) sur l'évolution du contour actif. De gauche à droite : initialisation, à $t = 0.4$, $t = 4$ et $t = 40$.

4.3.2.2. Force d'expansion ou de contraction de la zone 'Area expansion or contraction force' (F_A) :

Une simple force est ajoutée à l'équation (4.2), pour provoquer les zones de texte d'être entourées par le contour de niveau zéro défini par $(\{x, y\} | \phi(\mathbf{r}) = 0)$. Cette force permet au contour actif de s'étendre, ou de se contracter, pour détecter la région d'intérêt dans l'image. La force résultante utilise la fonction indicatrice de bord (4.6), et elle peut être écrite comme suit [Rivest-Hénault et al., 2012]:

$$\mathbf{F}_A = D\mathbf{g}(\phi) \quad (4.5)$$

Où $\mathbf{g} \in \{-1, 1\}$, définit la direction, et \mathbf{g} est la fonction indicatrice de bord 'edge indicator function', donnée par:

$$\mathbf{g}(I) = \frac{1}{1 + \nabla(K_\sigma * I)} \quad (4.6)$$

$K_\sigma * I$ est la convolution de l'image I avec une gaussienne, pour atténuer l'influence du bruit ; K_σ est un noyau de lissage gaussien, avec un paramètre d'échelle $\sigma > 0$, défini par:

$$K_\sigma(I) = \frac{1}{(2\pi)^{n/2}\sigma^n} e^{-|I|^2/2\sigma^2} \quad (4.7)$$

Le rôle de la fonction \mathbf{g} , dans le terme d'énergie F_A , est de ralentir le rétrécissement ou l'expansion du contour de niveau zéro, lorsqu'il atteint les contours d'objet, en prenant des valeurs plus petites. Voir la figure 4.5 pour une illustration de l'effet de F_A sur l'évolution d'un contour actif.

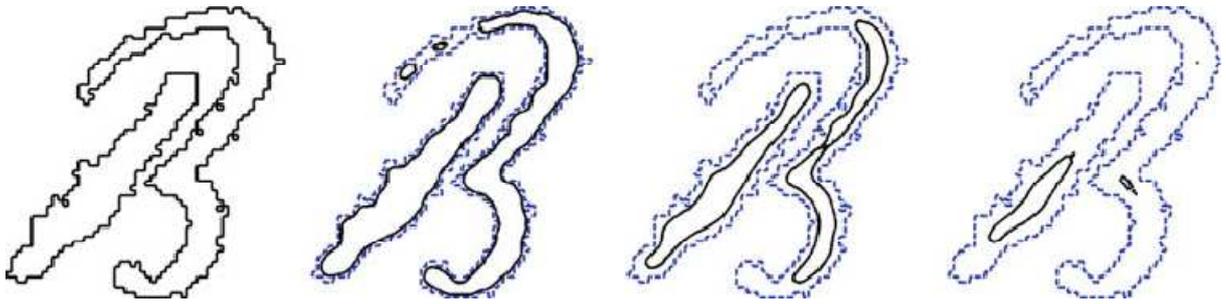


Figure 4.5. Effet de la force d'expansion ou de contraction de la zone de (4.5) avec $D = -1$ (rétrécissement) sur l'évolution du contour actif. De gauche à droite : initialisation, à $t = 0.4$, $t = 1.4$ et $t = 2$.

4.3.2.3. Terme d'ajustement de données 'Data fitting term' (F_{DF}) :

Les inhomogénéités d'intensité se produisent souvent dans les images de documents réelles, et peuvent entraîner des difficultés considérables lors de la binarisation. Afin de surmonter les difficultés, causées par ces inhomogénéités d'intensité, nous proposons d'ajouter un terme d'ajustement de données 'data fitting term', à l'équation d'évolution du contour actif. Ce terme était proposé par Li et al. dans [Li et al., 2008a], afin de gérer les difficultés causées par les inhomogénéités d'intensité dans les images. Le terme d'ajustement de données, dans (4.8), est défini, en termes de contour et de deux fonctions d'ajustement, qui approximent localement les intensités d'image sur les deux côtés du contour. Ce terme est responsable de la conduite du contour actif vers les contours des objets.

$$F_{DF} = -(\lambda_1 e_1 - \lambda_2 e_2) \quad (4.8)$$

Où e_1 et e_2 sont les fonctions d'ajustement données par:

$$e_i(x) = \int K_\sigma(y-x) |I(x) - f_i(y)|^2 dy \quad (4.9)$$

$$i = 1, 2,$$

Où f_1 et f_2 sont donnés par:

$$f_i(x) = \frac{K_\sigma(x) * [M_i^\varepsilon(\phi(x))I(x)]}{K_\sigma(x) * M_i^\varepsilon(\phi(x))} \quad (4.10)$$

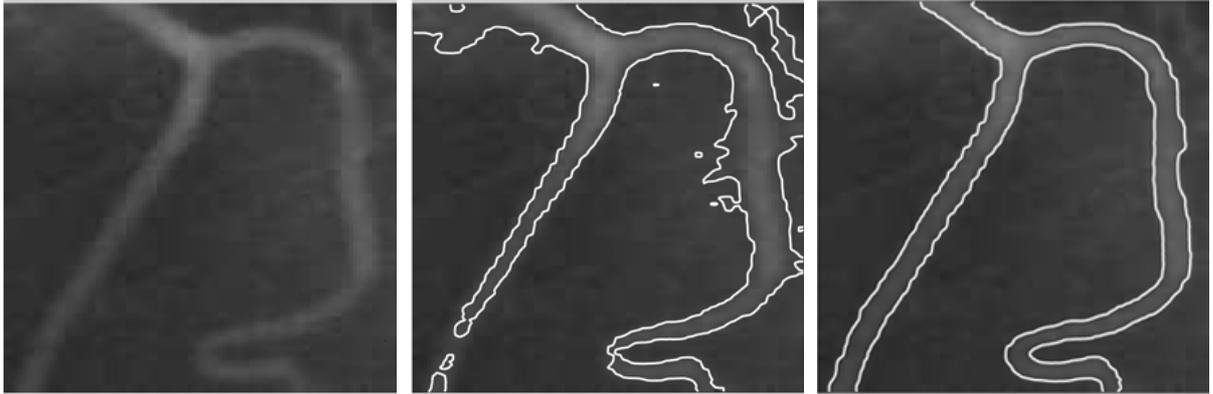
$$i = 1, 2,$$

$$M_1^\varepsilon(\phi) = H_\varepsilon(\phi),$$

$$M_2^\varepsilon(\phi) = 1 - H_\varepsilon(\phi).$$

En pratique, la fonction de Heaviside H (également fonction échelon unité, fonction marche d'escalier), dans les fonctionnelles d'énergie ci-dessus, est approximée par une fonction lisse H_ε définie par :

$$H_\varepsilon(x) = \frac{1}{2} \left[1 + \frac{2}{\pi} \arctan \left(\frac{x}{\varepsilon} \right) \right] \quad (4.11)$$



(a)

(b)

(c)

Figure 4.6. Effet du terme d'ajustement de données 'Data fitting term' sur l'évolution du contour actif :

- (a) Image originale avec inhomogénéité d'intensité (contours d'objet faibles) ;
- (b) Résultat du modèle de Chan et Vese [Vese et Chan, 2002] (qui n'utilise pas le 'data fitting term') ;
- (c) Résultat du modèle proposé (qui utilise le 'data fitting term').

Selon la figure 4.6 qui illustre l'effet de la force F_{DF} sur l'évolution du contour actif, en raison de cette force, l'information d'intensité dans les régions locales est extraite, pour guider le mouvement du contour actif, ce qui permet au modèle de faire face à l'inhomogénéité d'intensité, contrairement à d'autres modèles, qui n'utilisent pas le 'data fitting term' (Figure 4.6).

La force d'ajustement des données 'data fitting force' est brièvement présentée ici, pour plus de détails, nous renvoyons le lecteur intéressé au travail de Li et al. qui ont proposé cette force dans [Li et al., 2008a].

Le terme d'ajustement des données 'data fitting term' est ensuite incorporé dans une formulation level set avec les autres forces.

4.3.2.4. Formulation avec les Level Set :

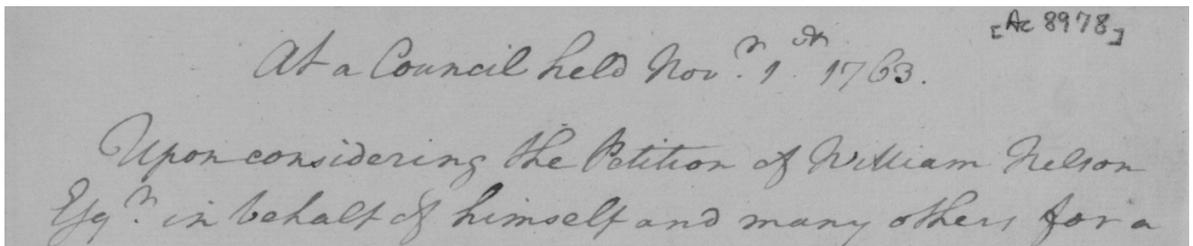
Dans la méthode des levels set [Osher et Sethian, 1988], un contour est représenté par le niveau zéro d'une fonction appelée fonction level set. Dans ce chapitre, nous laissons la fonction level set ϕ prendre des valeurs positives et négatives, à l'intérieur et à l'extérieur du contour actif, respectivement. L'évolution de la fonction level set, établie dans notre modèle, est donnée par:

$$\frac{\partial \phi}{\partial t} = -\delta_\varepsilon(\phi)(\lambda_1 e_1 - \lambda_2 e_2) + \nu \delta_\varepsilon(\phi) \operatorname{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) + \beta Dg \delta_\varepsilon(\phi) \quad (4.12)$$

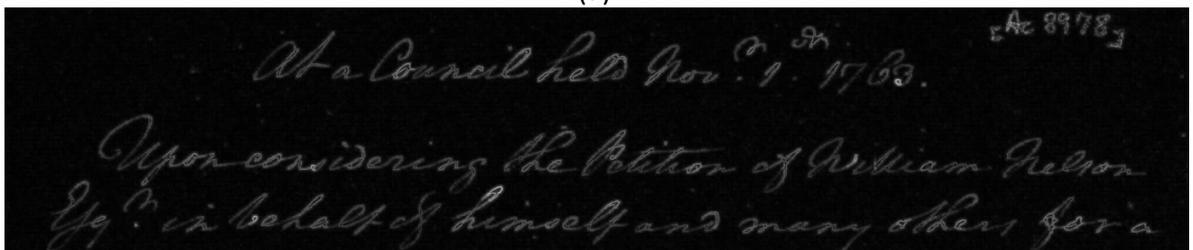
L'équation (4.12) représente l'équation d'évolution de la fonction level set, à résoudre dans la méthode proposée. Le terme $-\delta_\varepsilon(\phi)(\lambda_1 e_1 - \lambda_2 e_2)$ est appelé terme d'ajustement des

données 'data fitting term'. Ce terme joue un rôle clé dans le modèle proposé, car il est responsable de la conduite du contour actif vers les contours des objets.

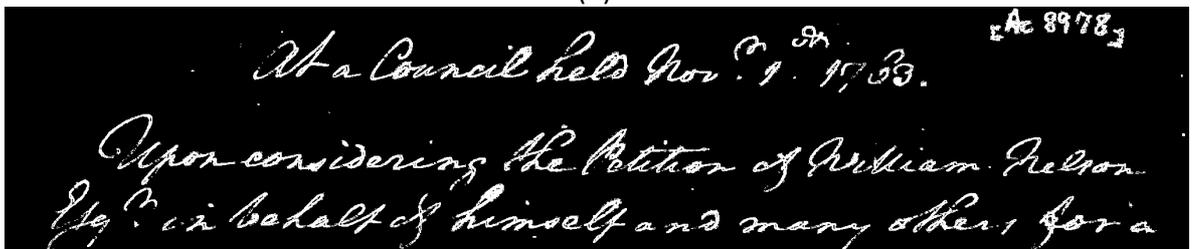
Le deuxième terme $\nu \delta_\epsilon(\phi) \operatorname{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right)$ a un effet de raccourcissement de la longueur ou du lissage, sur le contour de niveau zéro, ce qui est nécessaire pour produire des résultats plus naturels. Ce terme est appelé aussi terme de longueur de l'arc 'arc length term'. Étant donné que la carte du texte, extraite lors de l'initialisation, peut être imprécise, le troisième terme dans (4.12), appelé 'area reduction force' F_A , est inclus, pour favoriser l'élimination des pixels texte les moins probables de la carte d'initialisation. Pour que cela fonctionne comme prévu, le paramètre D de (4.5) est mis à -1.



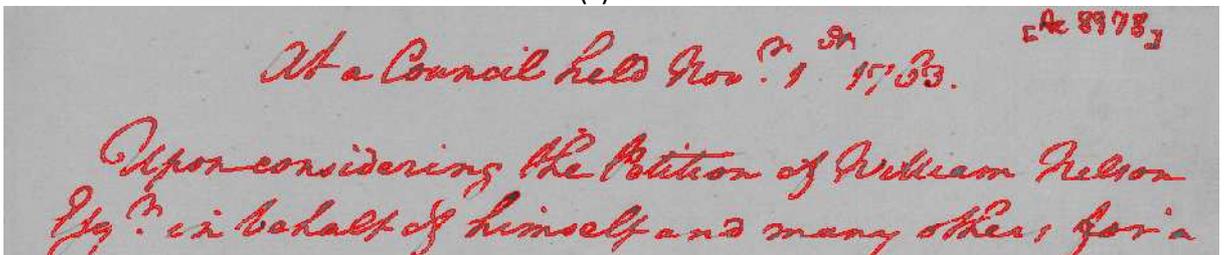
(a)



(b)



(c)



(d)

Figure 4.7. Résultat de la deuxième étape :

- (a) Image d'entrée ; (b) Image de contraste ; (c) Image des pixels de haut contraste ;
- (d) Contour final obtenu, en utilisant le modèle de contour actif proposé.

Selon la figure 4.7.d, nous pouvons voir que, à la fin de l'évolution du contour actif, les quelques pixels de l'arrière-plan détectés comme texte à la figure 4.7.c, sont éliminés grâce au terme F_A . Cependant, certains caractères, faiblement contrastés, ne sont pas correctement détectés. Dans la prochaine étape, nous proposons une étape d'optimisation de la binarisation, pour résoudre ce problème.

4.3.3. Etape 3: Optimisation de la binarisation

Le but de cette étape est de détecter les caractères du texte faiblement contrastés, qui ne sont pas détectés à l'étape précédente (Figure 4.8), et de produire la binarisation finale.

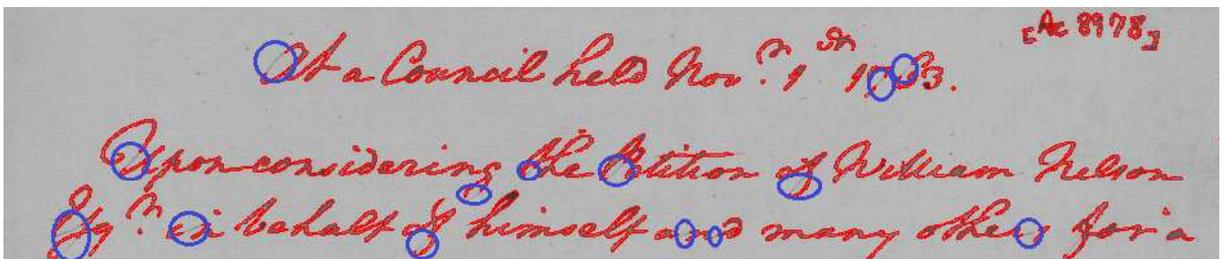


Figure 4.8. Les caractères du texte faiblement contrastés non détectés à l'étape 2 (entourés par des cercles bleus).

À cette fin, nous proposons d'utiliser l'évolution du contour actif, dans l'image du seuillage adaptatif moyen de [Fisher et al., 2003], correspondante à l'image d'entrée (Figure 4.9.a). Nous choisissons le seuillage adaptatif moyen, car il s'agit d'un bon détecteur de tous les objets, mal et bien contrastés, qui sont présents dans la même image du document. Par conséquent, il permet au contour actif de se propager dans les caractères du texte, et de détecter même les parties les plus faibles. Le seuillage adaptatif moyen a l'avantage de détecter tout le texte, mais il présente beaucoup de bruit de fond. Nous choisissons ensuite les valeurs des paramètres les plus adéquates, afin d'éviter l'émergence de nouveaux contours actifs, loin du contour initial, lors de l'évolution de la courbe.

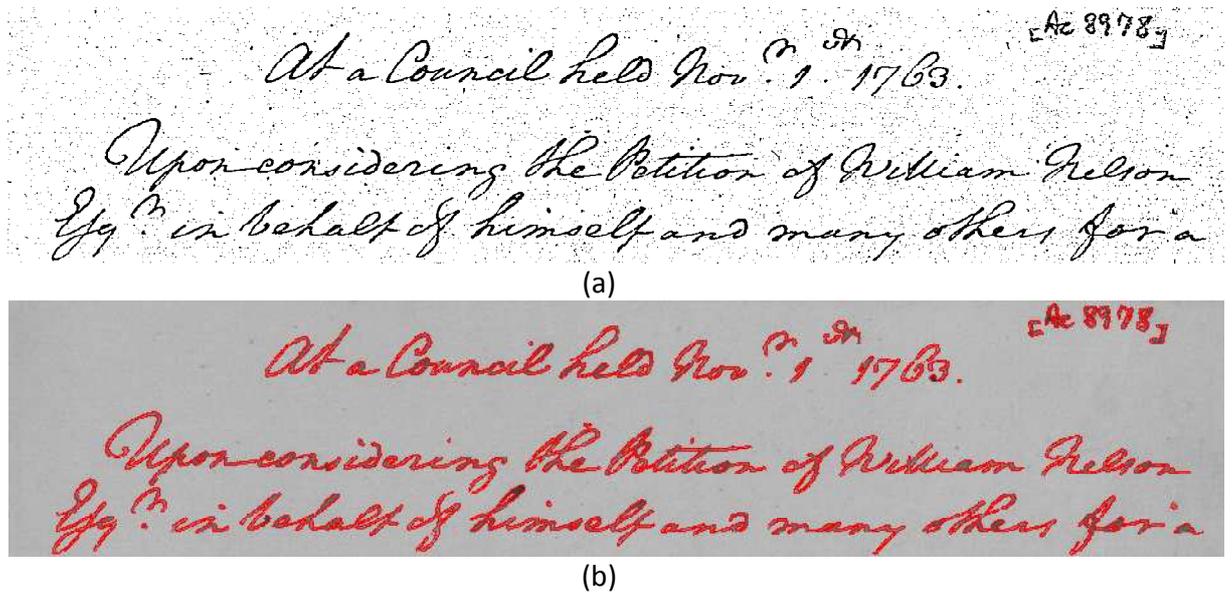


Figure 4.9. Le contour actif final du modèle proposé (lorsque le contour actif stagne) :
 (a) Image du seuillage adaptatif moyen de [Fisher et al., 2003] ;
 (b) Contour actif final.

D'après la figure 4.9.b, on peut voir que la plupart des caractères faiblement contrastés, sont bien détectés à la fin de cette étape. Lorsque le contour actif stagne, la binarisation finale (Figure 4.10) est obtenue par un simple seuillage de la fonction level set ϕ , les pixels dans l'image du document sont étiquetés avant-plan si $\phi > 0$ et arrière-plan sinon.

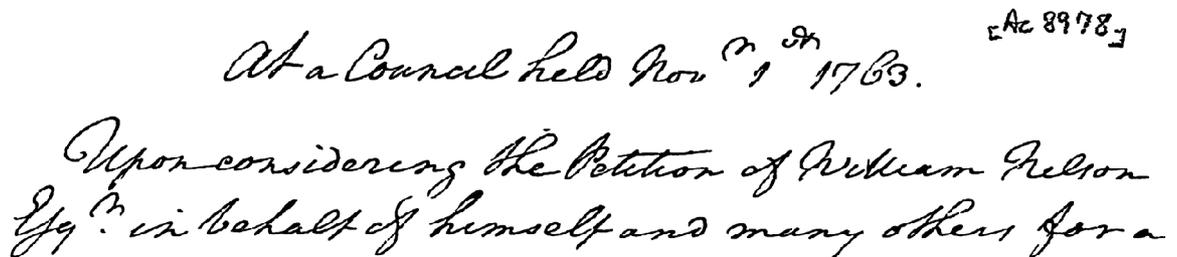


Figure 4.10. Résultat de la binarisation finale par la méthode proposée.

La méthode de binarisation proposée décrite ici a été mise en œuvre sur **Matlab** (voir Annexe 1). Les résultats sont présentés et discutés dans la section suivante.

4.4. Expérimentations, résultats et discussion

4.4.1. Bases de test

Deux ensembles d'images ont été utilisées pour l'évaluation des performances de la méthode proposée dans ce chapitre, comparativement aux méthodes de la littérature. Le premier est un ensemble publique, composé d'images de documents, provenant des

collections, proposées dans le cadre des compétitions DIBCO 2009², DIBCO 2011³ et H-DIBCO 2012⁴. Ces collections contiennent des images de documents réels (imprimées et manuscrites), avec les images de vérité terrain correspondantes. Les images de ces bases de données incluent différents types de dégradations, ce qui rend leur binarisation une tâche difficile. Des exemples d'images de ces collections sont montrés dans les figures suivantes.

Les résultats de la méthode proposée sont comparés à un certain nombre de techniques de binarisation connues, y compris la méthode d'Otsu [Otsu, 1979], de Sauvola [Sauvola et Pietikainen, 2000], de Gatos [Gatos et al., 2006], de Su [Su et al., 2010, Su et al., 2013], de Lu [Lu et al., 2010] (le gagnant de DIBCO'09) et celle de Rivest-Hénault [Rivest-Hénault et al., 2012]. Nous accordons une attention particulière à la technique de Rivest-Hénault dans nos comparaisons car elle est également basée sur les contours actifs.

Le deuxième ensemble de test est composé d'images de documents dégradés réelles en langue arabe, issues de la Bibliothèque Nationale d'Algérie⁵ et du Centre National des Manuscrits d'Adrar. Des exemples d'images de cet ensemble sont illustrés dans les prochaines figures.

Les résultats de la binarisation dans les prochaines figures montrent la performance supérieure de notre technique de binarisation, par rapport à d'autres techniques de l'état de l'art.

4.4.2. Sélection des paramètres

Les différents tests effectués sur un grand nombre d'images de documents ont démontré que les paramètres : $\sigma = 3$, $\Delta(t) = 0.1$, $\beta = 1$, $\nu = 0.001 \times 255 \times 255$, et $\varepsilon = 1$, donnent les meilleurs résultats de binarisation. Les coefficients λ_1 et λ_2 sont les poids sur les régions de l'extérieur, et de l'intérieur de (C) , respectivement. À la deuxième étape de la méthode de binarisation proposée, nous définissons les coefficients comme $\lambda_1 = \lambda_2 = 1$ (à l'équation (4.8)), ce qui conduit à une concurrence équitable entre les régions de l'intérieur et de l'extérieur du contour de niveau zéro pendant l'évolution. À la troisième étape, nous définissons $\lambda_1 = 1$, et $\lambda_2 = 2$ dans (4.8) ; notez que nous choisissons une valeur plus grande pour λ_2 pour éviter l'émergence de nouveaux contours, loin du contour initial. Les itérations sont terminées lorsque le contour actif stagne. Cette condition est définie formellement comme suit : si le nombre de pixels qui ont changé leur étiquette n'a pas augmenté après une période de N itérations, les itérations sont arrêtées.

L'influence du paramètre le plus important sur le comportement de notre méthode sera discutée dans la sous-section 4.4.5.

² <http://users.iit.demokritos.gr/~bgat/DIBCO2009/benchmark/>.

³ <http://utopia.duth.gr/~ipratika/DIBCO2011/benchmark/>

⁴ <http://utopia.duth.gr/~ipratika/HDIBCO2012/benchmark/>

⁵ http://www.patrimoineculturel.algerien.com/bibliotheque.php?id=1#top_title

4.4.3. Implémentation de la méthode proposée

Comme nous avons dit précédemment, nous avons utilisé deux ensembles d'images pour la validation de cette méthode. Le premier ensemble contient des images de documents réels, prises des collections publiques, utilisées dans le contexte des compétitions internationales de binarisation DIBCO 2009, H-DIBCO 2010, DIBCO 2011, H-DIBCO 2012 et DIBCO 2013. Le second ensemble contient des images de documents anciens arabes, issues de la Bibliothèque Nationale d'Algérie et du Centre National des Manuscrits d'Adrar.

4.4.3.1. Validation de notre méthode sur les bases de données des DIBCOs, et comparaison avec les méthodes de l'état de l'art

La figure 4.11.a illustre une image d'un document dégradé, avec des caractères faiblement contrastés de DIBCO'11. Notre méthode de binarisation construit d'abord une image de contraste, par l'image locale, maximale et minimale, puis extrait l'image à contraste élevé (Figure 4.11.b), qui est utilisée pour supprimer le bruit de fond et les artefacts, et générer la carte d'initialisation du contour actif. Une fonction level set \emptyset , est calculée à partir de l'image de contraste élevé obtenue, en utilisant un algorithme de transformation de distance. Ensuite, l'évolution itérative de notre modèle de contour actif commence. Les figures 4.11.c, 4.11.d, 4.11.e, 4.11.f, 4.11.g et 4.11.h illustrent le processus d'évolution de la courbe, à différentes étapes (itérations), du contour initial (Figure 4.11.c) au contour final (Figure 4.11.h). Comme on peut le voir sur la figure 4.11.i (résultat final), les caractères faiblement contrastés sont bien détectés à l'aide de la méthode proposée.

D'autre part, la méthode de Sauvola [Sauvola et Pietikainen, 2000] (Figure 4.11.j) produit beaucoup de bruit, en raison de la variation dans le fond du document, alors que la méthode de Su [Su et al., 2013] (Figure 4.11.l) n'arrive pas à détecter certains caractères à faible contraste, et la méthode de Howe [Howe, 2011] (Figure 4.11.k) introduit des bruits de fond.

Receipt from Miss Butterfull
for 2 beds, desks & benches.
March 24 - 1920.

(a)

Receipt from Miss Butterfull
for 2 beds, desks & benches.
March 24 - 1920.

(b)

Receipt from Miss Butterfull
for 2 beds, desks & benches.
March 24 - 1920.

(c)

Receipt from Miss Butterfull
for 2 beds, desks & benches.
March 24 - 1920.

(d)

Receipt from Miss Butterfull
for 2 beds, desks & benches.
March 24 - 1920.

(e)

Receipt from Miss Butterfull
for 2 beds, desks & benches.
March 24 - 1920.

(f)

Receipt from Miss Butterfull
for 2 beds, desks & benches.
March 24 - 1920.

(g)

Receipt from Miss Butterfull
for 2 beds, desks & benches.
March 24 - 1920.

(h)

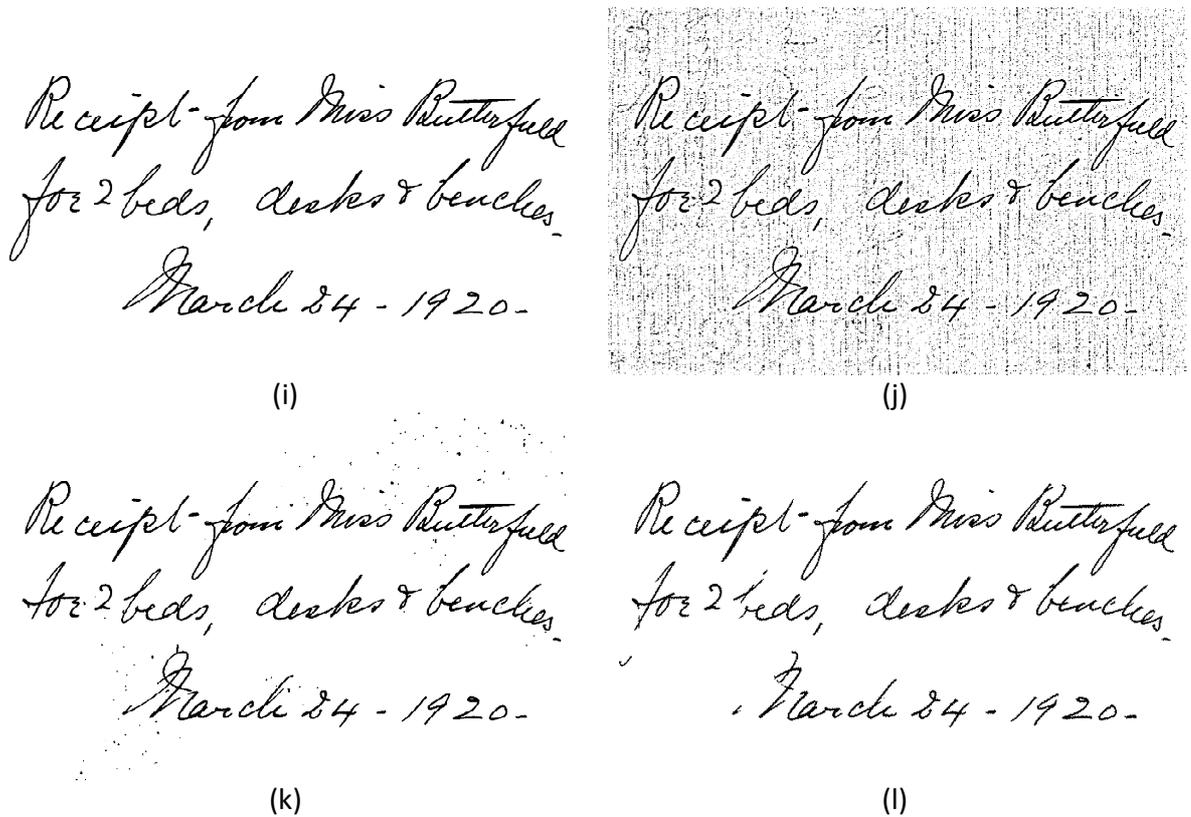


Figure 4.11. Résultats de la binarisation de l'image HW2 de la base de données de DIBCO'11 :

(a) Image à binariser ;

(b) Image de haut contraste ;

processus de l'évolution du contour actif à partir du contour initial (c) jusqu'au contour final

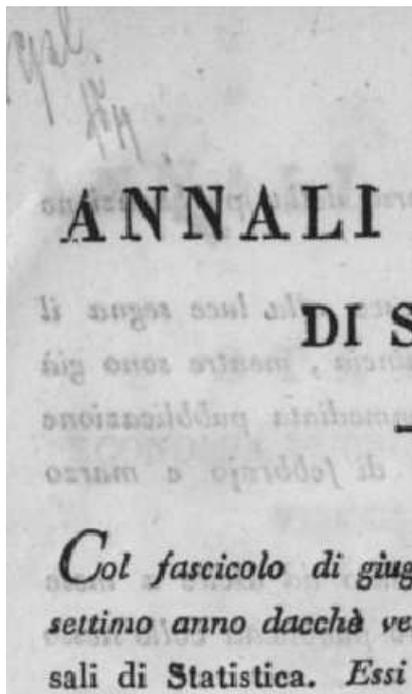
(h) ;

(i) Résultat final de la méthode proposée ;

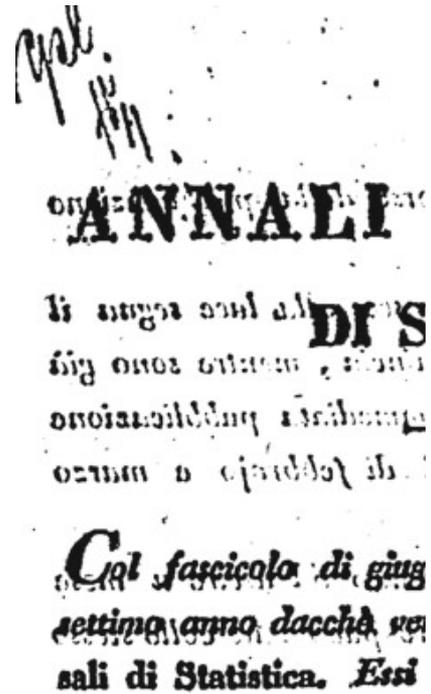
(j) Méthode de Sauvola [[Sauvola et Pietikainen, 2000](#)] ;

(k) Méthode de Howe [[Howe, 2011](#)] ;

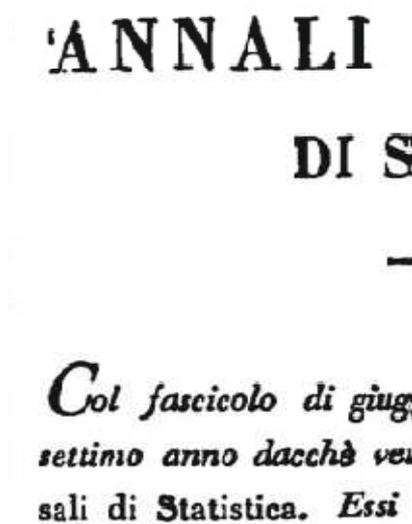
(l) Méthode de Su [[Su et al., 2013](#)].



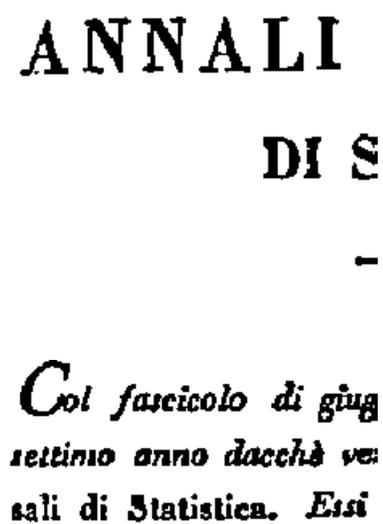
(a)



(b)



(c)



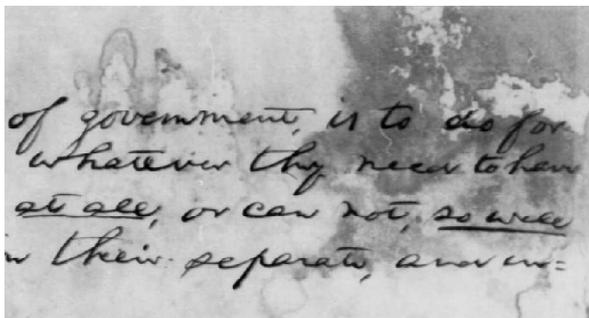
(d)

Figure 4.12. Résultats de la binarisation d'une image dégradée tirée de [Rivest-Hénault et al., 2012] :

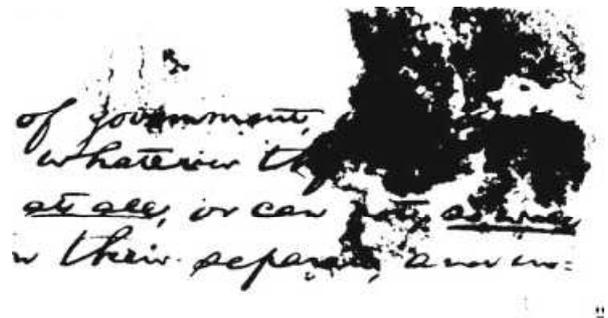
- (a) Image à binariser ;
- (b) Résultat de la binarisation par la méthode de Sauvola ;
- (c) Méthode de Rivest-Hénault [Rivest-Hénault et al., 2012] ;
- (d) Méthode proposée.

La figure 4.12.a illustre une image de document dégradé (où l'encre s'est infiltré d'un côté à l'autre et réduit la lisibilité), et la figure 4.13.a illustre une image de document dégradé de DIBCO'09, qui comprend une variation de contraste d'image et des taches. Notre méthode a extrait avec succès le texte des deux images de documents. D'après les figures 4.12 et 4.13, la méthode proposée est moins sensible au fort bruit de fond, et produit de meilleurs résultats de binarisation que d'autres approches de binarisation d'images de documents antérieures, car elle supprime le contraste du fond du document, par la normalisation, dans l'étape d'initialisation. En revanche, les méthodes d'Otsu [Otsu, 1979], de Sauvola [Sauvola et Pietikainen, 2000], de Gatos [Gatos et al., 2006] et de Rivest-Hénault [Rivest-Hénault et al., 2012] classent certains pixels d'arrière-plan sombres comme des pixels de texte, de façon incorrecte.

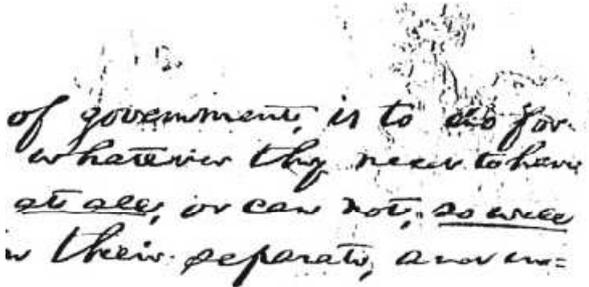
Dans la figure 4.14, la méthode de Rivest-Hénault, qui est également basée sur les contours actifs, classe plus de pixels sombres de l'arrière-plan comme des pixels de texte, de manière inappropriée, que notre méthode.



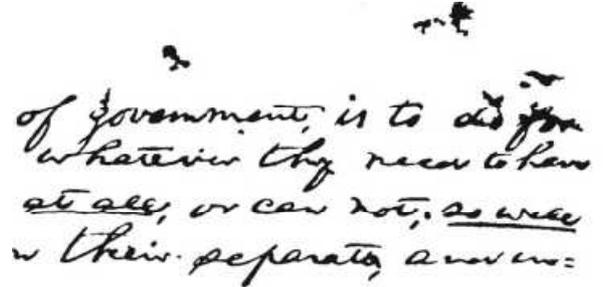
(a)



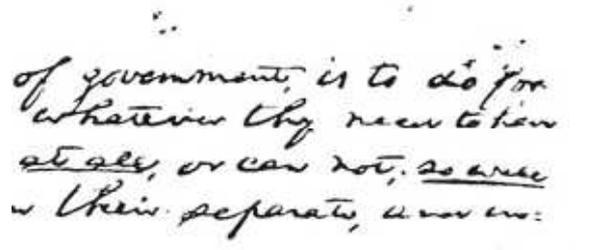
(b)



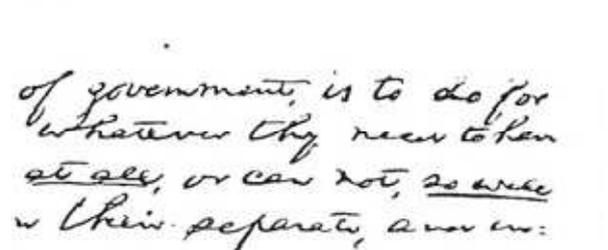
(c)



(d)



(e)



(f)

of government, is to do for
whatever they need to have
at all, or can not, so well
in their separate, answers

of government, is to do for
whatever they need to have
at all, or can not, so well
in their separate, answers:

(g)

(h)

Figure 4.13. Résultats de la binarisation de l'image H04 de la base de données de DIBCO'09 :

(a) Image à binariser ;

(b) Résultat de la méthode d'Otsu [Otsu, 1979] ;

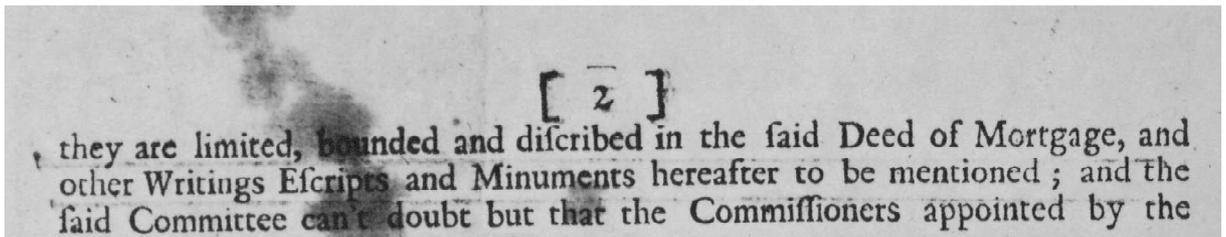
(c) Méthode de Sauvola [Sauvola et Pietikainen, 2000] ;

(d) Méthode de Rivest-Hénault [Rivest-Hénault et al., 2012] ;

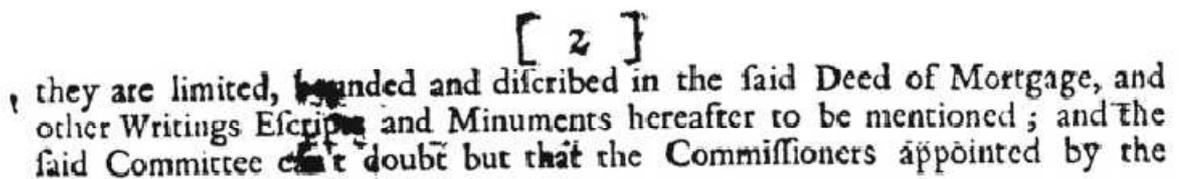
(e) Méthode de Gatos [Gatos et al., 2006] ;

(f) Méthode de Su [Su et al., 2010] ;

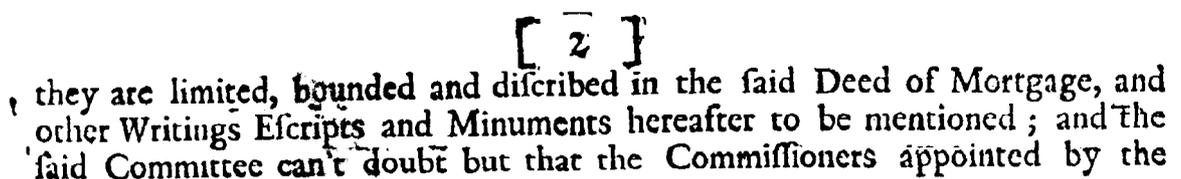
(g) Méthode de Lu [Lu et al., 2010] ; (h) Méthode proposée.



(a)



(b)



(c)

Figure 4.14. Résultats de la binarisation de l'image P04 de la base de données de DIBCO'09 :

(a) Image à binariser ;

(b) Résultat de la binarisation par la méthode de Rivest-Hénault [Rivest-Hénault et al., 2012] ;

(c) Méthode proposée.

Certains résultats de la méthode proposée sont présentés dans les figures précédentes, avec les résultats obtenus en utilisant d'autres méthodes. On peut noter que la méthode proposée conduit à des résultats plus cohérents, avec des contours plus continus. Elle extrait le texte correctement des images de documents, qui souffrent de différents types de dégradations. D'autre part, les autres méthodes produisent souvent une certaine quantité de bruit, en raison de la variation dans l'arrière-plan du document, ou ne peuvent pas détecter certains caractères de texte, faiblement contrastés.

4.4.3.2. Validation de la méthode proposée sur des images issues de la Bibliothèque Nationale d'Algérie et du Centre National des Manuscrits d'Adrar

La validité de la méthode proposée est aussi montrée sur des images de documents anciens arabes, issues de la Bibliothèque Nationale d'Algérie et du Centre National des Manuscrits d'Adrar. Les figures 4.15 et 4.17 illustrent les étiquettes des premières pages des livres, utilisés pour les tests.

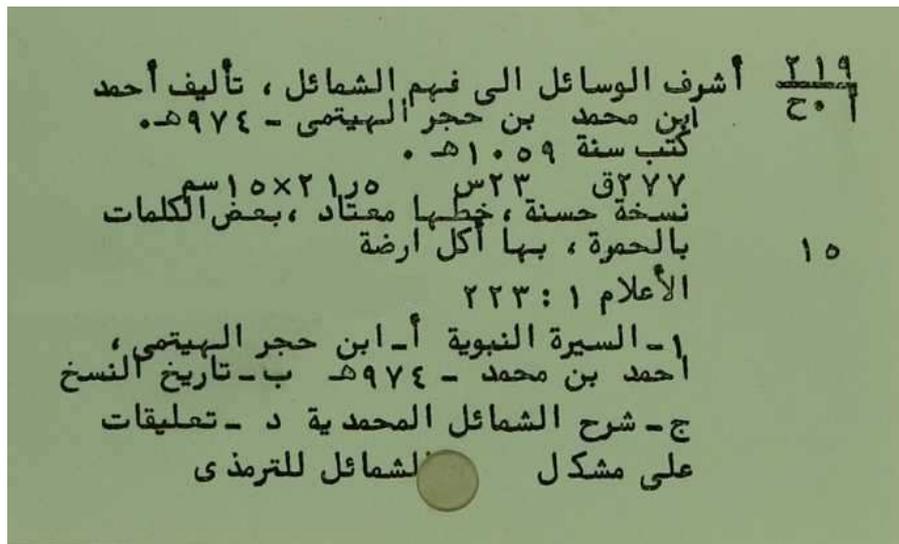


Figure 4.15. Etiquette de la première page du livre (أشرف الوسائل الى فهم الشمائل).

Les résultats de restauration de quelques pages, de dimension 2160×1440 pixels, de ces livres, sont présentés dans les figures 4.16 et 4.18. Théoriquement, La restauration des images de documents anciens a pour objectif de retrouver la qualité de l'image originale, avant sa dégradation physique ou numérique, mais pratiquement, les méthodes de restauration proposées dans la littérature, visent à séparer le texte du fond bruité et irrégulier, puis à restaurer le fond, par un lissage avec la couleur la plus dominante. La restauration présentée ici vise à séparer le texte du fond bruité par la méthode proposée dans ce chapitre, après avoir sélectionné les deux pages de l'image, puis à restaurer le fond par un lissage avec la couleur la plus dominante.



(a) Image de document à restaurer



(b) Image restaurée correspondante



(c) Image de document à restaurer



(d) Image restaurée correspondante



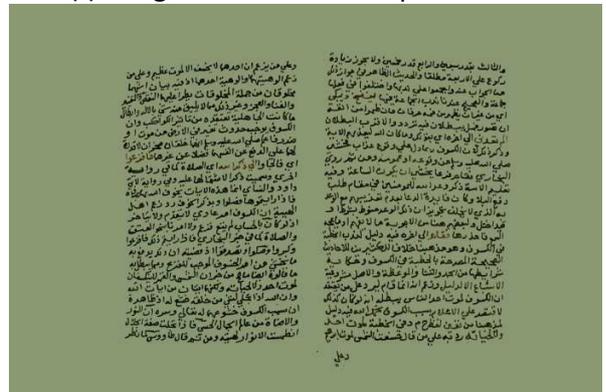
(e) Image de document à restaurer



(f) Image restaurée correspondante



(g) Image de document à restaurer



(h) Image restaurée correspondante

(g) Image de document à restaurer



(h) Image restaurée correspondante



(i) Image de document à restaurer



(j) Image restaurée correspondante



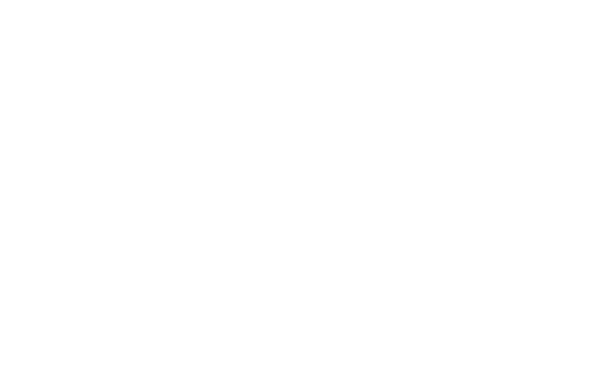
(k) Image de document à restaurer



(l) Image restaurée correspondante



(m) Image de document à restaurer



(n) Image restaurée correspondante





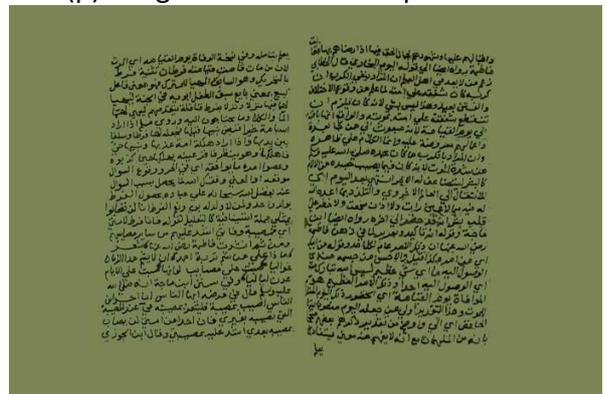
(o) Image de document à restaurer



(p) Image restaurée correspondante



(q) Image de document à restaurer



(r) Image restaurée correspondante

Figure 4.16. Résultats de la restauration de quelques images du livre (أشرف الوسائل الى فهم) أشرف (الشمائل) par notre méthode.

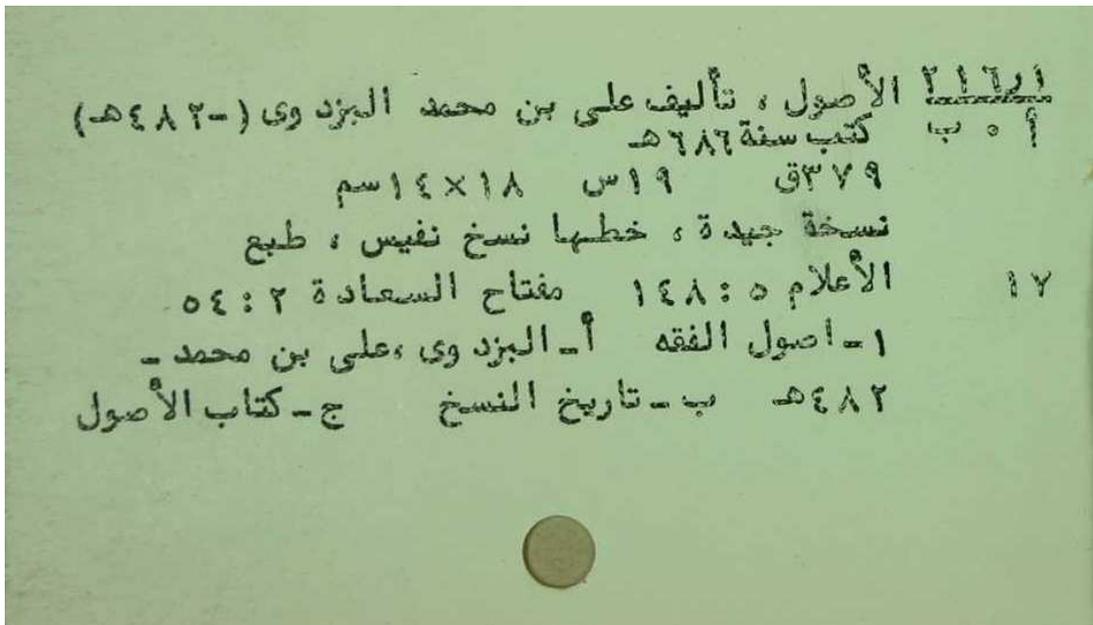


Figure 4.17. Etiquette de la première page du livre (الأصول).



(a) Image de document à restaurer



(b) Image restaurée correspondante



(c) Image de document à restaurer



(d) Image restaurée correspondante

Figure 4.18. Résultats de la restauration de quelques images du livre (الأصول) par notre méthode.

Plusieurs expérimentations ont été effectuées, afin d'évaluer les performances de la méthode proposée, et les résultats obtenus sont, à notre avis, encourageants. En effet, comme il est confirmé par les résultats expérimentaux, la méthode proposée a présenté un comportement fiable, pour la tâche complexe de séparation, entre l'avant-plan et arrière-plan, à partir des images de documents significativement dégradés.

4.4.4. Mesures de performance

Les mesures d'évaluation utilisées dans ce chapitre sont celles utilisées dans toutes les éditions des compétitions DIBCO/ H-DIBCO, et qui sont adoptées dans le rapport de DIBCO'09 [Gatos et al., 2009], y compris $F - measure$ (%), rapport signal maximum sur bruit 'Peak Signal to Noise Ratio' (PSNR), métrique de taux négatif 'Negative Rate Metric' (NRM) et Métrique de pénalité de fausse-classification 'Misclassification Penalty Metric' (MPM).

Ces mesures estiment l'efficacité des algorithmes de binarisation au niveau pixel. Elles sont calculées, en comparant le résultat de la binarisation, avec l'image de vérité terrain. Dans ce qui suit, nous notons I_{GT} l'image de vérité terrain, et I_b l'image résultante de la binarisation.

On note aussi TP , TN , FP , et FN le nombre de : vrais positifs, vrais négatifs, faux positifs, et faux négatifs, respectivement définis comme suit (Figure 4.19) :

- Un pixel noir est classé comme vrai positif, s'il est présent dans les deux images I_b et I_{GT} .
- Un pixel est classé comme vrai négatif, s'il n'est pas présent dans les deux images.
- Un pixel est classé comme faux positif, s'il est présent seulement dans I_b .
- Un pixel est classé comme faux négatif, s'il est présent seulement dans I_{GT} .

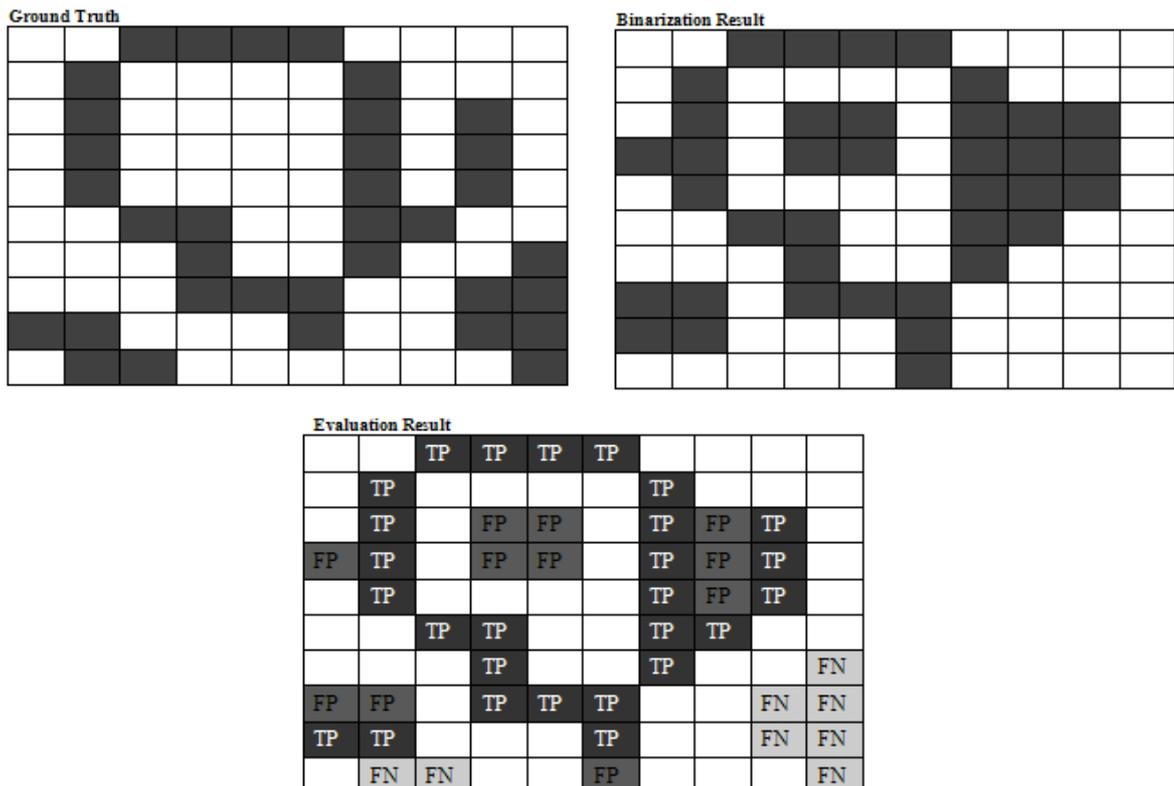


Figure 4.19. Exemple d'évaluation d'un résultat.

A partir de ces quatre valeurs, plusieurs mesures ont été définies :

- Le **F – measure** (%) évalue comment un algorithme peut récupérer les pixels désirés ;
- Le **PSNR** mesure la proximité de l'image résultante par rapport à l'image de vérité terrain ;
- Le **NRM** mesure le taux d'incompatibilité des pixels entre l'image de vérité terrain et l'image résultante ;
- Le **MPM** mesure à quel point l'image résultante représente le contour de l'image de vérité terrain.

Les formules qui permettent le calcul de ces mesures sont décrites par les équations (4.13), (4.14), (4.15) et (4.16) dans les paragraphes suivants.

4.4.4.1. *F – measure* 'F – measure' (FM)

F – measure a été introduit, la première fois par Chinchor [Chinchor, 1992]. Du fait de sa simplicité, *F – measure* est considérée comme l'une des mesures les plus utilisées pour l'évaluation quantitative de la binarisation. Elle a été utilisée dans toutes les éditions précédentes des compétitions DIBCO/ H-DIBCO. Ainsi, *F – measure* est défini par :

$$F - measure = \frac{2 \times Recall \times Precision}{Recall + Precision} \quad (4.13)$$

Avec :

$$Recall = \frac{TP}{TP + FN} \quad Precision = \frac{TP}{TP + FP}$$

4.4.4.2. Rapport signal maximum sur bruit (*PSNR*)

PSNR est une mesure de similarité entre deux images. Cependant, plus la valeur de *PSNR* est grande, plus la similarité entre les deux images est grande. Il est défini comme le rapport du signal maximum, sur le bruit moyen carré [Damera-Venkata et al., 2000]. *PSNR* a été utilisée dans toutes les éditions des compétitions DIBCO/ H-DIBCO. Pour une image, de taille $M \times N$, elle est donnée par :

$$PSNR = 10 \log_{10} \left(\frac{C^2 \times M \times N}{\sum_{x=1}^M \sum_{y=1}^N [I_{GT}(x, y) - I_b(x, y)]^2} \right) \quad (4.14)$$

C est la différence entre l'avant plan et l'arrière plan (ici 255).

4.4.4.3. Métrique de taux négatif (*NRM*)

NRM est l'une des mesures, utilisées dans toutes les éditions de DIBCO/ H-DIBCO. *NRM* est basée sur les disparités, entre les pixels de l'image de vérité terrain et les pixels de l'image binarisée [Aguilera et al., 2005]. Elle combine le taux de faux négatifs (NR_{FN}) et le taux de faux positifs (NR_{FP}) comme suit:

$$NRM = \frac{NR_{FN} + NR_{FP}}{2} \quad (4.15)$$

Avec :

$$NR_{FN} = \frac{FN}{FN + TP} \quad NR_{FP} = \frac{FP}{FP + TN}$$

La meilleure qualité de binarisation est obtenue avec une valeur basse de *NRM*.

4.4.4. Métrique de pénalité de fausse-classification (*MPM*)

La métrique de pénalité de mauvaise-classification évalue le résultat de la binarisation, par rapport à la vérité terrain, sur une base objet par objet [Aguilera et al., 2005]. Les pixels mal classés sont pénalisés par leur distance à la frontière de l'objet de vérité terrain. *MPM* a été aussi utilisée dans les compétitions DIBCO/ H-DIBCO.

$$MPM = \frac{MP_{FN} + MP_{FP}}{2} \quad (4.16)$$

Avec :

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

d_{FN}^i et d_{FP}^j désignent la distance du $i^{\text{ème}}$ faux négatif, et du $j^{\text{ème}}$ faux positif pixel, au contour de l'image de vérité terrain. Le facteur de normalisation D est la somme de toutes les distances pixels-à-contour de l'objet de vérité terrain. Une valeur basse de *MPM* indique que l'algorithme est efficace dans l'identification des frontières des objets.

Notez que toutes les mesures de performance citées précédemment sont implémentées sous **Matlab** par Moghaddam et sont disponibles sur le lien suivant :

<http://www.mathworks.com/matlabcentral/fileexchange/27652-objective-evaluation-of-binarization-methods-for-document-images>

Il existe également d'autres mesures d'évaluation qui sont citées dans [Kefali, 2015].

Tableau 4.1. Résultats d'évaluation sur la base de données de DIBCO 2009.

Méthode	<i>F – measure</i> (%)	<i>PSNR</i>	<i>NRM</i> ($\times 10^{-2}$)	<i>MPM</i> ($\times 10^{-3}$)
Lu ⁶	91.24	18.66	4.31	0.55
Fabrizio ⁷	90.06	18.23	4.75	0.89
Rivest-Hénault ⁸	89.34	17.79	5.32	1.90
Proposée	92.13	18.87	4.54	0.57

Les performances de toutes les méthodes du tableau 4.1 sont rapportées à partir de [Gatos et al., 2009].

⁶ classée 1^{ère} à DIBCO'09

⁷ classée 2^{ème} à DIBCO'09

⁸ classée 3^{ème} à DIBCO'09

Tableau 4.2. Résultats d'évaluation sur la base de données de H-DIBCO 2012.

Méthode	<i>F – measure</i> (%)	<i>PSNR</i>	<i>DRD</i>
Otsu	80.18	15.03	26.45
Sauvola	82.89	16.71	6.59
Howe ⁹	89.47	21.80	3.44
Lelore ¹⁰	92.85	20.57	2.66
Su ¹¹	91.54	20.14	3.04
Proposée	91.67	19.96	2.76

Les performances de toutes les méthodes du tableau 4.2 sont rapportées à partir de [Pratikakis et al., 2012].

Des résultats expérimentaux sont présentés aux tableaux 4.1 et 4.2. La méthode proposée a été testée sur deux bases de données publiques, qui ont été utilisés dans la compétition de binarisation des images de documents (DIBCO'09), et la compétition de binarisation des images de documents manuscrits (H-DIBCO'12). Elle réalise des précisions de **92.13%** et de **91.67%**, respectivement, et sont significativement plus élevées ou proches de celles des méthodes les plus performantes, rapportées dans les deux compétitions.

Par rapport à d'autres méthodes, notre méthode se comporte mieux, en terme de *PSNR*, de *NRM* et de *MPM*, ce qui signifie qu'elle produit une précision plus élevée, et préserve mieux les traits du texte. Notez que, dans l'évaluation de notre méthode, nous avons suivi le même protocole de DIBCO, pour une comparaison équitable avec les méthodes participantes à DIBCO.

Malgré les bons résultats présentés dans ces expériences, il est important de noter qu'une bonne initialisation peut être décisive dans de nombreux cas. Par exemple, lorsque l'image à binariser contient de grandes zones de taches, l'image de contraste peut ne pas fournir suffisamment d'informations, pour une binarisation correcte de l'image du document. En outre, lorsque le texte de l'arrière plan est aussi sombre, ou même plus foncé, que le texte de l'avant plan, la méthode proposée ne peut pas classer correctement les textes des deux cotés. Nous laissons cette amélioration aux travaux futurs.

4.4.5. Influence de la valeur du paramètre d'échelle σ

Il est nécessaire d'examiner l'influence du paramètre d'échelle σ (4.7), sur les résultats de binarisation de la méthode proposée, bien que le même paramètre d'échelle ait été utilisé pour toutes les images étudiées précédemment dans ce chapitre. Pour cela, nous appliquons notre méthode, en utilisant différents paramètres d'échelle, **0.5**, **1**, **3** et **5**, pour la même image de document, utilisée dans la figure 4.13.

⁹ classée 1^{ère} à DIBCO'12

¹⁰ classée 2^{ème} à H-DIBCO'12

¹¹ classée 3^{ème} à H-DIBCO'12

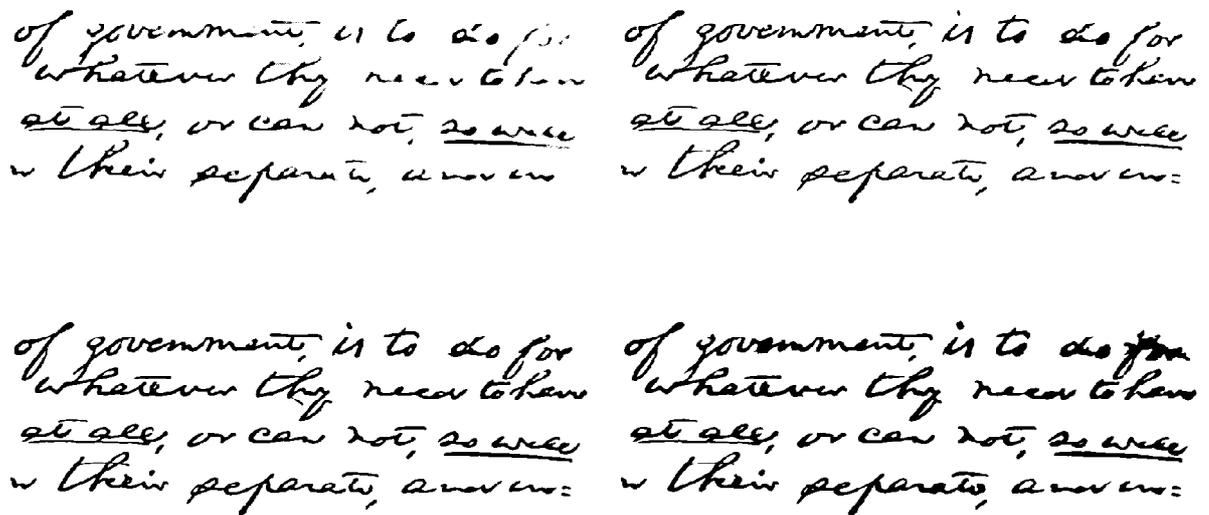


Figure 4.20. Résultats de la binarisation de notre méthode avec le paramètre d'échelle $\sigma = 0.5, 1, 3$ et 5 (de gauche à droite et de haut en bas).

Les résultats de différents paramètres d'échelle σ , sur la figure 4.20, montrent que la binarisation donne des contours épais qui se chevauchent parfois, avec une grande valeur de σ , et des contours minces, brisés et peu clairs, avec une petite valeur de σ . La valeur du paramètre σ définit la qualité de la binarisation, comme on peut le voir sur la figure 4.20. Les différents tests effectués sur un grand nombre d'images de documents ont démontré qu'avec $\sigma = 3$, nous obtenons le résultat de binarisation le plus précis, et les caractères faiblement contrastés sont raisonnablement bien détectés.

4.4.6. Matériel utilisé, coûts de calculs et complexité de la méthode proposée

Malgré les nombreux avantages des modèles de contour actif, par rapport aux méthodes classiques de segmentation d'images, ils sont plus coûteux en termes de calcul.

Le tableau 4.3 montre les résultats des temps de calcul obtenus, dans nos expériences, avec un code **Matlab** exécuté sur un Intel (R) Core (TM) i5-2400, 3.10 GHz, 4,00 Go RAM, avec Matlab R2008a.

Selon le tableau 4.3, lorsque la taille de l'image est trop grande, la méthode proposée, basée sur les contours actifs, est plus coûteuse, en termes de complexité de calcul, par rapport aux méthodes de seuillage, comme celle de Sauvola. Avec la méthode de Rivest-Hénault, le coût de calcul est encore plus élevé ; Par exemple, le traitement d'une image de taille 512×512 pixels prend 626 s [Rivest-Hénault et al., 2012].

Tableau 4.3. Comparaison des temps de calcul de l'implémentation.

Taille de l'image (pixels x pixels)	Méthode proposée	Méthode de Sauvola
2025 x 426 (image H01 de DIBCO'2009)	206 s	0.02 s
1218 x 781 (image HW2 de DIBCO'2011)	203 s	0.02 s
1493 x 1613 (image HW5 de DIBCO'2013)	445 s	0.04 s

4.5. Conclusion

Dans ce chapitre, nous avons présenté notre première contribution à la binarisation d'images de documents historiques, connues par leur mauvaise qualité. Dans cette contribution, nous avons proposé une nouvelle méthode de binarisation, basée sur les contours actifs, et le contraste de l'image, évalué en fonction du maximum et du minimum locaux.

Une telle méthode combinée conduit, comme montré dans les expériences, à une grande précision, lorsqu'elle est appliquée aux images de documents historiques, avec une variété de dégradations, telles que la présence de caractères faiblement contrastés '**faint characters**', et d'illumination non uniforme '**uneven illumination**'. Elle surpasse plusieurs autres méthodes de binarisation de documents, en termes de ***F – measure*** (%), ***PSNR***, ***NRM*** et ***MPM***. La méthode proposée réussit effectivement à capturer des caractères, de texte faiblement et correctement contrastés dans la même image de document.

La méthode proposée dans ce chapitre traite l'image de document indépendamment de sa langue et son type d'écriture, comme il est montré dans les expériences, elle traite des images de manuscrits avec des langues et des types d'écriture différents : l'anglais et le grec pour les bases de données DIBCOs et HDIBCOs, et l'arabe pour les images de manuscrits du patrimoine.

Cependant, la méthode de binarisation proposée est coûteuse en termes de temps de calcul, et sa performance est limitée, lorsque l'image à binariser est fortement dégradée, par l'effet de transparence du verso '**bleed-through**' ou des taches '**smear**'.

En tant que perspective pour les travaux futurs, une amélioration sur les forces motrices (termes énergétiques), pour rendre la méthode plus adaptable aux fortes dégradations sera considérée.

Le chapitre suivant décrit notre deuxième contribution, à savoir une méthode de binarisation d'images de documents dégradés.

CHAPITRE 5

Contributions à la binarisation d'images de documents dégradés par l'algorithme ISauvola 'Improved Sauvola'

5.1. Introduction

La problématique liée au prétraitement des documents anciens reste encore d'actualité. Des approches multiples ont été proposées pour la binarisation des images de documents anciens. Cependant, la présence simultanée, dans un même document, de deux ou plusieurs types de dégradations, laisse le problème encore ouvert.

Ce chapitre présente notre deuxième contribution, qui est une nouvelle méthode de binarisation d'images de documents anciens. La méthode est proposée dans le but de parvenir à une binarisation simple, rapide et efficace en même temps. La méthode proposée ISauvola, pour « Improved Sauvola », est une amélioration de la méthode de binarisation de Sauvola, qui est largement utilisée en pratique.

La méthode de Sauvola donne de bons résultats sur des images de documents classiques, cependant :

- Elle sur-segmente les défauts et la texture du papier, sur l'arrière plan de l'image ;
- Elle a tendance à ne pas segmenter les parties intérieures des caractères épais, puisque la taille de la fenêtre d'analyse étant fixe durant tout le traitement, ce qui rend difficile le traitement des documents dont la taille des caractères est très variable.
- ...

Au cours de ce chapitre, nous rappelons tout d'abord le principe de l'algorithme de Sauvola, tout en mentionnant ses limites. Nous présentons par la suite les différentes étapes de la méthode proposée, et nous exposons quelques résultats obtenus, en les comparant à ceux des méthodes ultérieures.

5.2. Motivations et objectifs

Au lieu de concevoir une nouvelle méthode de binarisation, nous essayons, au cours de ce chapitre, d'apporter des améliorations sur la méthode de binarisation existante de Sauvola, dont le but est d'améliorer, non seulement la performance de la méthode originale de Sauvola, mais aussi la robustesse sur différents types d'images de documents dégradés.

La méthode est proposée dans le but de parvenir à une technique de binarisation rapide et efficace en même temps.

5.3. Algorithme de Sauvola et ses limites

5.3.1. Algorithme de Sauvola

La méthode de Sauvola [Sauvola et Pietikainen, 2000] utilise une image en niveaux de gris en entrée. Comme la plupart des images de documents sont des images en couleurs, la conversion en images en niveaux de gris est nécessaire [Moghaddam et Cheriet, 2010].

Pour la binarisation d'une image en niveaux de gris I , la méthode de Sauvola traite chaque pixel de l'image localement et indépendamment du reste des pixels ; elle calcule un seuil T pour chaque pixel, en fonction des informations contenues dans son voisinage local (Figure 5.1). L'idée de la méthode consiste à faire varier le seuil, en fonction des valeurs de la moyenne locale et de l'écart type local, dans une fenêtre centrée sur chaque pixel, suivant la formule suivante:

$$T = m \times \left[1 + k \times \left(\frac{s}{R} - 1 \right) \right] \quad (5.1)$$

où k est un paramètre positif, défini par l'utilisateur, m et s sont respectivement la moyenne et l'écart type locaux, calculés dans une fenêtre, de taille ω , centrée sur le pixel courant et R est la plage dynamique de l'écart-type ($R = 128$ pour images de niveaux de gris à 8 bits). La taille de la fenêtre utilisée pour calculer m et s reste à définir par l'utilisateur, dans l'article d'origine [Sauvola et Pietikainen, 2000].

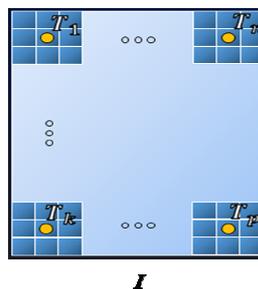


Figure 5.1. Principe de la méthode de Sauvola.

Les principaux avantages de la méthode de Sauvola sont :

- Elle fonctionne relativement bien sur des images de documents bruitées et floues [Sezgin et Sankur, 2004].
- Elle est rapide.
- ...

D'après l'équation de binarisation (5.1), l'utilisateur doit fournir deux paramètres (ω ; k). Certaines techniques ont été proposées pour les estimer :

- ✓ Sezgin et Sankur [Sezgin et Sankur, 2004] et Sauvola et Pietikainen [Sauvola et Pietikainen, 2000] ont utilisé $\omega = 15$ et $k = 0.5$.

- ✓ Badekas et Papamarkos [[Badekas et Papamarkos, 2005](#)] indiquent que $\omega = 14$ et $k = 0.34$ est le meilleur compromis, entre la suppression du bruit de l'effet de transparence et la qualité de l'extraction des objets dans les documents classiques.
- ✓ Rangoni et al. [[Rangoni et al., 2009](#)] ont basé la recherche des paramètres optimaux, sur la qualité des résultats de reconnaissance optique de caractères (OCR), et ont trouvé $\omega = 60$ et $k = 0.4$.

L'ajustement de ces paramètres nécessite généralement une connaissance à priori de l'ensemble des documents, pour obtenir les meilleurs résultats. Par conséquent, il n'y a pas de consensus dans la communauté de recherche, en ce qui concerne la valeur de ces paramètres.

5.3.2. Limites de la méthode de Sauvola

La méthode de Sauvola souffre de différentes limites qui sont présentées dans les paragraphes suivants [[Lazzara et Géraud, 2014](#)].

5.3.2.1. Mauvaise binarisation des objets faiblement contrastés

Les objets faiblement contrastés peuvent être considérés comme texture du fond, ou comme bruit de l'effet de transparence, en raison de la formule du seuil (5.1), et peuvent être supprimés ou détectés partiellement. La figure 5.2 illustre ce problème ; en effet, la région d'intérêt considérée montre les valeurs prises en compte dans une fenêtre, de taille $\omega = 51$, centrée au point central représenté en vert, où les contrastes sont très faibles. L'histogramme correspondant illustre la sensibilité de la méthode de Sauvola, à la valeur du paramètre k . Les pixels de l'objet ne peuvent pas être détectés correctement, si k est supérieur à 0.034 . Une faible valeur de ce paramètre peut aider à récupérer des objets à faible contraste, mais comme il est défini pour l'ensemble du document, il modifie également d'autres parties du résultat : les objets correctement contrastés sont plus épais dans ce cas, ce qui peut entraîner des connexions involontaires entre les composants. Dans ce cas, le bruit de fond, qui est généralement mal contrasté, sera détecté comme objet.

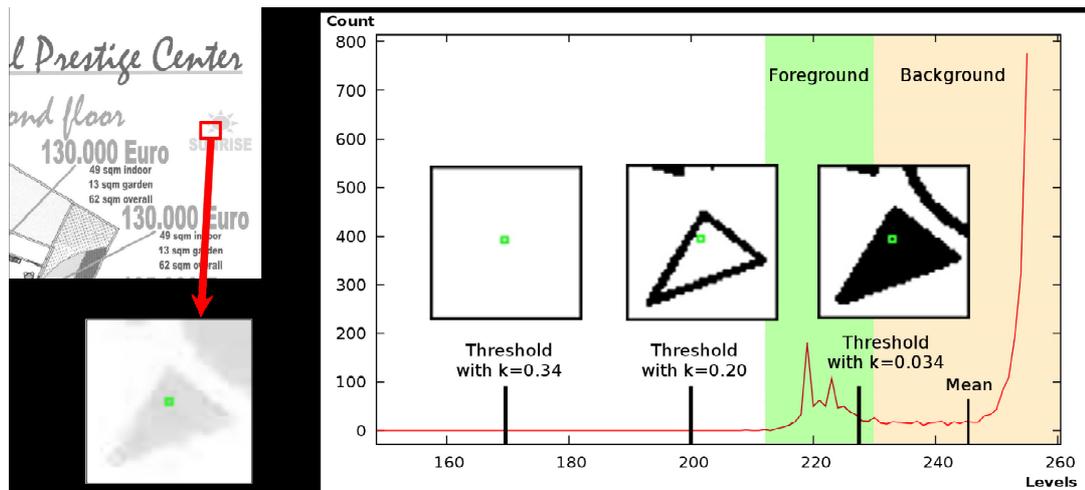


Figure 5.2. Influence du paramètre k sur la binarisation des objets faiblement contrastés.

5.3.2.2. Mauvaise binarisation des documents, dont la taille des objets est très variable

En présence d'objets de taille très variable dans l'image à binariser, la méthode de Sauvola ne pourra pas détecter tous les objets correctement. Dans la plupart des cas, une petite fenêtre ω peut être utilisée, pour détecter le texte dans une image de document ; le petit texte sera détecté parfaitement dans ce cas, mais le grands texte ne sera pas ou sera détecté partiellement. La figure 5.3.h illustre ce qui se produit, lorsque la fenêtre sélectionnée est trop petite, par rapport aux objets du document. A l'intérieur des grands caractères, les valeurs des pixels sont localement identiques, donc, si la fenêtre sélectionnée est trop petite, les statistiques à l'intérieur des grands caractères se comportent comme en arrière-plan, car la formule de Sauvola (5.1) repose sur le fait qu'il existe un minimum de contraste dans la fenêtre, pour définir un pixel du premier plan. Au même temps, si une grande fenêtre est sélectionnée pour le traitement, les grands caractères seront bien détectés, mais les petits seront détectés en gras.

Des travaux ont été proposés récemment pour améliorer la performance de la méthode de Sauvola, et palier à ce problème, comme :

- ✓ Lazzara et Géraud ont décrit dans [Lazzara et Géraud, 2014] une implémentation multi-échelle efficace de l'algorithme de Sauvola, afin de garantir une bonne binarisation pour les petits objets et même les grands, dans le même document, sans ajuster manuellement la taille de la fenêtre du traitement au contenu.
- ✓ Hadjadj et Meziane proposent dans [Hadjadj et Meziane, 2017] une amélioration de la méthode de Sauvola. La méthode proposée utilise les contours actifs, pour détecter les parties intérieures des grands caractères, qui sont mal détectées par la méthode de Sauvola, et réduire le bruit de fond détecté par cette même méthode, pour garantir une bonne binarisation, pour les objets de petite taille, ainsi que ceux de grande taille, dans la même image de document.

5.3.2.3. Mauvaise binarisation du texte texturé

Les textures sont vraiment sensibles à la taille de la fenêtre ω . Les figures 5.3.a et 5.3.d montrent un texte texturé et non texturé, respectivement avec la même taille de police. Bien que le texte texturé soit gras, les parties internes des caractères sont manquantes, après la binarisation (voir la figure 5.3.b). Sur la figure 5.3.e, le texte est toujours bien conservé, et adapté pour l'OCR. Dans la figure 5.3.c, l'utilisation d'une fenêtre plus grande a amélioré le résultat de la binarisation sur le texte texturé. Cependant, cette solution ne peut pas être appliquée, si le texte texturé est mélangé avec du texte brut, puisque, comme le montre la figure 5.3.f, le texte détecté sera en gras.

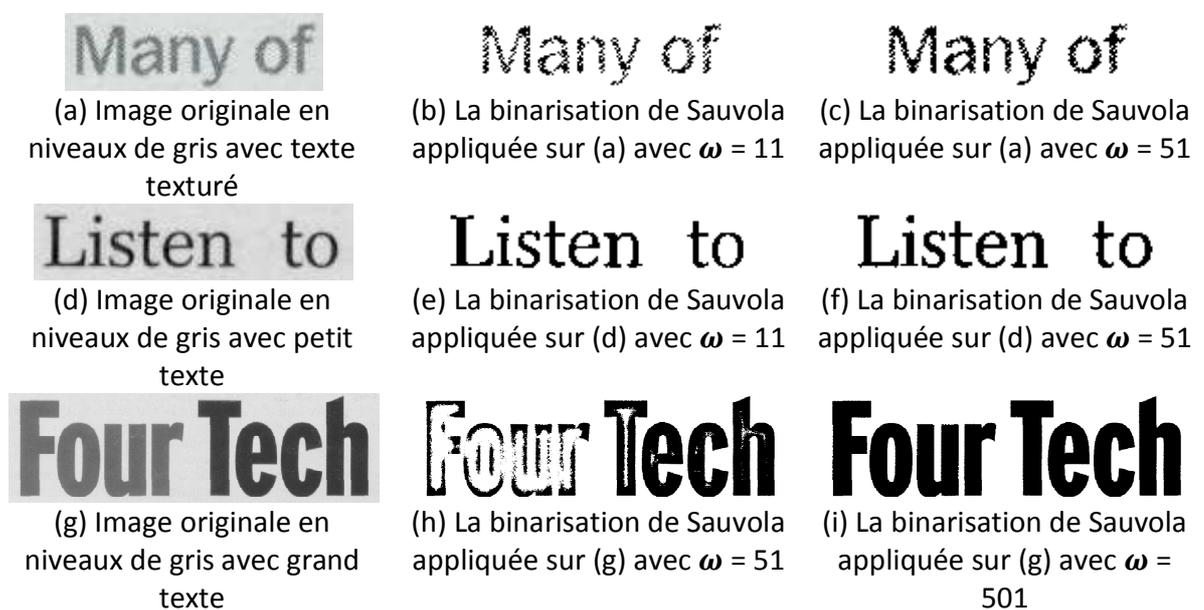


Figure 5.3. Influence du paramètre ω sur le résultat de la binarisation de Sauvola [[Lazzara et Géraud, 2014](#)].

5.3.2.4. Interférence d'objet spatial

Ce problème apparaît principalement avec les légendes des images, comme dans la figure 5.4. Des fenêtres trop grandes peuvent inclure des données, provenant d'objets de nature différente. Dans la figure 5.4, les données de l'image, situées au-dessus de la légende, sont prises en compte, conduisant à des statistiques non pertinentes, et à une binarisation incorrecte.



Figure 5.4. Influence d'une grande fenêtre ω (interférence d'objet).

Dans le reste de ce chapitre, la méthode proposée « ISauvola : Improved Sauvola », publiée au proceeding de la conférence ICIAR en 2016 [Hadjadj et al., 2016], sera présentée ; elle est proposée pour surmonter une des quatre limitations de la binarisation de Sauvola, mentionnées précédemment, qui est la « Mauvaise binarisation des objets faiblement contrastés ».

5.4. Description de la technique proposée

La méthode proposée 'ISauvola' est une amélioration de l'algorithme de Sauvola [Sauvola et Pietikainen, 2000]. Elle a pour but d'atteindre une meilleure détection des caractères bien contrastés, et même ceux faiblement contrastés, présents dans le même document, sans ajuster manuellement les valeurs des paramètres utilisateur.

L'organigramme global de la méthode proposée est illustré à la figure 5.5. La méthode se décompose en trois étapes :

- (1) Une carte d'initialisation est extraite, à partir de l'image du document d'entrée, pour identifier les pixels de l'avant-plan les plus probables ;
- (2) L'algorithme de Sauvola est appliqué à l'image du document d'entrée ;
- (3) Pour produire la binarisation finale, il suffit de détecter, dans l'image de binarisation de Sauvola, l'ensemble des objets se chevauchant avec chaque pixel du texte de la carte d'initialisation.

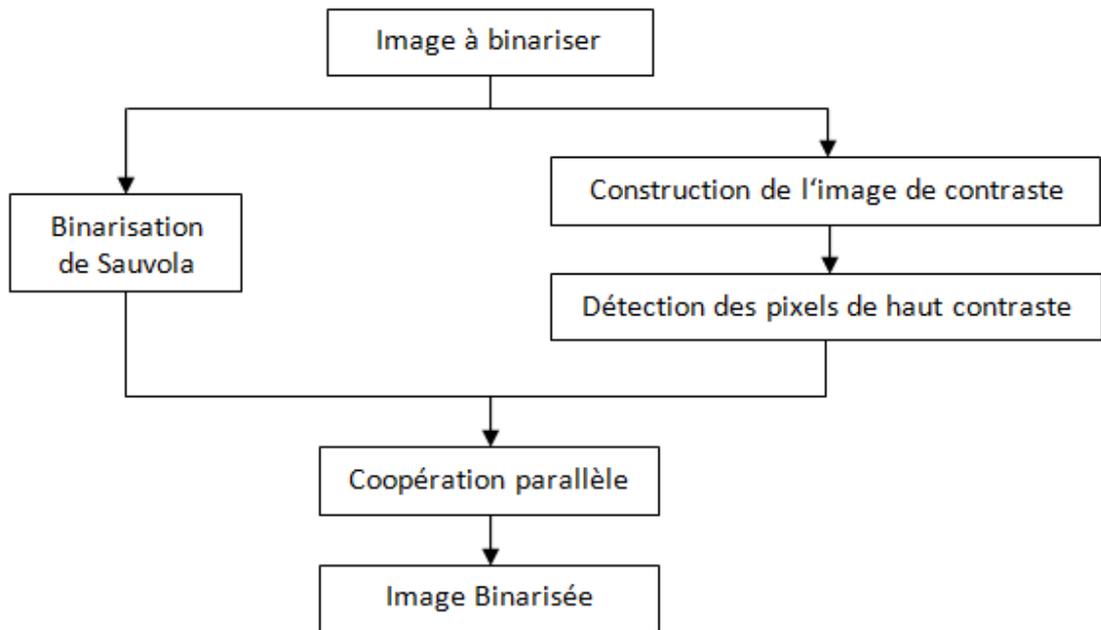


Figure 5.5. Organigramme de la méthode proposée.

Dans la prochaine sous-section, une binarisation initiale est estimée à l'aide du contraste de l'image.

5.4.1. Etape 1 : Initialisation par la méthode du contraste

Dans cette étape, une approche d'initialisation, basée sur le contraste de l'image, est utilisée pour identifier les pixels de l'avant plan les plus probables. L'initialisation utilisée construit d'abord une image de contraste, évaluée par le maximum et le minimum locaux [van Herk, 1992, Su et al., 2010], puis détecte les pixels de haut contraste, qui se situent habituellement autour des contours du texte, voir les sous-sections 4.3.1.1 et 4.3.1.2.

5.4.2. Etape 2 : Binarisation de Sauvola

Dans cette étape, le seuillage de Sauvola, décrit dans la sous-section 5.3.1, est effectué sur l'image du document d'entrée.

Dans nos expérimentations, nous avons constaté que la valeur de R de l'équation (5.1) a une faible influence sur la qualité de la binarisation, alors que les valeurs de k , et la taille de la fenêtre ω , l'affectent de manière significative. Les objets faiblement contrastés peuvent être considérés comme texture du fond, ou comme bruit de l'effet de transparence, en raison de la formule du seuil (5.1), et peuvent être supprimés ou détectés partiellement. Une faible valeur de ce paramètre peut aider à récupérer des objets à faible contraste, mais, comme il est défini pour l'ensemble du document, il modifie également d'autres parties du résultat : les objets correctement contrastés sont plus épais dans ce cas, ce qui peut entraîner des connexions involontaires entre les composants. Dans ce cas, le bruit de fond, qui est généralement mal contrasté, sera détecté comme objet.

La taille de la fenêtre est un paramètre important, pour obtenir de bons résultats ; une valeur trop faible peut entraîner des caractères cassés et/ou des caractères avec des trous, alors qu'une valeur trop élevée peut entraîner des caractères gras. Cette taille doit dépendre du contenu du document.

Une combinaison optimale de k et de ω produira une bonne image binaire. Dans nos expérimentations, nous avons choisi une valeur faible pour le paramètre k , ceci pour détecter tous les pixels de texte (faiblement et correctement contrastés), et une valeur faible pour le paramètre ω , pour réduire le chevauchement entre les caractères.

5.4.3. Etape 3 : Coopération parallèle

Dans cette étape, nous avons utilisé une coopération parallèle, entre l'image de haut contraste et l'image de binarisation de Sauvola, pour obtenir la binarisation finale.

La combinaison parallèle consiste à détecter, dans l'image de binarisation de Sauvola, l'ensemble des pixels qui se chevauchent avec chaque pixel texte de l'image à contraste élevé, comme décrit dans l'Algorithme 1.

Algorithme 1: Step 3 - Parallel combination between the high contrast image and Sauvola's binarization image.

Require: The high contrast image I_C (constructed at step 1) and Sauvola binarization image I_S (constructed at step 2)

Ensure: The final binary result I_B

1: for all pixel p in I_C : do

2: if $I_C(p)=\text{true}$ then // p is part of an object in I_C

3: Detect the set of pixels overlapping with p in I_S .

4: end if

5: end for

6: Store the new binary result to I_B .

La méthode proposée décrite a été implémentée sur **Matlab**, et les résultats sont présentés et discutés dans la section suivante.

5.5. Expérimentations, résultats et discussion

5.5.1. Validation de la méthode proposée sur les bases de données des DIBCOs, et comparaison avec les méthodes de l'état de l'art

La méthode décrite précédemment a été testée sur les images de documents, utilisées dans **les Concours de Binarisation d'Images de Documents (DIBCOs)**, qui subissent différents types de dégradations. Nous comparons également notre méthode avec d'autres méthodes de binarisation bien connues, y compris la méthode de seuillage de Sauvola [[Sauvola et Pietikainen, 2000](#)].

Les différents tests effectués sur un grand nombre d'images de documents ont démontré que les paramètres suivants : $\omega = 15 \times 15$, $R = 128$, comme recommandé dans dans l'article d'origine [[Sauvola et Pietikainen, 2000](#)], et $k = 0,01$, donnent les meilleurs résultats de binarisation. Une faible valeur du paramètre k peut aider à récupérer des objets

à faible contraste, puisqu'il est défini pour l'ensemble de l'image, et peut modifier également d'autres parties du résultat : beaucoup de bruits de fond et d'artefacts sont récupérés comme objets, mais la coopération parallèle proposée à la troisième étape peut supprimer le bruit efficacement, car la première étape supprime le contraste du fond du document, par la normalisation, comme décrit dans la section 5.4.1.

Les résultats de la binarisation dans les prochaines figures montrent l'efficacité de la technique de seuillage proposée.

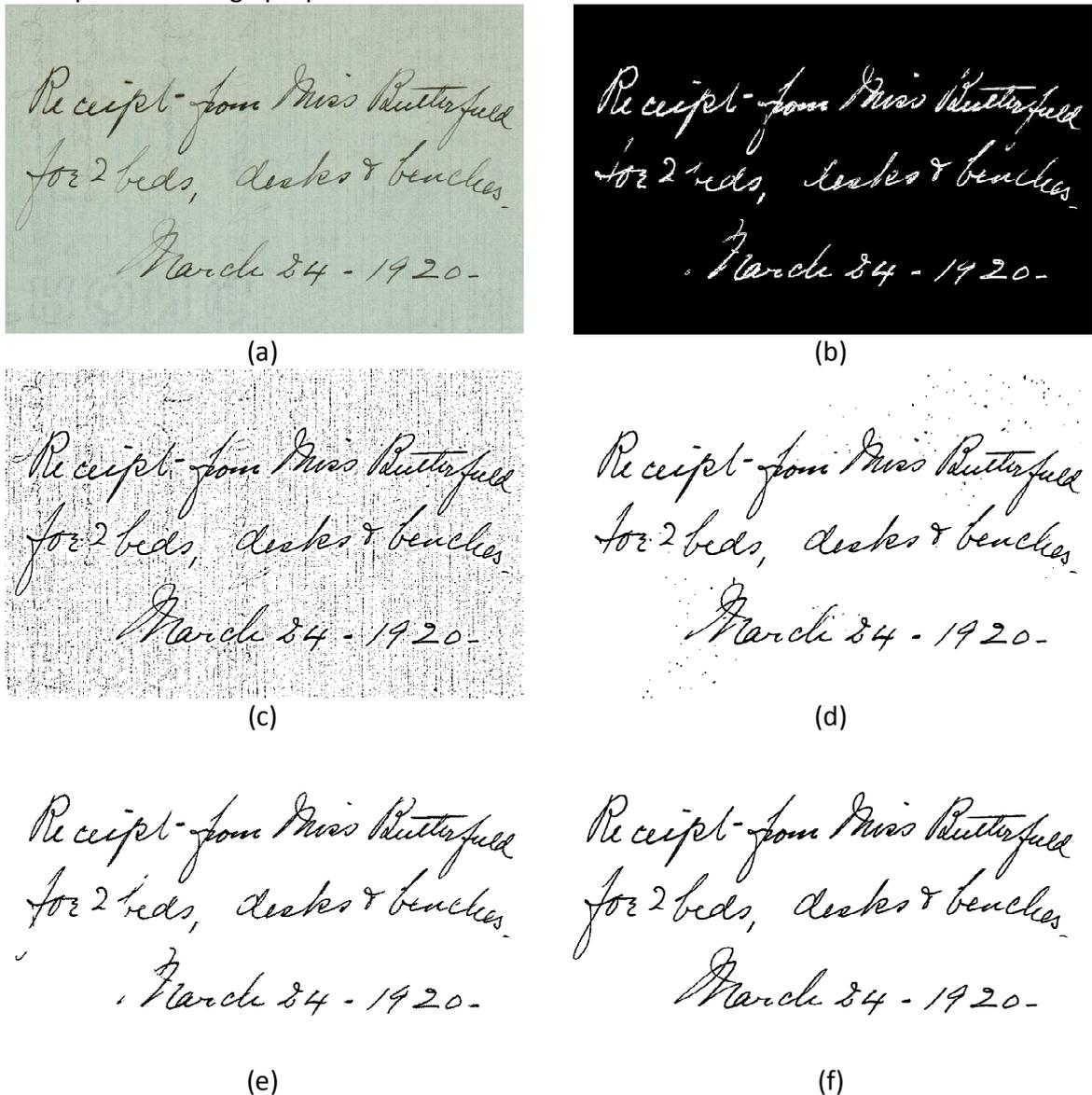
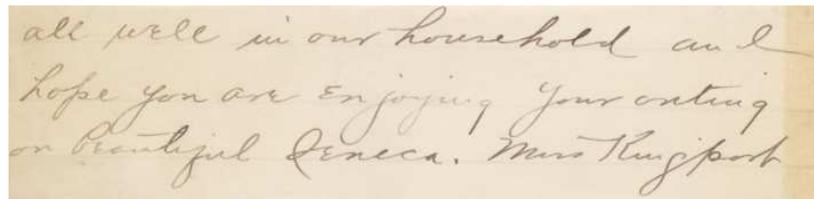


Figure 5.6. Résultats de la binarisation de l'image HW2 de la base de données de DIBCO'11 : (a) Image à binariser ; (b) Image de haut contraste ; (c) Méthode de Sauvola [Sauvola et Pietikainen, 2000] ; (d) Méthode de Howe [Howe, 2011]¹ ; (e) Méthode de Su [Su et al., 2013]² ; (f) Méthode proposée.

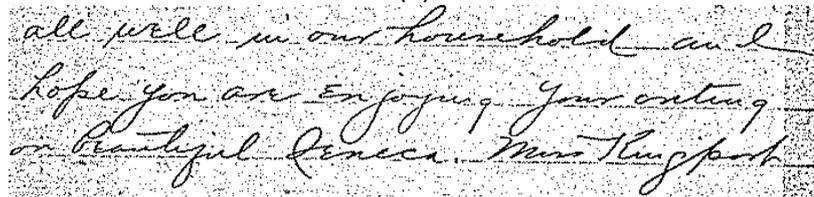
¹ classée 3^{ème} à DIBCO'11

La figure 5.6.a illustre une image de document dégradée, avec des caractères faiblement contrastés. Notre technique de binarisation construit d'abord une image de contraste, par l'image locale maximale et minimale, puis extrait l'image à contraste élevé (Figure 5.6.b), qui sert à supprimer le bruit de fond et les artefacts. Ensuite, l'algorithme de Sauvola est appliqué sur l'image d'entrée (Figure 5.6.c), pour détecter tous les pixels de texte (faiblement et bien contrastés). Ensuite, nous combinons parallèlement entre les deux résultats, en cherchant dans l'image de binarisation de Sauvola l'ensemble des pixels qui se chevauchent, avec chaque pixel texte de l'image à contraste élevé, pour produire la binarisation finale.

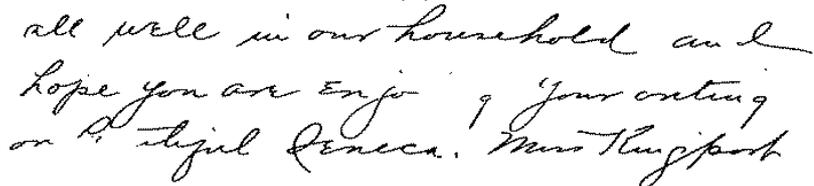
Comme on peut le voir sur la figure 5.6.f (résultat final), les caractères faiblement contrastés sont raisonnablement bien détectés en utilisant notre méthode. D'autre part, la méthode de Sauvola [Sauvola et Pietikainen, 2000] produit beaucoup de bruit, en raison de la variation dans l'arrière-plan du document. La méthode de Su [Su et al., 2013] et la méthode de Howe [Howe, 2011], qui sont des méthodes bien classées à DIBCO'11, ne permettent pas de détecter certains caractères, faiblement contrastés.



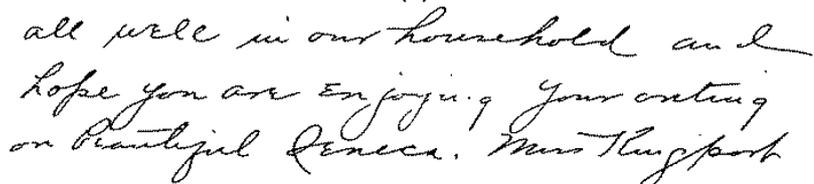
(a)



(b)



(c)



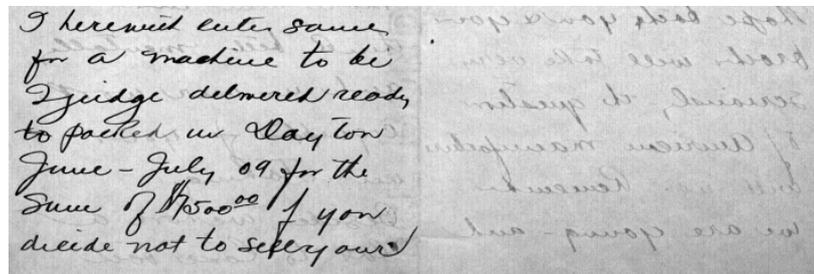
(d)

all well in our household and I
 hope you are enjoying your outing
 on beautiful Seneca. Mrs. Thurgood

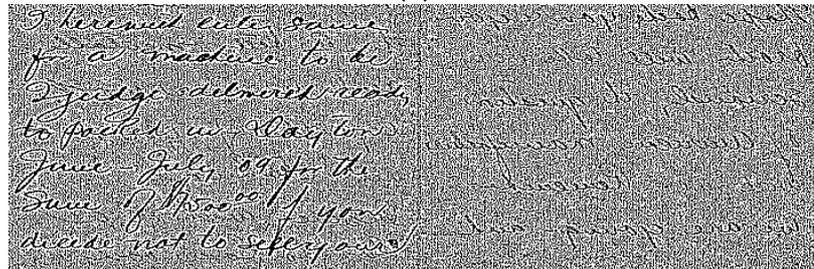
(e)

Figure 5.7. Résultats de la binarisation de l'image H12 de la base de données de HDIBCO'12 :

- (a) Image à binariser ;
- (b) Méthode de Sauvola ;
- (c) Méthode de Lore³ ;
- (d) Méthode de Howe⁴ ;
- (e) Méthode proposée.



(a)



(b)

I herewith enter same
 for a machine to be
 I judge delivered ready
 to packed in Clayton
 June - July, 09 for the
 sum of \$500.00 if you
 decide not to select one

(c)

Figure 5.8. Résultats de la binarisation de l'image HW4 de la base de données de DIBCO'13:

- (a) Image à binariser ;
- (b) Méthode de Sauvola ; (c) Méthode proposée.

³ classée 2^{ème} à H-DIBCO'12

⁴ classée 1^{ère} à H-DIBCO'12

La figure 5.8.a montre une image de document dégradée. La transparence du verso est un problème très difficile à aborder en traitement d'images de documents historiques, et en particulier en binarisation. Ce problème se produit lorsque l'encre s'infiltré, du côté verso du document vers le côté recto (et vice versa). Pendant une longue période de stockage, il devient difficile de distinguer les écritures des deux cotés, lorsque l'encre s'est complètement infiltrée dans la page, de sorte qu'il ressemble au premier plan, ou qu'il devienne plus sombre. Comme on peut le voir sur la figure 5.8.c, le bruit de l'effet de transparence est raisonnablement bien éliminé, en utilisant notre méthode. Cette dernière peut supprimer plus de bruit que la méthode de Sauvola, car elle supprime le contraste du fond du document, par la normalisation dans l'étape d'initialisation. À titre de comparaison, la méthode de Sauvola classe simplement, les pixels sombres de l'arrière-plan, comme pixels du texte, de manière incorrecte.

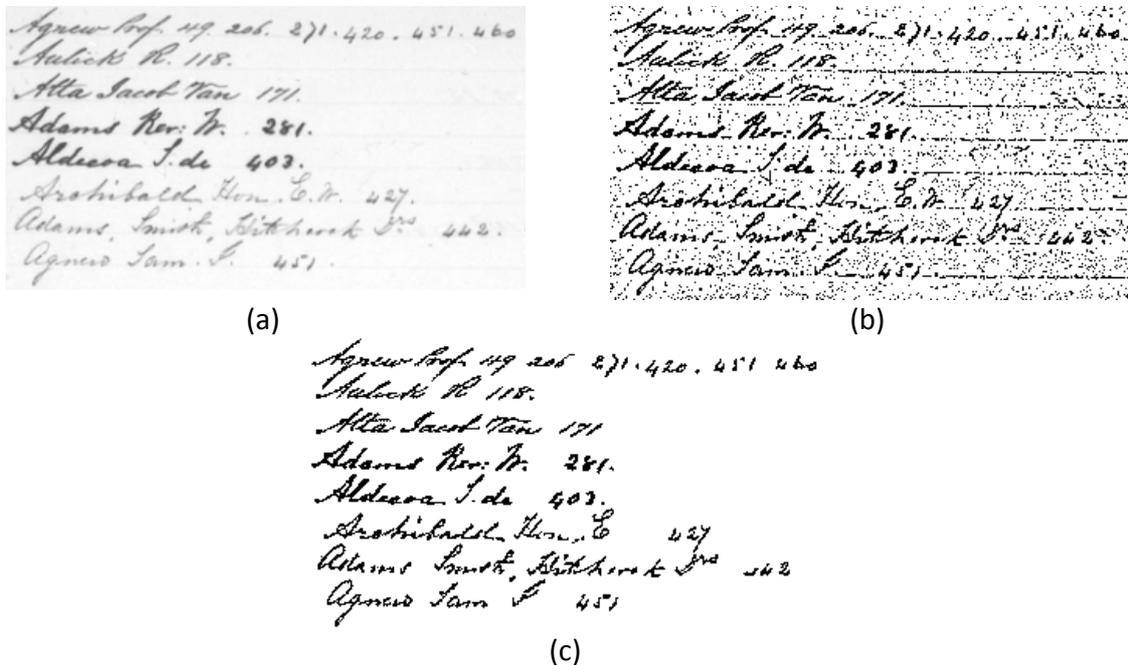


Figure 5.9. Résultats de la binarisation de l'image H04 de la base de données de HDIBCO'10:

- (a) Image à binariser ;
- (b) Méthode de Sauvola ;
- (c) Méthode proposée.

Les figures 5.7.a et 5.9.a montrent deux images de document dégradées, avec des caractères faiblement contrastés. La figure 5.9 montre que la technique proposée est tolérante aux variations de contraste du document, et capable de binariser des caractères faiblement contrastés, et des images mal éclairées, avec peu de bruit de fond, alors que d'autres méthodes peuvent, soit introduire une certaine quantité de bruit, soit ne pas détecter les pixels du texte faiblement contrastés (Figure 5.7.c).

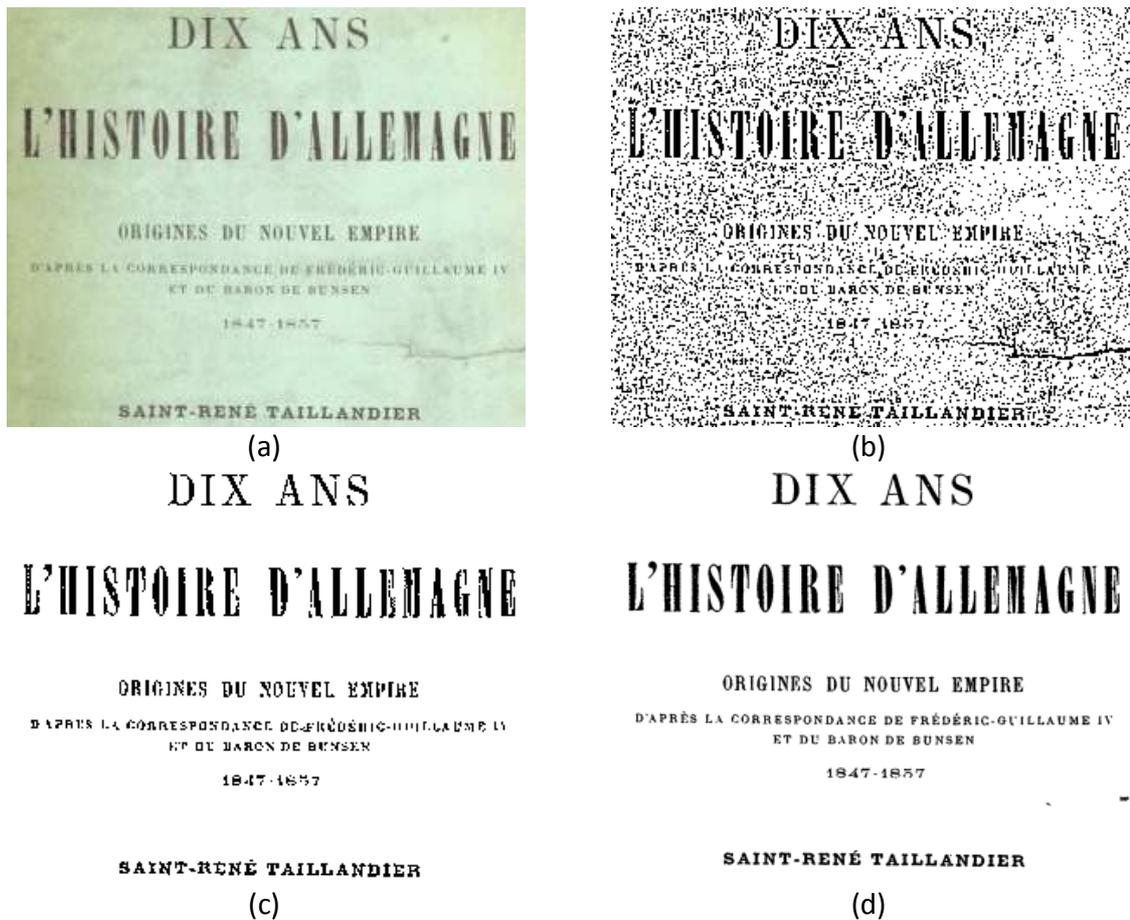


Figure 5.10. Résultats de la binarisation de l'image PR3 de la base de données de DIBCO'13 :
 (a) Image à binariser (uneven illumination degraded document image) ;
 (b) Méthode de Sauvola ;
 (c) Méthode de Su⁵ ;
 (d) Méthode proposée.

Les figures 5.6, 5.7, 5.8, 5.9 et 5.10 montrent cinq exemples de binarisation de documents. Comme indiqué, notre méthode extrait le texte correctement des images de documents, qui présentent différents types de dégradations, sans ajuster manuellement la valeur des paramètres k et ω . D'autre part, la méthode de Sauvola produit souvent une certaine quantité de bruit, en raison de la variation en arrière-plan.

⁵ classée 1^{ère} à DIBCO'13



Figure 5.11. Résultats de la binarisation de l'image P03 de la base de données de DIBCO'09 :

- (a) Image à binariser ;
- (b) Résultat de la méthode de Sauvola à partir de [Rivest-Hénault et al., 2011] ;
- (c) Méthode proposée.

La figure 5.11 montre que la performance de notre méthode de binarisation proposée dans ce chapitre est limitée à la présence de petits et grands objets dans un même document ; La méthode de Sauvola ne parvient pas à récupérer tous les objets correctement car son paramètre de fenêtre est défini pour toute l'image (problème présenté à la section 5.3.2.2), et comme la méthode proposée est basée sur le résultat de la méthode de Sauvola donc elle a hérité ce problème.

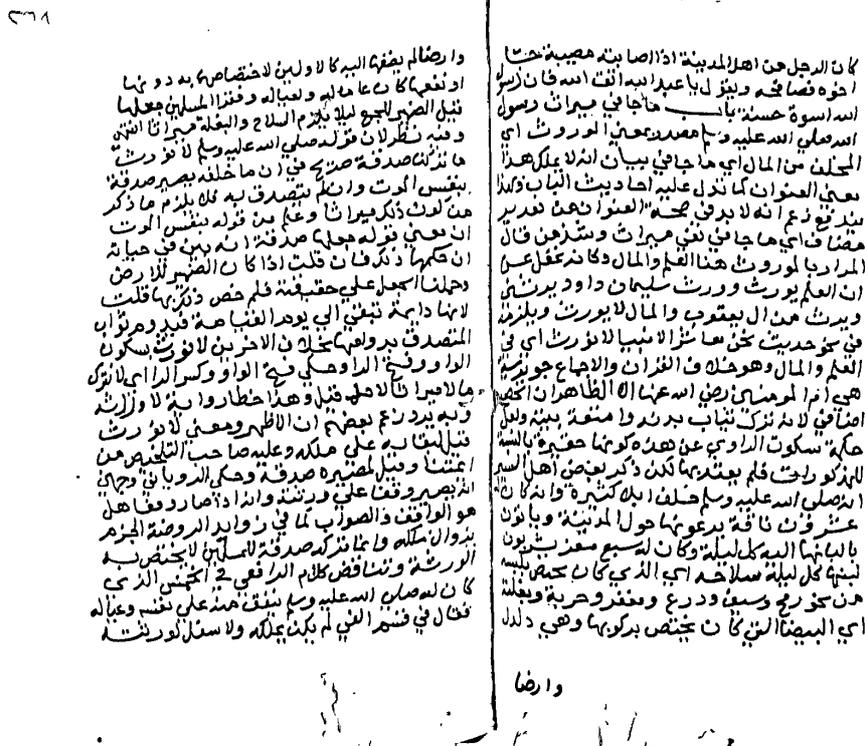
5.5.2. Validation de la méthode proposée sur des images issues de la Bibliothèque Nationale d'Algérie et du Centre National des Manuscrits d'Adrar

La validité de la méthode proposée est aussi montrée sur des images de documents anciens arabes, issues de la Bibliothèque Nationale d'Algérie et du Centre National des Manuscrits d'Adrar.

Les résultats de binarisation de quelques images de documents, de dimension 2160×1440 pixels, sont présentés à la figure 5.12.



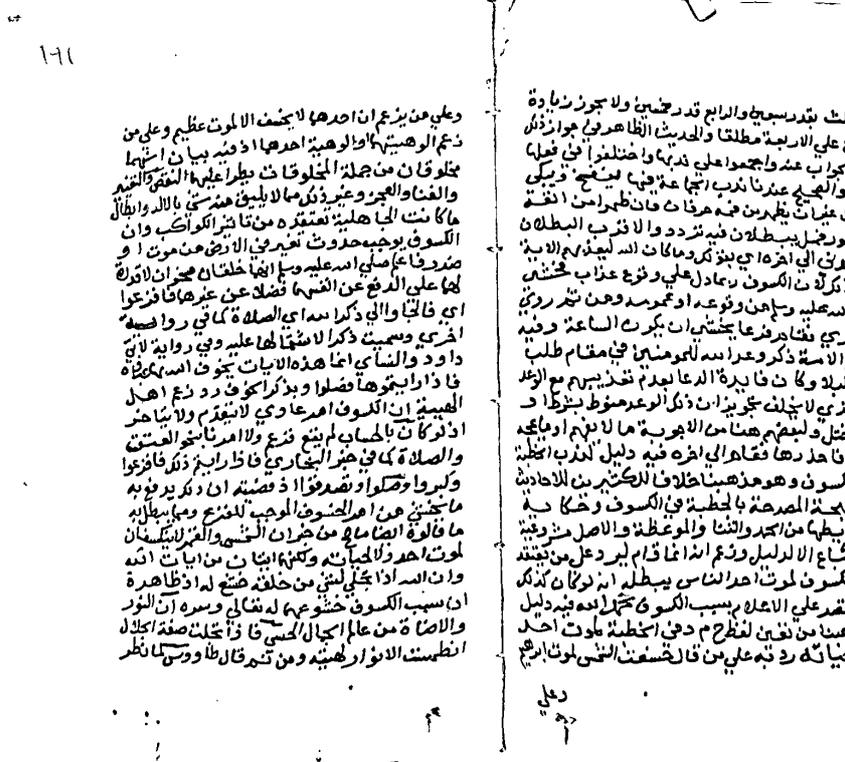
(a) Image de document à binariser



(b) Image binarisée de (a)



(c) Image de document à binariser



(d) Image binarisée de (c)



(e) Image de document à binariser



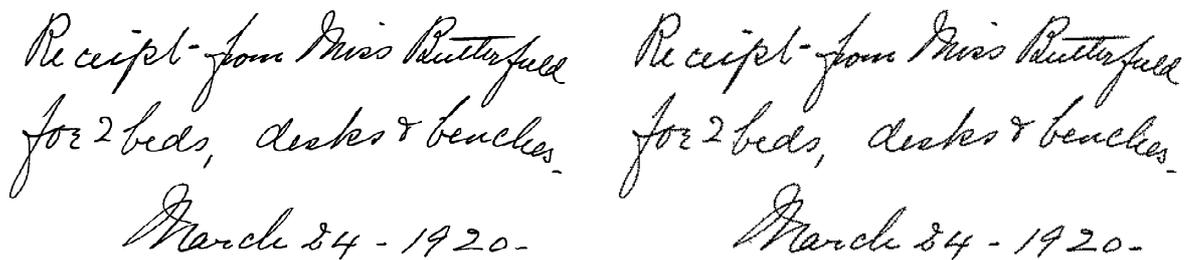
(f) Image binarisée de (e)

Figure 5.12. Résultats de la binarisation de quelques images de documents par la méthode proposée.

Plusieurs expérimentations ont été effectuées afin d'évaluer les performances de la méthode proposée, les résultats obtenus sont, à notre opinion, encourageants en considérant le principe simple de la méthode proposée.

En effet, comme il été confirmé par les résultats expérimentaux, la méthode proposée a présenté un comportement fiable pour la tâche complexe de séparation entre l'avant-plan et arrière-plan à partir des images de documents significativement dégradées.

5.5.3. Comparaisons avec la première méthode de contribution



(a) Résultat de la première méthode de contribution



(c) Résultat de la première méthode de contribution

(b) Résultat de la deuxième méthode de contribution



(d) Résultat de la deuxième méthode de contribution

Figure 5.13. Comparaison entre les résultats des deux méthodes de contribution.

La figure 5.13 montre que les deux méthodes proposées aux chapitre 4 et 5 réussissent effectivement à capturer des caractères de texte faiblement et correctement contrastés dans la même image de document, comme montré aux figures 5.13.a et 5.13.b, sans ajuster manuellement la valeur des paramètres utilisés dans les deux méthodes de binarisation. Cependant, la performance de la deuxième méthode de binarisation proposée est limitée à la présence de petits et grands objets dans un même document (Figure 5.13.d) et la première méthode de binarisation proposée est plus couteuse en termes de temps de calcul.

5.6. Conclusion

Ce chapitre présente une nouvelle méthode de binarisation d'images de documents historiques, efficace contre différents types de dégradation de documents, tels que le faible contraste et l'éclairage irrégulier.

La méthode présentée dans ce chapitre coopère, de façon parallèle la méthode de Sauvola et la méthode de contraste, définie par le maximum et le minimum locaux. Les deux méthodes sont intégrées dans le but d'atteindre une meilleure détection des caractères bien contrastés, et même ceux faiblement contrastés, présents dans le même document, sans ajuster manuellement les valeurs des paramètres utilisateur.

La méthode combinée conduit, comme montré dans les expériences, à une grande précision, lorsqu'elle est appliquée à des images de documents historiques, avec une variété de dégradations. La méthode proposée réussit effectivement à capturer des objets faiblement et correctement contrastés, dans un seul document.

La méthode proposée dans ce chapitre traite l'image de document indépendamment de sa langue et son type d'écriture, comme il est montré dans les expériences, elle traite des images de manuscrits avec des langues et des types d'écriture différents : l'anglais et le grec pour les bases de données DIBCOs et HDIBCOs, et l'arabe pour les images de manuscrits du patrimoine.

Cependant, la performance de notre méthode de binarisation est limitée, à la présence de petits et grands objets dans un même document. La méthode ISauvola ne parvient pas à récupérer tous les objets correctement, car son paramètre de fenêtre est défini pour toute l'image. Dans de futurs travaux, nous nous concentrerons sur la gestion de différentes tailles d'objet.

CONCLUSION GÉNÉRALE

Cette étude a présenté quelques contributions, au domaine émergent de la restauration des images de documents, et plus particulièrement de documents anciens du patrimoine. Contrairement aux images naturelles, les images de documents sont des images particulières, qui nécessitent des traitements adaptés. Leur restauration ne peut pas être basée sur la simple transposition de méthodes existantes. Les images de documents anciens du patrimoine possèdent de nombreuses dégradations, dues à la mauvaise conservation physique, et au vieillissement naturel des supports, auxquels s'ajoutent des dégradations numériques, provenant de la chaîne d'acquisition et du traitement informatique. Ces dégradations limitent considérablement l'accès de ces images au grand public. Il existe donc actuellement une forte demande, pour le développement d'outils de traitements informatiques, adaptés à la restauration des images de documents. Nous nous sommes concentrés sur la binarisation, considérée comme l'une des opérations de prétraitement les plus importantes.

1. Bilan

En résumé, nos principales contributions sont les suivantes :

Contribution à la binarisation d'images de documents dégradés par contours actifs

Au cours de l'état de l'art, il est apparu que les contours actifs ou modèles déformables, malgré leurs aptitudes à traiter efficacement des images médicales et satellitaires, n'étaient pas beaucoup représentés dans la binarisation de documents manuscrits. Nos recherches ont donc été orientées vers cette approche.

Un objet déformable consiste en un ensemble de points, qui vont se déplacer au fur et à mesure du déroulement de l'algorithme de déformation. Les fondements mathématiques de ces modèles proviennent de différents domaines, tels que la géométrie, la physique, ou encore la théorie de l'approximation. La géométrie est utilisée pour représenter la forme de l'objet; la physique impose des contraintes sur la variation de la forme dans l'espace et le temps; enfin, la théorie de l'approximation optimale fournit les justifications formelles des outils, permettant d'adapter les modèles aux données mesurées. La formulation des contours actifs par la méthode des ensembles de Niveaux (Level Sets), permet de les implémenter, tout en gérant le problème de changements de topologie.

Bien que les avantages des contours actifs soient multiples, ils souffrent de certains points négatifs, notamment la sensibilité à l'initialisation du contour initial. Afin de résoudre ce problème, la méthode des maximum et minimum locaux est utilisé pour générer automatiquement une carte d'initialisation du contour actif; un seuillage adaptatif moyen est également utilisé pour produire la binarisation finale.

Contribution à la binarisation d'images de documents dégradés par l'algorithme ISauvola

Au lieu de concevoir une nouvelle méthode de binarisation, nous avons essayé, au chapitre 5, d'apporter des améliorations sur la méthode de binarisation existante de Sauvola, dont le but d'améliorer, non seulement la performance de la méthode originale de Sauvola, mais aussi la robustesse sur différents types d'images de documents dégradées.

La méthode proposée coopère parallèlement la méthode de Sauvola et la méthode du contraste, définie par les maximum et minimum locaux. Les deux méthodes de binarisation sont intégrées, dans le but d'atteindre une meilleure détection des caractères bien contrastés et même ceux faiblement contrastés, présents dans le même document, sans ajuster manuellement les valeurs des paramètres utilisateur. La méthode est proposée dans le but de parvenir à une technique de binarisation rapide et efficace en même temps.

Notre recherche sur la méthode de Sauvola et sur les méthodes de seuillage en général, nous a permis de percevoir l'impact du choix des valeurs des paramètres utilisateur sur les résultats de la binarisation. Une mauvaise initialisation de ces paramètres engendre des résultats de mauvaise qualité. Usuellement, les paramètres de la méthode de Sauvola sont estimés à l'aide de l'utilisateur qui détermine visuellement la valeur la plus adéquate pour chaque paramètre et pour chaque image par essaie-erreur. Notre méthode proposée **ISauvola** ne nécessite pas l'ajustement manuel des valeurs des paramètres utilisateurs pour chaque image et pour chaque type de dégradation et la qualité du résultat obtenu ne dépend pas du choix des valeurs des paramètres utilisateur grâce à l'enchaînement de ces trois étapes.

Les deux méthodes proposées, dans les chapitres 4 et 5, ont été validées sur les bases d'images des DIBCOs, avec une variété de dégradations, avec lesquelles elles ont été quantifiées et comparées aux autres algorithmes existants. La validité des deux méthodes proposées est aussi montrée sur des images de documents anciens arabes, issues de la Bibliothèque Nationale d'Algérie et du Centre National des Manuscrits d'Adrar. Différentes expérimentations ont été effectuées, et les résultats obtenus par les deux méthodes sont encourageants.

Les deux méthodes proposées traitent l'image de document indépendamment de sa langue et son type d'écriture, comme il est montré dans les expériences, elles traitent des images de manuscrits avec des langues et des types d'écriture différents : l'anglais et le grec pour les bases de données DIBCOs et HDIBCOs, et l'arabe pour les images de manuscrits du patrimoine.

Les deux méthodes proposées dans cette thèse ont été citées par différents chercheurs et dans différents travaux de recherche, tels que : [Hu et al., 2015, Liu et al., 2016, Chakraborty et Blumenstein, 2016] et autres.

2. Perspectives

Comme tous les travaux de recherche, plusieurs extensions sont envisageables, pour améliorer le travail proposé dans cette thèse.

Les méthodes proposées utilisent la méthode des maximum et minimum locaux (MML). Cette méthode peut traiter la dégradation de l'effet de transparence du verso '**Bleed-through**', lorsque les traits du texte du côté arrière sont beaucoup plus faibles que ceux du côté avant de la page. Mais lorsque les traits du texte du côté arrière sont aussi sombres que les traits du texte du côté avant de la page, la méthode MML ne peut pas classer correctement les deux types de traits de caractères. En outre, la méthode MML dépend fortement des pixels d'image à contraste élevé. En conséquence, elle peut introduire une erreur, si l'arrière-plan des images de document dégradées contient une certaine quantité de pixels qui sont denses, et en même temps qui ont un contraste d'image assez élevé. Nous étudierons ces problèmes dans nos travaux futurs.

La méthode de binarisation proposée au chapitre 4 est coûteuse en termes de calcul, à cause du contour actif utilisé, en plus du fait que sa performance est limitée, lorsque l'image à binariser est fortement dégradée, par l'effet de transparence du verso '**bleed-through**', ou par la présence de taches '**smear**'. Comme perspective pour les travaux futurs, une amélioration sur les forces motrices (termes énergétiques), du contour actif, pour rendre la méthode plus adaptable aux fortes dégradations, sera considérée.

Annexe A Interface graphique développée

Dans cette thèse, nous avons utilisé l'environnement de développement "Matlab" parce qu'il offre aux chercheurs scientifiques des outils en traitement d'image et du signal. Cet environnement m'a fait gagner beaucoup de temps et m'a permis de réaliser mes tests d'une manière très large et très rapide. Aussi, j'ai développé une interface graphique nommée "**(ADIP) Arabic Document Images Processing**" qui contient toutes les fonctions de bases de traitement d'image, et mes algorithmes de binarisation des images de documents. Cette annexe a pour objectif de présenter "Matlab", et l'interface graphique "**(ADIP) Arabic Document Images Processing**".

A.1. Notre système

(ADIP) « Arabic Document Images Processing » est un système pour le traitement d'images et pour la binarisation des images de documents. Il contient toutes les fonctions de bases pour le traitement d'images. Dans la barre de menu, on trouve des mots cliquables permettant de :

- Effectuer des opérations de filtrage sur l'image (Moyenneur, gaussien, Médian, ...).
- Utiliser plusieurs filtres pour effectuer la détection de contours de l'image concernée (Sobel, Canny, Laplacien, ...).
- Appliquer à l'image des opérations morphologiques de base (Dilatation, Erosion, ...).
- ...

Le mot cliquable « Traitement Complet » du menu, offre à l'utilisateur la possibilité de faire tout les traitements, et d'afficher l'image restaurée finale par l'une des méthodes proposées aux chapitres 4 et 5.

Les figures [A.1](#) et [A.2](#) montrent notre interface graphique.

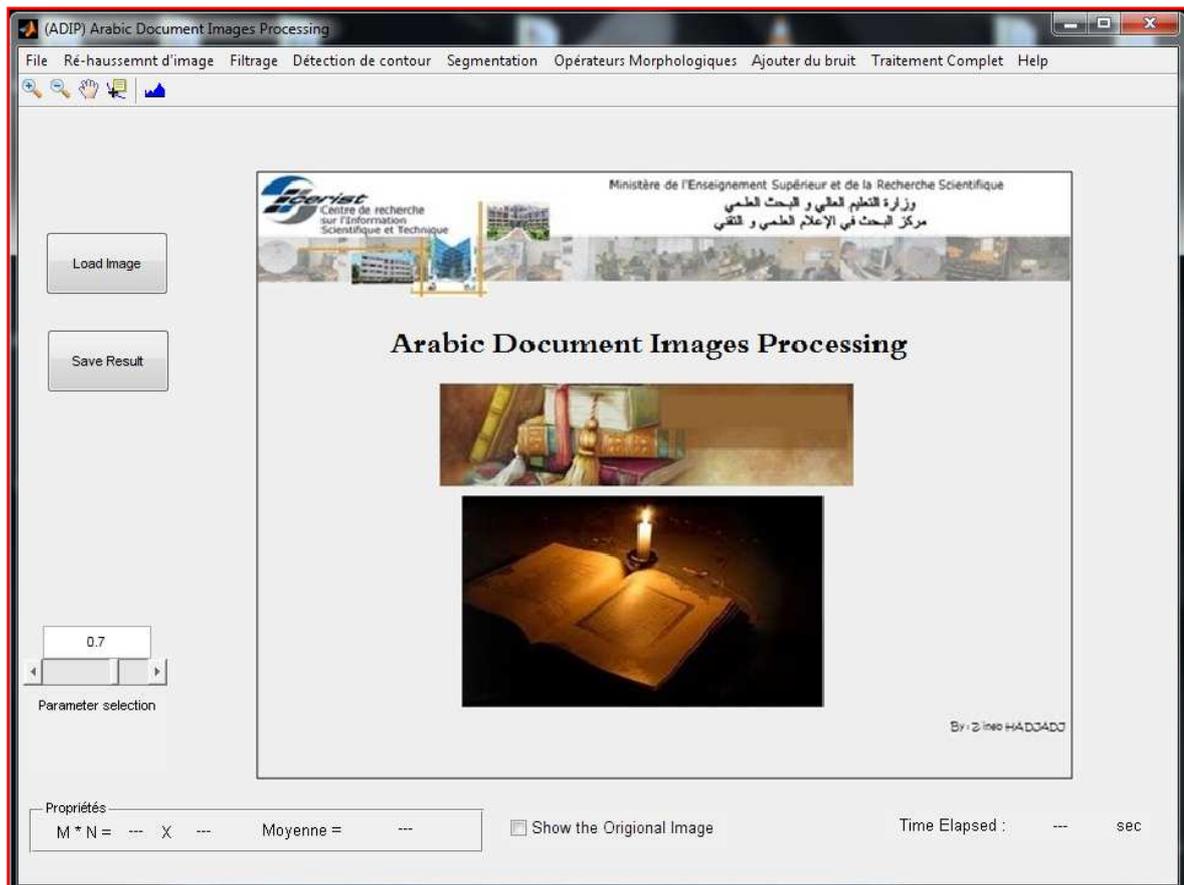
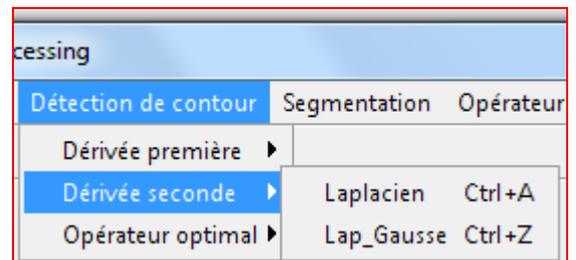
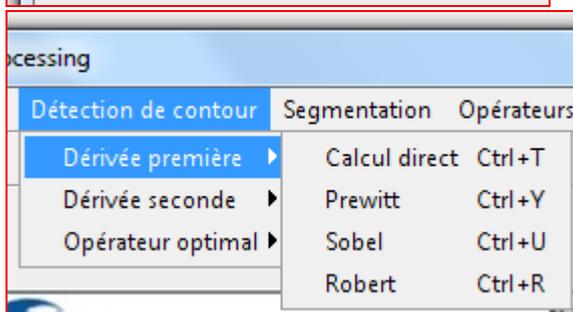
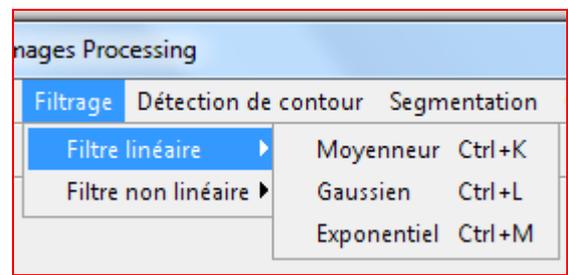
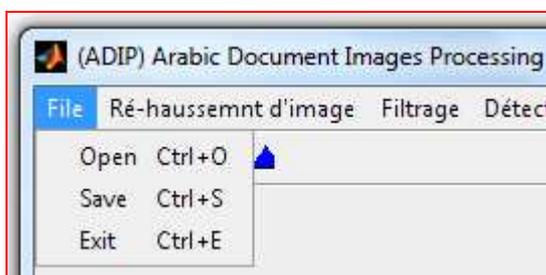


Figure A.1. Notre système de traitement d'image et de binarisation des images de documents " (ADIP) Arabic Document Images Processing ".



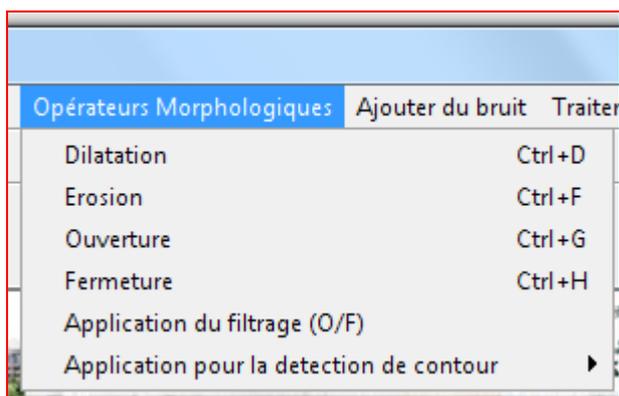


Figure A.2. Menu de fonctions permettant d'appliquer des traitements de base à l'image.

Annexe B
LISTE DES PUBLICATIONS DE L'AUTEUR

Publications dans des revues internationales

- Z. Hadjadj, M. Cheriet, A. Meziane, Y. Cherfa, **A new efficient binarization method: application to degraded historical document images**, Signal, Image and Video Processing (SIVP), vol. 11, n° 6, pp. 1155-1162, 2017.

Publications dans des conférences internationales

- Z. Hadjadj, A. Meziane, **Binarization of Document Images with Various Object Sizes**, 1st IEEE International Workshop on Arabic Script Analysis and Recognition (ASAR 2017), April 03-05, Nancy, France, pp. 21-25 (2017). [978-1-5090-6628-5/17/\\$31.00 ©2017 IEEE](https://doi.org/10.1109/ASAR.2017.7991115)
- I. Setitra , Z. Hadjadj, A. Meziane, **A Tracking Approach for Text Line Segmentation in Handwritten Documents**, 6th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2017), February 24-26, Porto, Portugal, pp. 193 - 198 (2017).
- Z. Hadjadj, A. Meziane, Y. Cherfa, M. Cheriet, Insaf Setitra, **ISauvola: Improved Sauvola's Algorithm for Document Image Binarization**, 13th International Conference on Image Analysis and Recognition (ICIAR 2016), July 13-15, Póvoa de Varzim, Portugal, pp. 737–745 (2016). doi: [10.1007/978-3-319-41501-7_82](https://doi.org/10.1007/978-3-319-41501-7_82)
- Z. Hadjadj, A. Meziane, M. Cheriet, Y. Cherfa, **An active contour based method for image binarization: application to degraded historical document images**, 14th International Conference on Frontiers in Handwriting Recognition (ICFHR 2014), Crete, Greece, pp. 655–660 (2014). doi:[10.1109/ICFHR.2014.115](https://doi.org/10.1109/ICFHR.2014.115)
- Z. Hadjadj, M. Cheriet, A. Meziane, **Binarization of Degraded Historical Document Images**, International Conference on Artificial Intelligence and Information Technology (ICA2IT 2014), March 10-12, Ouargla, Algeria (2014).

Bibliographie

- [Abutaleb, 1989] ABUTALEB, A. S. (1989). Automatic thresholding of gray-level pictures using two-dimensional entropy. *Computer Vision, Graphics, and Image Processing*, 47(1):22–32.
- [Agrawal et Doermann, 2009] AGRAWAL, M. et DOERMANN, D. S. (2009). Clutter noise removal in binary document images. *In ICDAR*, pages 556–560. IEEE Computer Society.
- [Aguilera et al., 2005] AGUILERA, J., WILDENAUER, H., KAMPEL, M., BORG, M., THIRDE, D. et FERRYMAN, J. (2005). Evaluation of motion segmentation quality for aircraft activity surveillance. *In Proceedings of the 14th International Conference on Computer Communications and Networks, ICCCN '05*, pages 293–300. IEEE Computer Society.
- [Allier et al., 2006] ALLIER, B., BALI, N. et EMPTOZ, H. (2006). Automatic accurate broken character restoration for patrimonial documents. *IJDAR*, 8(4):246–261.
- [Allier et Emptoz, 2004] ALLIER, B. et EMPTOZ, H. (2004). Segmentation and typography extraction in document images using geodesic active regions. *In ICPR (1)*, pages 409–412. IEEE Computer Society.
- [Badekas et Papamarkos, 2005] BADEKAS, E. et PAPAMARKOS, N. (2005). Automatic evaluation of document binarization results. *In SANFELIU, A. et LAZO-CORTÉS, M., éditeurs : CIARP*, volume 3773 de *Lecture Notes in Computer Science*, pages 1005–1014. Springer.
- [Badekas et Papamarkos, 2007] BADEKAS, E. et PAPAMARKOS, N. (2007). Optimal combination of document binarization techniques using a self-organizing map neural network. *Eng. Appl. of AI*, 20(1):11–24.
- [Bechar, 2009] BECHAR, I. (2009). Flux du vecteur du gradient. <http://www.tsi.enst.fr/pages/enseignement/ressources/mti/gvf/gvf2.htm>.
- [Bernsen, 1986] BERNSEN, J. (1986). Dynamic thresholding of grey-level images. *In Proceedings of the Eighth International Conference on Pattern Recognition*, pages 1251–1255.
- [Bertalmio et al., 2000] BERTALMIO, M., SAPIRO, G., CASELLES, V. et BALLESTER, C. (2000). Image inpainting. *In Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '00*, pages 417–424.
- [Bottou et al., 1998] BOTTOU, L., HAFFNER, P., HOWARD, P. G., SIMARD, P. Y., BENGIO, Y. et LECUN, Y. (1998). High quality document image compression with "djvu". *J. Electronic Imaging*, 7(3):410–425.

- [Bresson *et al.*, 2007] BRESSON, X., ESEDOGLU, S., VANDERGHEYNST, P., THIRAN, J.-P. et OSHER, S. (2007). Fast global minimization of the active contour/snake model. *Journal of Mathematical Imaging and Vision*, 28:151–167.
- [Brink, 1992] BRINK, A. D. (1992). Thresholding of digital images using two-dimensional entropies. *Pattern Recognition*, 25(8):803–808.
- [Bukhari *et al.*, 2009a] BUKHARI, S. S., SHAFAIT, F. et BREUEL, T. M. (2009a). Coupled snakelet model for curled textline segmentation of camera-captured document images. *In ICDAR*, pages 61–65. IEEE Computer Society.
- [Bukhari *et al.*, 2009b] BUKHARI, S. S., SHAFAIT, F. et BREUEL, T. M. (2009b). Script-independent handwritten textlines segmentation using active contours. *In ICDAR*, pages 446–450. IEEE Computer Society.
- [Canny, 1986] CANNY, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698.
- [Cao et Govindaraju, 2009] CAO, H. et GOVINDARAJU, V. (2009). Preprocessing of low-quality handwritten documents using markov random fields. *IEEE Trans. Pattern Anal. Mach. Intell.*, 31(7):1184–1194.
- [Caselles *et al.*, 1993] CASELLES, V., CATTÉ, F., COLL, T. et DIBOS, F. (1993). A geometric model for active contours in image processing. *Numerische Mathematik*, 66(1):1–31.
- [Caselles *et al.*, 1997] CASELLES, V., KIMMEL, R. et SAPIRO, G. (1997). Geodesic active contours. *International Journal of Computer Vision*, 22(1):61–79.
- [Chakraborty et Blumenstein, 2016] CHAKRABORTY, A. et BLUMENSTEIN, M. (2016). Marginal noise reduction in historical handwritten documents - a survey. *In DAS*, pages 323–328. IEEE Computer Society.
- [Chen *et al.*, 2008] CHEN, Q., SUN, Q.-S., HENG, P.-A. et XIA, D.-S. (2008). A double-threshold image binarization method based on edge detector. *Pattern Recognition*, 41(4):1254–1267.
- [Chen *et al.*, 2002] CHEN, Y., TAGARE, H. D., THIRUVENKADAM, S. R., HUANG, F., WILSON, D. C., GOPINATH, K. S., BRIGGS, R. W. et GEISER, E. A. (2002). Using prior shapes in geometric active contours in a variational framework. *International Journal of Computer Vision*, 50(3):315–328.
- [Chenoune, 2008] CHENOUNE, Y. (2008). *Estimation des déformations myocardiaques par analyse d'images*. Thèse de doctorat, Université Paris EST, France.
- [Chinchor, 1992] CHINCHOR, N. (1992). Muc-4 evaluation metrics. *In MUC*, pages 22–29.
- [Chou *et al.*, 2010] CHOU, C.-H., LIN, W.-H. et CHANG, F. (2010). A binarization method with learning-built rules for document images produced by cameras. *Pattern Recognition*, 43(4):1518–1530.
- [Cohen *et al.*, 2012] COHEN, R., KEDEM, K., DINSTEIN, I. et EL-SANA, J. (2012). Occluded character restoration using active contour with shape priors. *In ICFHR*, pages 497–502.
- [Damera-Venkata *et al.*, 2000] DAMERA-VENKATA, N., KITE, T. D., GEISLER, W. S., EVANS, B. L. et BOVIK, A. C. (2000). Image quality assessment based on a degradation model. *IEEE Trans. Image Processing*, 9(4):636–650.

- [Diday, 1971] DIDAY, E. (1971). Une nouvelle méthode en classification automatique et reconnaissance des formes la méthode des nuées dynamiques. *Revue de Statistique Appliquée*, 19(2):19–33.
- [Drira, 2007] DRIRA, F. (2007). *Contribution à la restauration des images de documents anciens*. Thèse de doctorat, Institut National des Sciences Appliquées de Lyon, France.
- [Du et al., 2009] DU, X., PAN, W. et BUI, T. D. (2009). Text line segmentation in hand-written documents using mumford-shah model. *Pattern Recognition*, 42(12):3136–3145.
- [Eikvil et al., 1991] EIKVIL, L., TAXT, T. et MOEN, K. (1991). A fast adaptive method for binarization of document image. *In ICDAR*, pages 435–443.
- [Elomary, 1994] ELOMARY, Y. (1994). *Modèles déformables et multirésolution pour la détection de contours de traitement d’images. (Deformable models and multiresolution for edge detection in image processing)*. Thèse de doctorat, Joseph Fourier University, Grenoble, France.
- [Farahmand et al., 2013] FARAHMAND, A., SARRAFZADEH, A. et SHANBEHZADEH, J. (2013). Document image noises and removal methods. *In Proceedings of the International MultiConference of Engineers and Computer Scientists*, volume 1, pages 436–440.
- [Fernando et al., 2010] FERNANDO, B., KARAOGLU, S. et TRÉMEAU, A. (2010). Extreme value theory based text binarization in documents and natural scenes. *In Proceedings of the 3rd International Conference on Machine Vision (ICMV)*, pages 144–151.
- [Fisher et al., 2003] FISHER, R., PERKINS, S., WALKER, A. et WOLFART, E. (2003). Adaptive thresholding. <http://homepages.inf.ed.ac.uk/rbf/HIPR2/adpthrsh.htm>.
- [Garain et al., 2006] GARAIN, U., PAQUET, T. et HEUTTE, L. (2006). On foreground - background separation in low quality document images. *IJDAR*, 8(1):47–63.
- [Garcia et Delakis, 2004] GARCIA, C. et DELAKIS, M. (2004). Convolutional face finder : A neural architecture for fast and robust face detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(11):1408–1423.
- [Gatos et al., 2009] GATOS, B., NTIROGIANNIS, K. et PRATIKAKIS, I. (2009). Icdar 2009 document image binarization contest (dibco 2009). *In ICDAR*, pages 1375–1382. IEEE Computer Society.
- [Gatos et al., 2006] GATOS, B., PRATIKAKIS, I. et PERANTONIS, S. J. (2006). Adaptive degraded document image binarization. *Pattern Recognition*, 39(3):317–327.
- [Gatos et al., 2008] GATOS, B., PRATIKAKIS, I. et PERANTONIS, S. J. (2008). Improved document image binarization by using a combination of multiple binarization techniques and adapted edge information. *In ICPR*, pages 1–4. IEEE Computer Society.
- [Geman et Geman, 1984] GEMAN, S. et GEMAN, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:721–741.
- [Hadjadj et al., 2017] HADJADJ, Z., CHERIET, M., MEZIANE, A. et CHERFA, Y. (2017). A new efficient binarization method : application to degraded historical document images. *Signal, Image and Video Processing*, 11(6):1155–1162.

- [Hadjadj *et al.*, 2014] HADJADJ, Z., MAZIANE, A., CHERIET, M. et CHERFA, Y. (2014). An active contour based method for image binarization : Application to degraded historical document images. *In ICFHR*, pages 655–660. IEEE Computer Society.
- [Hadjadj et Meziane, 2017] HADJADJ, Z. et MEZIANE, A. (2017). Binarization of document images with various object sizes. *In 1st International Workshop on Arabic Script Analysis and Recognition, ASAR 2017, Nancy, France, April 3-5, 2017*, pages 21–25.
- [Hadjadj *et al.*, 2016] HADJADJ, Z., MEZIANE, A., CHERFA, Y., CHERIET, M. et SETITRA, I. (2016). Isauvola : Improved sauvola’s algorithm for document image binarization. *In CAMPILHO, A. et KARRAY, F., éditeurs : ICIAR*, volume 9730 de *Lecture Notes in Computer Science*, pages 737–745. Springer.
- [Hamza *et al.*, 2005] HAMZA, H., BELAÏD, A. et SMIGIEL, E. (2005). Neural based binarization techniques. *In ICDAR*, pages 317–321. IEEE Computer Society.
- [Hedjam, 2013] HEDJAM, R. (2013). *Visual image processing in various representation spaces for documentary heritage preservation*. Thèse de doctorat, École de Technologie Supérieure, Université du Québec, Montréal, Québec, Canada.
- [Hedjam *et al.*, 2010] HEDJAM, R., MOGHADDAM, R. F. et CHERIET, M. (2010). Text extraction from degraded document images. *In EUVIP*, pages 247–252. IEEE.
- [Hedjam *et al.*, 2011] HEDJAM, R., MOGHADDAM, R. F. et CHERIET, M. (2011). A spatially adaptive statistical method for the binarization of historical manuscripts and degraded document images. *Pattern Recognition*, 44(9):2184–2196.
- [Howe, 2011] HOWE, N. R. (2011). A laplacian energy for document binarization. *In ICDAR*, pages 6–10. IEEE Computer Society.
- [Hu *et al.*, 2015] HU, H., LIU, Y. et LIU, M. (2015). Genetic algorithm based image binarization approach and its quantitative evaluation via pooling. *In Proceedings Volume 9811, MIPPR 2015 : Multispectral Image Acquisition, Processing, and Analysis*.
- [Kapur *et al.*, 1985] KAPUR, J. N., SAHOO, P. K. et WONG, A. K. C. (1985). A new method for gray-level picture thresholding using the entropy of the histogram. *Computer Vision, Graphics, and Image Processing*, 29(3):273–285.
- [Kass *et al.*, 1988] KASS, M., WITKIN, A. et TERZOPOULOS, D. (1988). Snakes : Active contour models. *INTERNATIONAL JOURNAL OF COMPUTER VISION*, 1(4):321–331.
- [Kefali, 2015] KEFALI, A. (2015). *Plate-forme pour l’analyse et la restauration des documents anciens*. Thèse de doctorat, Université de Badji Mokhtar- Annaba, Algérie.
- [Ketata et Khemakhem, 2010] KETATA, D. et KHEMAKHEM, M. (2010). Un survol sur l’analyse et la reconnaissance de documents : imprimé, ancien et manuscrit. *In Colloque International Francophone sur l’Ecrit et le Document (CIFED2010)*, page 12pages, sousse, Tunisia.
- [Khashman et Sekeroglu, 2007] KHASHMAN, A. et SEKEROGLU, B. (2007). Global binarization of document images using a neural network. *In YÉTONGNON, K., CHBEIR, R. et DIPANDA, A., éditeurs : SITIS*, pages 665–672. IEEE Computer Society.

- [Khurshid *et al.*, 2009] KHURSHID, K., SIDDIQI, I., FAURE, C. et VINCENT, N. (2009). Comparison of niblack inspired binarization methods for ancient documents. In BERKNER, K. et LIKFORMAN-SULEM, L., éditeurs : *DRR*, volume 7247 de *SPIE Proceedings*, page 72470U. SPIE.
- [Kim *et al.*, 2002] KIM, I. K., JUNG, D.-W. et PARK, R.-H. (2002). Document image binarization based on topographic analysis using a water flow model. *Pattern Recognition*, 35(1):265–277.
- [Kita et Wakahara, 2010] KITA, K. et WAKAHARA, T. (2010). Binarization of color characters in scene images using k-means clustering and support vector machines. In *ICPR*, pages 3183–3186. IEEE Computer Society.
- [Kittler et Illingworth, 1986] KITTLER, J. et ILLINGWORTH, J. (1986). Minimum error thresholding. *Pattern Recognition*, 19(1):41–47.
- [Kricha *et al.*, 2006] KRICHA, A., GHARDALLOU LASMAR, A. et ESSOUKRI BEN AMARA, N. (2006). Exploration des Ondelettes en Prétraitement des Documents Anciens. pages 157–162. SDN06.
- [Kumar *et al.*, 2013] KUMAR, D., PRASAD, M. N. A. et RAMAKRISHNAN, A. G. (2013). Evaluation of document binarization using eigen value decomposition. In ZANIBBI, R. et COÜASNON, B., éditeurs : *DRR*, volume 8658 de *SPIE Proceedings*, page 86580X. SPIE.
- [Lazzara et Géraud, 2014] LAZZARA, G. et GÉRAUD, T. (2014). Efficient multiscale sauvola’s binarization. *IJDAR*, 17(2):105–123.
- [Leventon *et al.*, 2000] LEVENTON, M. E., GRIMSON, W. E. L. et FAUGERAS, O. D. (2000). Statistical shape influence in geodesic active contours. In *CVPR*, pages 1316–1323. IEEE Computer Society.
- [Leydier *et al.*, 2004] LEYDIER, Y., BOURGEOIS, F. L. et EMPTOZ, H. (2004). Serialized k-means for adaptative color image segmentation : Application to document images and others. In MARINAI, S. et DENGEL, A., éditeurs : *Document Analysis Systems*, volume 3163 de *Lecture Notes in Computer Science*, pages 252–263. Springer.
- [Li *et al.*, 2008a] LI, C., KAO, C.-Y., GORE, J. C. et DING, Z. (2008a). Minimization of region-scalable fitting energy for image segmentation. *IEEE Trans. Image Processing*, 17(10):1940–1949.
- [Li *et al.*, 2008b] LI, Y., ZHENG, Y., DOERMANN, D. S. et JAEGER, S. (2008b). Script-independent text line segmentation in freestyle handwritten documents. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30(8):1313–1329.
- [Likforman-Sulem, 2003] LIKFORMAN-SULEM, L. (2003). Apport du traitement des images à la numérisation des documents manuscrits anciens. *Document Numérique*, 7(3-4):13–26.
- [Liu *et al.*, 2016] LIU, M., LIU, Y., HU, H. et NIE, L. (2016). Genetic algorithm and mathematical morphology based binarization method for strip steel defect image with non-uniform illumination. *J. Visual Communication and Image Representation*, 37:70–77.

- [Lu *et al.*, 2010] LU, S., SU, B. et TAN, C. L. (2010). Document image binarization using background estimation and stroke edges. *IJDAR*, 13(4):303–314.
- [Lázaro *et al.*, 2010] LÁZARO, J., MARTÍN, J. L., ARIAS, J., ASTARLOA, A. et CUADRADO, C. (2010). Neuro semantic thresholding using ocr software for high precision ocr applications. *Image Vision Comput.*, 28(4):571–578.
- [Messaoud *et al.*, 2011] MESSAOUD, I. B., AMIRI, H., ABED, H. E. et MÄRGNER, V. (2011). New binarization approach based on text block extraction. *In ICDAR*, pages 1205–1209. IEEE Computer Society.
- [Meziou, 2013] MEZIOU, L. I. (2013). *Active contours segmentation based on alpha-divergences : Segmentation of medical and biomedical images*. Theses, Université de Cergy Pontoise.
- [Moghaddam et Cheriet, 2009a] MOGHADDAM, R. F. et CHERIET, M. (2009a). Application of multi-level classifiers and clustering for automatic word spotting in historical document images. *In ICDAR*, pages 511–515. IEEE Computer Society.
- [Moghaddam et Cheriet, 2009b] MOGHADDAM, R. F. et CHERIET, M. (2009b). Rslidi : Restoration of single-sided low-quality document images. *Pattern Recognition*, 42(12): 3355–3364.
- [Moghaddam et Cheriet, 2010] MOGHADDAM, R. F. et CHERIET, M. (2010). A multi-scale framework for adaptive binarization of degraded document images. *Pattern Recognition*, 43(6):2186–2198.
- [Moghaddam et Cheriet, 2012] MOGHADDAM, R. F. et CHERIET, M. (2012). Adotsu : An adaptive and parameterless generalization of otsu’s method for document image binarization. *Pattern Recognition*, 45(6):2419–2431.
- [Moghaddam *et al.*, 2009] MOGHADDAM, R. F., RIVEST-HÉNAULT, D. et CHERIET, M. (2009). Restoration and segmentation of highly degraded characters using a shape-independent level set approach and multi-level classifiers. *In ICDAR*, pages 828–832. IEEE Computer Society.
- [Mumford et Shah, 1989] MUMFORD, D. et SHAH, J. (1989). Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. on Pure and Applied Mathematics*, 42(5):577–685.
- [Niblack, 1986] NIBLACK, W. (1986). *An Introduction to Digital Image Processing*. Prentice Hall, Englewood Cliffs.
- [Ntirogiannis *et al.*, 2009] NTIROGIANNIS, K., GATOS, B. et PRATIKAKIS, I. (2009). A modified adaptive logical level binarization technique for historical document images. *In ICDAR*, pages 1171–1175. IEEE Computer Society.
- [Ntirogiannis *et al.*, 2014] NTIROGIANNIS, K., GATOS, B. et PRATIKAKIS, I. (2014). A combined approach for the binarization of handwritten document images. *Pattern Recognition Letters*, 35:3–15.
- [Osher et Sethian, 1988] OSHER, S. et SETHIAN, J. A. (1988). Fronts propagating with curvature-dependent speed : Algorithms based on hamilton-jacobi formulations. *Journal of Computational Physics*, 79(1):12–49.

- [Otsu, 1979] OTSU, N. (1979). A thresholding selection method from gray-level histograms. 9(1):62–66.
- [Papamarkos, 2003] PAPAMARKOS, N. (2003). A neuro-fuzzy technique for document binarisation. *Neural Computing and Applications*, 12(3-4):190–199.
- [Papamarkos *et al.*, 2000] PAPAMARKOS, N., STROUTHOPOULOS, C. et ANDREADIS, I. (2000). Multithresholding of color and gray-level images through a neural network technique. *Image Vision Comput.*, 18(3):213–222.
- [Parker *et al.*, 1993] PARKER, J. R., JENNINGS, C. et SALKAUSKAS, A. G. (1993). Thresholding using an illumination model. *In ICDAR*, pages 270–273. IEEE.
- [Pratikakis, 2011] PRATIKAKIS, I. (2011). Dibco 2011 binarization results. <http://utopia.duth.gr/~ipratika/DIBC02011/dibco2011results.htm/>.
- [Pratikakis *et al.*, 2010] PRATIKAKIS, I., GATOS, B. et NTIROGIANNIS, K. (2010). H-dibco 2010 - handwritten document image binarization competition. *In ICFHR*, pages 727–732. IEEE Computer Society.
- [Pratikakis *et al.*, 2011] PRATIKAKIS, I., GATOS, B. et NTIROGIANNIS, K. (2011). Icdar 2011 document image binarization contest (dibco 2011). *In ICDAR*, pages 1506–1510. IEEE Computer Society.
- [Pratikakis *et al.*, 2012] PRATIKAKIS, I., GATOS, B. et NTIROGIANNIS, K. (2012). Icfhr 2012 competition on handwritten document image binarization (h-dibco 2012). *In ICFHR*, pages 817–822.
- [Rabeux, 2013] RABEUX, V. (2013). *Évaluation de la qualité des documents anciens numérisés*. Thèse de doctorat, University of Bordeaux, France.
- [Rangoni *et al.*, 2009] RANGONI, Y., SHAFAIT, F. et BREUEL, T. (2009). Ocr based thresholding. *In Proceedings of IAPR Conference on Machine Vision Applications*, pages 98–101.
- [Rivest-Hénault *et al.*, 2012] RIVEST-HÉNAULT, D., MOGHADDAM, R. F. et CHERIET, M. (2012). A local linear level set method for the binarization of degraded historical document images. *IJDAR*, 15(2):101–124.
- [Saidane et Garcia, 2007] SAIDANE, Z. et GARCIA, C. (2007). Robust binarization for video text recognition. *In ICDAR*, pages 874–879. IEEE Computer Society.
- [Sauvola et Pietikäinen, 2000] SAUVOLA, J. J. et PIETIKÄINEN, M. (2000). Adaptive document image binarization. *Pattern Recognition*, 33(2):225–236.
- [Sezgin et Sankur, 2004] SEZGIN, M. et SANKUR, B. (2004). Survey over image thresholding techniques and quantitative performance evaluation. *J. Electronic Imaging*, 13(1):146–168.
- [Singh *et al.*, 2012] SINGH, T. R., ROY, S., SINGH, O. I., SINAM, T. et SINGH, K. M. (2012). A new local adaptive thresholding technique in binarization. *CoRR*, abs/1201.5227.
- [Stathis *et al.*, 2008] STATHIS, P., KAVALLIERATOU, E. et PAPAMARKOS, N. (2008). An evaluation technique for binarization algorithms. *J. UCS*, 14(18):3011–3030.

- [Su *et al.*, 2010] SU, B., LU, S. et TAN, C. L. (2010). Binarization of historical document images using the local maximum and minimum. *In* DOERMANN, D. S., GOVINDARAJU, V., LOPRESTI, D. P. et NATARAJAN, P., éditeurs : *Document Analysis Systems*, ACM International Conference Proceeding Series, pages 159–166. ACM.
- [Su *et al.*, 2011] SU, B., LU, S. et TAN, C. L. (2011). Combination of document image binarization techniques. *In* *ICDAR*, pages 22–26. IEEE Computer Society.
- [Su *et al.*, 2013] SU, B., LU, S. et TAN, C. L. (2013). Robust document image binarization technique for degraded document images. *IEEE Trans. Image Processing*, 22(4):1408–1417.
- [Trier et Taxt, 1995] TRIER, i. D. et TAXT, T. (1995). Evaluation of binarization methods for document images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 17(3):312–315.
- [van Herk, 1992] van HERK, M. (1992). A fast algorithm for local minimum and maximum filters on rectangular and octagonal kernels. *Pattern Recognition Letters*, 13(7):517–521.
- [Vese et Chan, 2002] VESE, L. A. et CHAN, T. F. (2002). A multiphase level set framework for image segmentation using the mumford and shah model. *International Journal of Computer Vision*, 50(3):271–293.
- [Wakahara et Kita, 2011] WAKAHARA, T. et KITA, K. (2011). Binarization of color character strings in scene images using k-means clustering and support vector machines. *In* *ICDAR*, pages 274–278. IEEE Computer Society.
- [Wolf, 2010] WOLF, C. (2010). Document ink bleed-through removal with two hidden markov random fields and a single observation field. *IEEE Trans. Pattern Anal. Mach. Intell.*, 32(3):431–447.
- [Xu et Prince, 1997] XU, C. et PRINCE, J. L. (1997). Gradient vector flow : A new external force for snakes. *In* *CVPR*, pages 66–71. IEEE Computer Society.
- [Xu et Prince, 1998] XU, C. et PRINCE, J. L. (1998). Snakes, shapes, and gradient vector flow. *IEEE Trans. Image Processing*, 7(3):359–369.
- [Yang *et al.*, 1994] YANG, J.-D., CHEN, Y.-S. et HSU, W.-H. (1994). Adaptive thresholding algorithm and its hardware implementation. *Pattern Recognition Letters*, 15(2):141–150.
- [Yang et Yan, 2000] YANG, Y. et YAN, H. (2000). An adaptive logical method for binarization of degraded document images. *Pattern Recognition*, 33(5):787–807.
- [Yosef *et al.*, 2009] YOSEF, I. B., MOKEICHEV, A., KEDEM, K., DINSTEIN, I. et EHR- LICH, U. (2009). Adaptive shape prior for recognition and variational segmentation of degraded historical characters. *Pattern Recognition*, 42(12):3348–3354.