

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

وزارة التعليم العالي و البحث العلمي

BADJI MOKHTAR-ANNABA UNIVERSITY

UNIVERSITY BADJI MOKHTAR-ANNABA



جامعة باجي مختار - عنابة

Faculté des Sciences de l'Ingéniorat
Département d'Informatique

Année : 2014/2015

THESE

Présentée en vue de l'obtention du diplôme de

Doctorat en Sciences

Intitulée :

**Plate-forme pour l'analyse et la restauration des
documents anciens**

Option

Intelligence Artificielle

Par

Abderrahmane KEFALI

Devant le Jury

Pr. Mokhtar SELLAMI	Professeur	Université d'Annaba	Président
Pr. Halima BAH-ABIDET	Professeur	Université d'Annaba	Directrice de Thèse
Pr. Hamid SERIDI	Professeur	Université de Guelma	Examineur
Pr. Mohamed Chawki BATOUCH	Professeur	Université de Constantine 2	Examineur
Pr. Abdelouahab MOUSSAOUI	Professeur	Université de Sétif	Examineur
Dr. Toufik SARI	MCB	Université d'Annaba	Invité

ملخص

تتكسر اليوم مجموعات كبيرة من الوثائق والمستندات بمختلف أنواعها في المكتبات والمتاحف بالإضافة إلى مؤسسات أخرى ذات طبيعة بيروقراطية اجتماعية أو سياسية. الوثائق التاريخية الخاصة بالحضارات القديمة والأرشفات الوثائقية تعتبر أحسن مثال على هذا التراث الذي يمثل التراث، التاريخ وهوية الشعوب. المحافظة إذن على هاته الوثائق تعتبر اليوم حاجة ملحة. تعد رقمنة الوثائق وحفظها بشكل صور واحدة من طرق المحافظة عليها، لكن الرقمنة وحدها ليست كافية، بل يجب أن تكون مرفوقة بأدوات تسمح بمعالجتها وتحليلها آلياً. كذلك، عملية الرقمنة تتم غالباً صور ذات نوعية رديئة، بخلفية غير متجانسة، مع آثار شفافية الورق وهذا بالرغم من التقدم الكبير في تقنيات الرقمنة والتصوير. في الحقيقة أغلبية التشوهات في الصور ترجع إلى تلف الوثائق الأصلية نفسها. جميع هاته المشاكل شكلت تحدياً للباحثين في مجال تحليل الوثائق. بغض النظر عن الهدف من التحليل سواء تحديد المحتوى (التعرف على الكتابة، الرموز أو الأشياء، تحديد الكاتب)، الفهرسة والبحث، أو بكل مساهمة الترتيب، التنظيم والتبويب، يجب علينا المرور بمرحلة المعالجة الأولية. هذه المرحلة تضم مجموعة من العمليات التي تطبق على صور الوثائق المرقمنة من أجل تحضيرها للمراحل التالية في مسار التحليل.

العمل المقدم في هذه المذكرة يدخل في إطار مسعى المحافظة وتقييم الوثائق الموروثة التي أدرك المجتمع الدولي أهميتها. نهتم في هذا العمل بالمعالجة الأولية لصور الوثائق التاريخية، وبدرجة أكثر فصل الأمامية عن الخلفية والتي تعتبر واحدة من أهم عمليات المعالجة الأولية. قدمنا في هذه المذكرة أربع مشاركات منظمة حسب محورين أساسيين. المشاركة الأولى هي مشاركة من مشاركتين هما مشاركة في تقييم تقنيات فصل الأمامية عن الخلفية، وبأكثر دقة تقنيات التثنية. المشاركة الأولى تشمل القيام بدراسة مقارنة مع تقييم كمي لعدد مهم من تقنيات التثنية من أجل تحديد التقنية الأحسن لاستخراج النصوص من صور الوثائق التاريخية. المشاركة الثانية هي اقتراح طريقة جديدة لتقييم تقنيات التثنية من أجل البحث في صور الوثائق العربية القديمة. في كل واحدة من المشاركتين الأخيرتين تم اقتراح طريقة جديدة لتثنية صور الوثائق القديمة. الأولى هي طريقة تثنية تعتمد على تصنيف البكسل إلى أسود وأبيض باستعمال شبكة عصبية آلية، في حين أن الطريقة الثانية عبارة عن تقنية هجينة للتثنية تنفذ في مرحلتين وتضم عدة طرق تثنية من أجل الحصول على نتائج أفضل. في الأربع أعمال السابقة قمنا باستعمال مجموعتين للتجربة. المجموعة الأولى مؤلفة من صور حقيقية عامة، والثانية عبارة عن مجموعة صور تركيبية. قمنا بعدة تجارب والنتائج المتحصل عليها مشجعة.

كلمات مفاتيح: المعالجة الأولية للوثائق القديمة، التثنية، فصل الأمامية عن الخلفية، تقييم النتائج، الوثائق العربية القديمة، الشبكات العصبية الآلية.

ABSTRACT

Currently, important numerical documentary collections exist in the libraries, museum and other institutions in pedagogic or sociopolitical nature. Historical documents of old civilizations and public archives are the typical example of such richness which represent the heritage, history and dignity of nations. Preserving these documents is now an essential need. A method of conservation consists to digitize and to save them in image format, but only it will not be sufficient, it must be supported by computer programs allowing there analysis and recognition. However, the digitization process often generates images of poor quality with heterogeneous background, transparency effects, etc. despite of the spectacular advances in the capturing technology. In fact, most of these distortions in the image are due to the deterioration suffered by the original items. All these problems yielded a new challenge to researchers in the field of document analysis. Whatever you are interested in the content recognition (handwriting, symbols or objects, writer ...), indexing and search, or simply classification, organization and cataloging, you will inevitably passe through a preprocessing step. This step consists of a set of procedures applied on the images of scanned documents in order to prepare them to the next steps in the analysis.

The work presented in this manuscript integrates the attempt of saving and valorizing patrimonial documents of witch the international community renewed with their interests. We are interested in this work to the preprocessing of historical document images, specifically the foreground/ background separation considered as one of the most important preprocessing steps. Four contributions arranged according to two axes are presented in this thesis. The two firsts are contributions to the performance evaluation of binarization techniques. Thus, the first contribution is an objective comparative study of a large number of binarization algorithms to determine the most efficient one for text extraction in historical document images. The second contribution consists to the proposition of a new evaluation methodology of binarization methods for searching images of degraded Arabic documents. In each one of the two lasts contributions, a new binarization method of degraded document images is proposed. The first method is a technique based on the classification of the pixels into black and white using an artificial neural network, especially an MLP (multi-layer Perceptron). The second method is a hybrid binarization technique that proceeds on two steps and combines several thresholding methods in order to have a more accurate result. Two test sets have been used in the four contributions. The first is a public set of real images, and the second is a collection of synthetic images. Various experiments have been carried out and the results are encouraging.

Keywords: degraded document preprocessing, binarization, foreground/ background separation, performance evaluation, old Arabic documents, artificial neural networks.

RESUME

D'importants fonds documentaires existent actuellement dans les bibliothèques, musées et autres institutions à caractères pédagogiques ou sociopolitiques. Les documents historiques des civilisations anciennes et les archives nationales sont l'exemple typique de telles richesses qui représentent le patrimoine, l'histoire et la dignité des nations. La conservation de ces documents constitue aujourd'hui un besoin incontournable. Une méthode de conservation consiste à les numériser et les stocker ainsi sous format image, mais seule, la numérisation n'est pas suffisante, elle doit être accompagnée d'outils informatiques permettant leur analyse et reconnaissance. Cependant, le processus de numérisation génère souvent des images de mauvaise qualité, à fond hétérogène, avec effet de transparence, etc. et ce malgré les avancées spectaculaires de la technologie des outils de capture numérique. En fait la plupart de ces déformations au niveau de l'image sont dues aux détériorations subies par les originaux eux-mêmes. Tous ces problèmes constituent un nouveau défi aux chercheurs dans le domaine de l'analyse de documents. Que l'on fait de l'identification du contenu (écriture, symboles ou objets, scripteur), l'indexation et la recherche, ou plus simplement la classification, l'organisation et le 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.

Le travail présenté dans ce mémoire s'intègre dans la démarche de sauvegarde et de valorisation de documents patrimoniaux dont la communauté internationale a pris récemment conscience de l'intérêt. Nous nous intéressons dans le présent travail aux prétraitements d'images de documents historiques, et plus précisément à la séparation avant-plan/ arrière-plan considérée comme l'une des opérations de prétraitement les plus importantes. Quatre contributions arrangées selon deux axes sont présentées dans ce mémoire. Les deux premières sont des contributions à l'évaluation des performances des techniques de binarisation. Ainsi, la première contribution s'agit d'une étude comparative d'un grand nombre d'algorithmes de binarisation afin de déterminer l'algorithme le plus performant pour l'extraction du texte dans les images de documents historiques. La deuxième contribution consiste à la proposition d'une nouvelle méthodologie d'évaluation des méthodes de binarisation pour la recherche dans les images de documents arabes dégradés. Dans chacune des deux dernières contributions, une nouvelle méthode de binarisation d'images de documents dégradés est proposée. La première est une technique basée sur la classification des pixels en noirs et blancs en utilisant un réseau de neurones artificiel de type PMC (Perceptron multi couches). La deuxième s'agit d'une technique de binarisation hybride qui procède en deux étapes et qui combine plusieurs méthodes de seuillage dont le but d'avoir un résultat plus sûr. Deux ensembles d'images de test ont été utilisés dans les quatre contributions. Le premier est un ensemble public d'images réelles, et le deuxième est une collection d'images synthétiques. Différentes expérimentations ont été effectuées et les résultats obtenus sont encourageants.

Mots clés : prétraitement de documents dégradés, binarisation, séparation avant plan/ arrière plan, évaluation des performances, documents arabes anciens, réseaux de neurones artificiels.

REMERCIEMENTS

Mes premiers remerciements vont à Pr. Halima BAHI-ABIDET, la directrice de ce travail. Je suis heureux d'avoir eu la chance durant ces dernières années de partager ses réflexions, de disposer de son expertise scientifique et de son expérience pour mener à bien ce travail.

Ce fût vraiment un très grand plaisir d'avoir été co-encadré pour la troisième fois, par Dr. Toufik SARI, qu'il trouve ici l'expression de ma plus profonde reconnaissance et gratitude pour son aide précieux. Je le remercie pour tout le temps qu'il m'a consacré, pour ses conseils, pour sa disponibilité à tout moment, pour sa gentillesse, pour son aide illimité, pour son soutien moral, et pour la confiance qu'il m'a témoignée.

Je remercie vivement Mr. Mokhtar SELLAMI, Professeur à l'université Badji Mokhtar – Annaba pour m'avoir fait l'honneur de présider mon jury de thèse. Je tiens également à remercier Mr. Hamid SERIDI, Professeur à l'université de Guelma, Mr. Mohamed Chawki BATOUCHE, Professeur à l'université de Constantine 2, et Mr. Abdelouahab MOUSSAOUI, Professeur à l'université de Sétif qui m'ont tous fait l'honneur d'accepter de prendre part à ce jury et surtout de lire et d'expertiser mon travail.

Je remercie également mes chers amis : Ismail BOUACHA, Mohamed CHAOUI, Réda LAIFA, et Riad BOUACHA pour leur aide.

Les plus grands remerciements vont à mes parents pour leur soutien tout au long de mes études.

Merci à tous ceux qui ont contribué de près ou de loin à l'élaboration de ce travail.

DEDICACES

Je dédie ce mémoire aux deux personnes que je chérisse le plus au monde :

Ma mère et mon père.

A la mémoire de mon grand-père REBIAI.

A mon frère Fateh et à ma sœur Mouna.

A mon cher oncle Faouzi LOUNDI

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

Liste des Tableaux

Tableau 3. 1. Evaluation analytique des méthodes de binarisation dans [KAM 93]	72
Tableau 3. 2. Travaux antérieurs sur l'évaluation des techniques de binarisation.....	91
Tableau 4. 1. Répartition d'images de documents dans les collections publiques de DIBCO et H-DIBCO	99
Tableau 4. 2. Valeurs optimales des paramètres des méthodes de binarisation inclus dans l'évaluation	110
Tableau 4. 3. Résultats d'évaluation moyens des algorithmes de binarisation inclus dans l'étude comparative	112
Tableau 4. 4. Classement final et partiel des algorithmes de binarisation inclus dans l'étude comparative	114
Tableau 4. 5. Performances de la méthode de Khashman et Sekeroglu [KHA 07] pour différentes images	119
Tableau 4. 6. Résultats de binarisation d'une image exemple par différentes méthodes de binarisation, et les signatures extraites des images binarisées.....	136
Tableau 4. 7. Résultats d'évaluation (visuelle et en termes de distance d'édition) des méthodes de binarisation sur une image exemple.....	137
Tableau 4. 8. Résultats d'évaluation (visuelle et en termes de distance d'édition) des algorithmes de binarisation sur le premier ensemble de test.....	138
Tableau 4. 9. Résultats d'évaluation (visuelle, supervisé-pixel, et en termes de distance d'édition) des algorithmes de binarisation sur le deuxième ensemble de test	138
Tableau 4. 10. Résultats d'évaluation (supervisée-pixel et en termes de distance d'édition) de la qualité des deux images de la figure 4.29.....	141
Tableau 5. 1. Résultats moyens obtenus du 1 ^{er} MLP sur les quatre collections publiques d'images...	155
Tableau 5. 2. Résultats moyens obtenus par le 1 ^{er} MLP sur la collection d'images synthétiques	155
Tableau 5. 3. Résultats moyens obtenus par le 2 ^{ème} MLP sur les collections publiques et synthétique	156
Tableau 5. 4. Résultats moyens des différentes méthodes de binarisation sur l'ensemble d'images publiques	157
Tableau 5. 5. Résultats moyens des différentes méthodes de binarisation sur l'ensemble d'images synthétiques	158
Tableau 5. 6. Résultats moyens des différentes méthodes de binarisation entre les deux ensembles..	159
Tableau 5. 7. Classement final des méthodes comparées sur les deux ensembles de test.....	159
Tableau 5. 8. Performances moyennes obtenues par les différentes méthodes de binarisation en utilisant notre méthodologie d'évaluation	163
Tableau 5. 9. Résultats moyens des différentes méthodes de binarisation sur l'ensemble d'images publiques	172
Tableau 5. 10. Résultats moyens des différentes méthodes de binarisation sur l'ensemble d'images synthétiques	173
Tableau 5. 11. Résultats moyens des différentes méthodes de binarisation entre les deux ensembles	173
Tableau 5. 12. Classement final des méthodes comparées sur les deux ensembles de test.....	174
Tableau 5. 13. Performances moyennes obtenues par les différentes méthodes de binarisation en utilisant notre méthodologie d'évaluation	178

Liste des Figures

Figure 1. Exemples de documents dégradés	3
Figure 2. Schéma général du processus d'analyse et de reconnaissance de documents [MAR 08].....	4
Figure 3. Exemples de défauts engendrés par la numérisation [RAB 13]	5
Figure 4. Séparation idéale avant-plan / arrière-plan d'une image de document dégradé	6
Figure 1. 1. Exemples de dégradations affectant les documents anciens	13
Figure 1. 2. Exemples de dégradations dues à la numérisation	14
Figure 1. 3. Résultat du prétraitement d'une image de document dégradé tirée de [DRI 07] ..	15
Figure 1. 4. Position du prétraitement dans le processus d'analyse et de reconnaissance de documents selon [KET 10]	16
Figure 1. 5. Suppression du bruit marginal.....	19
Figure 1. 6. Suppression de l'effet de transparence	20
Figure 1. 7. Restauration de la forme d'un caractère [ALL 03].....	21
Figure 1. 8. Résultat de correction de la courbure [RAB 13]	23
Figure 1. 9. Correction de l'inclinaison d'un document mal orienté	24
Figure 1. 10. Segmentation texte/ graphique d'un document ancien	28
Figure 1. 11. Extraction des lignes d'un texte manuscrit [OUW 10].....	29
Figure 2. 1. Exemple de binarisation d'une image en niveaux de gris.....	35
Figure 2. 2. Séparation avant-plan/ arrière-plan d'un document couleur [GAR 06]	36
Figure 3. 1. Catégorisation des techniques d'évaluation des algorithmes de segmentation selon [ZHA 96].....	70
Figure 3. 2. Hiérarchie des méthodes d'évaluation des algorithmes de segmentation selon [ZHA 08].....	70
Figure 4. 1. Influence de la qualité de binarisation sur le résultat de l'OCR.....	96
Figure 4. 2. Image extraite de la collection DIBCO 2009 avec l'image de vérité terrain correspondante	99
Figure 4. 3. Image extraite de la collection H-DIBCO 2010 avec l'image de vérité terrain correspondante	100
Figure 4. 4. Image extraite de la collection DIBCO 2011 avec l'image de vérité terrain correspondante	100
Figure 4. 5. Image extraite de la collection H-DIBCO 2012 avec l'image de vérité terrain correspondante	101
Figure 4. 6. Image extraite de la collection DIBCO 2013 avec l'image de vérité terrain correspondante	101
Figure 4. 7. Exemples d'images de référence utilisées pour la génération d'images de documents synthétiques	102
Figure 4. 8. Quelques images de fonds anciens utilisées pour la génération d'images de documents synthétiques	103
Figure 4. 9. Image de document synthétique obtenue par la combinaison de l'image de référence de la figure 4.7.b et de l'image du fond de la figure 4.8.b	105

Figure 4. 10. Image de document synthétique obtenue de la fusion par étirement du masque	106
Figure 4. 11. Performances des méthodes de binarisation incluses dans l'étude comparative	117
Figure 4. 12. Résultats de binarisation de différentes images par la méthode de Khashman et Sekeroglu [KHA 07].....	118
Figure 4. 13. Résultats de binarisation de différentes images en utilisant la méthode de Tabatabai et al. [TAB 10]	121
Figure 4. 14. Schéma général de la méthodologie d'évaluation proposée	124
Figure 4. 15. Caractéristiques structurelles présentes dans un mot arabe	125
Figure 4. 16. Schéma général de la méthodologie d'évaluation proposée en utilisant des images synthétiques.....	127
Figure 4. 17. Pixel courant P_0 et ses voisins	128
Figure 4. 18. Lissage d'une image de mot binarisée	128
Figure 4. 19. Résultat de redressement d'une image de document incliné.....	129
Figure 4. 20. Résultat de segmentation en lignes d'une image de document	129
Figure 4. 21. Ligne de texte segmentée en composantes connexes	130
Figure 4. 22. Ligne de base et zone médiane détectées sur une ligne de texte	131
Figure 4. 23. Suivi de contours d'un objet.....	131
Figure 4. 24. Caractéristiques extraites sur une ligne de texte.....	132
Figure 4. 25. Exemples d'images de documents historiques réelles utilisées pour tester la méthodologie d'évaluation proposée.....	134
Figure 4. 26. Exemples d'images de documents historiques synthétiques utilisées pour tester la méthodologie d'évaluation proposée	135
Figure 4. 27. Performances des méthodes de binarisation sur le premier ensemble.....	139
Figure 4. 28. Performances des méthodes de binarisation sur le deuxième ensemble.....	139
Figure 4. 29. Exemple d'une image de référence, de sa dilatation et de son érosion	141
Figure 5. 1. Livres et pages historiques	146
Figure 5. 2. MLP pour classifier un pixel p dans une fenêtre de 3×3 pixels	147
Figure 5. 3. Schéma général de notre première méthode de binarisation proposée	148
Figure 5. 4. Vecteur d'entrées pour un pixel dans une fenêtre 3×3	149
Figure 5. 5. Prétraitement et apprentissage du MLP	150
Figure 5. 6. Utilisation du MLP entraîné pour binariser une nouvelle image	152
Figure 5. 7. Erreurs d'apprentissage et de validation du MLP	154
Figure 5. 8. Résultats de binarisation de quelques images publiques en utilisant le 2 ^{ème} MLP	157
Figure 5. 9. Performances des méthodes de binarisation comparées en termes de FMeasure	160
Figure 5. 10. Performances des méthodes de binarisation comparées en termes de PSNR ...	161
Figure 5. 11. Performances des méthodes de binarisation comparées en termes de NRM ...	161
Figure 5. 12. Performances des méthodes de binarisation comparées en termes de MPM ...	162
Figure 5. 13. Graphe montrant les performances des méthodes de binarisation comparées en termes de DRD	162
Figure 5. 14. Seuillage global d'une image bimodale	166
Figure 5. 15. Seuillage global d'une image de mauvaise qualité	167


Figure 5. 16. Seuillage global en utilisant les deux seuils T_1 et T_2 de l'image de la figure 5.15.a	169
Figure 5. 17. Résultat final de binarisation de l'image de la figure 5.15.a en utilisant notre technique de seuillage hybride proposée	170
Figure 5. 18. Performances des méthodes de binarisation comparées en termes du temps d'exécution.....	174
Figure 5. 19. Performances des méthodes de binarisation comparées en termes de Fmeasure.	175
Figure 5. 20. Performances des méthodes de binarisation comparées en termes de PSNR ...	175
Figure 5. 21. Performances des méthodes de binarisation comparées en termes de NRM	176
Figure 5. 22. Performances des méthodes de binarisation comparées en termes de MPM	176
Figure 5. 23. Graphe montrant les performances des méthodes de binarisation comparées en termes de DRD	177

Table des matières

ملخص	i
ABSTRACT	ii
RESUME	iii
REMERCIEMENTS	iv
DEDICACES	v
Liste des Tableaux	vi
Liste des Figures	vii
Table des matières	x
Introduction générale	1
Contexte et problématique.....	2
Contribution.....	5
Organisation du mémoire	7
Partie I : Etat de l'art	9
Chapitre 1. Prétraitement d'images de documents	10
1.1. Introduction	11
1.2. Dégradations d'images de documents	11
1.2.1. Qu'est ce qu'une dégradation.....	11
1.2.2. Les différents types de dégradations	12
1.2.2.1. Dégradations des documents originaux	12
1.2.2.2. Dégradations dues à la numérisation	13
1.3. Prétraitement d'images de documents.....	15
1.4. Les opérations du prétraitement les plus courantes	16
1.4.1. Identification du bruit	17
1.4.2. Débruitage et Restauration.....	18
1.4.2.1. Suppression du bruit Sel et poivre	18
1.4.2.2. Suppression du bruit marginal.....	18
1.4.2.3. Séparation entre le recto et le verso	20
1.4.2.4. Restauration des formes de caractères	21
1.4.2.5. Correction de la courbure.....	22
1.4.3. Séparation avant-plan/ arrière-plan	23
1.4.4. Correction de l'inclinaison (redressement)	24
1.4.5. Segmentation.....	26
1.4.5.1. Familles des techniques de segmentation.....	27
1.4.5.2. Segmentation texte/ graphique	28
1.4.5.3. Segmentation du texte en lignes	28
1.4.5.4. Segmentation du mot en caractères	31
1.5. Conclusion.....	32
Chapitre 2 : Séparation avant-plan/ arrière-plan	33
2.1. Introduction	34
2.2. Séparation avant-plan/ arrière-plan, binarisation, et seuillage.....	34
2.3. Différentes approches de séparation avant-plan/ arrière-plan.....	36
2.4. Binarisation basée sur le seuillage	38
2.4.1. Méthodes de seuillage global	38
2.4.1.1. Méthodes de seuillage naïves	39
2.4.1.2. Méthodes basées sur l'entropie	39

2.4.1.3. Méthodes basées sur la similarité d'attributs.....	43
2.4.1.4. Méthodes basées sur la forme de l'histogramme.....	44
2.4.1.5. Méthodes de seuillage basées sur le regroupement des niveaux de gris	44
2.4.1.6. Méthodes de seuillage spatial.....	47
2.4.2. Méthodes de seuillage local	48
2.4.2.1. Méthodes basées sur la variance locale.....	48
2.4.2.2. Méthodes basées sur l'intervalle de gris local.....	52
2.4.2.3. Techniques utilisant les informations du voisinage.....	53
2.4.2.4. Seuillage Hiérarchique.....	55
2.4.3. Méthodes de seuillage hybride.....	56
2.4.3.1. Méthodes basées sur la morphologie.....	56
2.4.3.2. Méthodes de binarisation à deux niveaux	57
2.5. Binarisation basée sur le regroupement.....	59
2.5.1. Méthodes de regroupement basées sur la classification	59
2.5.1.1. Binarisation en utilisant les ANNs	59
2.5.1.2. Binarisation en utilisant l'algorithme K-means.....	60
2.5.1.3. Binarisation en utilisant les SVMs	61
2.5.1.4. Binarisation en utilisant les arbres de décision.....	62
2.5.2. Méthodes de regroupement basées sur la région.....	62
2.6. Conclusion.....	62
Chapitre 3 : Evaluation et comparaison des techniques de binarisation.....	64
3.1. Introduction	65
3.2. Définitions	66
3.2.1. Evaluation	66
3.2.2. Comparaison	67
3.2.3. Différence entre l'évaluation de la binarisation et la comparaison des binarisations.....	68
3.3. Les différentes approches d'évaluation de la qualité de binarisation.....	68
3.3.1. Catégorisation des approches d'évaluation.....	68
3.3.2. Approches d'évaluation de la binarisation.....	70
3.3.2.1. Evaluation subjective.....	70
3.3.2.2. Evaluation au niveau système	71
3.3.2.3. Evaluation analytique	72
3.3.2.4. Evaluation empirique non-supervisée.....	72
3.3.2.5. Evaluation empirique supervisée.....	73
3.4. Mesures d'évaluation empiriques	74
3.4.1. Mesures empiriques non-supervisées	74
3.4.2. Mesures empiriques supervisées	78
3.5. Travaux antérieurs d'évaluation.....	86
3.6. Conclusion.....	92
Partie II : Contributions	93
Chapitre 4 : Contributions à l'évaluation des techniques de binarisation	94
4.1. Introduction	95
4.2. Evaluation objective de plusieurs techniques de binarisation d'images de documents dégradés ...	96
4.2.1. Motivations et objectifs	96
4.2.2. Base d'images utilisée	98
4.2.3. Méthodes de binarisation incluses dans l'évaluation.....	106
4.2.4. Mesures d'évaluation utilisées	108
4.2.5. Expérimentations et résultats	109

4.2.6. Conclusion	122
4.3. Une nouvelle méthode d'évaluation des techniques de binarisation pour la recherche dans les images de documents arabes dégradés	122
4.3.1. Motivations et objectifs	122
4.3.2. Description de l'approche proposée	123
4.3.2.1. Préparation des signatures de vérité terrain.....	125
4.3.2.2. Traitement des images de documents binarisées	127
4.3.2.3. Comparaison.....	132
4.3.3. Expérimentations et résultats	133
4.3.3.1. Méthodes de binarisation évaluées	133
4.3.3.2. Images de test	134
4.3.3.3. Procédure d'évaluation	135
4.3.3.4. Résultats obtenus	136
4.3.4. Conclusion	141
4.4. Conclusion du chapitre.....	142
Chapitre 5 : Contributions à la binarisation d'images de documents dégradés	143
5.1. Introduction	144
5.2. Séparation avant-plan/ arrière-plan dans les images de documents dégradés en utilisant un réseau de neurones artificiel.....	144
5.2.1. Motivations et objectifs	144
5.2.2. Description de la technique proposée.....	145
5.2.2.1. Prétraitement de l'image.....	147
5.2.2.2. Définition du MLP.....	151
5.2.2.3. Apprentissage du MLP	151
5.2.2.4. Binarisation	152
5.2.3. Expérimentations et résultats	152
5.2.3.1. Première expérimentation	153
5.2.3.2. Deuxième expérimentation	155
5.2.3.3. Troisième expérimentation.....	163
5.2.4. Conclusion	164
5.3. Une nouvelle méthode hybride de seuillage pour la séparation texte/ fond dans les images de documents dégradés	165
5.3.1. Motivations et objectifs	165
5.3.2. Description de la technique proposée	168
5.3.2.1. Seuillage global	168
5.3.2.2. Seuillage local des pixels restants.....	170
5.3.3. Expérimentations et résultats	171
5.3.3.1. Base de test.....	171
5.3.3.2. Première expérimentation	172
5.3.3.3. Deuxième expérimentation	178
5.3.4. Conclusion	179
5.4. Conclusion du chapitre.....	179
Conclusion générale et perspectives	180
Bilan	181
Perspectives	183
Bibliographie.....	185
Liste des publications.....	205

A scroll of aged parchment with a slightly textured, yellowish-brown surface. The top and bottom edges are rolled up, showing the thickness of the paper. The text is centered and written in a white, gothic-style font. The background is a plain white surface.

Introduction générale

Contexte et problématique

L'homme a vécu depuis sa création en groupes et en sociétés. Puisque il est sociable par sa nature, il a eu toujours besoin de communiquer avec les autres, et par conséquent il a eu besoin des moyens d'expression et de communication qui lui permettent de s'exprimer, de présenter ses idées et besoins, et de transmettre des informations, des connaissances, et des savoirs. Le premier mode de communication que l'être humain a connu est le langage oral qui est un système de signaux vocaux de perception auditive [SOU 02]. Alors que l'oral constitue à nos jours le mode de communication le plus utilisé, il comporte des limites relatives au caractère éphémère de sa matérialisation. La communication ne peut avoir lieu que dans la mesure où les sujets communicants, émetteur et récepteur, sont réunis dans des limites précises de l'espace et du temps, et en plus elle est destinée à un public nécessairement restreint [SOU 02][ALL 03].

Pour ces raisons, l'homme a été obligé de chercher d'autres moyens de communication permettant non seulement une diffusion plus large des informations, mais aussi une conservation dans le temps des informations diffusées. Cependant, il a remplacé les signaux vocaux éphémères par des signes plus durables parce qu'elles sont matérialisées sur des supports non volatiles, et inventa ainsi l'*écriture* vers l'an 3200 avant J.C [SOU 02]. L'écriture est devenue alors via l'usage à la fois un moyen de communiquer, de transmettre, et surtout de sauvegarder et d'archiver la mémoire collective. Il est nécessaire de signaler que l'écriture n'est qu'une transcription de la parole sur un support, et qu'un moyen de matérialiser l'information, comme expliqua Anne-Marie Christian dans [CHR 01] : « *L'art ne reproduit pas le visible, il rend visible* » a écrit Paul Klee. *L'Histoire de l'écriture repose sur le même principe. L'écriture ne reproduit pas la parole, elle la rend visible* ». L'émergence de l'écriture conduit tout naturellement à la notion de document qui peut être défini comme un jeu d'informations consignées par écrit [DUO 05]. Pour matérialiser l'écriture, l'homme a utilisé différents supports : pierres, marbre, galets, tablettes de cire, de bois ou d'argile, papyrus, parchemins. Tous ces supports ont été remplacés de nos jours par le papier inventé en Chine au 2^{ème} siècle et qui constitua un essor dans le développement de documents. De la même manière ont été développées les techniques de l'écriture, qui a resté manuscrite, jusqu'à l'invention de l'imprimerie au milieu du 15^{ème} siècle. Cette invention a marqué le début d'une nouvelle ère dans la production et la diffusion de documents écrits [JOU 06]. La production d'un livre entier est devenue plus rapide, plus facile et moins chère, ce qui a produit, quelques siècles plus tard, une quantité énorme de documents.

Actuellement, une grande partie des documents (manuscrits et imprimés) produits dans le passé sont principalement stockés dans les bibliothèques, les musées, les archives, ainsi que dans d'autres institutions à caractères pédagogiques ou sociopolitiques. Ces documents historiques hérités de nos parents et ancêtres nous ont transmis les traces d'une culture passée qui servira de façonner celle du présent puis celle du future. Ils constituent ainsi une part importante de notre patrimoine et présente une richesse culturelle et scientifique inestimable. Les collections historiques sont depuis de nombreuses années, le support de travaux de recherche en sciences humaines et sociales. Bien que la valeur des documents patrimoniaux ne soit pas seulement portée par leur contenu mais aussi par l'objet lui-même [LEY 06], ces documents ont d'intérêts pour différentes personnes qui ont besoins d'étudier les originaux. Ils sont utilisés par les historiens pour identifier l'origine des livres (date d'impression, lieu d'impression, ...) et retracer sa vie. Les biologistes peuvent utiliser les notes manuscrites d'observation sur le terrain pour comparer l'état courant d'un écosystème avec les conditions du passé. Les manuscrits anciens pourraient aussi servir à étudier l'évolution de la forme et du style de l'écriture d'une société au fil du temps, ce qui reflète les changements historiques et

punctuation, les abréviations et la façon avec laquelle ils ont évolué, permet aux paléographes et historiens d'identifier les périodes au cours desquelles un manuscrit a été écrit.

En effet, les documents historiques stockés dans les bibliothèques et les archives encourent une dégradation progressive (jaunissement du papier, fragilisation et déchirements des feuilles, arrachements ou l'amincissement des pages, ...) parce qu'ils ne sont pas conservés dans des bonnes conditions, et lorsqu'ils ne se détériorent pas naturellement, souffrent d'être trop souvent manipulés pour consultation ce qui peut risquer de les endommager. Ces documents sont donc menacés d'un réel danger de disparition. La figure 1 présente quelques exemples de dégradations fréquemment trouvées dans les documents historiques. Afin de réduire le nombre de manipulations, les bibliothèques limitent l'accès aux documents fragiles à certains spécialistes, solution inconcevable à long terme.



Figure 1. Exemples de documents dégradés

Les conservateurs de documents historiques ont donc un défi important à relever : comment préserver les documents originaux (qui sont fragiles par nature), et rendre accessibles au public ces millions de pages contenant des informations importantes? Pour atteindre ces objectifs, les bibliothèques ont commencé à numériser les collections de documents historiques qu'ils détiennent. Alors que l'idée de mener des campagnes de numérisation de documents a commencé à émerger dans les années 60, c'est dans les années 70 que datent les premières campagnes de numérisation significatives (USA, France, etc.) [JOU 06]. L'UNESCO, responsable de la protection du patrimoine culturel mondial, a lancé en 1992 le programme «Mémoire du monde», pour sauvegarder et promouvoir ce patrimoine par la numérisation [STE 95]. Cette initiative, conjointement avec l'évolution des performances techniques liées à la numérisation et la démocratisation d'Internet dans les années 90, ont engendré la création de nombreux projets et bibliothèques numériques à travers le monde pour sauvegarder le contenu de dizaines de milliers d'ouvrages anciens. Parmi les projets de numérisation nous pouvons citer le projet *Passe-Partout (Recherche d'ornement Passe-Partout)* [PAS 96], *Gallica* [GAL 97], *Google books* [GOO 04], *Europeana* [EUR 08], *Koninklijke Bibliotheek* [BIB 11], *BLSB (British Library Search Beta)* [BLS 10], *OCA (Open Content Alliance)* [OCA], *IMPACT (Improving Access to Text)* [IMP], *NaviDoMass (NAVigation into DOcument MASSes)* [NAV 10], *IDP (The International Dunhuang Project)* [IDP].

Au sens strict du terme, la numérisation consiste seulement à effectuer la capture des images des documents, on parle alors de capture d'images [DRI 07]. Elle présente l'intérêt d'éviter les contacts directs avec l'ouvrage et ainsi une meilleure conservation, un partage des images avec un grand nombre de personnes via Internet, CDs ou autres supports électroniques, un accès simultané et un feuilletage virtuel. Cependant, cette simple numérisation n'est pas suffisante pour répondre aux besoins des utilisateurs des bibliothèques numériques. Elle doit être accompagnée d'outils informatiques destinés à améliorer les conditions d'analyse et de recherche des documents. L'analyse d'images de documents est un domaine de recherche actif à la frontière de la reconnaissance de formes et de l'analyse d'images. C'est une discipline qui étudie les possibilités algorithmiques de reconstituer une information structurée à partir de la forme visuelle brute (à partir de l'image de document) [BAP 98].

Pendant longtemps, les efforts dans le domaine de l'analyse de documents ont été concentrés sur la problématique de la reconnaissance optique de caractères (OCR), ainsi que sur quelques domaines très spécifiques qui l'exploitent, tel que la lecture de chèques, le tri du courrier postal, et l'acquisition de formulaires. Cependant, le principal intérêt d'un document ne se trouve pas dans sa forme physique, mais plutôt dans son contenu [BEL 01], ce qui fait sortir l'analyse de documents de la simple reconnaissance à des objectifs beaucoup plus larges pour répondre aux besoins actuels, en profitant des progrès réalisés dans des domaines proches, tel que le traitement du signal, l'analyse d'images, la reconnaissance de formes, ou l'intelligence artificielle. Les chercheurs se sont attaqués à une multitude de sous-problèmes, explorant autant les fondements théoriques que les approches empiriques. Chaque tentative d'application a mis en évidence de nouvelles difficultés pratiques, et des moyens pour les contourner [BAP 98].

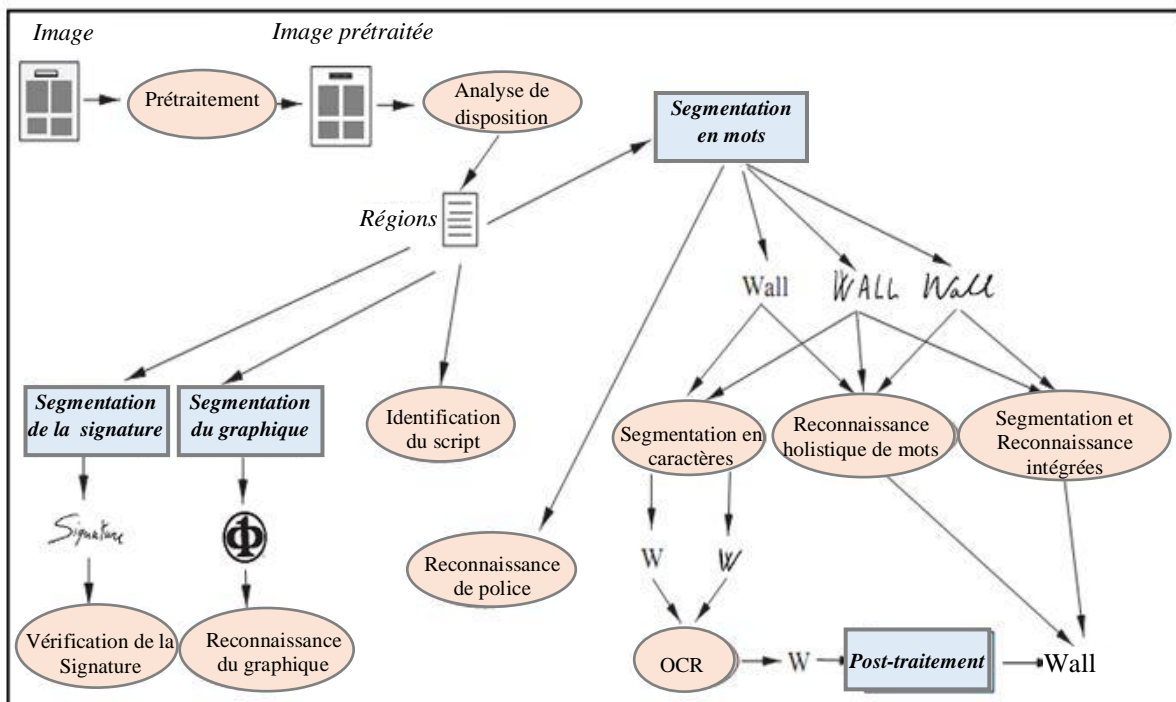


Figure 2. Schéma général du processus d'analyse et de reconnaissance de documents [MAR 08]

Il convient de rappeler que l'analyse de documents regroupe plusieurs traitements (voir la Figure 2) et que ces traitements automatiques ne fonctionnent correctement que si les images reçues du capteur possèdent les qualités suffisantes, dénuées de bruit, corrigées de leurs erreurs éventuelles, homogénéisées, normalisées et réduites à l'essentiel ce qui n'est pas toujours vérifié. Comme nous avons déjà mentionné, la numérisation permet de garder une

«copie numérique» des documents ; mais pour un grand nombre de documents historiques, la numérisation arrive trop tard et beaucoup de documents ont été déjà dégradés et par conséquent ils ont été numérisés en mauvais état. En plus des détériorations naturelles des anciens documents, les images reçues du capteur peuvent comporter des dégradations dues à une mauvaise numérisation. On parle de mauvaise numérisation quand l'image de document résultante n'est pas suffisamment fidèle à la page originale [RAB 13]. L'utilisation d'un capteur de numérisation à usage quotidien qui est incompatible avec une numérisation de qualité (scanneur de bureau, appareil photo numérique, Webcam, etc.), un mauvais réglage du capteur (résolution, éclairage, couleur, etc.), des mauvaises manipulations (mouvements de la page pendant sa numérisation, mauvais placement de la vitre, etc.), etc. peuvent produire des images numériques de qualité inacceptable (Figure 3). En cherchant donc à préserver les documents originaux, les images numériques obtenues présentent des défauts (géométriques, photométriques, etc.) qui nécessitent des prétraitements.

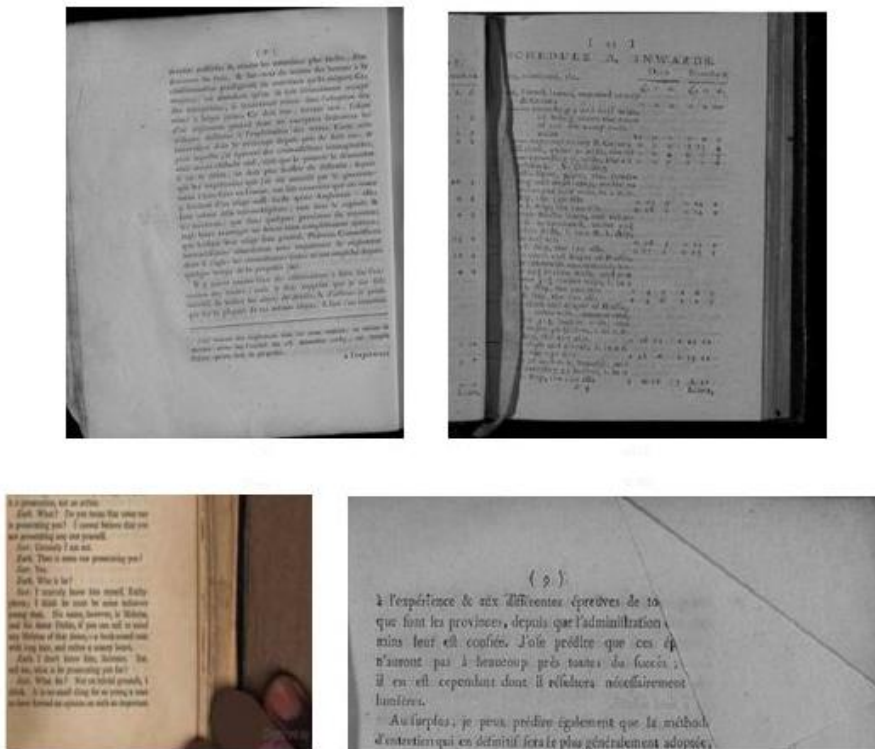


Figure 3. Exemples de défauts engendrés par la numérisation [RAB 13]

Le prétraitement d'images de documents est un domaine clé qui doit être privilégié et traité en priorité car les performances des systèmes d'analyse et de reconnaissance de documents sont fortement limitées par la qualité d'images traitées. Le présent travail de cette thèse s'inscrit dans ce contexte.

Contribution

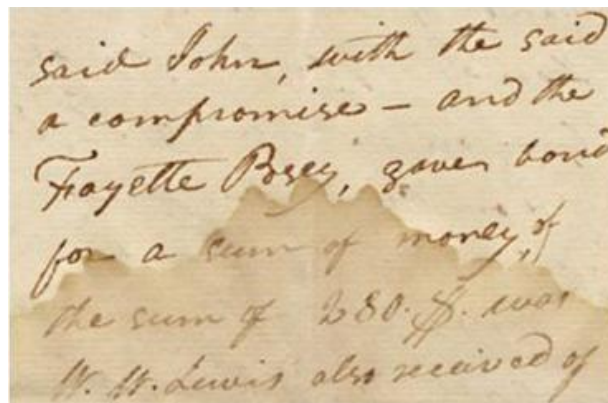
Il existe actuellement des images de documents historiques numérisés de différentes origines, et de qualité très variable. Ces images subissent généralement deux types de dégradations : les dégradations dues au processus de numérisation, et les détériorations des documents originaux par l'usure du temps. Ce sont ces deux sources de dégradation qui sont au cœur des problèmes actuels et qui rendent difficile l'exploitation directe des images de documents numérisés [DRI 07]. Le prétraitement de ces images est alors indispensable. Il regroupe un ensemble de traitements de bas niveau (au niveau pixel) ayant comme but d'éliminer (au maximum) le

bruit incrusté aux données reçues du capteur et de ne garder autant que possible que l'information pertinente, afin de préparer le terrain aux étapes suivantes dans le processus d'analyse et de reconnaissance de documents.

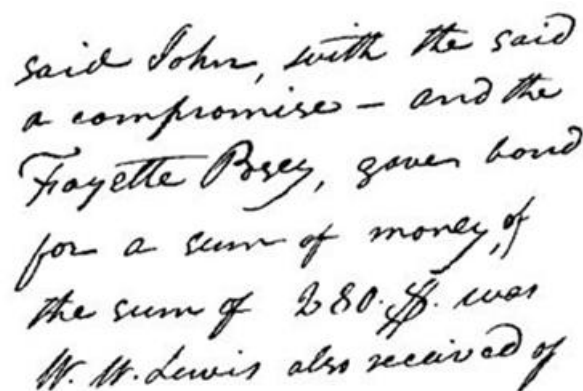
En raison de la diversité des sources de bruit, et de la multitude de ses effets sur l'image, il n'existe pas une seule technique de prétraitement adaptée à toutes les situations. En effet, plusieurs opérations peuvent être employées pour le rehaussement de la qualité des images numérisées : *séparation avant-plan/ arrière-plan, redressement, restauration, filtrage, etc.*

Dans le présent travail, notre objectif est de prétraiter les images de documents anciens, en suivant les recommandations de [OGI 08] [BOU 08b]. Il ne faut plus regarder le document ancien comme un cas particulier de documents de mauvaise qualité mais plutôt comme une catégorie à part entière. Il possède ses propres spécificités mais bien évidemment il partage plusieurs caractéristiques des documents dégradés.

Nous nous intéressons dans ce travail de thèse à une des opérations de prétraitement les plus usuelles et les plus importantes, à savoir *la séparation entre l'avant-plan et l'arrière-plan*, pour l'amélioration de la qualité des images de documents historiques numérisés. Cette opération a comme but d'extraire le texte et les autres éléments de l'avant-plan à partir de l'image de document. De ce fait, une séparation parfaite entre l'avant-plan et l'arrière-plan permet d'éliminer le bruit et le fond et de ne garder que les informations pertinentes (texte, figures, tableaux, etc.). La technique la plus usuelle pour atteindre cette séparation est de binariser l'image [LEE 02] et dans ce cas l'avant-plan sera coloré en noir dans l'image binaire résultante, et l'arrière-plan en blanc, comme il est montré par la Figure 4. Cependant cette opération a une grande influence sur les étapes suivantes dans le processus d'analyse et de reconnaissance de documents et donc elle doit être effectuée le plus fidèlement possible.



(a) Image de document dégradé



(b) Résultat de séparation parfaite avant-plan/ arrière-plan

Figure 4. Séparation idéale avant-plan / arrière-plan d'une image de document dégradé

Le présent manuscrit décrit les travaux que nous avons élaboré dans le cadre de la séparation avant-plan / arrière-plan dans les images de documents historiques. Les travaux réalisés peuvent être résumés en quatre contributions regroupées en deux : contributions à l'évaluation des performances des techniques de binarisation, et contributions à la séparation avant-plan/arrière-plan dans les images de documents dégradés. La première contribution consiste à procéder une étude comparative avec une évaluation objective quantitative d'un grand nombre de techniques de binarisation afin de déterminer la technique la plus performante en moyenne pour l'extraction du texte dans les images de documents historiques. A notre connaissance, notre étude comparative est plus large que les travaux d'évaluation procédés dans la littérature pour plusieurs raisons ; pour le nombre de techniques comparées, pour la taille de la base de test, et pour le nombre de critères d'évaluation utilisés. Dans la deuxième contribution, une nouvelle méthodologie d'évaluation des performances de méthodes de binarisation est proposée pour la recherche dans les images de documents arabes dégradés. Les deux autres contributions s'inscrivent dans le cadre de la proposition de nouvelles méthodes de binarisation d'images de documents dégradés. Ainsi deux méthodes sont proposées. La première est une technique de binarisation basée sur la classification de pixels en *noirs* ou *blancs* en utilisant un réseau de neurones artificiel de type PMC (*Perceptron multi couches* ou MLP pour *Multi-layer Perceptron*), tandis que la deuxième méthode s'agit d'une technique de binarisation hybride qui utilise des informations globales et locales sur l'image et combine plusieurs méthodes de binarisation dont le but d'avoir un résultat plus sûr au bout d'un temps raisonnable.

Organisation du mémoire

Cette thèse est structurée en cinq chapitres répartis en deux parties. La première partie contient les chapitres 1, 2 et 3 et 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 nos différentes contributions, la mise en œuvre, ainsi que les résultats obtenus à chaque niveau.

Chapitre 1 : Prétraitement d'images de documents

Ce chapitre est consacré à la présentation des principales notions relatives au prétraitement d'images de documents. Il présente d'abord les notions de bruit et de dégradation, pour bien montrer leurs effets sur les images, et leurs influences sur les étapes d'analyse. Puis, il aborde la tâche de prétraitement elle-même, ses objectifs, et ses résultats sur les images de documents. Il donne par la suite un aperçu sur les opérations usuelles utilisées pour le rehaussement de la qualité des images de documents numérisés. Il mentionne ainsi pour chaque opération de prétraitement les différentes approches et techniques proposées dans la littérature.

Chapitre 2 : Séparation avant-plan / arrière-plan

Le deuxième chapitre propose un état de l'art sur la séparation avant-plan / arrière-plan dans les images de documents, sans se limiter à un type particulier de documents. Nous concentrons dans ce chapitre à la binarisation en tant que technique la plus connue de séparation et nous focalisons sur la binarisation d'images en niveaux de gris car dans la plupart de 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 : Evaluation et comparaison des techniques de binarisation

L'objectif de ce chapitre est de dresser un état de l'art des principaux travaux d'évaluation de la qualité des images binarisées et de comparaison des méthodes de binarisation d'images. Le chapitre commence par mettre le point sur la différence entre les deux concepts : évaluation et comparaison. Ensuite, il expose les différentes approches d'évaluation des techniques de binarisation proposées dans la littérature. Après, il cite et détaille les principales mesures et critères utilisés par les chercheurs pour accomplir l'évaluation. Le chapitre se termine par la description de quelques travaux antérieurs relatifs à l'évaluation des méthodes de binarisation que nous avons jugés intéressants.

Chapitre 4 : Contribution à l'évaluation des techniques de binarisation

Ce chapitre se détache des aspects théoriques abordés dans les trois premiers chapitres et s'oriente vers la présentation de nos deux premières contributions qui s'intéressent à l'évaluation quantitative et objective des performances des techniques de binarisation d'images de documents. La première contribution s'agit d'une étude comparative étendue d'un grand nombre de techniques de seuillage sur des images de documents dégradés. La deuxième consiste à une nouvelle méthodologie d'évaluation des méthodes de binarisation pour la recherche de mots dans les images de documents arabes dégradées. Pour chacune des deux contributions, nous mentionnons tout d'abord ses motivations et ses objectifs, nous présentons la démarche suivie, nous décrivons les bases d'images utilisées et nous exposons les expérimentations effectuées et les résultats obtenus.

Chapitre 5 : Contribution à la binarisation d'images de documents dégradés

Le cinquième chapitre décrit nos deux dernières contributions, qui s'agissent de deux nouvelles méthodes de binarisation d'images de documents dégradés. La première méthode est une technique de binarisation basée sur la classification et non pas sur le seuillage comme la majorité des méthodes de la littérature. Dans cette méthode la classification est effectuée à l'aide d'un réseau de neurones artificiel (*ANN pour Artificial Neural Network*). Cette proposition peut être justifiée par le fait que la binarisation est en fait un processus de classification où les pixels de l'image sont divisés en deux classes «Noirs» et «Blancs», la tâche qui peut être accomplie efficacement par les ANNs. La deuxième méthode est une technique de seuillage hybride et elle est proposée dans le but de parvenir à une technique de binarisation rapide et efficace en même temps. La technique proposée combine plusieurs méthodes de binarisation afin d'avoir un résultat plus sûr. Pour chacune des deux méthodes proposées, nous mentionnons tout d'abord ses motivations et ses objectifs, nous présentons la démarche suivie, nous décrivons les bases d'images utilisées et nous exposons les expérimentations effectuées et les résultats obtenus.

PARTIE I :
ETAT DE L'ART

Chapitre 1.

Prétraitement d'images de documents

1.1. Introduction

Un très grand nombre de documents de différentes catégories : journaux, formulaires, cartes géographiques, dessins techniques, partitions musicales, documents historiques, etc. existent actuellement dans les bibliothèques, les musées, et les archives nationales. La conservation de ces documents et leur accès au grand public est devenu aujourd'hui un besoin incontournable. En plus des documents stockés, la gigantesque quantité de documents qui circulent dans les postes, les banques et les entreprises, a obligé les chercheurs à développer des outils pour les traiter et les analyser automatiquement. Afin de surmonter ces problèmes, la numérisation a été mise en place. Dans un souci de préservation du document, il est impératif de garantir que l'image résultant de la numérisation soit la plus fidèle possible au document original. Cependant, le processus de numérisation génère souvent des images de mauvaise qualité, à fond hétérogène, avec effet de transparence, etc. et ce malgré les avancées spectaculaires de la technologie des outils de capture numérique. En fait la plupart de ces déformations au niveau de l'image sont dues aux détériorations subies par les originaux eux-mêmes. Tous ces problèmes rendent difficile l'exploitation directe de ces images et constituent ainsi un nouveau défi aux chercheurs dans le domaine de l'analyse de documents.

Que l'on fait de l'identification du contenu (reconnaissance de l'écriture, des symboles ou des objets, identification du scripteur), l'indexation et la recherche, ou plus simplement la classification, l'organisation et le catalogage, on passe inévitablement par un ensemble de traitements appliqués sur les images de documents numérisés afin de les préparer aux étapes suivantes dans le processus d'analyse. Cet ensemble de traitements souvent appelé prétraitements (aussi appelé traitement de bas niveau dans d'autres littératures) fait l'objet du présent chapitre.

Le chapitre est organisé de la manière suivante. Dans un premier temps, nous mettons le point sur les différents défauts et dégradations qui apparaissent souvent dans les images de documents pour bien comprendre leurs effets nuisibles sur la qualité des images et afin de mieux sentir la difficulté de traitement de des documents dégradés. Dans un deuxième temps, nous abordons le prétraitement, ses objectifs, et ses résultats sur les images de documents. Le reste du chapitre est consacré à la présentation des opérations de prétraitement les plus usuelles.

1.2. Dégradations d'images de documents

Avant d s'intéresser aux améliorations à apporter aux documents numérisés pour faciliter leur interprétation, il est important de comprendre quels types de dégradations, ou défauts, pouvant intervenir dans les images e documents.

1.2.1. Qu'est ce qu'une dégradation

Les images de documents historiques laissent apparaître souvent de nombreux défauts appelés dégradation rendant les algorithmes de traitement classiques inefficaces.

F. Drira [DRI 07] a défini la dégradation comme « *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 leurs séparations. Les dégradations ont des effets plus destructeurs sur les images de documents que sur toutes autres images dites "naturelles"* ».

1.2.2. Les différents types de dégradations

Différents types de dégradations peuvent apparaître sur les images de documents numérisés, et elles proviennent de diverses causes et origines. Certaines dégradations proviennent du processus même de numérisation (résolution insuffisante, format de compression inadéquate, etc.). D'autres sont dues au mauvais état des documents originaux causé par le processus de vieillissement, les conditions de conservation des documents, ou à cause des modifications apportées par l'humain (annotations, corrections, etc.).

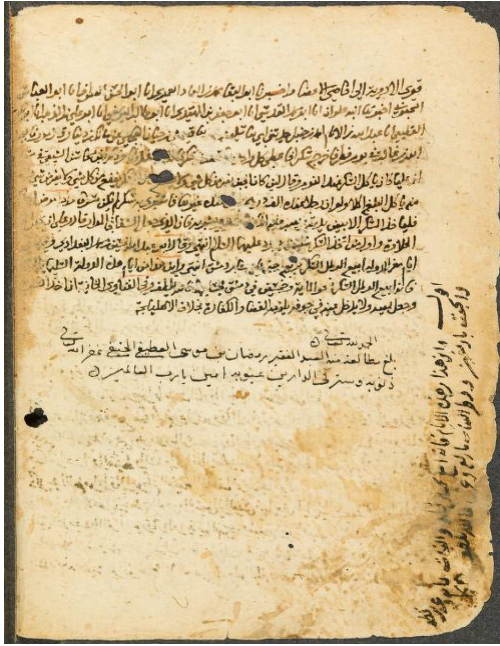
De bonnes études sur les dégradations et leur catégorisation ont été menées dans [DRI 07] et [RAB 13]. F. Drira [DRI 07] a établi une classification non-causale orientée traitements des dégradations des documents en deux familles : dégradations uniformes et dégradations non-uniformes. Les premières affectent tout le document de façon uniforme, tandis que les deuxièmes varient suivant la position dans l'image. Dans les dégradations non-uniformes on distingue les dégradations du fond des dégradations des formes. V. Rabeux [RAB 13] a réalisé une taxonomie des principales dégradations d'un document ancien en deux catégories : les dégradations inhérentes à l'ouvrage lui-même et celles issues de la numérisation.

Nous présentons dans cette section les altérations couramment rencontrées dans les images de documents numérisés tout en adoptant la catégorisation la plus usuelle des dégradations en deux classes : les dégradations relatives aux documents originaux et les dégradations résultantes d'une mauvaise numérisation.

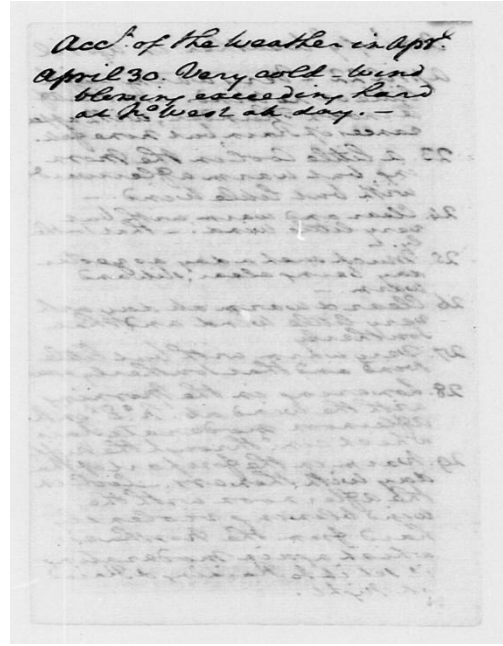
1.2.2.1. Dégradations des documents originaux

Cette catégorie de dégradations regroupe tous les défauts qui proviennent essentiellement de l'état du document original [RAB 13]. Elle constitue ainsi la majeure partie des défauts affectant les images de documents anciens numérisées. Parmi les dégradations les plus usuelles nous trouvons :

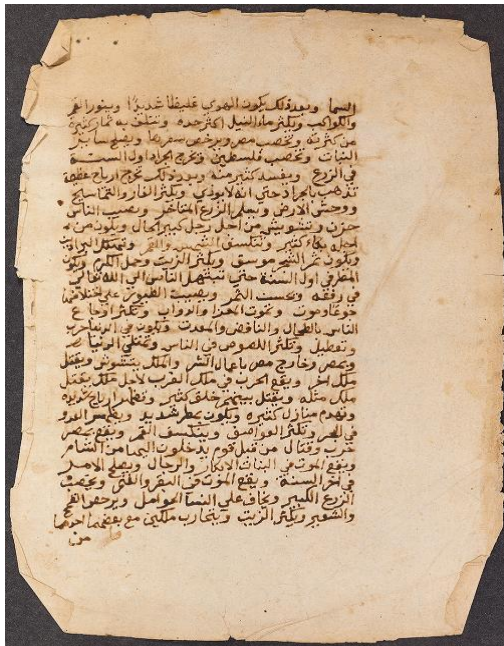
- **Les défauts touchant la forme des caractères** : ce genre de défauts apparaissent fréquemment dans les documents anciens. Ils modifient la continuité des traits et déforment leurs contours, rendant les textes illisibles ou difficiles à segmenter. Parmi les dégradations les plus courantes nous trouvons : la rupture des traits, l'effacement de parties de caractères, la fusion de plusieurs caractères, etc. Des exemples de dégradations touchant les caractères sont montrés par les figures 1.1.a et 1.1.d.
- **L'effet de transparence** : qui se produit généralement dans les papiers trop fins laissant apparaître l'impression du verso à travers le recto. Un exemple est illustré dans la figure 1.1.b.
- **Les taches** : souvent dues à l'acidité de l'encre, à l'humidité du lieu de stockage, mais peuvent avoir d'autres causes. La figure 1.1.a montre l'effet d'une tache d'humidité sur une page d'un livre ancien.
- **Les pliures et ondulations** : comme le montre la figure 1.1.c.
- **Les déchirures et trous** : qui peuvent se produire à cause d'une mauvaise utilisation, de l'usure, des insectes, etc. Voir les figures 1.1.a et 1.1.c.
- **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 1.1.d).



(a) Déformation des caractères, taches d'humidité et trous



(b) Effet de transparence



(c) Plis et déchirures

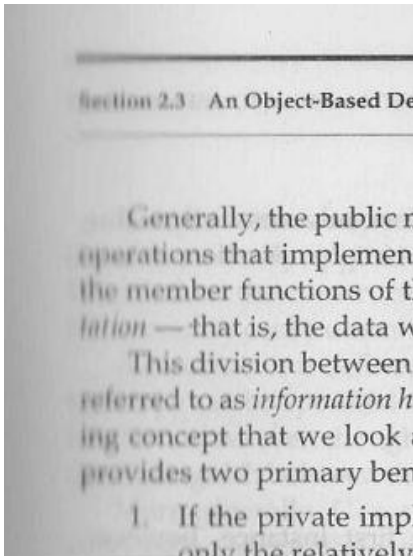


(d) Tampon [DRI 07]

Figure 1. 1. Exemples de dégradations affectant les documents anciens

1.2.2.2. Dégradations dues à la numérisation

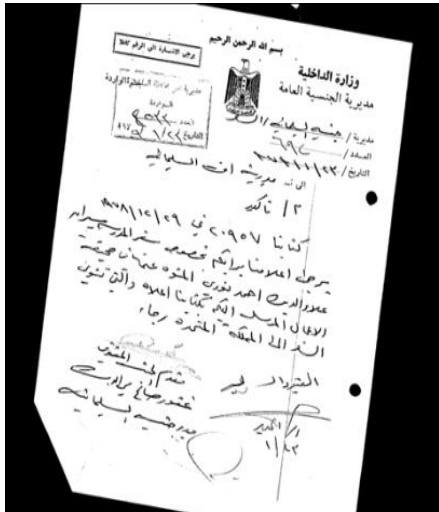
La numérisation, même de qualité, n'est qu'une représentation imparfaite, à un moment donné, d'un objet physique [DRI 07], et de ce fait elle peut introduire du bruit aux images reçues du capteur. Les défauts causés par la numérisation peuvent avoir plusieurs formes, citons [RAB 13]:



(a) Effet du flou [LIN 10]



(b) Bruit marginal [FAR 13]



(c) Inclinaison et bruit de fouillis [AGR 09]



(d) Courbure [DRI 07]



(e) Apparition des doigts tenant le document [RAB 13]

Figure 1. 2. Exemples de dégradations dues à la numérisation

- **L'effet du flou** : qui se produit d'un mauvais réglage des paramètres du capteur (Figure 1.2.a).
- **Le bruit marginal** : formé de l'ensemble d'ombres qui apparaissent dans les marges verticales ou horizontales d'une image [FAR 13]. 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 1.2.b montre ce genre de bruit.
- **Les défauts géométriques**: tels que l'inclinaison résultante d'un mauvais positionnement du document sur la vitre du scanner (Figure 1.2.c), et la courbure du texte sur les bords due à l'épaisseur du livre (Figure 1.2.d).
- **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 [RAB 13].

- **L'illumination non uniforme** : due aux défauts d'éclairage.
- **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 1.2.e.
- **Le bruit des images numériques**: causé par le capteur du scanner. Le bruit le plus courant est le bruit impulsionnel 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 [FAR 13]. Il en résulte de nombreuses sources telles que des perforations, de l'inclinaison du document, ou de la connexion d'énormes quantités de bruit poivre. Un exemple est donné par la figure 1.2.c.

1.3. Prétraitement d'images de documents

Comme nous avons déjà mentionné, les images de documents numérisées contiennent divers types de bruit et 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.

D'une façon générale, le mot « Prétraitement » désigne les opérations secondaires effectuées sur des informations avant le traitement principal. En informatique, le « Prétraitement » est défini comme la première étape d'un traitement de données, ayant pour objet de corriger les défauts dus à la chaîne de saisie de données [OFF 14].

En analyse et reconnaissance d'images de documents, le prétraitement regroupe un ensemble de techniques et opérations de bas niveaux qu'on applique sur l'image numérisée avant toute autre traitement. Ces techniques ont pour but de réduire le bruit superposé aux données, d'éliminer les défauts liés à l'image numérisée, d'atténuer les dégradations, et d'améliorer ainsi la qualité de l'image bruitée. Les prétraitements peuvent être effectués à des fins de visualisation, et/ou en vue de préparer le terrain aux traitements ultérieurs [LIK 03]. Cependant, les prétraitements n'augmentent pas notre connaissance sur le contenu du document, mais elle peut aider à l'extraire [NAG 00]. Le résultat des prétraitements est une version épurée dépourvue de bruit de l'image qui serait la meilleure possible pour les traitements suivants dans la chaîne [RAB 13][HAD 06]. La figure 1.3. présente le résultat de prétraitement d'une image de document dégradé par une tache.

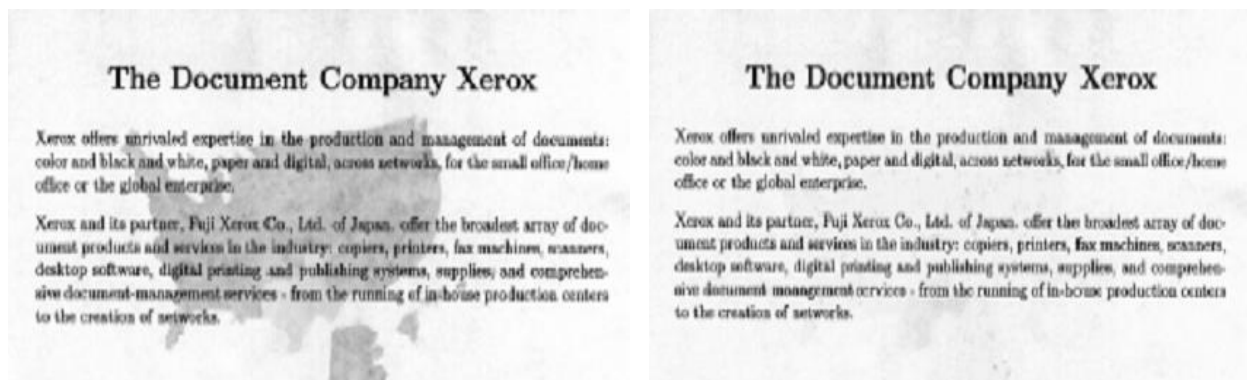


Figure 1.3. Résultat du prétraitement d'une image de document dégradé tirée de [DRI 07]

La figure 1.4 illustre la position du prétraitement dans le processus d'analyse et de reconnaissance de documents.

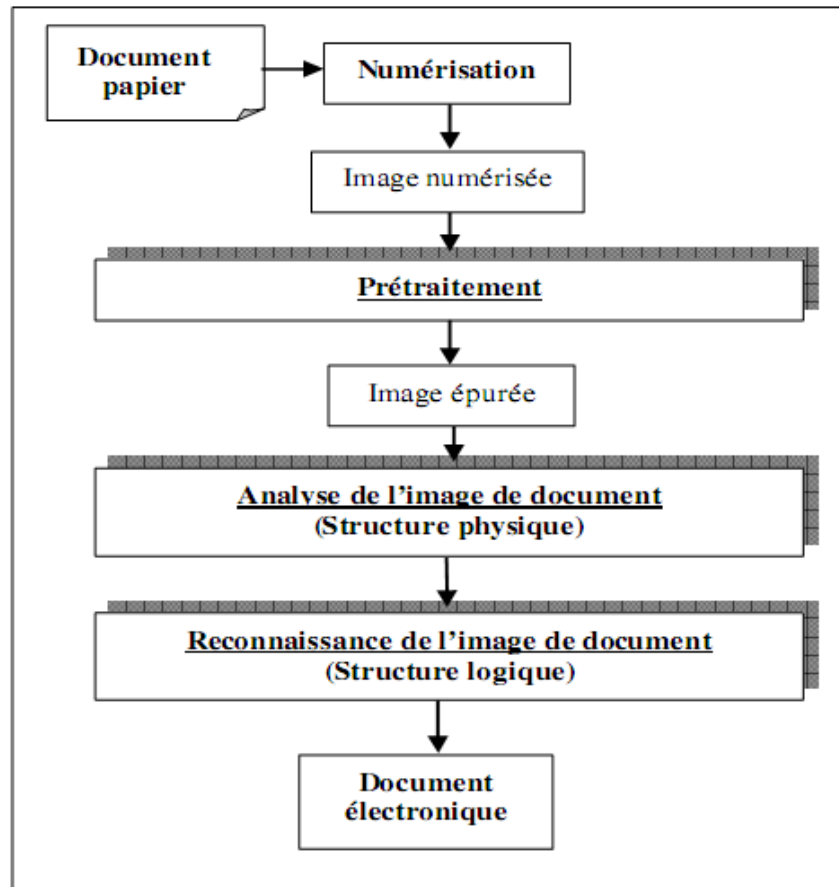


Figure 1. 4. Position du prétraitement dans le processus d'analyse et de reconnaissance de documents selon [KET 10]

1.4. Les opérations du prétraitement les plus courantes

Les dégradations présentes dans les images de documents sont à l'origine d'un grand nombre d'erreurs et elles influencent considérablement les performances des systèmes d'analyse et de reconnaissance de documents [RAB 13]. Afin de se débarrasser de toute sorte de bruit intervenant sur l'image et en vue de limiter à l'essentiel les informations contenues dans le document, le prétraitement doit être mis en place.

Le prétraitement est généralement composé d'une série d'opérations dont chacune est destinée soit au rehaussement de l'image bruitée, soit à la réduction des informations inutiles pour préparer le terrain aux étapes suivantes dans la chaîne d'analyse. Par conséquent, les étapes du prétraitement jouent un rôle important dans l'analyse d'images de documents puisque elles affectent directement les performances des traitements subséquents, comme l'extraction des primitives, et la reconnaissance optique de caractères (OCR). Notons qu'il est impossible d'être exhaustif sur toutes les opérations de prétraitement appliquées. Toute opération permettant de réduire le bruit, de rehausser l'image et d'améliorer sa visibilité, ou de préparer des données aux tâches suivantes peut être utilisée comme prétraitement.

Les opérations de prétraitement les plus couramment utilisées sont : la suppression du bruit et des points parasites, la séparation avant-plan/ arrière plan, la séparation recto/verso, la correction de l'inclinaison et de la courbure, le rehaussement du contraste. Dans le cas de l'écriture on trouve également : la segmentation, la squelettisation, la normalisation de caractères, la sélection des zones d'intérêt, etc. [NAG 95][NAG 00] [BOU 08][BEL 01] [KET 10].

En raison de la variété de types de bruit et de dégradations intervenant sur l'image, il n'existe aucune technique de prétraitement universelle qui s'occupe de toutes les dégradations. De multiples techniques ont été développées pour le rehaussement des documents dont chacune est spécifique à un seul défaut. La combinaison de plusieurs techniques de rehaussement est souvent nécessaire car comme a déclaré [BOU 08], 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. Ainsi, les techniques exactes qui doivent être appliquées pour le prétraitement d'une image donnée dépendent entièrement du type de document traité, de la nature des dégradations intervenantes, et de l'objectif visé de l'application.

Notons que l'application des techniques de prétraitement doit être effectuée attentivement car elles font souvent apparaître des effets secondaires sur d'autres éléments de l'image considérée [BOU 08][LIK 03]. Un filtrage passe-bas par exemple élimine certains bruits mais rend l'écriture floue.

Comme l'application des techniques de rehaussement est relative au type de bruit intervenant sur l'image, et dans le but d'automatiser la détermination des techniques à appliquer, des travaux sur l'identification des bruits dans les images de documents ont été effectués.

Nous allons développer dans la suite les principales étapes de prétraitement en mentionnant les techniques les plus utilisées dans chaque étape, mais avant ça nous décrivons quelques travaux sur l'identification du bruit dans les images de documents.

1.4.1. Identification du bruit

L'identification du bruit est une tâche préalable aux prétraitements. Elle peut être effectuée visuellement par un humain ou automatiquement en employant diverses techniques de classification et d'apprentissage. Bien que l'identification visuelle humaine est largement utilisée [SIL 09], elle est trop couteuse en temps et devient donc non-pratique pour une large collection d'images. L'automatisation de cette tâche permet de diminuer considérablement les interventions humaines dans le processus de prétraitements. En effet l'identification et la classification du bruit est un champ de recherche relativement récent et n'a pas suscité beaucoup d'études [LIN 09].

Lins et al. [LIN 10] ont proposé d'identifier le bruit présent dans les documents en utilisant un classifieur de type forêt aléatoire [BRE 01] entraîné sur un ensemble d'exemples étiquetés manuellement par un humain. Ainsi, huit caractéristiques ont été choisies pour l'apprentissage du classifieur. Le classifieur reporté est capable de détecter cinq types de bruit, à savoir : la transparence, le bruit marginal, l'inclinaison, le bruit d'orientation, et le flou. L'architecture de la détection de bruit est formée de classifieurs qui opèrent en parallèles pour la détection des différents types de bruit. Pour la transparence, le classifieur détermine également la force du bruit : fort, moyen, faible. Cette approche a été utilisée dans [ESI 10] dans le cadre de la conception d'un outil de traitement d'images de documents historiques. Dans [BEN 11], Ben Messaoud et al. ont proposé une approche de détection et de classification du bruit présent dans l'image, dans le cadre de la conception d'une méthodologie de paramétrage automatique de techniques de binarisation. L'approche proposée est limitée à l'identification de trois classes de bruit : transparence, avant-plan et arrière-plan similaires, et fond variable. L'identification des deux premières classes de bruit est effectuée en se basant sur une seule caractéristique de l'image, à savoir son histogramme de niveaux de gris. L'effet de transparence est détecté lorsque l'histogramme présente deux pics dans la région du fond. Si la distance entre les deux pics représentant l'avant-plan et l'arrière-plan est inférieure à un seuil déterminé, on remarque une similarité entre l'avant-plan et l'arrière-plan. Finalement, en

analysant les variances locales obtenues sur plusieurs fenêtres, on peut déduire si le document contient un fond variable ou non.

1.4.2. Débruitage et Restauration

La restauration a pour objectif de retrouver la qualité de l'image originale avant sa dégradation physique ou numérique. On parle aussi de débruitage comme on cherche à réduire le bruit présent dans l'image. La restauration est un domaine proche de l'amélioration des images et la frontière entre ces deux domaines est floue. En effet, l'amélioration et la restauration partagent des méthodes communes mais leur finalité reste différente. L'amélioration cherche à rehausser la qualité visuelle de l'image et non à récupérer sa forme avant la dégradation. Une image fidèle au document original, nettoyée de toutes les dégradations, n'est pas forcément une belle image [DRI 07].

Malgré la grande masse de documents numérisés qui présentent plusieurs types de défauts, le nombre de travaux spécifiques à la restauration d'images de document dégradées est encore réduit, et la plupart d'entre eux portent sur les corrections de dégradations bien particulières [DRI 07]. Un bon état l'art sur les techniques de restauration de documents les plus fréquemment cités dans la littérature a été exposé par F. Drira dans sa thèse de doctorat [DRI 07]. Nous présentons dans cette section les opérations les plus courantes de restauration et les techniques les plus utilisées de suppression du bruit.

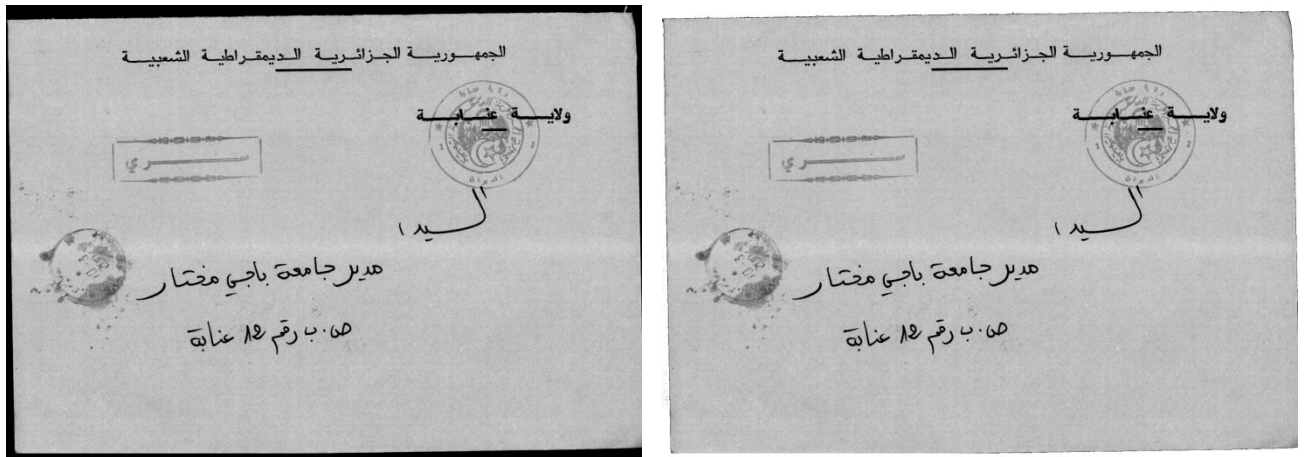
1.4.2.1. Suppression du bruit Sel et poivre

Le bruit "Poivre" peut être introduit dans l'image de document pendant sa numérisation. Lorsque le bruit inséré dans l'image soit périodique et donc se trouve en dehors des informations utiles, sa suppression est souvent aisée par des techniques classiques de filtrage et ne cause aucune destruction [BEL 92b]. Les composants isolés du bruit peuvent être supprimés en utilisant un simple filtre comme le filtre médian [STO 92]. Lorsque les composants de bruit soient plus larges, d'autres algorithmes plus efficaces tels que l'algorithme '*k-fill*' [PRE 10] ou des opérateurs morphologiques [SER 83] semblent être plus efficaces pour l'élimination du bruit [FAR 13].

Le bruit "Sel" se montre comme un manque d'encre dans l'image de document. Du bruit Sel isolé peut être supprimé par des filtres simples comme le filtre médian. En 2007, une méthode basée sur la morphologie a été proposée [GRA 07]. Cette méthode résolve un des problèmes les plus importants des approches basées sur la morphologie en utilisant une phase d'apprentissage pour trouver les paramètres de l'élément structurant le plus adéquat. Ensuite, l'opérateur de dilatation est utilisé pour remplir les emplacements où un manque d'encre est remarqué. Cependant cette méthode est lente à cause de la phase d'apprentissage, et produit des connections indésirables entre quelques caractères [FAR 13].

1.4.2.2. Suppression du bruit marginal

Le bruit marginal apparaît souvent dans les images de documents numérisés sous forme d'un cadre noire large entourant l'image. Diverses techniques ont été employées pour l'élimination et la suppression du bruit marginal. La figure 1.5 montre le résultat de cette opération.



(a) Image originale avec bruit marginal

(b) Image débruitée

Figure 1. 5. Suppression du bruit marginal

1.4.2.2.1. Méthodes basées sur le filtrage des composantes connexes

L'approche la plus répandue de suppression du bruit marginal est d'effectuer un filtrage des composantes connexes en se basant sur leur taille et leur forme. Ces techniques sont efficaces pour la suppression des bordures noires mais inefficaces pour la suppression du bruit textuel [WAG 14]. Fan et al. [FAN 02] ont proposé une méthode de suppression des bordures noires dans les images de documents numérisés qui opère en deux étapes : détection du bruit et suppression du bruit. La détection du bruit marginal s'effectue par la réduction de la résolution de l'image de document, et par l'analyse de la forme et la position des régions. La suppression du bruit est réalisée par un seuillage local pour les documents en niveaux de gris et par croissance de région pour les images binaires. Dans [AVI 04b], la suppression de bordures noires a été effectuée à l'aide d'un algorithme basé sur l'étiquetage des composantes connexes et les graphes d'adjacence de régions.

1.4.2.2.2. Méthodes basées sur l'analyse des profils de projection

L'analyse des profils de projection est aussi largement utilisée en suppression du bruit marginal. Cette technique est rapide et facile à implémenter mais elle échoue pour les documents inclinés [WAG 14]. Le et Thoma [LE 96] ont proposé un algorithme utilisant plusieurs heuristiques pour l'élimination du bruit marginal. L'algorithme repose sur la classification des lignes et des colonnes blancs/textuels/ non-textuels, la segmentation des objets, et l'analyse des profils de projections. En se basant sur la projection verticale, Zhang et al. [ZHA 01b] ont proposé de déterminer la position du bruit marginal et de détecter sa frontière avec le texte pour nettoyer l'image. De même, Stamatopoulos et al. [STA 07] ont proposé un algorithme de détection et de suppression du bruit marginal qui repose sur trois points : lissage RLSA (Run Length Smoothing Algorithm), profils de projection, et étiquetage des composantes connexes à partir de l'image binarisée.

1.4.2.2.3. Autres méthodes

Dans [CIN 02], des statistiques de l'image comme les vecteurs de différence horizontal/vertical et les luminosités des lignes ont été utilisées pour la suppression du bruit marginal textuel et non textuel à partir des images en niveaux de gris. Peerawit [PEE 04] a identifié le bruit à éliminer en comparant la densité des contours du bruit et du texte. Shafait et al. [SHA 07] ont utilisé des méthodes d'appariement géométriques pour la détection et la suppression des bordures dans les images de documents binaires. Une méthode basée sur la morphologie a été proposée dans [MEN 14]. Wagdy et al. [WAG 14] ont proposé un algorithme de

suppression du bruit marginal basé sur la théorie de Retinex et le seuillage global. [SHA 09] utilise l'analyse des profils de projections conjointement avec le filtrage des composantes connexes.

1.4.2.3. Séparation entre le recto et le verso

L'apparition des marques en transparence ou des éléments du verso sur le recto constitue une des dégradations les plus courantes des documents anciens. Diverses approches ont été utilisées pour la suppression de l'effet de transparence (Figure 1.6). Nous présentons ici quelques approches.



(a) Image dégradée

(b) image restaurée

Figure 1. 6. Suppression de l'effet de transparence

1.4.2.3.1. Méthodes basées sur le recalage

Le recalage est particulièrement utilisé pour la suppression des marques en transparence. Le principe des méthodes utilisant le recalage est de mettre en correspondance les deux faces d'une même page pour en extraire le texte original associé à chaque face. Sharma [SHA 00] applique un filtre linéaire adaptatif sur les deux faces du document simultanément pour séparer le texte des marques en transparence. Dans [DUB 01], les intensités des pixels du recto sont comparées avec son correspondant du verso. Si l'intensité du pixel du recto est moins élevée que celle du verso, sa valeur est conservée.

1.4.2.3.2. Méthodes basées sur l'analyse de l'orientation

Ces méthodes sont basées sur l'analyse de l'orientation de l'écriture dans les documents caractérisés par des styles d'écriture très orientés pour distinguer le texte du recto de celui du verso. Wang et al. [WAN 03] ont proposé une analyse de l'orientation à base de la transformée en ondelettes. Ainsi, l'image est filtrée selon deux directions : 45° pour obtenir le recto et 135° pour obtenir le verso.

1.4.2.3.3. Méthodes basées sur l'analyse des formes

La méthode de Wolf [WOL 06], comme représentant de cette classe, est basée sur une régularisation de la segmentation par les champs de Markov pour la séparation entre le recto et le verso. Cette séparation entre le recto et le verso repose sur la distribution spatiale des pixels et non pas sur des informations sur la couleur ou sur les niveaux de gris.

1.4.2.3.4. Méthodes basées sur l'analyse de la couleur

Smigiel [SMI 04] a utilisé une carte de Kohonen pour la classification des pixels en quatre classes (fond, texte, texte coloré et texte du verso) et par conséquent pour supprimer les taches

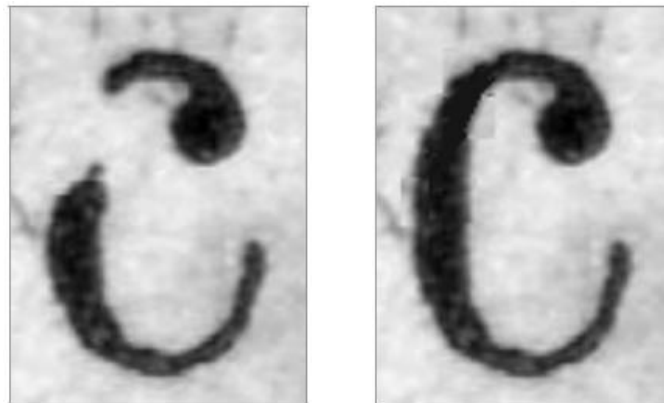
en transparence. L'inconvénient de cette méthode est sa nécessité d'une fastidieuse phase d'apprentissage sur une portion représentative d'une page. Dans [LEY 04], une méthode basée sur une sérialisation de l'algorithme des nuées dynamiques (*k-means*) a été proposée. Il s'agit d'une classification automatique des couleurs sur une fenêtre glissante parcourant toute l'image. Son principal inconvénient est la détermination manuelle du nombre de classes et des couleurs appartenant à chaque classe [DRI 07]. Garain [GAR 06] a proposé une méthode de séparation recto/ verso basée sur la segmentation en région de couleur qui sont ensuite agrégées suivant des règles prédéfinies. Cependant, cette méthode engendre de nombreuses fragmentations des traits. La méthode proposée dans [SIL 09] consiste à appliquer deux fois un algorithme de seuillage entropique pour séparer les trois composants du document (avant-plan, recto, et verso). La première application sépare le recto du reste du document et la seconde application dissocie le verso de l'arrière plan.

1.4.2.3.5. Méthodes basées sur la séparation de sources

En considérant que les taches de transparence et le texte du verso ont généralement une couleur différente du texte original du recto, il semble que la séparation de sources permet de séparer ces différentes couches d'information. Dans ce contexte, Tonazzini et al. [TON 04] ont utilisé l'Analyse en Composantes Indépendantes (ACI) [HYV 99] pour faire une séparation aveugle entre le recto et le verso. Cette approche ne nécessite ni l'information du verso, ni un modèle, ni un apprentissage. C'est donc une approche très intéressante du point de vu de l'utilisateur [DRI 07].

1.4.2.4. Restauration des formes de caractères

Les documents historiques numérisés présente diverses dégradations affectant la forme des caractères : rupture des traits, vide dans les traits, fusion de caractères, effet d'escalier, etc. Plusieurs approches ont été proposées pour reconstruire la forme d'origine de caractères. La figure 1.7 illustre le résultat de restauration d'un caractère cassé.



(a) Image dégradée de la lettre C (b) Image restaurée

Figure 1. 7. Restauration de la forme d'un caractère [ALL 03]

1.4.2.4.1. Méthodes basées sur l'établissement d'un modèle

Ces méthodes se basent sur la définition d'un modèle pour la restauration des formes. Un des premiers modèles de dégradations de caractères a été proposé par H. Baird [BAI 93]. Ce modèle inclut un certain nombre de paramètres, à savoir, la taille du texte, la résolution, la translation, la rotation, le changement d'échelle, etc. Un autre modèle de dégradation proposé par Kanungo [KAN 00] simule certains défauts, comme le flou, dus à l'impression et la numérisation, par des opérations de fermeture morphologique. Dans [ZHA 01b], le modèle

proposé est défini à partir de la numérisation d'une image idéale de caractères. La restauration par modèle conduit en général à des résultats intéressants. Mais, elle exige des hypothèses de déformation homogène sur l'ensemble de la collection [DRI 07]. Allier [ALL 03] a proposé l'utilisation du contour actif (*Snakes*) pour la restauration des caractères cassés. Cette restauration consiste à faire évoluer un contour actif, guidé par le modèle d'une lettre entière, sur la lettre cassée. L'inconvénient de cette méthode est que la réparation ne peut être effectuée que si l'on dispose du modèle de lettre à réparer et si la lettre possède une topologie simple [DRI 07].

1.4.2.4.2. Méthodes basées sur la redondance de formes

Cette approche consiste à collectionner toutes les formes similaires d'une page, ensuite on construit le modèle de la forme idéale de la lettre à partir de la superposition de toutes les formes dégradées de la même lettre. Hobby et Baird [HOB 97] ont proposé de superposer plusieurs occurrences du caractère dégradé et à moyenniser l'image résultante pour obtenir une approximation de l'image idéale du caractère. Chaque type de caractère dans le document sera ensuite substitué par son prototype. Le problème de cette méthode est qu'elle n'est efficace que pour des documents imprimés qui contiennent de faibles dégradations n'affectant la structure générale des formes des caractères [DRI 07].

1.4.2.4.3. Méthodes basées sur la récupération de liens dans les caractères cassés

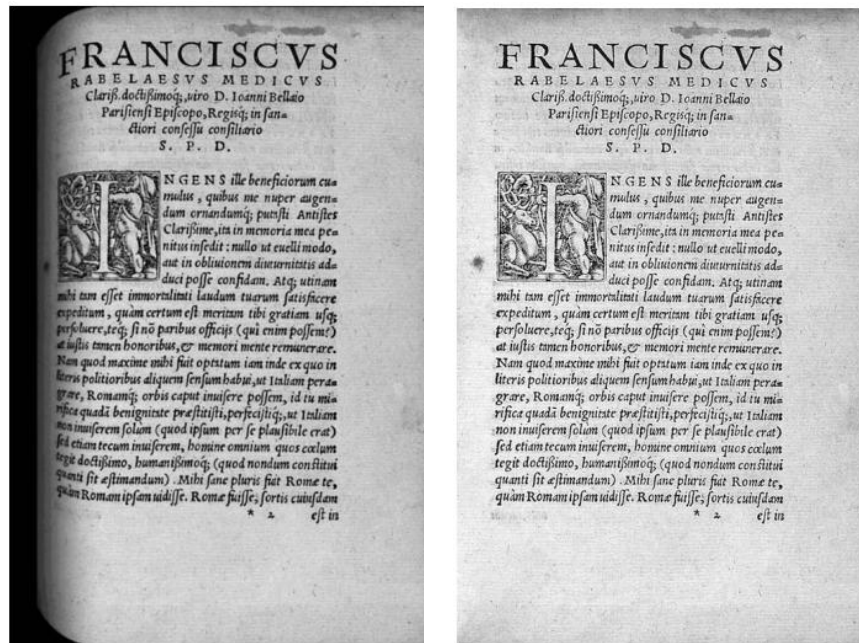
L'utilisation de la morphologie mathématique, notamment les opérations de dilations et d'érosions conditionnelles, pour la restauration des formes est possible pour les petites dégradations qui concernent quelques pixels seulement. Elle permet de remplir les vides et/ou de corriger les ruptures des traits. Wichello et Yan [WHI 96] ont proposé une méthode qui combine la morphologie avec un suivi de contours pour rétablir une certaine continuité des contours dégradés. Le suivi de contours qui se fait à l'aide de masques de tailles variables, autorise la connexion des parties du caractère dans une direction privilégiée, mais, il risque de relier les éléments qui ne devraient pas l'être s'ils ont une distance inférieure à la taille du masque. Yu et Yan [YU 01] ont résolu ce problème en ajoutant l'information sur la distance entre deux composantes connexes.

1.4.2.5. Correction de la courbure

Un autre défaut souvent produit lors de la numérisation des livres volumineux est l'apparition de la courbure à proximité de la reliure du livre. La mise à plat de l'image du document peut être effectuée par des corrections géométriques utilisant une analyse en deux ou en trois dimensions de l'orientation des traits ou des lignes de texte. La figure 1.8 illustre le résultat de correction de la courbure pour une image de document dégradée.

1.4.2.5.1. Correction par un modèle 2D

Les méthodes de cette classe procèdent en deux étapes : calcul de la courbure et traitement géométrique pour la corriger. Le calcul de la courbure peut être fait à partir des lignes de textes [WU 02] [ZHA 03b] ou par extrapolation des bords de la page [TAK 01][HIR 01]. Une combinaison des deux solutions a été aussi proposée [TRI 03]. Cette approche est limitée pour la segmentation des lignes de textes des documents dégradés ou des manuscrits pour lesquels les alignements ne sont pas naturellement horizontaux. En plus, la présence des taches dans le fond de l'image complexifie les opérations de segmentation des lignes de texte [DRI 07].



(a) Image dégradée

(b) courbure corrigée

Figure 1. 8. Résultat de correction de la courbure [RAB 13]

1.4.2.5.2. Correction par un modèle 3D

Les méthodes de correction basées sur un modèle 3D, peuvent être réalisées avec une prise de vue soit normale (un appareil photographique), soit particulière (deux appareils photographiques ou deux photographies du document prises sous des éclairages différents). Pour une prise de vue normale, les déformations des feuilles sont généralement modélisées par une surface cylindrique aux abords de la reliure [CAO 03][KAN 93]. La surface est définie par son rayon de courbure. Pour une prise de vue particulière avec l'utilisation de plusieurs vues sous des angles ou des éclairages différents d'un même document, une mise en correspondance est calculée entre les images pour reconstruire le relief de la page [YAM 04].

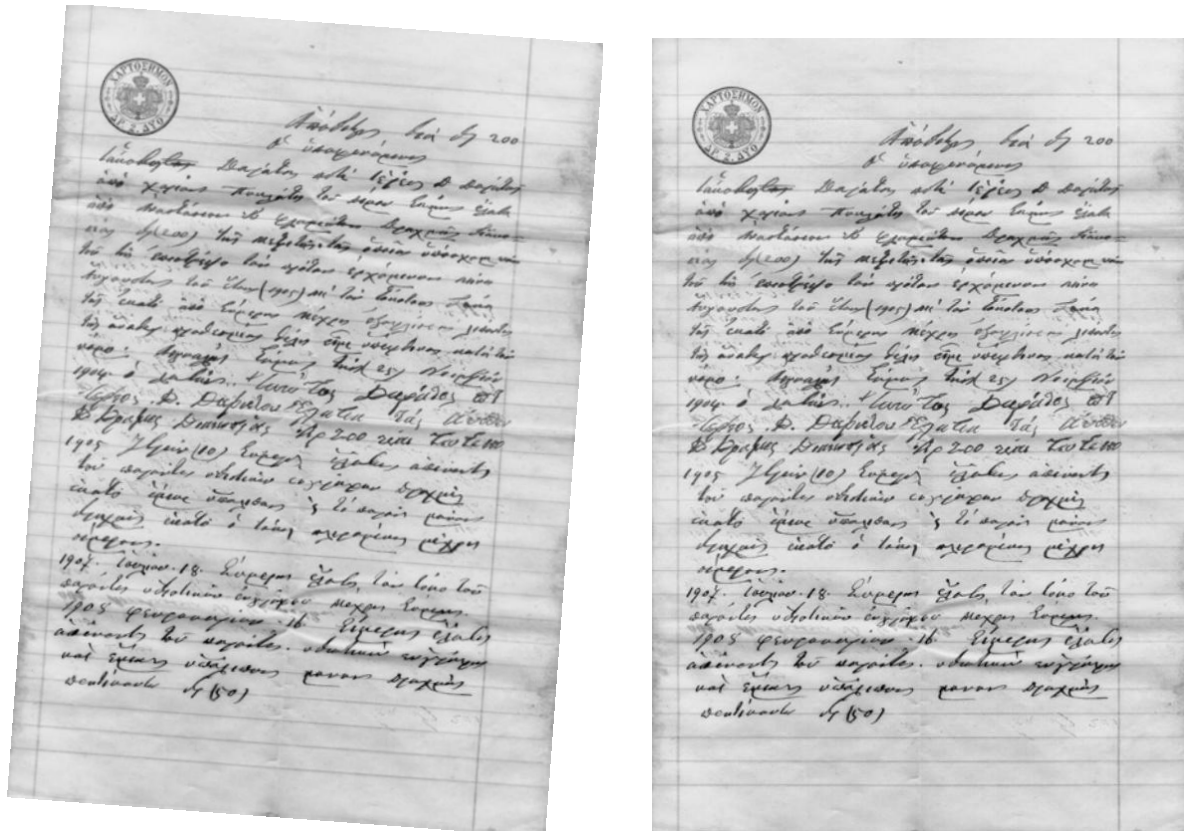
1.4.3. Séparation avant-plan/ arrière-plan

La séparation entre l'avant-plan et l'arrière-plan est l'une des opérations les plus importantes du prétraitement. L'image de document est considérée comme formée de deux couches superposées d'informations : l'avant-plan composé de tous les éléments pertinents qui constituent l'essentiel du document (texte, figures, dessins, tableaux, formules, etc.), et l'arrière-plan représentant le reste de l'image. La séparation avant-plan/ arrière-plan a pour but de séparer les deux couches d'informations composant le document. Elle permet ainsi d'éliminer l'arrière-plan jugé inutile pour les applications d'analyse et de reconnaissance de documents, et de ne garder que les informations pertinentes (les éléments de l'avant-plan).

Cependant, cette opération a attiré l'attention des chercheurs depuis longtemps. Nous trouvons dans la littérature un très grand nombre de méthodes de séparation avant-plan/ arrière-plan utilisant diverses techniques et ce nombre ne cesse d'augmenter chaque année. A cause de l'importance de cette opération et son influence sur les étapes suivantes dans le processus d'analyse et de reconnaissance, nous l'avons consacré tout un chapitre (chapitre 2).

1.4.4. Correction de l'inclinaison (redressement)

Les traitements suivants dans le processus d'analyse et de reconnaissance de documents sont très sensibles à l'inclinaison et nécessitent que l'écriture soit correctement orientée. Si ce n'est pas le cas, on doit s'occuper de la correction de l'inclinaison (Figure 1.9). En effet, l'inclinaison peut être introduite à l'image de deux manières possibles: soit lors de la saisie optique par un mauvais positionnement des pages par exemple, soit par le scripteur lui-même lors de l'écriture. La correction de l'inclinaison consiste à estimer tout d'abord l'angle de l'inclinaison, puis de ré-échantillonner l'image.



(a) Document incliné

(b) document bien orienté

Figure 1. 9. Correction de l'inclinaison d'un document mal orienté

Plusieurs méthodes ont été proposées dans la littérature pour la détection de l'inclinaison et sa correction par la suite. Les meilleurs algorithmes proposés sont les moins affectés par la présence de parties graphiques, de zones noires dans le texte, ou de zones d'ombre près des marges (bruit marginal) [BEL 01]. O'Gorman [OGO 93] a proposé une catégorisation des différentes méthodes de redressement selon les techniques employées en trois catégories : méthodes basées sur le profil de projection, méthodes basées sur la transformé de Hough, et méthodes utilisant le regroupement des voisins. D'autres catégories de techniques peuvent être distinguées, à savoir, les méthodes basées sur la transformée de Fourier, sur la corrélation inter-ligne, sur la morphologie mathématique, et sur l'Analyse en Composantes Principales (ACP) [ARA 06][SHI 05][SHI 05].

Nous présentons dans cette section, quelques méthodes de détection de l'inclinaison appartenant à différentes classes.

1.4.4.1. Méthodes basées sur la Transformée de Hough

La transformée de Hough est une technique bien connue pour la détection des formes paramétriques notamment des lignes et des courbes dans une image. Le défaut majeur de cette méthode est sa lenteur de calcul. Les différents algorithmes qui se basent sur la transformée de Hough essaient de minimiser son temps de calcul.

Amin et Fisher [AMI 00] ont proposé un algorithme de détection de l'inclinaison dans les textes complexes. L'algorithme commence par l'étiquetage des composantes connexes, qui seront regroupés par la suite en blocs. En fonction de leur taille, les blocs seront divisés en trois catégories : bruit, petit ou grand. Si un bloc est suffisamment proche d'un groupe de blocs de même catégorie, il y est ajouté. A la fin du regroupement, on calcule la transformée de Hough sur les centres des blocs les plus bas dans le groupe (la dernière ligne dans le cas d'un paragraphe). Les résultats sont ensuite stockés pour chaque bloc dans un histogramme pondéré du nombre de composantes connexes. Il ne reste plus qu'à en choisir le maximum. D'une manière analogue, Yu et Jain [YU 96] ont suggéré d'appliquer la transformée de Hough sur les centres des BAG (Block Adjacency Graph) [YU 92] des composantes connexes afin de diminuer le grand temps de calcul de la transformée de Hough. Dans [ARA 06], la détection de l'angle de l'inclinaison se fait en quatre étapes ; détection des composantes connexes, filtrage des composantes connexes en éliminant les caractères possédant des hampes ou des jambages et les petits composants tels que les points diacritiques et les marques de ponctuation ; amincissement des blocs, et finalement application de la transformée de Hough sur les points restants. [SEH 05] décrit une méthode de détection de l'inclinaison applicable sur des images de documents arabes imprimés. Dans le but de réduire le temps de calcul de la transformée de Hough, les auteurs ont proposé de réduire le nombre de points à soumettre à la transformée de Hough et d'utiliser les centres de liaisons inter-caractères. Ainsi, la technique proposée opère en cinq étapes : extraction des composantes connexes, filtrage des composantes connexes, détection des liaisons, filtrage des liaisons, et application de la transformée de Hough.

1.4.4.2. Méthodes basées sur le profil de projection

Les méthodes basées sur le profil de projection consistent à parcourir l'image selon des directions (des angles) proches de l'horizontale, et à calculer le profil de projection selon chaque direction. Ensuite, des caractéristiques extraites à partir de chaque profil de projection sont comparées pour déterminer l'angle d'inclinaison [HUL 98].

Une méthode de détection de l'inclinaison basée sur le profil de projection a été proposée par Baird dans [BAI 87]. Cette méthode calcule plusieurs projections selon différents angles proches d'une direction probable et l'angle d'inclinaison est celui avec lequel la projection présente le pic maximum. Newman et al. [NEW 99] ont proposé une méthode de détection de l'angle d'inclinaison dont la projection s'effectue sur une fenêtre contenant quelques fragments de lignes de texte et non sur l'image entière. La technique utilisée pour détecter l'angle d'inclinaison consiste à tourner l'image dans un petit intervalle de degrés, et à choisir l'angle qui maximise la variance du profil. Le profil de projection a été utilisé aussi par Postl [POS 86] en combinaison avec la transformée de Fourier. Dans cette méthode, l'angle d'inclinaison est déterminé à partir de l'espace de Fourier. Motivé par le fait que la majorité des caractères latins possèdent des traits verticaux, Papandreou et Gatos [PAP 11] ont proposé une méthode de détection de l'inclinaison dans les images de documents latins basée sur les projections verticales en combinaison avec un critère de minimisation du rectangle englobant.

1.4.4.3. Méthodes basées sur le Regroupement

Une des méthodes basées sur le regroupement est la méthode FNNC (*Focused Nearest-Neighbor Clustering*). Cette méthode [JIA 99] prend comme entrée, un ensemble de points caractéristiques qui peuvent être les centres des composantes connexes, les codes de passage, ou n'importe quelles caractéristiques qui indiquent l'orientation dominante du document. A partir de ces points caractéristiques, on calcule l'angle d'inclinaison locale en trois étapes : Détermination des plus-proches voisins, calcul de la ligne d'inclinaison locale, et groupement des plus proches voisins focalisés. Okun et al [OKU 99] ont utilisé un assemblage simple et efficace des lignes de texte, basée sur les moments d'ordre 1 et 2 pour la détection de l'inclinaison. Cette méthode est applicable sur des images de faible résolution (50 dpi), ce qui réduit considérablement le temps d'exécution, et elle est capable de détecter n'importe quel angle d'inclinaison. La méthode consiste en quatre étapes : Réduction de la résolution de l'image, détection des composantes connexes, classification des composantes, et estimation de l'inclinaison. L'algorithme proposé par Avila et Lins [TEN 05] commence par l'étiquetage des composantes connexes, en les délimitant par un bloc englobant. Il associe ensuite à chaque bloc, son voisin le plus proche, puis l'ensemble des blocs dans la même direction que le premier et respectant certaines contraintes (de taille et de distance). En calculant l'inclinaison de cette série de blocs, on obtient au final l'angle d'inclinaison du document.

1.4.4.4. Méthodes basées sur la morphologie

La morphologie mathématique notamment l'ouverture (RCT) et la fermeture (ROT) morphologiques récursives ont été utilisées dans [CHE 95] pour l'estimation de l'angle d'inclinaison dans les images de document. L'algorithme proposé est résumé en cinq étapes : réduction de la résolution de l'image, remplissage de l'espace inter-caractères en utilisant la RCT, élimination des ascendants et des descendants par l'application de la ROT, optimisation de ligne, et estimation de l'inclinaison. La morphologie mathématique a été utilisée aussi dans [HON 02] comme prétraitement avant l'application de la transformée de Hough. L'opérateur de dilatation morphologique est d'abord appliqué sur l'image binaire dans le but de remplir l'espace inter-caractères. Un amincissement est ensuite procédé afin d'extraire le squelette de l'image dilatée. Finalement, l'angle d'inclinaison est détecté en appliquant la transformée de Hough sur l'image amincie.

1.4.4.5. Méthodes basées sur la corrélation croisée

Dans ces méthodes, la corrélation croisée est utilisée pour la détection de l'angle de l'inclinaison. La corrélation peut être calculée pour l'image entière comme dans [YAN 93], ou pour des petites régions de l'image, déterminées aléatoirement comme procédé dans [CHA 97]. Chen et Ding [CHE 99] ont ajouté une étape de vérification pour déterminer la convenance des régions choisies.

1.4.5. Segmentation

La segmentation, est l'une des opérations les plus essentielles de prétraitement et d'extraction de l'information, préalables à toute reconnaissance. La segmentation d'une image de document consiste à trouver sa structure physique en délimitant les différentes entités structurelles homogènes : colonnes, blocs graphiques, lignes de texte, mots, caractères, etc. [KET 10][TRU 05]. La structure physique du document peut être représentée par un arbre hiérarchique décrivant les relations spatiales entre les composantes physiques du document de manière hiérarchique [TRU 05]. Si les documents structurés comme les pages des journaux, les formulaires, les brochures, les courriers postaux, etc. possèdent une structuration riche,

cette structure se limite dans les documents manuscrits à une hiérarchie simple: les lignes de textes et parfois les mots qui la composent [LIK 08]. Dans les documents calligraphiés, la structure des colonnes ou des blocs en marge est aussi recherchée.

Cependant, l'extraction de la structure physique d'un document commence dans la plupart des cas par la séparation des entités textuelles de tous les objets non textuels (illustrations, dessins, figures, graphiques, cadres, etc.). Dans une deuxième étape, les zones de textes sont segmentées en colonnes, paragraphes, lignes, mots, caractères et ponctuations.

Nous présentons dans cette section les différentes approches utilisées pour la segmentation d'images de documents tout en nous limitant à la segmentation texte/graphique, la segmentation du texte en lignes, et en caractères.

1.4.5.1. Familles des techniques de segmentation

Les états de l'art séparent usuellement les techniques de segmentation de documents en deux familles : méthodes classiques et méthodes à base de texture [JOU 06][KET 10].

1.4.5.1.1. Les méthodes classiques de segmentation

Ces méthodes traitent généralement les documents imprimés à prédominance textuelle et présentant une structure simple [KET 10]. Elles sont applicables dans la plupart des cas sur des images binaire et elles peuvent être classées en trois approches :

a) Approche descendante

L'approche descendante est souvent utilisée pour des documents à structure bien définie. La segmentation dans cette approche s'appuie sur une forte connaissance a priori du modèle du document pour le découper en blocs de plus en plus fins.

b) Approche ascendante

Les méthodes ascendantes sont guidées par les données et n'incluent donc pas de connaissances sur le modèle. Ces méthodes partent des données de bas niveaux (couleur, position...) relatives aux pixels. Les pixels sont regroupés en caractères, les caractères en mots, les mots en lignes puis en paragraphes pour former la structure physique du document.

c) Approche mixte (hybride)

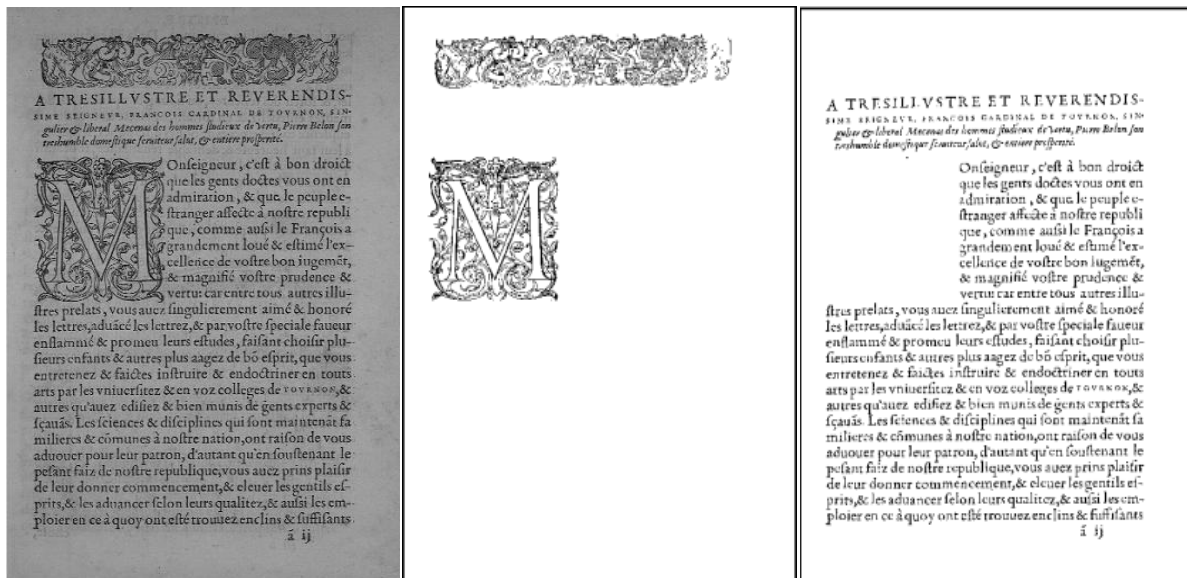
L'approche mixte ou hybride consiste à combiner le principe des deux approches descendantes et ascendantes.

1.4.5.1.2. Les méthodes de segmentation à base de texture

Les outils de caractérisation de textures sont traditionnellement issus de l'analyse d'images naturelles, mais ils se montrent bien adaptés à l'analyse de structure de documents [JOU 06]. Ces méthodes sont applicables sur des images couleurs et ne nécessitent pas une binarisation. En effet, les aspects fortement texturés du texte, des photos et des dessins rendent possible la description très fine de fontes, de polices mais aussi des éléments non textuels. On trouve : les approches statistiques (les matrices de Co-occurrence proposées par Haralick), les approches géométriques, les approches à base de modèles probabilistes (les champs de Markov, les fractales) et les approches d'ordre fréquentiel (les transformées de Fourier, les transformées de Gabor, les transformées d'ondelettes). Veiller voir [JOU 06] pour plus d'informations.

1.4.5.2. Segmentation texte/ graphique

Les traitements suivants dans le processus d'analyse de document, opèrent sur une image dont on a séparé le texte des graphiques. La distinction entre les éléments textuels (caractères, symboles) et les éléments graphiques est alors nécessaire. Selon le besoin de l'application visée, la sortie de cette étape peut être deux images (une contenant la partie textuelle et l'autre la partie graphique) dans le cas où l'analyse des parties graphiques est intéressante, ou bien une seule image propre, débarrassée d'éléments non textuels, si la partie graphique n'a pas d'importance pour le processus d'analyse. La figure 1.10 présente un exemple.



(a) Document original

(b) zone graphique

(c) zone textuelle

Figure 1.10. Segmentation texte/ graphique d'un document ancien

Plusieurs approches ont été proposées dans la littérature pour la séparation entre le texte et le graphique. Dans les images binaires, la taille des composantes connexes ou leur aspect constituent un critère important pour la distinction entre les composantes textuelles et les composantes graphiques [KHU 09b][WAK 98][BOS 08]. Des connaissances a priori sur la position des éléments non-textuels peuvent aussi aider à leur élimination, notamment pour les lettrines dans les manuscrits médiévaux [GUS 99]. Les densités de pixels noirs et blancs pour chaque composante connexe sont utilisées dans [HAD 06] pour la séparation texte/image dans les images à structure complexes et plus particulièrement les images de journaux. Les techniques basées sur la morphologie mathématique ont été aussi utilisées [GRA 00] [MEN 00]. Andersen et Zhang [AND 03] ont proposé une segmentation texte/dessin basée sur l'apprentissage automatique en utilisant un Perceptron Multicouche entraîné sur un certain nombre d'images segmentées manuellement.

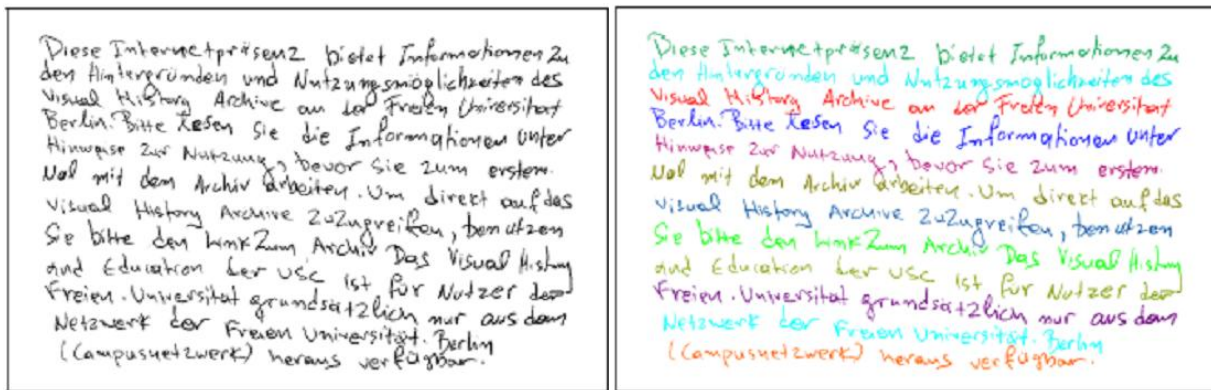
Pour les images couleurs, des méthodes qui se basent sur l'utilisation d'outils comme les filtres de Gabor, l'algorithme des C-moyen flous, ou les ondelettes et en s'appuyant sur l'étude des fréquences et des orientations des différentes parties de l'image ont été proposées [LEE 01] [BOU 06].

1.4.5.3. Segmentation du texte en lignes

L'extraction des lignes de textes est un préalable à tous les processus d'analyse et reconnaissance de mots ou des caractères. Elle a comme objectif d'assigner chaque composant du texte à une ligne appropriée ; ce qui permet de préparer les données pour les

traitements ultérieurs tel que la segmentation en mots, en caractères et l'extraction des caractéristiques (Figure 1.11). La segmentation du texte en lignes suppose l'élimination préalable des composantes graphiques apparaissant dans le document [LIK 03].

Cette opération est relativement facile quand le texte est régulier, non incliné, mono-orienté, et ne comportant pas de chevauchement. Ces conditions sont sans doute réunies pour des textes imprimés mais pas souvent pour des textes manuscrits. Ces derniers sont caractérisés par une variation de la distance interligne, la présence de plusieurs lignes de base générant des orientations différentes du texte et la fragmentation des caractères (due à la binarisation ou au non homogénéité de l'encre) et les caractères de deux lignes de texte peuvent se toucher ou se chevaucher. Tous ces problèmes compliquent considérablement la segmentation en lignes. Dans le cas de l'écriture arabe, ces situations existent fréquemment à cause de la présence de caractères ascendants et descendants.



(a) Image initiale

(b) Lignes séparées

Figure 1. 11. Extraction des lignes d'un texte manuscrit [OUW 10]

Les techniques d'extraction des lignes de texte se divisent comme toutes les méthodes de segmentation en deux classes : descendante et ascendante.

1.4.5.3.1. Méthodes descendantes

Les méthodes descendantes sont intéressantes lorsque des connaissances sur la structure du document sont disponibles, le modèle sous-jacent guide alors la segmentation [DEF 95]. Ces méthodes divisent le document pour aboutir aux lignes. Dans cette catégorie, nous trouvons la technique de la projection et la méthode guidée par le modèle de document [OUW 10].

a) Méthodes basées sur la projection

Le profil de projection d'une image est obtenu en additionnant les valeurs des pixels sur l'axe de l'orientation du document binaire. Le principe des méthodes basées sur le profil de projection est d'analyser les maximas et les minimas du profil de projection pour localiser les lignes de texte [NAG 84] [BAI 91] [ELL 90]. Ainsi, les minimas de l'histogramme correspondent aux zones de séparation entre les lignes du texte. Cependant ces techniques sont sensibles au chevauchement et à l'inclinaison du document. Pour surmonter le problème de l'inclinaison, Benasri et al. [BEN 99] ont proposé une méthode d'extraction des lignes d'un texte manuscrit arabe basée sur la projection partielle et le suivie de contour partiel. Cette méthode présente l'avantage qu'elle n'est pas sensible à l'inclinaison de l'écriture, mais elle impose que les lignes adjacentes ne doivent pas être collées. La projection horizontale a été utilisée aussi par Antonacopoulos et Karatzas dans [ANT 04]. Dans cette méthode, chaque minimum dans le profil de projection est considéré un point potentiel pour la segmentation en lignes. Ces minimas sont ensuite filtrés par analyse de distance.

Dans [NIC 09], les auteurs ont utilisé des traceurs (axes) des minima estimés à partir de l'histogramme de projection verticale pour segmenter les lignes des documents manuscrits latins. Ils ont parti de l'hypothèse topologique que, pour chaque ligne du texte, il existe un chemin allant d'un côté de l'image à l'autre qui ne traverse qu'une seule ligne. Ces chemins correspondent aux alignements des minima locaux des profils de projections des lignes.

b) Méthodes utilisant un modèle de document

Parmi les techniques les plus connues qui utilisent des modèles, nous trouvons le découpage en XY et le diagramme de Voronoï. Le principe du découpage en XY (XY-CUT) introduit par Nagy et al. dans [NAG 88] est de découper un document horizontalement et verticalement en plusieurs rectangles récursivement jusqu'à la satisfaction d'une condition. La condition d'arrêt est définie selon l'application souhaitée. Notons que l'algorithme de découpage en XY a été utilisé pour des documents qui ne contiennent pas beaucoup des variations [OUW 10].

L'algorithme de Voronoï est applicable sur des documents contenant beaucoup de variations. Il consiste à appliquer un pavage pour un document afin de trouver les zones souhaitées et permet ainsi de partitionner l'espace en régions à partir de points caractéristiques disposés sur un plan [OUW 10] [JOU 06]. Le regroupement des pavés se base sur la distance entre les composantes connexes [LU 04b], à l'aide du graphe de voisinage [KIS 98], etc.

1.4.5.3.2. Méthodes ascendantes

Les méthodes ascendantes partent des pixels ou des composantes connexes pour former les lignes. Nous trouvons sous cette catégorie le regroupement de composantes connexes par proximité, les méthodes basées sur la transformation de Hough, les méthodes basées sur la technique de lissage, etc. [OUW 10].

a) Méthodes basées sur le regroupement de composantes connexes

Dans [LIK 94], Likforman-Sulem et al. ont proposé une méthode itérative d'extraction des lignes adaptée aux documents manuscrits non-contraints, basée sur le groupement perceptif des composantes connexes. La méthode commence par l'étiquetage des composantes connexes. Ensuite, elle applique des masques sur les composantes connexes de l'image pour sélectionner celles qui ont une direction fiable. A partir des points d'ancrage¹, les composantes connexes qui satisfont certains critères de proximités, de similarité et continuité de direction sont regroupées dans un alignement. Feldbach et Tönnies [FEL 01] ont proposé une méthode utilisant les lignes de base et les lignes de centre pour extraire les lignes des documents de registres des anciennes églises. Les lignes de base sont extraites en regroupant les points minima voisins du squelette des composantes connexes. Ensuite, la ligne de centre de l'écriture est estimée en utilisant l'orientation de la ligne de base et les points terminaux dans le squelette. Ces alignements sont ensuite regroupés selon divers critères.

b) Méthodes basées sur la transformée de Hough

Likforman-Sulem et al. ont proposé dans [LIK 95] une autre méthode basée sur la transformée de Hough pour l'extraction des lignes des documents manuscrits. La transformée de Hough est utilisée ici pour la détection des alignements. Ces dernières sont validées par la suite en utilisant les informations contextuelles comme les voisines internes et externes. Une méthode basée sur la transformée de Hough et sur le regroupement des composantes connexes a été proposée dans [MAL 09]. La méthode consiste à analyser les orientations estimées en projetant les bords des composantes connexes afin de regrouper les composantes connexes qui ont les mêmes orientations. Dans [LOU 09], la transformée de Hough a été appliquée sur les

¹ Ce sont des composantes connexes ayant une direction privilégiée

composantes connexes en utilisant les centres de gravité de leurs rectangles englobants pour l'extraction des lignes et des mots des anciens manuscrits grecs.

c) Méthodes basées sur le lissage

Ces méthodes sont basées sur le lissage notamment sur l'algorithme RLSA (*Run Length Smearing Algorithm*) pour l'extraction des lignes de texte. Dans [BOU 97], les lignes sont extraites en appliquant l'algorithme RLSA adapté à une image en niveaux de gris. Au lieu de connecter une série de pixels noirs et blancs, le gradient de l'image est dilaté selon la direction horizontale. Shi et Govindaraju [SHI 04] ont proposée une méthode qui utilise une variété de l'algorithme RLSA, à savoir le RLSA flou. Dans cette méthode, les lignes de texte sont extraites en analysant les alignements obtenues du lissage RLSA flou. Ainsi, les alignements dont la taille n'excède pas un certain seuil sont supprimés.

d) Méthodes utilisant d'autres techniques

Plusieurs autres méthodes basées sur diverses techniques ont été proposées dans la littérature. Nous trouvons des méthodes fondées sur le réseau répulsif-attractif [OZT 99], des méthodes utilisent le regroupement par l'arbre recouvrant minimal [YIN 08][NIC 04], d'autres reposent sur le regroupement par la déformation de contour localisé et propagation [SUM 04] [LI 08] [DU 09], etc. Pour plus d'informations sur ces méthodes et autres, on vous invite à consulter l'état de l'art présenté par N. Ouwayed [OUW 10] dans sa thèse de doctorat.

1.4.5.4. Segmentation du mot en caractères

Quelques applications comme la reconnaissance de caractères ou bien de mots manuscrits (surtout dans le cas de grands vocabulaires) nécessitent une segmentation des mots de texte en caractères pour faciliter leur analyse. Cependant, l'extraction des primitives est plus facile sur un caractère que sur un mot. En effet cette séparation n'est pas toujours possible : les limites entre caractères surtout dans les mots manuscrits sont parfois difficiles à déterminer même pour un être humain, ce qu'affecte fortement le processus de reconnaissance. Cette difficulté a été clairement évoquée par Sayre en 1973 et peut être résumée par le dilemme suivant : « pour reconnaître les lettres, il faut segmenter le tracé et pour segmenter le tracé, il faut reconnaître les lettres » [SAY 73].

La majorité des travaux sur la segmentation en caractères ont été focalisés sur les mots manuscrits. Nous pouvons citer deux approches de segmentation des mots:

1.4.5.4.1. Segmentation explicite

Les approches de segmentation explicite ou discrète effectuent une sélection des points de segmentation les plus probables par une analyse des composantes. La solution communément adoptée consiste à segmenter le mot manuscrit en parties inférieures aux lettres appelés graphèmes et à retrouver les lettres puis le mot par combinaison de ces graphèmes. Parmi les travaux basés sur la segmentation en graphème, nous citons [ALB 06], [WSH 09], et [CHA 12]. D'autres points de segmentation ont été utilisés par les chercheurs pour la séparation des caractères. Ainsi, les minimas locaux du contour supérieur ont été utilisés dans [SAR 02] et [DIN 08], et les mesures de plages blanches entre les caractères ou bien les sous-mots ont été appliquées dans [KUN 07], [ABD 06] et [ALK 08].

A partir des points de segmentation potentiels, deux méthodes peuvent être utilisées pour choisir la segmentation finale de la séquence de lettres: les méthodes dites de segmentation puis reconnaissance et les méthodes de segmentation/reconnaissance [CHE 13]. Les premières choisissent les meilleurs points de segmentation sans recourir à la reconnaissance des

caractères, alors que les deuxièmes, beaucoup plus répandues, sont basées sur l'utilisation d'un moteur de reconnaissance de caractères pour valider la segmentation précédée.

Plusieurs techniques peuvent être classées dans cette catégorie : méthodes utilisant l'espace blanc et hauteur [HEN 68], analyse de projection verticale [NOM 05] [ANT 05], analyse de composantes connexes [WAN 94], et points de repère [HAR 72].

1.4.5.4.2. Segmentation implicite

Les approches à segmentation implicite considèrent tous les points du tracé comme points de segmentation potentiels à l'aide d'une fenêtre glissante qui parcourt la séquence de lettres à l'aide d'une fenêtre de taille fixe, en extrayant des caractéristiques de bas niveau. L'analyse des fenêtres est effectuée par un classifieur qui détermine la classe d'appartenance de chaque fenêtre en fonction des fenêtres voisines. Cependant, différents classifieurs ont été utilisés : les HMM (Hidden Markov Models) dans [ALH 09], [XIA 10], et [BEN 08] ; les réseaux de neurones artificiels dans [GRA 09], etc.

1.5. Conclusion

Nous avons vu dans ce chapitre que la mauvaise qualité des images de documents numérisés peut provenir de causes multiples. Elle provient de la qualité des documents originaux, de leurs conditions de préservation, ou du processus même de numérisation. Tout système d'analyse et de reconnaissance automatiques d'images de documents numérisés doit donc prendre en considération ces défauts en intégrant des opérations permettant de les contourner en vue d'atteindre des résultats satisfaisants. Ces opérations ayant pour but d'éliminer le bruit superposé aux données et d'améliorer la qualité des images sont communément appelées *Prétraitements, traitements de bas niveau* ou *traitements au niveau-pixel*.

Compte tenu de la variété des dégradations et de leurs causes, en plus de l'absence d'un objectif universel à toutes les applications d'analyse et de reconnaissance de documents, il est difficile de s'accorder sur un schéma universel de prétraitement. En effet tout traitement visant à nettoyer les images ou à améliorer leur qualité peut être incluse. Nous nous sommes intéressés dans ce chapitre à l'étude des principales opérations de prétraitement communément appliquées pour le rehaussement d'images documents. A chaque niveau de prétraitement, nous avons présenté les différentes approches proposées dans la littérature et les différentes techniques utilisées.

Chapitre 2 :
Séparation avant-
plan/ arrière-plan

2.1. Introduction

Usuellement, le traitement et l'analyse automatique d'images de documents doit s'occuper de la séparation entre l'avant-plan et l'arrière-plan de l'image. L'objectif principal de ce traitement est d'extraire les informations pertinentes (texte, figures, tableaux, etc.) ou « l'avant-plan » à partir du fond du document ou bien « l'arrière-plan ». Extraire l'avant-plan de l'arrière-plan permet d'éliminer les informations inutiles présentes dans l'image et par conséquent d'employer des techniques d'analyses simples vis-à-vis des images originales. La séparation avant-plan/arrière-plan est communément appelé « *binarisation* » lorsque le résultat de la séparation est une image à deux niveaux seulement, généralement noir et blanc, dont le noir représente l'avant-plan et le blanc représente l'arrière-plan.

En effet, les étapes suivantes dans le processus de traitement et d'analyse de documents sont applicables directement sur des images binaires, et les résultats de ces étapes dépendent fortement de la qualité des images binarisées obtenues. De telles façons, la séparation avant-plan/ arrière-plan est une étape critique ; une mauvaise séparation peut entraîner la perte d'informations pertinentes ou l'introduction du bruit dans l'image et dans les deux cas des résultats erronés. Cette difficulté augmente dans le cas des documents historiques caractérisés par des détériorations et des dégradations subies au cours de leur cycle de vie rendant leur traitement automatique difficile à plusieurs niveaux.

Cependant, un grand nombre de techniques ont été proposées dans la littérature pour la séparation entre l'avant-plan et l'arrière-plan dont la majorité d'entre elles sont applicables sur des images en niveaux de gris. D'autre part, avec le développement des dispositifs d'acquisition pour les images couleurs, des documents comme des livres, des magazines, des journaux, des carnets personnels, des documents historiques, etc. sont actuellement stockés telles qu'ils. Par conséquent, les recherches portant sur des documents couleurs ont également attiré une attention considérable dans le passé récent.

Dans le présent chapitre, nous nous intéressons à la séparation entre l'avant-plan et l'arrière-plan dans les images de documents. Nous étudions la binarisation en tant que la technique la plus usuelle de séparation et nous focalisons sur la binarisation d'images 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.

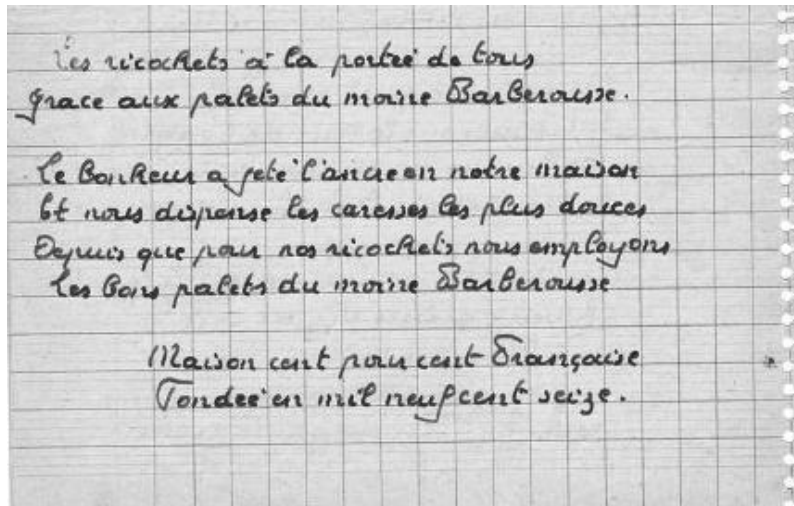
La suite de ce chapitre est organisée comme suit. Dans un premier temps, nous soulignons la différence entre les concepts de : *séparation avant-plan/ arrière-plan*, *binarisation*, et *seuillage*. Dans un deuxième temps, nous exposons les différentes approches de binarisation existantes et les diverses catégorisations proposées. Après, nous présentons un tour d'horizon d'un nombre important de méthodes de binarisation les plus fréquemment citées dans la littérature, et nous terminerons par une conclusion.

2.2. Séparation avant-plan/ arrière-plan, binarisation, et seuillage

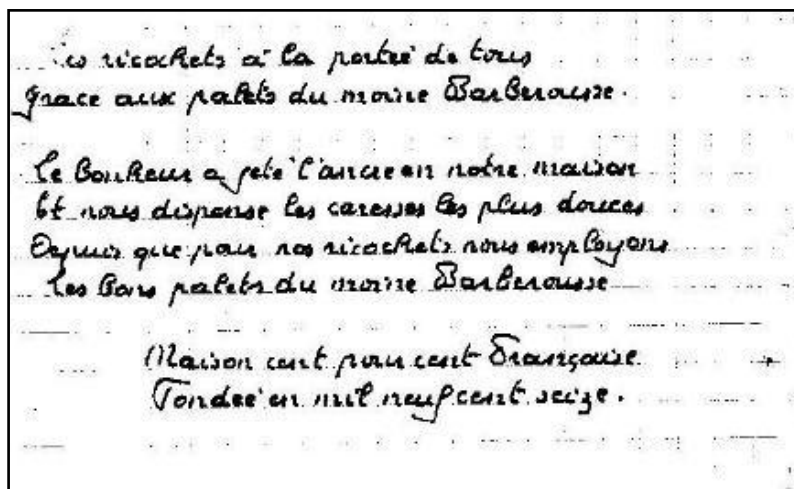
L'image issue de la numérisation d'un document comporte deux types d'informations, le fond (l'arrière-plan) et les objets de l'avant-plan (regroupant le texte, les symboles, les figures, les tableaux, etc.). Une des étapes clés dans le traitement et l'analyse d'images de documents est de séparer les deux types d'informations [NAG 00]. La séparation entre l'avant-plan et l'arrière-plan dans une image de document consiste à éliminer le fond, et à garder les

informations pertinentes de l'avant-plan, à savoir le texte, les figures, et les tableaux. Cette opération permet de réduire la quantité d'information à traiter, tout en conservant le signal dans sa quasi-intégralité.

En effet, la technique la plus usuelle d'atteindre la séparation avant-plan/ arrière-plan est de binariser l'image [LEE 02]. La binarisation est un traitement irréversible qui permet de transformer une image en niveaux de gris ou en couleurs, en une image noir et blanc [GAD 00] [DRI 07]. La binarisation d'une image de document doit donc nous fournir, dans le cas idéal, l'avant-plan en noir et l'arrière-plan en blanc. La figure 2.1. présente un exemple de binarisation d'une image en niveaux de gris.



(a) Image en niveaux de gris



(b) Image binarisée

Figure 2. 1. Exemple de binarisation d'une image en niveaux de gris

Notons que la binarisation est considérée aussi comme une technique simple de segmentation, où les pixels de l'image sont répartis en deux classes : les pixels qui appartiennent au fond et ceux appartenant à l'avant-plan (l'objet).

La plupart des méthodes de binarisation existantes sont basées sur un calcul de(s) seuil(s) afin de séparer les deux classes de pixels (seuillage). Cependant, plusieurs confondent entre la binarisation, l'opération de transformation d'une image en niveaux de gris ou en couleurs en une image en noir et blanc, et le seuillage, la technique la plus répandue de binarisation où la transformation s'effectue suite à un (des) seuil (s) à définir. Cette confusion peut être sentie clairement dans les deux définitions suivantes :

- « L'opération de binarisation vise à séparer l'information manuscrite du fond de l'image à l'aide d'une méthode de seuillage » [CHA 06].
- « Le principe commun à toutes les méthodes de binarisation est de calculer un seuil τ et d'affecter 1 si la valeur du pixel est supérieur à τ et 0 sinon » [BAU 05].

Cependant, comme a signalé [GAR 06], l'introduction des documents couleurs a induit une différence très fine entre la binarisation et la séparation avant-plan/ arrière-plan. Bien que la binarisation essentiellement accomplit la séparation avant-plan/ arrière-plan dans les images de documents, Garain et al. [GAR 06] ont considéré que pour les documents couleurs, la séparation avant-plan / arrière-plan réfère à un aspect plus général parce que les documents couleurs contiennent très souvent des éléments du premier plan (et peut être de l'arrière-plan aussi) en différentes couleurs. En conséquence, pour eux, la séparation entre l'avant-plan et l'arrière-plan ne doit pas uniquement classifier les pixels de l'image en avant-plan et arrière-plan, mais il est intéressant aussi d'étiqueter les différentes entités de l'avant-plan en les attribuant des couleurs différentes (Figure 2.2).

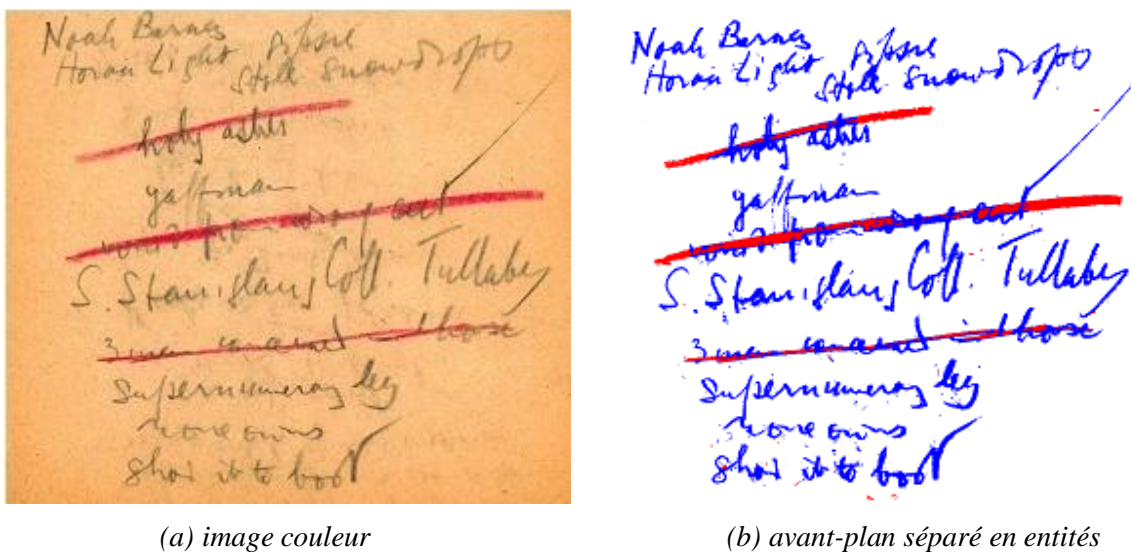


Figure 2. 2. Séparation avant-plan/ arrière-plan d'un document couleur [GAR 06]

Dans ce qui suit, nous considérons uniquement la binarisation qui est la solution la plus usuelle de séparation entre l'avant-plan et l'arrière-plan, et nous focalisons sur les techniques de binarisation d'images en niveaux de gris.

2.3. Différentes approches de séparation avant-plan/ arrière-plan

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. D'autre part, très peu d'études ont été intéressées à l'extraction du premier-plan dans les images de documents couleurs. Nous trouvons dans la littérature plusieurs articles de synthèse sur les méthodes de binarisation qui ont été publiés aussi non seulement pour leur présentation, mais aussi pour leur catégorisation, et évaluation.

Ainsi, 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 de seuillage global* et *méthodes de seuillage local* [TRI 95][KHU 09]. Les méthodes de seuillage global calculent un seuil unique pour l'image entière en utilisant des informations globales sur l'image. Ces méthodes utilisent diverses manières

pour trouver le seuil optimal de binarisation. D'après [STA 08], cinq groupes de méthodes globales peuvent être distingués ; les méthodes basées sur des procédures de classification, sur l'histogramme, sur le regroupement (*clustering*), sur l'entropie, et sur la distribution des Gaussiennes. Les méthodes de seuillage local assignent un seuil à chaque pixel ou petite région de l'image en se basant sur des informations locales. Cinq groupes de méthodes locales existent aussi selon [STA 08]. Nous trouvons 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 segmenter l'image en avant-plan et arrière-plan [SAU 00][STA 08][SIN 11]. En effet, les méthodes globales ont été largement utilisées dans plusieurs applications à cause de leur simplicité. Cependant ces méthodes sont performantes uniquement lorsque les documents originaux sont de bonne qualité, ayant un histogramme bimodal qui sépare l'avant-plan et l'arrière-plan. Lorsque les documents soient de mauvaise qualité, contenant différents types de bruits et de dégradations (taches d'encre et d'humidité, effet de transparence, etc.) avec un fond texturé et illumination inégale ; 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, une analyse plus détaillée est nécessaire, et nous faisons recours aux méthodes locales. Ces méthodes sont plus précises et peuvent être appliquées aux fonds variables, très foncés ou avec un faible contraste, mais elles sont très lentes et coûteuses en temps à cause du calcul qui s'effectue pour chaque pixel séparément en se basant sur des informations de son voisinage.

Comme nous avons signalé précédemment, les méthodes basées sur le seuillage constituent la grande partie des techniques de binarisation proposées dans la littérature, mais il existe d'autres techniques de binarisation n'utilisant pas le seuillage. De ce fait, d'autres catégorisations plus générales ont été proposées. Kumar et al. [KUM 13], ont considéré que les techniques de binarisation d'images peuvent être réparties en *méthodes supervisées* et *méthodes non-supervisées*. Les méthodes supervisées utilisent un algorithme d'apprentissage automatique entraîné en se basant sur un ensemble d'apprentissage pour accomplir la binarisation. Dans les méthodes non supervisées, des informations statistiques sont utilisées pour segmenter l'image en deux parties. Les informations statistiques employées peuvent être globales ou locales et en ce sens les méthodes non supervisées sont des méthodes basées sur le seuillage.

Une autre catégorisation a été décrite dans [FER 10]. Dans cette catégorisation, les méthodologies de binarisation d'images sont divisées en deux catégories principales : *méthodes basées sur le seuillage*, et *méthodes basées sur le regroupement*. Les méthodes basées sur le seuillage utilisent un seuil global ou local pour binariser l'image. Dans les méthodes basées sur le regroupement nous distinguons deux classes : *méthodes de regroupement basées sur la région*, et *méthodes de regroupement basées sur la classification (clustering)*. Les *méthodes de regroupement basées sur la région* sont principalement basées sur la *croissance de région*, ou sur la *division et la fusion (splitting and merging)*, alors que les *méthodes de regroupement basées sur la classification* reposent sur la classification des valeurs d'intensités ou de couleurs en fonction d'un critère d'homogénéité.

Concernant la binarisation d'images couleurs, Garain et al. [GAR 06] ont mentionné que les techniques de binarisation ou de séparation avant-plan/ arrière-plan proposées pour les images de documents couleurs sont principalement fondées sur le principe de *classification des couleurs* ou de *segmentation des couleurs*, et en ce sens, ces techniques sont très différentes des algorithmes de binarisation traditionnels basés sur le seuillage. De même, des algorithmes de binarisation génériques, ont été personnalisés de plusieurs façons pour atteindre une

séparation satisfaisante entre l'avant-plan et l'arrière-plan dans les images de document couleurs [GAP 06].

Sezgin et Sankur [SEZ 04] ont établi une classification des méthodes de binarisation selon l'information qu'elles exploitent en six catégories :

- **Méthodes basées sur la forme de l'histogramme** : les méthodes de cette classe effectuent un seuillage en se basant sur la forme de l'histogramme, par exemple, les pics, les vallées, et les courbures.
- **Méthodes basées sur le regroupement (clustering)** : dans ces méthodes, les niveaux de gris sont regroupés en deux parties : avant-plan et arrière-plan, ou alternativement sont modélisés comme une mixture de deux gaussiennes.
- **Méthodes basées sur l'entropie** : les méthodes de cette classe utilisent l'entropie de l'avant-plan et de l'arrière-plan de l'image, l'entropie croisée entre l'image originale et l'image binarisée, etc. pour calculer le seuil.
- **Méthodes basées sur les attributs d'objets** : ces méthodes tendent à trouver un seuil en se basant sur quelques mesures de similarité entre l'image binarisée et l'image en niveaux de gris comme la similarité des formes, etc.
- **Méthodes de binarisation spatiales** : les méthodes spatiales utilisent la distribution de probabilité de haut niveau et/ou la corrélation entre les pixels pour trouver le seuil.
- **Méthodes locales** : ces méthodes adoptent une valeur de seuil pour chaque pixel.

Le reste du chapitre est réservé à la présentation des différentes techniques de binarisation les plus fréquemment citées dans la littérature sous une notation uniforme. Nous adoptons la catégorisation décrite dans [FER 10] en *méthodes basées sur le seuillage* et *méthodes basées sur le regroupement*, et la classification des méthodes basées sur le regroupement en *méthodes basées sur la région* et *méthodes basées sur la classification*. Nous considérons la répartition des méthodes de seuillage en *méthodes globales*, *méthodes locales*, et *méthodes hybrides* en tenant compte de la classification de Sezgin et Sankur [SEZ 04] et celle de Stathis et al. [STA 08] dans chaque catégorie.

2.4. Binarisation basée sur le seuillage

Notons I l'image en niveaux de gris dont les intensités varient de 0 (noir) à 255 (blanc), et I_b l'image binaire résultante. $I(x,y)$ et $I_b(x,y)$ dénotent l'intensité du pixel (x,y) dans l'image I et I_b respectivement.

2.4.1. Méthodes de seuillage global

Notons h l'histogramme d'intensités correspondant à l'image en niveaux de gris. Le nombre de pixels ayant le niveau de gris i est noté h_i .

Les méthodes de seuillage global calculent un seul seuil global T_{opt} pour séparer l'image en deux classes *Objet (ou avant-plan)* et *Fond (ou arrière-plan)*. Ainsi, le seuillage s'effectue comme suit :

$$I_b(x, y) = \begin{cases} 255, & \text{si } I(x, y) > T_{opt} \\ 0, & \text{si } I(x, y) \leq T_{opt} \end{cases}$$

Notons $p_i = p(i)$, la probabilité d'occurrence du niveau de gris i dans l'image.

On note la probabilité cumulative, $P_g = \sum_{i=0}^g p_i$

La probabilité des classes *Objet* et *Fond* séparées par le niveau de gris t sont notées $P_f(t)$ et $P_b(t)$ respectivement, avec :

$$P_f(t) = P_t = \sum_{i=0}^t p_i \quad \text{et} \quad P_b(t) = 1 - P_t = \sum_{i=t+1}^{255} p_i$$

Notons $\mu_f(t)$ et $\mu_b(t)$ les moyennes des deux classes en fonction du niveau de gris t .

$$\mu_f(t) = \frac{\sum_{i=0}^t i \times h_i}{\sum_{i=0}^t h_i} \quad \text{et} \quad \mu_b(t) = \frac{\sum_{i=t+1}^{255} i \times h_i}{\sum_{i=t+1}^{255} h_i}$$

Nous distinguons six groupes de méthodes de seuillage global.

2.4.1.1. Méthodes de seuillage naïves

Ce sont des méthodes qui utilisent des procédés simples pour le choix du seuil de binarisation.

2.4.1.1.1. Seuillage global fixe

C'est la technique de binarisation la plus simple, elle consiste à comparer le niveau de gris de chaque pixel de l'image avec un seuil global fixe T_{opt} (par exemple 127) [GUP 07].

2.4.1.1.2. Seuillage percentile (méthode P-tile)

La méthode *P-tile* est l'une des techniques de seuillage les plus anciennes. Cette méthode considère l'image comme composée d'objets foncés sur un fond clair [DOY 62]. En supposant que le pourcentage des pixels d'objets est connu (p), le seuil est défini comme le niveau de gris le plus haut qui sélectionne au moins p % de pixels comme des pixels d'objets dans l'image binarisée.

2.4.1.1.3. Méthode de Khashman et Sekeroglu

Cette méthode calcule un seuil global ou un point de séparation entre l'avant-plan et l'arrière-plan pour chaque document en utilisant la relation entre la valeur de luminance et la moyenne d'intensités de l'image [KHA 07]. Le seuil de binarisation T_{opt} est donné par :

$$T_{opt} = 2 \cdot \mu - G_{max}$$

Tel que : μ est le niveau de gris moyen, G_{max} est le niveau de gris le plus haut dans l'image.

2.4.1.2. Méthodes basées sur l'entropie

Ces méthodes exploitent l'entropie de distribution des niveaux de gris pour l'estimation d'un seuil global de binarisation.

2.4.1.2.1. Méthode de Pun

Dans cette méthode, le seuil optimal T_{opt} est choisi comme le niveau de gris qui maximise la somme de l'entropie de l'avant-plan (H_f) et l'entropie du fond (H_b) données par [PUN 80]:

$$H_f(t) = - \sum_{i=0}^t p_i \times \log(p_i) \quad \text{et} \quad H_b(t) = - \sum_{i=t+1}^{255} p_i \times \log(p_i)$$

2.4.1.2.2. Méthode de Johannsen et Bille

L'algorithme de Johannsen et Bille [JOH 82] essaye de trouver le seuil T_{opt} qui minimise la fonction : $S(t) = S_f(t) + S_b(t)$, tel que :

$$S_f(t) = \log(P_f(t)) + (1/P_f(t)) [E(p_t) + P_{t-1}]$$

$$S_b(t) = \log(P_b(t)) + (1/P_b(t)) [E(p_t) + E(P_b(t))]$$

Avec : $E(x) = -x \log(x)$.

2.4.1.2.3. Méthode de Kapur et al.

La méthode de Kapur et al. [KAP 85] est considérée comme une extension de la méthode de Pun [PUN 80]. A la déference de cette dernière, la méthode de Kapur et al. prend en compte la distribution de probabilité de l'objet P_f et la distribution de probabilité du fond P_b dans la détermination de l'entropie de division. Le seuil de binarisation T_{opt} est choisi dont l'entropie $H = H_f + H_b$ soit maximale, avec :

$$H_f(t) = -\sum_{i=0}^t \frac{P_i}{P_f(t)} \times \log\left(\frac{P_i}{P_f(t)}\right) \quad \text{et} \quad H_b(t) = -\sum_{i=t+1}^{255} \frac{P_i}{P_b(t)} \times \log\left(\frac{P_i}{P_b(t)}\right)$$

2.4.1.2.4. Méthode de seuillage par entropie croisée (Li et Lee)

Li et Lee ont proposé une méthode de seuillage où le regroupement des niveaux de gris en deux classes (objet et fond) est fondé sur la minimisation de l'entropie croisée [LI 93]. Le seuil optimal T_{opt} est calculé de façon à minimiser : $\eta(t) = \eta_f(t) + \eta_b(t)$, où :

$$\eta_f(t) = \sum_{i=0}^t i \times h_i \times \log\left(\frac{i}{\mu_f(t)}\right) \quad \text{et} \quad \eta_b(t) = \sum_{i=t+1}^{255} i \times h_i \times \log\left(\frac{i}{\mu_b(t)}\right)$$

2.4.1.2.5. Seuillage entropique flou (Shanhsbag)

Cette méthode considère les appartenances floues comme un indicateur de la puissance d'appartenance d'une valeur de gris à l'avant-plan ou à l'arrière-plan [SHA 94]. Ainsi, le seuil optimal T_{opt} est choisi comme le niveau de gris qui minimise la somme des entropies floues de l'avant-plan (H_f) et de l'arrière-plan (H_b).

$$H_f(t) = -\sum_{i=0}^t \frac{P_i}{P_f(t)} \times \log(\varphi_f(i)) \quad \text{et} \quad H_b(t) = -\sum_{i=t+1}^{255} \frac{P_i}{P_b(t)} \times \log(\varphi_b(i))$$

φ_f et φ_b sont les degrés d'appartenance d'un pixel ayant un niveau de gris t aux classes *Objet* et *Fond* respectivement.

$$\varphi(t-i) = 0.5 + \frac{p(t) + \dots + p(t-1-i) + p(t-i)}{2 \times P_f(t)}$$

$$\varphi(t+i) = 0.5 + \frac{p(t+1) + \dots + p(t-1+i) + p(t+i)}{2 \times P_b(t)}$$

2.4.1.2.6. Méthode de Yen et al.

Considérant l'avant-plan et l'arrière-plan de l'image comme deux sources de signaux différents, Yen et al. [YEN 95] choisissent le seuil comme le niveau de gris qui maximise la corrélation entropique (HC). Cette dernière est donnée par :

$$HC(t) = HC_f(t) + HC_b(t) = -\log\left(\sum_{i=0}^t \left(\frac{P_i}{P_f(t)}\right)^2\right) - \log\left(\sum_{i=t+1}^{255} \left(\frac{P_i}{P_b(t)}\right)^2\right)$$

2.4.1.2.7. Méthode de Cheng et al.

La méthode de Cheng et al. est basée sur le principe d'entropie maximale et la c -partition floue, pour le choix du seuil [CHE 98]. Considérons deux ensembles flous *Objet* et *Fond* dont les fonctions d'appartenance sont définis par :

$$\varphi_f = \begin{cases} 1, & \text{si } x \leq a \\ \frac{x-c}{a-c}, & \text{si } a < x < c \\ 0, & \text{si } x \geq c \end{cases} \quad \varphi_b = \begin{cases} 0, & \text{si } x \leq a \\ \frac{x-a}{c-a}, & \text{si } a < x < c \\ 0, & \text{si } x \geq c \end{cases}$$

Dans cette méthode, le seuil de binarisation est choisi comme le niveau de gris dont la fonction d'appartenance=0.5, et donc c'est le centre de l'intervalle $[a_{opt}, c_{opt}]$, tel que a_{opt} et c_{opt} sont les valeurs de a et c maximisant l'entropie de la division.

Cheng et al. ont proposé un algorithme pour trouver les valeurs optimales de a et c correspondants à l'entropie maximale de partitionnement. L'entropie de partitionnement est donnée par :

$$H(t) = -P_f(t) \log(P_f(t)) - P_b(t) \log(P_b(t))$$

2.4.1.2.8. Première méthode de Mello et Lins

Cette méthode a été proposée pour la binarisation d'images de documents historiques [MEL 00]. L'algorithme prend le niveau de gris le plus apparaissant t_l comme seuil initial et calcule pour ce seuil l'entropie de l'objet H_f et du fond H_b comme dans la méthode de Pun [PUN 80]. Le seuil optimal de binarisation est ainsi :

$$T_{opt} = 256(m_b \times H_b + m_f \times H_f)$$

Les deux facteurs m_b et m_f sont déterminés en évaluant l'entropie de l'image entière (H), et ils sont attribués comme suit :

- Si $0.25 < H < 0.30$, alors $m_b = 1$ et $m_f = 2.6$
- Si $H \leq 0.25$, alors $m_b = 2$ et $m_f = 3$
- Si $0.30 \leq H < 0.305$, alors $m_b = 1$ et $m_f = 2$
- Si $H \geq 0.305$, alors $m_b = 0.8$ et $m_f = 0.8$

2.4.1.2.9. Méthode de Portes de Albuquerque et al.

Portes de Albuquerque et al. [ALB 04] ont proposé un algorithme de seuillage entropique inféré du concept de l'entropie non-extensive de Tsallis [TSA 88]. Dans cette méthode, les entropies associés aux classes *Objet* (H_f) et *Fond* (H_b) séparées par le seuil t sont :

$$H_f(t) = \frac{1 - \sum_{i=0}^t \left(\frac{p_i}{P_f(t)} \right)^q}{q-1}, \quad H_b(t) = \frac{1 - \sum_{i=t+1}^{255} \left(\frac{p_i}{P_b(t)} \right)^q}{q-1}$$

Tel que : q est un indice entropique caractérisant le degré de non-extensivité.

Le seuil optimal T_{opt} est choisi comme le niveau de gris maximisant la fonction suivante :

$$H_f(t) + H_b(t) + (1-q)(H_f(t) \times H_b(t))$$

2.4.1.2.10. Méthode de Silva et al.

Dans cette méthode, le seuil optimal est choisi comme le niveau de gris qui minimise $|e(t)|$ définie par [SIL 06]:

$$|e(t)| = \left| \frac{H'(t)}{H/8} - \alpha(H/8) \right|$$

Avec : H l'entropie de l'histogramme donnée par : $H = -\sum_{i=0}^{255} p_i \times \log(p_i)$

$H'(t)$ est l'entropie à posteriorie. $H'(t) = -P_f(t) \log(P_f(t)) - P_b(t) \log(P_b(t))$

α est un facteur de perte.

$$\alpha(H_{a \text{ priori}}) = \begin{cases} -\frac{3}{7} H_{a \text{ priori}} + 0.8 & \text{si } H_{a \text{ priori}} < 0.7 \\ H_{a \text{ priori}} - 0.2 & \text{si } H_{a \text{ priori}} \geq 0.7 \end{cases}$$

2.4.1.2.11. Deuxième méthode de Mello et al.

Mello et al. ont proposé une méthode composée de deux algorithmes basés sur l'entropie de Tsallis pour la binarisation d'images de documents historiques [MEL 08]. Le premier algorithme calcule un seuil global comme suit : d'abord le document est attribué à une de trois classes en évaluant l'entropie de Shannon :

$$H = -\sum_{i=0}^{255} p_i \times \log(p_i)$$

Si $H \leq 0.26$, le document appartient à la classe 1. Si $0.26 < H < 0.30$, le document appartient à la classe 2. Si $H \geq 0.3$, le document est attribué à la classe 3.

Ensuite, on calcule l'entropie de Tsallis de l'avant-plan H_{qf} et de l'arrière-plan H_{qb} limités par le niveau de gris le plus fréquent t_l .

$$H_{qf} = \frac{X_f}{q-1} - \frac{1}{q-1} \sum_{i=0}^{t_l} (p_i)^q \quad \text{et} \quad H_{qb} = \frac{X_b}{q-1} - \frac{1}{q-1} \sum_{i=t_l+1}^{255} (p_i)^q$$

X_f et X_b sont le pourcentage des niveaux de gris au dessous et au dessus de t_l respectivement, et q un paramètre fixé 0.3 par les auteurs.

Le seuil final estimé par le premier algorithme est : $T_{opt} = m_f \times H_{qf} + m_b \times H_{qb}$

m_f et m_b sont des constantes définis pour chaque classe à part.

Si le seuil $T_{opt} > t_1$, le deuxième algorithme est appelé pour estimer une nouvelle valeur du seuil T_{opt} . Dans le deuxième algorithme le document est attribué à la classe 1: si $H \leq 0.23$; à la classe 2: si $0.23 < H < 0.28$; et à la classe 3: si $H \geq 0.28$.

La valeur du paramètre q est fixée pour chaque classe différemment. $q = 0.04$ pour la classe 1 ; $q = 0.05$ pour la classe 2 ; et $q=0.3$ pour la classe 3. Cette valeur est utilisée pour le calcul de l'entropie de Tsallis de l'avant-plan H_{qf} et de l'arrière-plan H_{qb} limités par le niveau de gris le plus fréquent t_1 .

$$H_{qf} = \frac{1 - \sum_{i=0}^{t_1} \left(\frac{P_i}{P_f(t)} \right)^q}{q-1} \quad \text{et} \quad H_{qb} = \frac{1 - \sum_{i=t_1+1}^{255} \left(\frac{P_i}{P_b(t)} \right)^q}{q-1}$$

Le seuil final est : $T = H_{qb} + H_{qf}$.

Notons que les documents de la troisième classe nécessite un prétraitement en appliquant un filtre de racine carrée.

2.4.1.3. Méthodes basées sur la similarité d'attributs

Les méthodes de cette classe choisissent la valeur du seuil en se basant sur des mesures de similarité entre l'image originale et sa l'image binarisée.

2.4.1.3.1. Seuillage par préservation des moments (Tsai)

Tsai [TSA 85] a considéré l'image en niveaux de gris comme une version floue d'une image binaire idéale. Il propose ainsi de déterminer le seuil optimal comme le niveau de gris avec lequel l'image binarisée préserve le mieux les trois moments mathématiques de l'image en niveaux de gris. Ainsi, le seuil optimal T_{opt} est le niveau de gris qui garantie :

$$\{m_1=b_1(t), m_2=b_2(t), m_3=b_3(t)\}$$

Tel que : m_k est le $k^{\text{ème}}$ moment de l'image en niveaux de gris et b_k le $k^{\text{ème}}$ moment de l'image binaire, et sont données par :

$$m_k(t) = \sum_{i=0}^{255} P_i \times i^k \quad \text{et} \quad b_k(t) = P_f(t) \times m_f^k + P_b(t) \times m_b^k$$

$$\text{Avec : } m_f^k = \sum_{i=0}^t P_i \times i^k \quad \text{et} \quad m_b^k = \sum_{i=t+1}^{255} P_i \times i^k$$

Cette méthode a été classée en première place dans des études comparatives précédentes [SAH 88] [OGO 94].

2.4.1.3.2. Méthode de Huang et Wang

Huang et Wang proposaient une méthode de seuillage basée sur la minimisation d'une mesure de floue [HUA 95]. La mesure de floue de l'image entière peut être obtenue par l'entropie de Shannon ou par la mesure de Yager.

L'entropie de Shannon pour une image de taille $M \times N$ est définie comme suit :

$$H = \frac{1}{MN \log(2)} \sum_{i=0}^{255} S(\varphi(i)) \times h_i$$

Avec : $S(t) = \sum_{i=0}^t h_i$

$\varphi(i)$: la fonction d'appartenance d'un pixel ayant le niveau de gris i et elle est donnée comme suit :

$$\varphi(i) = \begin{cases} \frac{1}{1 + |i - \mu_f|/C} & \text{si } i \leq t \\ \frac{1}{1 + |i - \mu_b|/C} & \text{si } i > t \end{cases}$$

C : la distance entre le niveau de gris minimum et le niveau de gris maximum dans l'image.

2.4.1.3.3. Méthode de Hertz et Schafer

Dans cette méthode, le seuil global est choisi comme le niveau de gris maximisant la similarité entre les contours obtenus à partir de l'image originale et les contours dérivés de l'image binarisée [HER 88]. Les contours des deux images sont obtenus en utilisant un opérateur de Sobel.

2.4.1.4. Méthodes basées sur la forme de l'histogramme

Cette catégorie de méthodes effectue un seuillage en se basant sur les propriétés de la forme de l'histogramme.

2.4.1.4.1. Méthode de Fisher

Fisher fait l'hypothèse que l'histogramme des niveaux de gris de l'image (h) est un mélange de deux gaussiennes [FIS 58]. Il faut donc déterminer le seuil comme le séparateur des deux distributions. Pour cela il utilise un critère de minimisation de la somme des inerties des classes.

$$Iner(t) = \sum_{i=0}^t h_i (i - \mu_f(t))^2 + \sum_{i=t+1}^{255} h_i (i - \mu_b(t))^2$$

2.4.1.4.2. Méthode du Mode (Prewitt et Mendelsohn)

Cette technique a été proposée par Prewitt et Mendelsohn [PRE 66]. En supposant que l'histogramme est bimodal, l'histogramme est lissé en utilisant un filtre mobile de trois points itérativement jusqu'à ce qu'il n'y en aura que deux maxima locaux j et k . Le seuil optimal T_{opt} est choisi égal à $(j + k) / 2$.

2.4.1.4.3. Méthode de Triangle (Zack et al.)

Le principe de cette méthode est simple : une ligne droite L est tirée du maximum (pic) de l'histogramme h_{max} à la fin de l'histogramme (niveau de gris le plus haut) pour construire la base triangulaire; le seuil est ensuite sélectionné au point de l'histogramme qui maximise la distance perpendiculaire d entre la ligne et l'histogramme [ZAC 77].

2.4.1.5. Méthodes de seuillage basées sur le regroupement des niveaux de gris

Dans cette classe d'algorithmes, les niveaux de gris subissent un processus de regroupement en deux classes : *Objet* et *Fond*.

2.4.1.5.1. Méthode d'Otsu

La méthode d'Otsu [OTS 79] essaye de trouver le seuil T_{opt} qui sépare l'histogramme de façon optimale en deux segments, ou en d'autres termes qui maximise la variance inter-segments ou bien qui minimise la variance intra-segments. Le calcul de la variance interclasses ou intra-classes est basé sur l'histogramme normalisé M de l'image d'où

$$\sum_{i=1}^{255} M_i = 1.$$

La variance interclasses pour chaque niveau de gris t est donnée par :

$$\sigma_{inter}^2 = q_f(t) \times q_b(t) \times [\mu_f(t) - \mu_b(t)]^2$$

Tel que:

$$\mu_f(t) = \frac{1}{q_f(t)} \sum_{i=0}^{t-1} i \times M_i \quad \text{et} \quad \mu_b(t) = \frac{1}{q_b(t)} \sum_{i=t}^{255} i \times M_i$$

$$q_f(t) = \sum_{i=0}^{t-1} M_i \quad \text{et} \quad q_b(t) = \sum_{i=t}^{255} M_i$$

La limite de cette méthode est qu'elle n'est applicable que lorsque l'image est bimodale (l'histogramme comporte deux pics).

Wolf et Jolion [WOL 02b] ont proposé d'utiliser une version deux-seuils de la méthode d'Otsu pour la binarisation d'images de la vidéo, dont le premier seuil k_{haut} correspond au seuil calculé par la méthode d'Otsu, et le deuxième seuil k_{bas} est calculé à partir du seuil k_{haut} et le premier mode m_0 de l'histogramme : $k_{bas} = m_0 + 0.87 (k_{haut} - m_0)$. Ainsi la décision d'appartenance à une des classes Objet et Fond est modifiée comme suit :

$$I(x,y) < k_{bas} \quad \rightarrow \quad I_b(x,y) = 0$$

$$I(x,y) > k_{haut} \quad \rightarrow \quad I_b(x,y) = 255$$

$$I(x,y) > I_{x,y} > k_{haut} \quad \rightarrow \quad I_b(x,y) = \begin{cases} 255 & \text{s'il existe un chemin à un pixel } I_b(u,v) > k_{haut} \\ 0 & \text{sinon} \end{cases}$$

2.4.1.5.2. Méthode ISODATA

Le seuillage par ISODATA consiste à trouver un seuil en séparant l'histogramme en deux classes itérativement avec la connaissance a priori des valeurs associées à chaque classe [VEL 80]. Cette méthode commence par la division de l'intervalle de valeurs non-nulles de l'histogramme en deux parties équidistantes, et ensuite prend μ_f et μ_b comme les moyennes arithmétiques des deux classes. Répéter jusqu'à la convergence, le calcul du seuil T_{opt} comme l'entier le plus proche de $(\mu_f + \mu_b) / 2$, et mettre à jour les deux moyennes μ_f et μ_b .

2.4.1.5.3. Seuillage par l'erreur minimale (Kittler et Illingworth)

Cette technique voit l'histogramme de niveaux de gris de l'image comme composé de deux classes (*Objet* et *Fond*) normalement distribuées [KIT 86]. L'histogramme est utilisé pour classer l'erreur d'un mélange de deux gaussiennes. L'algorithme permettra de repérer le seuil optimal qui minimise la probabilité d'erreur de classification.

Le seuil optimal est donc le niveau de gris qui minimise le critère $J(t)$ suivant :

$$J(t) = 1 + 2[P_f(t) \log(\sigma_f(t)) + P_b(t) \log(\sigma_b(t))] - 2[P_f(t) \log(P_f(t)) + P_b(t) \log(P_b(t))]$$

Tel que : $\sigma_f(t)$ et $\sigma_b(t)$ sont les écarts types des fonctions de densité Gaussienne des classes *Objet* et *Fond* respectivement.

2.4.1.5.4. Seuillage global itératif (Kavalieratou)

Kavalieratou a proposé un algorithme simple et rapide de seuillage appliqué sur les images de documents et les photos [KAV 05]. La méthode proposée sélectionne un seuil global à l'image entière en se basant sur une procédure itérative. A chaque itération i , les étapes suivantes sont exécutées :

- Calcul du niveau de gris moyen (T_i) de l'image,
- Soustraction de T_i de tous les pixels de l'image,
- Egalisation de l'histogramme pour étendre les pixels sur tout l'intervalle de niveaux de gris.

Après chaque itération, certains pixels sont transformés à l'arrière-plan. L'algorithme s'arrête lorsque : $|T_i - T_{i-1}| < 0.001$. L'auteur a annoncé que le nombre de répétitions ne dépasse jamais 20 répétitions.

2.4.1.5.5. Méthode de Ridler et Calvard

Cette méthode [RID 78] est basée sur une mixture de gaussienne de deux classes. A la $n^{\text{ième}}$ itération, un nouveau seuil T_n est établi en utilisant la moyenne de l'avant-plan (μ_f) et la moyenne de l'arrière-plan (μ_b):

$$T_{opt} = \lim_{n \rightarrow \infty} T_n \quad \text{tel que} \quad T_{n+1} = \frac{\mu_f(T_n) + \mu_b(T_n)}{2}$$

2.4.1.5.6. Méthode de Lloyd

Lloyd [LLO 85] a supposé que l'image est caractérisée par deux fonctions de densité gaussienne à variance égale. Le seuil de binarisation est choisi comme le niveau de gris qui minimise l'erreur totale de classification.

$$T_{opt} = \arg \min \left[\frac{\mu_f(t) + \mu_b(t)}{2} + \frac{\sigma^2}{\mu_f(t) - \mu_b(t)} \log \frac{P_b(t)}{P_f(t)} \right]$$

σ^2 est la variance de l'image entière.

2.4.1.5.7. Méthode de Jawahar et al.

Cette méthode [JAW 97] assigne des appartenances de regroupement floues aux pixels en fonction de leur distance aux moyennes des classes. Les moyennes des deux classes (μ_f) et (μ_b) et les fonctions d'appartenance (φ_f) et (φ_b) sont calculées comme suit:

$$\mu_f(t) = \frac{\sum_{i=0}^t i \times p(i) \times \varphi_f^\lambda(i)}{\sum_{i=0}^t p(i) \times \varphi_f^\lambda(i)} \quad \mu_b(t) = \frac{\sum_{i=t+1}^{255} i \times p(i) \times \varphi_b^\lambda(i)}{\sum_{i=t+1}^{255} p(i) \times \varphi_b^\lambda(i)}$$

$$\varphi_f^\lambda(i) = \frac{1}{1 + \left(\frac{d(i, \mu_f)}{d(i, \mu_b)} \right)^{2/\lambda-1}}, \quad \varphi_b^\lambda(i) = 1 - \varphi_f^\lambda(i)$$

d dénote la distance euclidienne, et λ est un indice de floue.

Le seuil est le niveau de gris pour lequel les deux fonctions d'appartenance floues sont égales.

2.4.1.6. Méthodes de seuillage spatial

Les algorithmes de cette classe n'utilisent pas seulement la distribution des niveaux de gris mais aussi la dépendance des pixels dans le voisinage.

2.4.1.6.1. Méthode de Boudra et al.

Boudraa et al. ont exploité les aspects ponctuel et contextuel des pixels, pour estimer le seuil de binarisation [BOU 07]. Pour l'aspect contextuel, la méthode intègre l'information de contours et au lieu d'utiliser un détecteur de contours classique elle utilise l'Opérateur d'Énergie de Teager-Kaiser (OETK) qui inclut en plus certains aspects du système visuel humain et qui reflète mieux l'activité en énergie du pixel qu'un détecteur de contours classique [BOU 07].

L'intensité du contour pour chaque pixel est calculée en utilisant un filtre quadratique basé sur l'OETK par les deux formules suivantes :

$$\Psi_A(x, y) = 2 \times I^2(x, y) - I(x-1, y) \times I(x+1, y) - I(x, y-1) \times I(x, y+1)$$

$$\Psi_B(x, y) = 2 \times I^2(x, y) - I(x-1, y+1) \times I(x+1, y-1) - I(x-1, y-1) \times I(x+1, y+1)$$

$\Psi_A(x, y)$ s'appelle filtre non-linéaire de type 1A qui agit suivant les directions horizontale et verticale, et $\Psi_B(x, y)$ filtre de type 1B qui agit suivant les directions diagonales.

L'activité en (x, y) s'exprime par: $\xi(x, y) = \max(\Psi_A(x, y), \Psi_B(x, y))$

Le seuil de binarisation T_{opt} est estimé comme suit :

$$T_{opt} = \frac{\sum_{x,y} w(x, y) \times I(x, y)}{\sum_{x,y} w(x, y)} \quad \text{tel que} \quad w(x, y) = \begin{cases} \xi(x, y) & \text{si } \xi(x, y) > \beta\sigma \\ 0 & \text{sinon} \end{cases}$$

β est un paramètre qui contrôle l'intensité de l'activité du pixel ($\beta = 5$ selon les auteurs), et σ l'écart type du bruit.

2.4.1.6.2. Méthode de Xiao et al.

Xiao et al. ont proposé une méthode de seuillage entropique dérivée de celle de Kapur et al. et basée sur l'histogramme de corrélation spatiale des niveaux de gris (GLSC pour *Gray-level spatial correlation histogram*) [XIA 08]. Ce dernier est défini par les auteurs eux-mêmes et utilisé pour décrire les propriétés locales de l'image.

L'histogramme GLSC est calculé comme suit. Considérant $f(x, y)$ le nombre de pixels ayant un niveau de gris proche de $I(x, y)$ dans une fenêtre $n \times n$ centrée sur le pixel (x, y) .

$$f(x, y) = \sum_{i=-\frac{n-1}{2}}^{\frac{n-1}{2}} \sum_{j=-\frac{n-1}{2}}^{\frac{n-1}{2}} \Gamma(|I(x+i, y+j) - I(x, y)| \leq \varepsilon)$$

$$\text{Tel que : } \Gamma(a \leq \varepsilon) = \begin{cases} 1 & \text{si } a \leq \varepsilon \\ 0 & \text{sinon} \end{cases}$$

L'histogramme GLSC est ensuite construit: $GLSC(k, m) = Prob(I(x, y) = k \text{ et } f(x, y) = m)$

Tel que k est un niveau de gris (entre 0 et 255) et m appartient à $[1 \dots n^2]$.

Après, les entropies des classes *Objet* (H_f) et *Fond* (H_b) séparées par un niveau de gris t sont calculées par :

$$H_f(t, n) = - \sum_{k=0}^t \sum_{m=1}^{n^2} \frac{GLSC(k, m)}{P_f(t)} \log \left(\frac{GLSC(k, m)}{P_f(t)} \right) W(m, n)$$

$$H_b(t, n) = - \sum_{k=t+1}^{255} \sum_{m=1}^{n^2} \frac{GLSC(k, m)}{P_b(t)} \log \left(\frac{GLSC(k, m)}{P_b(t)} \right) W(m, n)$$

Avec :

$$P_f(t) = \sum_{k=0}^t \sum_{m=1}^{n^2} GLSC(k, m), \quad P_b(t) = \sum_{k=t+1}^{255} \sum_{m=1}^{n^2} GLSC(k, m), \quad W(m, n) = \frac{1 + e^{-\frac{9m}{n^2}}}{1 - e^{-\frac{9m}{n^2}}}$$

Le seuil optimal est choisi pour qu'il maximise l'entropie totale : $H = H_f + H_b$.

Selon les auteurs, une fenêtre de 3×3 pixels et $\varepsilon = 4$ donnent les meilleurs résultats.

2.4.2. Méthodes de seuillage local

Les méthodes locales sont dites aussi adaptatives car elles calculent un seuil différent $T_{opt}(x, y)$ pour chaque pixel (x, y) de l'image en se basant sur les informations de son voisinage. Ainsi, le seuillage s'effectue comme suit :

$$I_b(x, y) = \begin{cases} 255 & \text{si } I(x, y) > T_{opt}(x, y) \\ 0 & \text{si } I(x, y) \leq T_{opt}(x, y) \end{cases}$$

2.4.2.1. Méthodes basées sur la variance locale

Les méthodes de cette classe calculent un seuil local à chaque pixel en se basant sur la variation locale de ses pixels voisins dans une fenêtre de taille $w \times w$ centrée sur lui.

On note $\mu(x, y)$ et $\sigma(x, y)$ la moyenne et l'écart type respectivement des intensités des pixels dans la fenêtre.

2.4.2.1.1. Méthode de Niblack

L'algorithme de Niblack calcule un seuil local à chaque pixel (x, y) en se basant sur la moyenne locale $\mu(x, y)$ et l'écart-type locale $\sigma(x, y)$ de se voisinage dans une fenêtre $w \times w$ centrée lui [NIB 86]. Ainsi le seuil local $T_{opt}(x, y)$ est donné par :

$$T_{opt}(x, y) = \mu(x, y) + k \times \sigma(x, y)$$

Avec k est un paramètre utilisé pour déterminer le nombre de pixels de contours considérés comme des pixels d'objets et prend des valeurs négatives (k est fixé -0.2 par les auteurs).

Dans l'étude comparative de Trier et Taxt [TRI 95], la méthode de Niblack (avec une étape de post-traitement de Yanowitz and Bruckstein [YAN 86]) se montre la plus performante, avec $w=15$. He et al. [HE 05] ont défini un algorithme en cinq étapes permettant un choix adaptatif des paramètres k et w :

- 1) Binariser l'image par une méthode globale,
- 2) Diviser l'image binaire en petites blocs,

- 3) Calculer le rapport r de nombre de pixels noirs à ceux blancs pour chaque bloc, et attribuer une valeur appropriée de k selon la valeur de r . Une grande valeur de r nécessite une valeur élevée de k ,
- 4) Calculer la hauteur h de chaque bloc. w vaut $h \times h$,
- 5) Binariser chaque bloc par l'algorithme de Niblack et en utilisant les valeurs locales de k et w .

2.4.2.1.2. Méthode de Sauvola et Pietikainen

L'algorithme de Sauvola et Pietikainen [SAU 00] est considéré comme une modification de celui de Niblack, pour donner plus de performance dans les documents présentant trop de variations et illumination inégal. Dans la modification de Sauvola et Pietikainen le seuil local de binarisation est donné par :

$$T_{opt}(x, y) = \mu(x, y) \times \left(1 - k \times \left(1 - \frac{\sigma(x, y)}{R} \right) \right)$$

Où R est le rang dynamique de l'écart-type σ , et le paramètre k prend des valeurs positives dans l'intervalle [0.2, 0.5]. Les valeurs $R=128$ et $k=0.5$ sont utilisées.

Dans [KHU 09], la méthode de Sauvola et Pietikainen se montrait plus efficace que la celle de Niblack dans le cas où le niveau de gris du texte est proche de 0, et celui du fond proche de 255. Cependant dans les images où les niveaux de gris des pixels du fond et du texte sont proches, les résultats sont peu satisfaisants.

2.4.2.1.3. Méthode de Wolf et al.

Pour s'adresser aux problèmes de l'algorithme de Sauvola et Pietikainen (faible contraste, écart des niveaux de gris etc.), Wolf et al. [WOL 02] ont proposé de normaliser le contraste et la moyenne de niveaux de gris de l'image, et de calculer le seuil par :

$$T_{opt}(x, y) = (1 - k) \times \mu(x, y) + k \times G_{min} + k \times \frac{\sigma(x, y)}{R} \times (\mu(x, y) - G_{min})$$

Ici k est fixé à 0.5, G_{min} est le niveau de gris minimum dans l'image et R l'écart-type maximum des niveaux de gris obtenu sur toutes les fenêtres.

Cette méthode donne généralement des meilleurs résultats que ses prédécesseurs, mais elle souffre des dégradations dans les performances lorsqu'un changement considérable dans les niveaux de gris du fond est existé [WOL 02].

2.4.2.1.4. Méthode de Zhang et Tan

Zhang et Tan [ZHA 01] ont proposé une modification de l'algorithme de Niblack dans le but d'enlever l'ombre qui peut se produit dans les pages de livres scannées. Dans cette modification, l'écart type local est normalisé en le divisant sur le rang dynamique R . De plus, la moyenne est utilisée pour multiplier (au lieu d'additionner) les termes de l'écart type. Ces modifications permettent d'augmenter la contribution de l'écart type, et de réduire la sensibilité du paramètre k . Ainsi le seuil local est donné par :

$$T_{opt}(x, y) = \mu(x, y) \times \left(1 + k \times \left(1 - \frac{\sigma(x, y)}{R} \right) \right)$$

Les auteurs utilisent les valeurs 100 pour R et 0.1 pour k .

2.4.2.1.5. Méthode de Bin Rais et al.

Cette technique est basée aussi sur la méthode de Niblack. Elle ne considère pas seulement les caractéristiques statistiques locales mais utilise également des statistiques globales.

Contrairement à Niblack qui fixe la valeur du paramètre k à -0.2, Bin Rais et al. calculent une valeur différente de k pour chaque pixel [BIN 04]. La valeur de k pour le pixel (x,y) est donnée par :

$$k(x, y) = -0.3 \times \frac{(\mu_g \times \sigma_g - \mu(x, y) \times \sigma(x, y))}{\max(\mu_g \times \sigma_g, \mu(x, y) \times \sigma(x, y))}$$

Tel que : μ_g et σ_g dénotent la moyenne et l'écart type globales de l'image entière respectivement. La moyenne $\mu(x,y)$ et l'écart type $\sigma(x,y)$ locales sont calculées sur une fenêtre de taille 75×75 .

Le seuillage s'effectue finalement par : $T_{opt}(x,y) = \mu(x,y) + k(x,y) \times \sigma(x,y)$.

2.4.2.1.6. Méthode de Feng et Tan

Cette méthode a été proposée dans le but d'extraire le texte du fond dans les images de documents de mauvaise qualité [FEN 04]. Pour déterminer le seuil local, la méthode de Feng et al. calcule la moyenne $\mu(x,y)$, l'écart type $\sigma(x,y)$ et l'intensité minimale $G_{min}(x,y)$ dans une première fenêtre de taille $w \times w$. Après, et en vue d'équilibrer l'effet d'illumination, le dynamique R de l'écart type est calculé dans une deuxième fenêtre plus large (de taille $w_2 \times w_2$) au lieu de l'image entière. Le seuil local est calculé ainsi par :

$$T_{opt}(x, y) = (1 - \alpha_1) \times \mu(x, y) + \alpha_2 \times \left(\frac{\sigma(x, y)}{R} \right) \times (\mu(x, y) - G_{min}(x, y)) + \alpha_3 \times G_{min}(x, y)$$

$$\text{Avec : } \alpha_2 = k_1 \times \left(\frac{\sigma}{R} \right)^\varphi \text{ et } \alpha_3 = k_2 \times \left(\frac{\sigma}{R} \right)^\varphi$$

α_1, k_1, k_2 et φ sont des constantes positives égales à 0.12, 0.25, 0.04 et 2 respectivement.

2.4.2.1.7. Méthode Nick

Cette méthode améliore considérablement la binarisation d'images de page claires et présentant de faible contraste, en déplaçant vers le bas, le seuil de binarisation [KHU 09]. Notons $d=w/2$, Le calcul du seuil est réalisé comme suit :

$$T_{opt}(x, y) = \mu(x, y) + k \sqrt{\frac{\left(\sum_{i=-d}^d \sum_{j=-d}^d I(x+i, y+j)^2 - (\mu(x, y))^2 \right)}{NP}}$$

Tel que : k est le facteur de Niblack et varie entre -0.1 à -0.2 selon les besoins de l'application, NP est le nombre total de pixels dans l'image.

Les auteurs ont déclaré que cette méthode marche très bien pour plusieurs (si pas tous) les types de documents anciens dégradés. Dans les tests réalisés, la taille de la fenêtre utilisée est $w=19$.

2.4.2.1.8. Méthode de Singh et al.

Singh et al. [SIN 11] ont proposé une nouvelle façon de calculer les seuils, basée sur les images de sommes intégrales (Integral sum images) ce qui accélère considérablement la binarisation et rend le temps de calcul fixe et indépendant de la taille de fenêtre glissante. La première étape de cette technique est le calcul de l'image de somme integrale g définie comme suit :

$$g(x, y) = \sum_{i=0}^x \sum_{j=0}^y I(i, j)$$

Le seuil local de binarisation $T(x,y)$ est déterminé par la suite en utilisant la moyenne locale $\mu(x,y)$ et l'écart moyenne $\delta(x,y)$ dans une fenêtre $w \times w$ centrée sur le pixel (x,y) .

$$T(x, y) = \mu(x, y) \times \left(1 + k \times \left(\frac{\delta(x, y)}{1 - \delta(x, y)} - 1 \right) \right)$$

Tel que : $\delta(x,y) = I(x,y) - \mu(x,y)$ et k un parametre dans $[0, 1]$ et suggéré 0.06 par les auteurs.

Notons $d=w/2$, la moyenne locale $\mu(x,y)$ est calculée en utilisant l'image de sommes integrales par:

$$\mu(x, y) = \frac{g(x+d-1, y+d-1) + g(x-d, y-d) - [g(x-d, y+d-1) + g(x+d-1, y-d)]}{w^2}$$

2.4.2.1.9. Méthode de Peng et al.

Dans le but d'extraire le text des anciens manuscrits, Peng et al. ont proposé une nouvelle méthode de binarisation adaptative [PEN 12]. Partant de la supposition que l'intensité de l'arrière-plan est consistante à l'interieur des zones locales de l'image de document, la méthode proposée calcule un seuil local $T_{opt}(x,y)$ pour chaque pixel en utilisant la moyenne locale $\mu(x,y)$ et la variance locale $\delta(x,y)$, calculées sur une fenêtre $m \times n$, et la variance minimale et maximale dans l'image entière δ_{min} et δ_{max} .

$$T_{opt}(x, y) = \mu(x, y) \left[\frac{1 - k}{\left(1 + e^{-B \left(\frac{\delta(x, y) - \delta_{min} - M}{\delta_{max} - \delta_{min}} \right)^{1/\nu}} \right)} + k \right]$$

Les paramètres k , B , M et ν prennent les valeurs 0.97, 25, 0.005 et 20 respectivement.

2.4.2.1.10. Méthode de Laskov

Cette méthode combine la simplicité de [GON 07] et la robustesse de [SAU 00] pour séparer avec le maximum d'exactitude possible le texte du fond dans les images de documents de faible contraste [LAS 12]. L'algorithme consiste en deux balayages ; horizontal et vertical ce qui permet d'enlever le bruit causé par le balayage unidimensionnel produit dans [GON 07]. Chacun des deux balayages produit une image binaire en calculant pour chaque pixel (x,y) un seuil local $T(x,y)$ par la formule de Sauvola [SAU 00]. La moyenne $\mu(x,y)$ et l'écart type $\sigma(x,y)$ locales sont calculées pour les pixels dans le rang du filtre. Notons B_h et B_v les images binaires résultantes des deux balayages.

Finalement les résultats des balayages horizontales et verticales sont combinés en utilisant une opération de disjonction logique pour produire l'image binaire finale I_b .

$$I_b(x, y) = B_h(x, y) \vee B_v(x, y).$$

2.4.2.2. Méthodes basées sur l'intervalle de gris local

Dans ces techniques, l'écart entre le niveau de gris maximum et minimum dans une fenêtre locale est utilisé pour la détermination des valeurs des seuils.

2.4.2.2.1. Méthode de Yasuda et al.

Cette méthode commence par étendre le rang dynamique de l'image suivi par un lissage non linéaire ce qui permet de préserver les contours fins [YAS 80]. Le lissage consiste à remplacer chaque pixel par la moyenne de ses huit voisins, dans le cas où le contraste locale (définie comme la l'écart entre les valeurs maximale et minimale locales) est inférieure à un seuil T_1 . Un premier seuillage adaptatif est ensuite appliqué. Un pixel est attribué à la classe *Fond* si son contraste locale est inférieure à un seuil T_2 ou si son niveau de gris est supérieur à la moyenne locale, calculés les deux sur une fenêtre $w \times w$. Sinon, la plage dynamique est élargie en conséquence. Finalement, l'image résultante est binarisée en attribuant à la classe *objet* tout pixel dont le minimum local sur une fenêtre 3×3 est inférieur à un seuil T_3 ou si la variance locale est supérieure à T_4 . Dans [SEZ 04], $T_1=50$, $T_2=16$, $T_3=128$, $T_4=35$.

2.4.2.2.2. Algorithme de seuillage dynamique (White et Rohrer)

Cette méthode a été proposée par White et Rohrer [WHI 83]. Elle consiste à comparer le niveau de gris de chaque pixel avec la moyenne des niveaux de gris de ses voisins dans une fenêtre $w \times w$ centrée sur lui ($w=15$ est suggéré dans [SEZ 04]). Ainsi la nouvelle valeur de chaque pixel (x, y) est attribuée comme suit :

$$I_b(x, y) = \begin{cases} 255 & \text{si } \mu(x, y) < I(x, y) \times \beta \\ 0 & \text{sinon} \end{cases}$$

Tel que : $\mu(x, y)$ est la moyenne locale dans une fenêtre $w \times w$, et β est un paramètre utilisé pour éliminer le bruit du fond et suggéré 2 dans [SEZ 04].

2.4.2.2.3. Méthode de Bernsen

Dans cette méthode, le seuil local est calculé comme la moyenne des intensités minimale et maximale dans la fenêtre [BER 86]. Ainsi pour chaque pixel de coordonnées (x, y) , le seuil local est donné par :

$$T_{opt}(x, y) = \frac{G_{min}(x, y) + G_{max}(x, y)}{2}$$

Tel que $G_{min}(x, y)$ et $G_{max}(x, y)$ sont le niveau de gris le plus bas et le plus haut respectivement, dans une fenêtre carrée $w \times w$ centrée sur le pixel (x, y) .

Cependant si le contraste local $C(x, y) = G_{max}(x, y) - G_{min}(x, y)$ est inférieur à un certain seuil l ($=15$), alors le voisinage consiste en une seule classe. Dans [TRI 95], le de $w=15$ donne les meilleurs résultats, tandis que dans [SEZ 04], la taille suggérée est $w=31$.

2.4.2.2.4. Seuillage en utilisant les moyennes mobiles (Gonzalez et Woods)

Cette méthode est proposée par Gonzalez et Woods dans [GON 07], chapitre 10, section « Using moving averages ». La méthode est basée sur un filtre moyen unidimensionnel (de taille $n \times 1$) qui se glisse sur les lignes de l'image dans une forme de zigzag (pour une ligne, de gauche à droite et pour la ligne suivante de droite à gauche et ainsi de suite).

Pour estimer un seuil local à un pixel (x, y) , la moyenne des intensités des pixels dans le rang du filtre ($\mu_n(x, y)$) est d'abord calculée. Ensuite, le seuil local $T_{opt}(x, y)$ est calculé par :

$$T_{opt}(x,y) = k \times \mu_n(x,y).$$

Avec k est un paramètre qui prend des valeurs dans l'intervalle $[0, 1]$

2.4.2.3. Techniques utilisant les informations du voisinage

Les méthodes de cette classe utilisent les informations de voisinage (contour, largeur de trait, etc.) pour la détermination des seuils locaux.

2.4.2.3.1. Méthode de Giuliano et al.

Cette technique a été proposée par Giuliano et al. dans le but d'extraire le texte du fond dans leur système de reconnaissance de caractères [GIU 77]. Dans cette technique la valeur binaire de chaque pixel est calculée en tenant compte des ses pixels voisins dans une fenêtre 9×9 découpée en 2 zones A_1 et A_2 . La zone A_1 contient les pixels dans un voisinage de 3×3 autour du pixel en question, tandis que la zone A_2 contient les quatre 3×3 voisinages adjacentes diagonalement à A_1 . Le seuillage est effectué ainsi comme suit :

Si $(I(x,y) < T_1)$, alors $I_b(x,y) := 0$

Sinon, $A_{2t} := \{(x,y) \mid (x,y) \in A_2 \text{ et } I(x,y) > T_2\}$

$a_1 :=$ moyenne des pixels dans A_1

$a_2 :=$ moyenne des pixels dans A_{2t}

Si $(T_3 \times a_2 + T_5 > T_4 \times a_1)$, alors $I_b(x,y) := 0$

Sinon $I_b(x,y) := 255$

Tel que : T_1, T_2, T_3, T_4 et T_5 sont des paramètres prédéterminés. Dans [PAL 86], les valeurs suivantes sont suggérées : $T_1=20, T_2=20, T_3=0.85, T_4=1$, et $T_5=0$.

2.4.2.3.2. Technique du niveau logique

Cette technique a été proposée par Kamel et Zhao [KAM 93]. Pour déterminer la nouvelle valeur de chaque pixel (x,y) , son niveau de gris est comparé avec les niveaux de gris moyens dans quatre fenêtres de taille $w \times w$, situées b pixels loin de (x,y) dans les quatre directions, avec $w=2b+1$ et b dénote la largeur du trait. La nouvelle valeur du pixel est attribuée comme suit :

$$I_b(x,y) = \begin{cases} 0 & \text{si } [L(x+w,y) \wedge L(x-w,y)] \vee [L(x,y+w) \wedge L(x,y-w)] \vee \\ & [L(x+w,y+w) \wedge L(x-w,y-w)] \vee [L(x+w,y-w) \wedge L(x-w,y+w)] \text{ est vrai} \\ 255 & \text{sinon} \end{cases}$$

Tel que :

$$L(i,j) = \begin{cases} \text{vrai} & \text{si } (\mu(i,j) - I(i,j)) \geq T_0 \\ \text{faux} & \text{sinon} \end{cases}$$

Avec $\mu(i,j)$ le niveau de gris moyen dans une fenêtre $w \times w$ centrée sur le pixel (i,j) . Dans [SEZ 04], les valeurs $b=8$ et $T_0=40$ sont les plus appropriées.

2.4.2.3.3. Algorithme de fonction intégrée (White et Rohrer)

Cette technique a été décrite par White et Rohrer dans [WHI 83]. Elle commence par l'application d'un opérateur de type Gradient appelé « activité dans l'image » afin de localiser les contours. L'activité en un pixel (x,y) notée $A(x,y)$ est la somme absolue des

dérivés rapprochées des deux directions horizontale et verticale prises sur une petite région (3×3).

$$A(x, y) = \sum_{i=-1}^1 \sum_{j=-1}^1 |d_x(x+i, y+j)| + |d_y(x+i, y+j)|$$

Avec d_x et d_y dénotent les dérivés horizontales et verticales respectivement et elles sont définies par Sobel comme suit :

$$d_x(x,y) = I(x-1, y) - I(x+1, y) \quad \text{et} \quad d_y(x,y) = I(x, y-1) - I(x, y+1).$$

Les pixels ayant une activité inférieure à un seuil prédéterminé T_A , sont marqués par un 0. Pour les autres pixels, l'opérateur Laplacien $dd_{xy}(x,y)$ est calculé et testé .

$$dd_{xy}(x,y) = I(x+2,y) + I(x-2,y) + I(x,y+2) + I(x,y-2) - 4 \cdot I(x,y).$$

Si le laplacien est positif, le pixel est étiqueté par '+' sinon par '-'.

Dans une séquence d'étiquettes, les contours sont identifiés par une transition '- +' ou '+ -'. Les pixels d'objets sont considérés comme étant les pixels étiquetés '0' et '+' contenus entre une paire '- +' et '+ -'. La distance entre cette paire représente la largeur du trait w . Les pixels du fond tendent à ne pas être inclus dans cette paire.

2.4.2.3.4. Méthode de Trier et Taxt

Trier et Taxt [TRI 95c], ont appliqué des modifications à l'*Algorithme de fonction intégrée* de White et Rohrer [WHI 83] afin d'améliorer la qualité de binarisation pour les images de document. Une de leurs modifications est la façon de classifier des pixels '0' dans l'image d'étiquettes à trois niveaux '+', '-', '0' qui est le résultat de l'opérateur gradient. Le pixel étiqueté '0' sera classé dans la classe à la quelle la majorité de ses 8 pixels voisins appartiennent. Un post-traitement est effectué pour supprimer les faux objets qui ont pu être détectés. La norme du gradient de l'image lissée est calculée. Les composantes 4-connexes dont les points de contours ont une valeur moyenne inférieure à un seuil T_P sont supprimées. Cette méthode donne de bons résultats sur des documents de bonne qualité, mais elle a l'inconvénient d'être régie par plusieurs paramètres qui restent difficiles à régler en pratique.

2.4.2.3.5. Technique de soustraction basée-masque (Kamel et Zhao)

Cette technique a été proposée dans [KAM 93] et elle consiste en quatre étapes. Dans la première étape, la plupart des pixels de l'arrière-plan sont détectés en utilisant un filtre logique de taille $w \times w$. Pour chaque pixel traité (x,y) , on note L_1, L_2, L_3 et L_4 les quatres lignes droites passant par le pixel (x,y) d'angles $0, \pi/4, \pi/2, 3\pi/4$ respectivement. Le pixel (x,y) sera coloré en noir, si au moins w pixels situés sur l'une des lignes L_1, L_2, L_3, L_4 et qui se trouvent à l'intérieur de la fenêtre $w \times w$, ont un niveau de gris supérieur à celui du pixel (x,y) , sinon il sera blanc. On note I_m l'image résultante. Dans la deuxième étape, l'image I_m est modifiée en détectant plus de pixels du *fond*.

Un pixel noir (x,y) devient blanc si : $\min(d_u+d_d, d_r+d_l) < w_{min}$ ou si $\min(d_u+d_d, d_r+d_l) > w$. Tels que : d_l, d_r, d_u et d_d dénotent la distance entre le pixel (x,y) et le premier pixel blanc rencontré sur chacune des quatre directions (gauche, droit, haut, bas) respectivement, notés P_l, P_r, P_u et P_d . w_{min} est la largeur minimale de trait. La troisième étape crée une nouvelle image de fond estimée (I_{m2}). Si le pixel (x,y) est noir dans I_m , son nouveau niveau de gris dans I_{m2} est estimé par interpolation linéaire des pixels P_l, P_r, P_u et P_d . Sinon $I_{m2}(x,y)=I(x,y)$.

Finalement, l'image binaire est obtenue par :

$$I_b(x, y) = \begin{cases} 0 & \text{si } I_{m2}(x, y) - I(x, y) > T \\ 255 & \text{sinon} \end{cases} \quad \text{avec } T \text{ un paramètre prédéterminé.}$$

2.4.2.3.6. Technique du moyen-gradient (Leedham et al.)

C'est une autre amélioration de la méthode de Niblack et elle est proposée par Leedham et al. [LEE 03]. Dans cette méthode, un seuil local est calculé pour chaque pixel de l'image en se basant sur la moyenne des niveaux de gris et la moyenne des valeurs du gradient dans une fenêtre $w \times w$ centrée sur le pixel. Le gradient de l'intensité du pixel $I(x, y)$ est défini par:

$$\nabla I_{(x,y)} = \left[\frac{\partial I(x, y)}{\partial x}, \frac{\partial I(x, y)}{\partial y} \right]$$

Le seuil local $T_{opt}(x, y)$ est calculé par la relation suivante :

$$\begin{aligned} \text{Si } G_{max}(x, y) - G_{min}(x, y) \geq R \text{ alors } T_{opt}(x, y) &= \mu(x, y) + k \times \mu_G(x, y) \\ \text{Sinon } T_{opt}(x, y) &= 0.5 \times \mu(x, y) \end{aligned}$$

Tel que : G_{max} et G_{min} dénotent le niveau de gris le plus haut et le plus bas dans la fenêtre, R : constante égal à 40, k : le facteur de Niblack = -1.5, $\mu(x, y)$ et $\mu_G(x, y)$: la moyenne des niveaux de gris et des gradients respectivement dans la fenêtre.

2.4.2.3.7. Méthode de Gatos et al.

Gatos et al. [GAT 06] ont proposé une méthode de seuillage adaptative pour la binarisation d'images de documents dégradés. La méthode se compose de plusieurs étapes. Un prétraitement en utilisant un filtre de Wiener est d'abord appliqué. Ensuite, une estimation initiale des régions du premier plan est effectuée, en procédant un premier seuillage par la méthode de Sauvola et Pietikainen [SAU 00]. Notons S l'image résultante. La troisième étape est le calcul de la surface du fond (B) à partir de l'image S . Ainsi, B est estimé comme suit :

$$B(x, y) = \begin{cases} I(x, y) & \text{si } S(x, y) = 255 \\ \text{résultat d'interpolation de voisinage de } (x, y) & \text{sinon} \end{cases}$$

Après, un seuillage final est procédé et l'image binaire I_b est obtenue par :

$$I_b(x, y) = \begin{cases} 0 & \text{si } B(x, y) - I(x, y) > d(x, y) \\ 255 & \text{sinon} \end{cases}$$

$d(x, y)$ est seuil local au pixel (x, y) .

Finalement une étape de post-traitement est appliquée en vue d'améliorer la qualité des régions textuelles et de préserver la connectivité du trait.

2.4.2.4. Seuillage Hiérarchique

2.4.2.4.1. Segmentation hiérarchique floue

Cette méthode est fondée sur un principe hiérarchique pour pallier le problème d'éclairage non uniforme que présente la plupart des images réelles [GAD 00]. Elle se compose de deux étapes. La première étape consiste à construire l'arbre quaternaire. Cette dernière est obtenue en décomposant l'image consécutivement en quatre sous images en évaluant le critère d'homogénéité. Chaque sous-image est associée à un nœud de l'arbre. Si une sous-image satisfait le critère d'homogénéité, la division n'est plus nécessaire, elle devient un nœud terminal dans l'arbre. Au cas contraire, cette sous-image est décomposée en quatre. Le processus continue jusqu'à ce que tous les nœuds dans l'arbre soient des terminaux. Dans la

deuxième étape dite *Décision et défuzzification*, elle calcule pour chaque niveau hiérarchique, le degré d'appartenance de chaque pixel (x, y) à la classe *Objet* et à la classe *Fond* en utilisant la fonction S de Zadeh. Le degré d'appartenance est donnée par :

$$\varphi(x, y) = S(I(x,y); \mu_l - \sigma_l, \mu_l, \mu_l + \sigma_l)$$

Tel que μ_l et σ_l représentent respectivement la moyenne et l'écart-type des niveaux de gris dans la sous-image donnée.

Ensuite, l'opérateur *t-conorme* de Zadeh est utilisé pour calculer le degré d'appartenance final de chaque pixel (x, y) à chacune des deux classes. Un pixel est affecté à la classe *Objet* si son degré d'appartenance agrégé à la classe *Objet* est supérieur à celle agrégé à la classe *Fond*. Sinon il est affecté à la classe *fond*.

2.4.3. Méthodes de seuillage hybride

Le seuillage hybride se définit par l'action conjuguée d'une analyse locale dans le voisinage d'un point et d'une analyse globale dans une région plus étendue autour de ce point (voir de l'image toute entière) [GAR 06].

2.4.3.1. Méthodes basées sur la morphologie

2.4.3.1.1. Seuillage par soustraction du fond

Cette technique a été proposée dans [LEE 03] et elle consiste en trois étapes. D'abord, l'arrière-plan est modelé en enlevant l'écriture par l'application d'une fermeture à l'image originale avec un élément structurant sous forme de disque. Après, l'arrière-plan est soustrait de l'image originale ce qui laisse uniquement l'avant-plan. Finalement, l'image résultante est segmentée en utilisant l'algorithme d'Otsu multiplié par une constante empirique.

2.4.3.1.2. Méthode de Tabatabaie et Bohlool

Tabatabaie et Bohlool [TAB 10] ont conçu une méthode de binarisation d'images de documents mal-illuminés basée sur le même principe. L'avant-plan est obtenu en soustrait le fond, estimé en utilisant l'opération de fermeture morphologique, de l'image originale. Ainsi, en considérant que la fermeture peut produire une estimation raisonnable de l'arrière-plan si on utilise l'élément structurant adéquat, les auteurs ont proposé une procédure de détermination de la taille convenable de l'élément structurant. Un seuillage global est d'abord appliqué sur l'image originale. Ensuite, on cherche la taille du plus grand carré noir pour chaque pixel et on stocke ces valeurs dans une matrice M . La plus grande valeur de M dans chaque ensemble connexe de pixels est calculée et attribuée aux autres éléments de l'ensemble. Après, on construit le *S-histogramme* (f) partant de la matrice M . La valeur de f au point i est égale au nombre d'éléments de M ayant la valeur i . Finalement, on détermine x_{max} , la plus grande valeur de x , avec x satisfait :

$$\sum_{i=1}^x f(i) > 0.02 \times \sum_{i=1}^{\infty} f(i) \text{ et } \sum_{i=x}^{2x} f(i) = 0$$

La taille de l'élément structurant sera $2 x_{max}$.

2.4.3.1.3. Méthode de Cavalcanti et al.

Cette méthode a été proposée pour le traitement des images de documents à fond complexe [CAV 06]. Elle commence par calculer deux seuils T_1 et T_2 . T_1 est choisi comme le premier niveau de gris dans l'histogramme ayant une fréquence supérieure à $(hauteur \times largeur)/n$. n est fixé 350 par expérimentations. $T_2 = T_1 - 40$. Ensuite, elle transforme tous les pixels de l'image

ayant un niveau de gris supérieur à T_1 en blanc. Mettant $ImageT_1$, l'image résultante. Les pixels dans $ImageT_1$ ayant un niveau de gris supérieur à T_2 sont colorés en blanc, générant une image $ImageT_2$. Dans l'image $ImageT_2$ résultante, certains pixels d'objet sont éliminés, produisant une discontinuité dans l'écriture. Pour préserver la continuité de trait, certains pixels sont récupérés de l'image $ImageT_1$. La récupération est implémentée par la dilatation de l'image $ImageT_2$ suivie par une opération AND binaire entre l'image dilatée et l'image $ImageT_1$.

2.4.3.1.4. Méthode de Gangamma et Srikanta Murthy

Gangamma et Srikanta Murthy [GAN 11] ont proposé une méthode basée sur une combinaison simple et efficace des filtres spacieux et opérateurs morphologiques pour éliminer l'arrière-plan et améliorer la qualité des images de documents indiens dégradés. La première étape de cette technique consiste à procéder une égalisation adaptative de l'histogramme (AHE) pour surmonter le problème de l'illumination inégale dans l'image de document. Notons E l'image résultante. Dans la deuxième étape, l'opération d'ouverture morphologique est appliquée sur E et l'image résultante est additionnée avec E résultant une image $E2$. Après, l'opération de fermeture morphologique est appliquée sur $E2$ en vue de la lisser. Notons $E3$ le résultat du lissage. L'image E est ensuite soustraite de $E3$ et le résultat est soustrait encore de l'image $E2$. Un filtre gaussien est appliqué par la suite afin d'éliminer le bruit. L'image améliorée finale est reconstruite en additionnant la dernière image avec l'image E . Finalement, un seuillage globale (l'algorithme d'Otsu) est nécessaire pour séparer le texte de l'arrière-plan.

2.4.3.2. Méthodes de binarisation à deux niveaux

2.4.3.2.1. Méthode de Wu et Amin

Sue Wu et Adnan Amin [WU 03] ont proposé une méthode hybride de seuillage des images d'enveloppes postales qui s'applique en deux étapes. La première étape applique un seuillage global en utilisant l'algorithme d'Otsu sur l'image originale. La deuxième étape ajuste la valeur de seuil en fonction des caractéristiques spatiales des composantes connexes formées par le seuillage global. Cependant, dans la deuxième étape, seules les composantes connexes présentant du faible contraste (dans l'image originale) sont binarisées localement. Chacune de ces composantes connexes est lissée en utilisant un filtre moyen 5×5 et son histogramme est ensuite calculé. Le seuil local de chaque composante est le premier minimum de l'histogramme. La méthode donne de bons résultats sur l'ensemble des images d'enveloppes simples, bien contrastées et de bonne qualité [WU 03].

2.4.3.2.2. Méthode IGT améliorée

Cette méthode et une amélioration de la technique de seuillage global itératif [KAV 05] et elle consiste en deux passes [KAV 06]. Dans le premier, un seuillage global est appliqué sur l'image entière, et dans le deuxième un seuillage local aux zones contenant encore du bruit. Pour ce faire l'image binaire résultante du seuillage global est divisée en plusieurs segments de taille $n \times n$, et pour chaque segment la fréquence f de pixels noirs est calculé. Les segments vérifiant le critère suivant sont gardés : $f > \mu + k \times \sigma$, tel que μ et σ dénotent la moyenne et l'écart type des fréquences des pixels noirs dans le segment, et k une constante (égal à 2 selon les auteurs). Pour chaque zone détectée, la méthode IGT est appliquée sur la zone correspondante dans l'image originale. Des zones de taille 50×50 donnent des bons résultats d'après les auteurs.

2.4.3.2.3. Méthode de Tabonne et al.

C'est une approche hiérarchique proposée par Tabonne et al. [TAB 06] pour la binarisation d'images de documents graphiques, basée sur un découpage en arbre quaternaire, dont l'intérêt est de traiter l'image à différents niveaux. La méthode proposée, qui peut être considérée comme la combinaison d'une approche globale et d'une approche locale, comprend deux étapes. La première est un seuillage global en utilisant la technique d'Otsu [OTS 79], qui élimine la plupart du fond pour garder la partie significative contenant l'objet. Ainsi, dans cette étape, les pixels du fond sont mis à *blanc* et ceux désignés comme faisant partie de l'objet conservent leur valeur initiale de niveau de gris. La seconde étape est une amélioration de la méthode de segmentation hiérarchique floue de Gadi et Benslimane [GAD 00], qui affine l'image résultante de l'étape précédente pour la rendre « plus nette ».

2.4.3.2.4. Méthode de Savakis

Savakis dans [SAV 98] a proposé un algorithme de seuillage hybride adapté à la numérisation de documents à haute vitesse. L'algorithme proposé est basé sur un seuillage adaptative et utilise les informations locales du contour pour décider l'application d'un seuillage global ou local. Il applique un seuillage global fixe sur les pixels à faible gradient qui appartiennent aux zones homogènes de fond. Les pixels possédant un gradient local répondant à une forte transition au niveau de tracé sont binarisés localement. Le seuil local à chaque pixel égale à la moyenne des intensités minimale et maximale dans une fenêtre centrée sur le pixel.

2.4.3.2.5. Méthode de Kamada et Fujimoto

Kamada et Fujimoto [KAM 99] ont proposé une méthode de binarisation séquentielle destinée aux images à faible résolution. La méthode s'effectue en deux étapes. D'abord un seuillage global est appliqué sur l'image d'entrée permettant de séparer l'avant-plan de l'arrière-plan et d'extraire ainsi les voisinages des caractères. Ensuite, les caractères sont extraits par un processus de seuillage local et une interpolation linéaire.

2.4.3.2.6. Binarisation en utilisant le maximum et le minimum locaux

Su et al. exploitent le contraste déterminé par le maximum et le minimum locaux de l'image, pour la binarisation d'images de documents historiques [SU 10]. La méthode proposée commence par la construction de l'image de contraste M_C en évaluant l'intensité minimale et maximale locales.

$$M_C(x, y) = \frac{G_{\max}(x, y) - G_{\min}(x, y)}{G_{\max}(x, y) + G_{\min}(x, y) + \varepsilon}$$

Tel que : $G_{\max}(x, y)$ et $G_{\min}(x, y)$ expriment l'intensité maximale et minimale respectivement dans une fenêtre 3×3 centrée sur le pixel (x, y) . ε est un entier positif très petit.

Ensuite, les pixels de haut contraste autour du contour du texte sont détectés, en binarisant l'image M_C par l'algorithme d'Otsu. Notons E l'image résultante. Finalement, l'image de document est binarisée localement comme suit :

$$I_b(x, y) = \begin{cases} 0 & N_e(x, y) \geq N_{\min} \text{ et } I(x, y) \leq \mu_{(x, y)} + \sigma_{(x, y)} / 2 \\ 255 & \text{sinon} \end{cases}$$

Avec : $N_e(x, y)$ le nombre de pixels de haut contraste, $\mu(x, y)$ la moyenne des intensités, $\sigma(x, y)$ l'écart type, dans une fenêtre $w \times w$ centrée sur le pixel (x, y) . N_{\min} est le nombre minimum de pixels de haut contraste dans la fenêtre. La moyenne et l'écart type sont calculés comme suit :

$$\mu_{(x,y)} = \frac{\sum_{w \times w} I(i,j) \times (1 - E(i,j))}{N_e}, \quad \sigma_{(x,y)} = \frac{\sum_{w \times w} ((I(i,j) - \mu_{w \times w}) \times (1 - E(i,j)))^2}{N_e}$$

2.5. Binarisation basée sur le regroupement

2.5.1. Méthodes de regroupement basées sur la classification

Dans ces méthodes, les pixels de l'image sont attribués à l'une des deux classes *Objet* ou *Fond* en fonction d'un critère d'homogénéité. Cependant plusieurs techniques ont été employées pour atteindre cette classification : l'algorithme *Kmeans* (*nuées dynamiques*), le Réseau de neurones artificiel (ANN pour *Artificial Neural Network*), la machine à support vectoriel (SVM pour *Support Vector Machine*), etc.

Nous présentons dans cette section quelques techniques de binarisation appartenant à cette classe.

2.5.1.1. Binarisation en utilisant les ANNs

Les ANNs sont des techniques de classification bien-connues. Récemment, les ANNs ont été utilisés pour la binarisation d'images.

Chen et Takagi [CHE 93] ont proposé de déterminer un seuil global optimal en utilisant un perceptron multicouche (MLP pour *multi-layer perceptron*) entraîné par un algorithme de rétro-propagation modifié. La technique commence par le calcul de l'histogramme de l'image contrastée. L'histogramme est ensuite mis à l'échelle dans [0, 1] et divisé en 128 parties. Le MLP utilisé a une couche d'entrée de 128 neurones recevant les 128 parties de l'histogramme normalisé, 2 couches cachées de 256 et 512 neurones respectivement, et une couche de sortie ayant 128 neurones.

N. Papamarkos et al. dans [PAP 00] 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 [PAP 03], 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 C-means*) 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.

[HAM 05] 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 sont regroupés, la SOFM est utilisé pour classifier les pixels de l'image entière. Chaque pixel est donné comme entrée à la SOFM en vue d'être classifié. La deuxième approche utilise un MLP 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 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 [KHA 07b], 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 du seuil global la plus optimale. 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 locales, et sur l'image entière résultant un seul seuil global. Après, le ANN supervisé est entraîné en utilisant les valeurs des seuils locaux comme entrées et la valeur du seuil global comme sortie. Le ANN entraîné est utilisé par la suite pour trouver le seuil global le plus optimal pour d'autres images.

Saidane et Garcia. [SAI 07] ont utilisé un réseau de neurones de convolution [GAR 04] 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.

Padekas et Papamarkos [BAD 07] 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 en utilisant une carte de Kohonen (SOFM). Premièrement, les valeurs les plus optimales de paramètres des 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 estimées. Après l'apprentissage, la SOFM est utilisée pour classifier chaque pixel de l'image en *Objet* ou *Fond*.

Làzaro et al. [LAZ 10] 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 et 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 du ANN et les valeurs de seuils optimales correspondantes représentent les sorties désirés.

2.5.1.2. Binarisation en utilisant l'algorithme K-means

L'algorithme *K-means* a été utilisé principalement pour l'extraction de l'avant-plan dans les images de documents couleurs.

Bottou et al. [BOT 98] 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 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. [LEY 04] ont proposée une méthode de séparation avant-plan/ arrière-plan basée sur une sérialisation de l'algorithme *k-means* appliqué séquentiellement sur l'image dans une fenêtre glissante. Cependant, l'approche implique une intervention manuelle pour décider le 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.

[GAR 06] décrit 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. Après, les composants du fond dominant sont déterminés 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, 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 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 du traitement de son parent. La combinaison des résultats du traitement de tous les blocs fournit l'image binarisée finale.

2.5.1.3. Binarisation en utilisant les SVMs

Kita et Wakahara [KIT 10] ont combiné l'algorithme *k-means* avec un 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 en 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, et le SVM estime le degré d'« apparence de caractère ». Dans la dernière étape, on sélectionne l'image binarisée ayant le maximum degré d'« apparence de caractère » comme le résultat de binarisation optimal. Dans [WAK 11], 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 binarisées provisoire obtenue dans la première étape est divisée en une séquence d'images d'un caractère chacune.

Une autre méthode basée sur les SVMs pour la binarisation d'images de documents a été décrite dans [CHO 10]. 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.

2.5.1.4. Binarisation en utilisant les arbres de décision

Une méthode de binarisation d'images de documents couleurs basée sur les arbres de décision a été décrite par Tsai et Lee dans [TSA 02]. La méthode proposée sélectionne diverses caractéristiques de couleur pour binariser les images documents couleurs. Ainsi, on utilise la caractéristique de saturation si les couleurs de l'image de document sont concentrées dans une plage limitée ou bien si le nombre total de pixels de faible luminance (moins de 60) est limité. De même, la luminance est utilisée lorsque les couleurs du premier plan de l'image sont considérables ou bien si les couleurs de l'arrière-plan de l'image sont concentrés dans une plage limitée. Dans les autres cas, la luminance et la saturation sont à fois employés.

2.5.2. Méthodes de regroupement basées sur la région

Sobottka et al. [SOB 00] ont proposé une approche pour extraire le texte à partir des documents couleurs et les couverts des journaux. La méthode proposée utilise deux techniques pour l'établissement d'hypothèses sur les régions de texte. La première est une technique d'analyse descendante qui extrait d'abord des blocs rectangulaires puis rejette les blocs ayant une couleur homogène. La seconde est une technique ascendante de croissance de région qui cherche des formes homogènes à partir de formes arbitraires. Les régions finales de texte sont obtenues en combinant les résultats des deux techniques précédentes.

Dans [STR 02], une nouvelle méthode de binarisation a été proposée dont le but d'extraire automatiquement le texte à partir de documents couleurs de types mixtes. La méthode proposée est basée sur une combinaison d'une technique de réduction de couleurs adaptative et d'une approche d'analyse de disposition. La technique de réduction de couleurs adaptative est d'abord utilisée pour obtenir le nombre optimal de couleurs et afin de convertir le document dans les couleurs principales. Après, en utilisant les couleurs principales, le document est découpé en surfaces de couleurs séparables. Ainsi, des images binaires sont obtenues, chacune correspond à une couleur principale. L'approche d'analyse de disposition est appliquée indépendamment sur chaque surface de couleurs ce qui permet d'identifier les régions de texte. Une procédure de fusion est finalement appliquée pour fusionner les régions de texte dérivées des champs de couleurs produisant l'image binarisée finale.

Badekas et al. [BAD 06] ont présenté une nouvelle méthode de binarisation d'images de documents couleurs. Premièrement, les couleurs de l'image de document sont réduites à un petit nombre en utilisant une nouvelle technique de réduction. Cette technique estime les couleurs dominants et ensuite les assigne les différentes couleurs de l'image de sorte que le fond et les composants du texte soient uniformes. Chaque couleur dominante définit une surface de couleurs dans laquelle les composantes connexes sont extraites. Après, dans chaque surface de couleurs, une procédure de filtrage des composantes connexes est appliquée suivie par une procédure de regroupement ce qui permet de construire des blocs de composantes connexes. Ces derniers sont ensuite redéfinis en obtenant la propriété de direction de connexion de chaque composante connexe. En utilisant la propriété de direction de connexion, les blocs de composantes connexes sont classifiés comme texte ou non-texte. Les blocs de texte identifiés dans les différentes surfaces de couleurs sont ensuite binarisés en utilisant des techniques de binarisation souhaitables, et les pixels restants sont considérés comme des pixels de fond.

2.6. Conclusion

Dans ce chapitre, nous avons abordé le problème de la séparation avant-plan/ arrière-plan considéré comme une des étapes clés dans le processus de traitement et d'analyse de

documents. Nous avons concentré sur la binarisation qui consiste à obtenir une image à deux niveaux (noir et blanc) uniquement dont le noir représente l'avant-plan et le blanc représente l'arrière-plan. En effet, une différence très fine existe entre la séparation avant-plan/ arrière-plan et la binarisation dans les images couleurs, c'est pourquoi nous avons commencé le chapitre par souligner la différence entre ces deux concepts. De même, nous avons mentionné la différence entre la binarisation et le seuillage, la technique la plus répandue de binarisation.

Nous avons mené une étude globale d'un grand nombre d'algorithmes de binarisation d'images tout en focalisant sur la binarisation d'images de documents en niveaux de gris. Nous avons considéré la répartition des différentes techniques de binarisation en techniques basées sur le seuillage et techniques basées sur le regroupement. Dans les techniques de seuillage, la binarisation est effectuée en fonction d'un ou plusieurs seuils calculés à partir des informations statistiques de l'image et nous distinguons trois groupes : seuillage global, seuillage local, et seuillage hybride. Les techniques basées sur le regroupement reposent soit sur la région (croissance de région, division et fusion), ou sur la classification des pixels en utilisant un algorithme de classification supervisé ou non-supervisé.

Un total de 80 méthodes de binarisation de différentes catégories ont été exposées : 61 méthodes de seuillage, et 19 méthodes de regroupement. Notons que seulement la présentation théorique des algorithmes de binarisation a été abordée dans ce chapitre. L'évaluation des performances des méthodes de binarisation et l'estimation de la qualité d'images binarisées constituent un autre problème important qui mérite une étude distincte. Cela est l'objet du prochain chapitre.

Chapitre 3 :

Evaluation et
comparaison des
techniques de
binarisation

3.1. Introduction

L'importance de l'étape de binarisation dans le processus d'analyse et de reconnaissance de documents ne peut être ignorée. Cette étape a pour objectif de séparer l'avant-plan de l'arrière plan de l'image, ce qui permet d'éliminer le fond et le bruit et de ne garder que les informations pertinentes (texte, figures, tableaux, etc.). Les étapes suivantes dans la chaîne d'analyse et de reconnaissance de documents sont applicables directement sur des images binaires (en noir et blanc), et les résultats de ces étapes dépendent fortement de la qualité de l'image binarisée obtenue. Par conséquent, il est impératif d'évaluer la qualité des images binarisées pour déterminer si la binarisation a été correctement accomplie ou pas.

Comme nous avons vu dans le chapitre précédent, un grand nombre de techniques ont été proposées dans la littérature pour la binarisation d'images et plus particulièrement d'images de documents. Malgré le nombre important de méthodes proposées, aucune d'entre elles ne se montre générique et applicable à tous les types d'images, et même pour les images du même type, la technique de binarisation qui fournit le meilleur résultat peut être différente d'une image à une autre. Cela rend la sélection de la méthode la plus adéquate pour l'appliquer dans le cadre d'un tel ou tel système de traitement d'images une tâche difficile qui doit être effectuée attentivement.

Cependant, pour que la sélection de l'algorithme le plus approprié à une image ou un type particulier d'images soit justifiée, il est nécessaire d'évaluer la qualité des images binarisées, de comparer les résultats de binarisation de plusieurs algorithmes et de déterminer l'algorithme qui produit les meilleurs résultats. Ainsi, l'évaluation et la comparaison des performances des algorithmes de binarisation sont devenues aujourd'hui un sujet important dans l'étude de la binarisation.

En effet, l'évaluation des algorithmes de binarisation n'est pas une tâche facile car il n'y a aucun processus d'évaluation standard permettant de comparer les résultats, surtout en tenant compte des différents objectifs pour lesquels l'évaluation est procédée : reconnaissance de texte, catégorisation, extraction des zones d'intérêt, etc. Cependant, comme a souligné George Nagy dans son article intitulé « *Document Image Analysis: Automated Performance Evaluation* » [NAG 95], sauf pour la reconnaissance Optique de Caractères (OCR), peu d'efforts ont tenté d'évaluer les performances des systèmes d'analyse d'images de documents. Peu de méthodes, mesures et critères spécifiques ont été proposés et utilisés par les chercheurs pour l'estimation de la qualité des images binarisées [LEE 03][KUM 13].

Récemment, l'évaluation des algorithmes de binarisation pour les images de documents a été organisée dans le cadre d'une compétition internationale : DIBCO (pour *Document Image Binarization Contest*) / H-DIBCO (pour *Handwritten Document Image Binarization Contest*). La série de compétitions DIBCO/ H-DIBCO a vu le jour en 2009 dans le contexte de la conférence ICDAR 2009 et elle a comme objectif de suivre les avancées dans le domaine de la binarisation de documents et d'évaluer les performances des méthodes comparées par des mesures quantitatives et sur la même base de test.

L'objectif de ce chapitre est de dresser un état de l'art des principaux travaux d'évaluation de la qualité des images binarisées et de comparaison des performances des méthodes de binarisation d'images, tout en focalisant sur les images de documents. Nous nous intéressons à l'exposition des différentes approches d'évaluation et de comparaison proposées dans la littérature, ainsi que les différentes mesures et critères utilisés.

La suite de ce chapitre est organisée de la manière suivante. Après avoir défini en quoi consistent l'évaluation et la comparaison, et souligné la différence entre les deux concepts,

nous exposons les différentes approches d'évaluation des techniques de binarisation existantes. Ensuite, nous citons et détaillons les principales mesures et critères utilisés par les chercheurs pour accomplir l'évaluation. Après, nous décrivons quelques travaux antérieurs relatifs à l'évaluation des techniques de binarisation, avant de conclure.

3.2. Définitions

Avant d'aborder le sujet de l'évaluation des performances des méthodes de binarisation, il est important de fixer le sens de certains concepts.

3.2.1. Evaluation

Au sens étymologique du terme, « évaluer » vient de «ex-valuere », qui signifie « extraire la valeur de », « faire ressortir la valeur de », déterminer l'importance de quelque chose et attribuer une valeur à une situation ou à un produit [ADI 13].

Continuons avec l'historique du terme. Le mot «évaluer» est apparu au XIV^{ème} siècle dans le petit Robert, il tirait son origine «d'avaluer» (1283), de l'ancien français valué, valeur, prix. En 1366, il s'agit « d'ésvaluer» qui veut dire « déterminer la valeur, le prix de quelque chose». Évaluer venait du latin «Valère » qui signifie : être fort, valoir.

Cependant, le Petit Robert [ROB 04], indique que le terme « Evaluer » peut avoir plusieurs significations:

- Porter un jugement sur la valeur, le prix de...
- Estimer, priser.
- Expertiser.
- Calculer, chiffrer.
- Surveiller, sous évaluer.
- Déterminer (une quantité) par le calcul sans recourir à la mesure directe.
- Jauger. Apprécier, estimer, juger.

A partir des définitions précédentes, [ADI 13] a déduit une distinction entre deux sortes de jugements portés par l'évaluation. Ainsi, les notions de calcul, de chiffre et de prix renvoient à un jugement objectif fondé sur des critères externes, alors que les notions d'estimation et d'appréciation renvoient à un jugement subjectif fondé sur les critères internes [ADI 13].

La même ambiguïté du terme se trouve dans d'autres définitions :

- 1) Selon le trésor de la langue française 2004 [TRE 04], « Evaluer » veut dire :
 - Déterminer, délimiter, fixer avec précision.
 - Conjecturer, faire l'estimation d'une quantité, d'une durée qui n'est pas encore véritable »
- 2) Dans le Larousse en ligne [LAR 14], « Evaluer » c'est :
 - Déterminer, fixer, apprécier la valeur, le prix de quelque chose, d'un bien, etc.
 - Déterminer approximativement la durée, la quantité, le nombre, l'importance de quelque chose.
- 3) Dans le dictionnaire de langue française en ligne de *Linternote* [LIN 14], « Evaluer » signifie :
 - Fixer la valeur de quelque chose. Synonyme de *chiffrer*.
 - Fixer approximativement une quantité, une qualité. Synonyme de *juger*.

D'autres définitions existent :

Pour Neil Postman, théoricien en communication, « l'évaluation est un élément inévitable, nécessaire et naturel dans toute communication humaine » [POS 79]. Cette évaluation évoquée par Postman est proche du jugement que nous portons les uns sur les autres, au cours de nos rencontres et de nos échanges. Elle repose souvent sur des préjugés, des idées reçues, des stéréotypes dont nous ne sommes pas toujours conscients mais auxquels nous échappons difficilement.

Selon [CHA 90], nous entendons évaluer comme une prise d'informations sur un observable en regard de références pour éclairer des décisions.

Le grand dictionnaire terminologique de l'office québécois de la langue française [OFF 14] présente plusieurs définitions du mot « Evaluation » relatives à différents domaines, dont seulement une d'entre elles est liée à l'informatique. Cependant dans cette définition, « l'évaluation est une technique qui consiste à vérifier ou à tester, à l'aide d'un logiciel approprié, l'état d'un système informatique ou d'un réseau, afin d'en évaluer le niveau de sécurité et d'en déterminer les vulnérabilités ».

3.2.2. Comparaison

Nous présentons ici quelques définitions terminologiques du mot « comparaison » :

1) Dans Larousse en ligne [LAR 14] :

- La « comparaison » est l'action de comparer, de rapprocher des personnes ou des choses pour examiner leurs ressemblances ou leurs différences ; rapprochement.
- « Comparer » signifie :
 - Mettre en parallèle des choses, des personnes pour faire apparaître les similitudes ou les différences.
 - Établir entre des personnes ou des choses un lien, un rapport d'égalité, les rapprocher par la pensée, le jugement, l'appréciation, les placer au même rang, sur le même plan.
 - Mettre des personnes ou des choses en parallèle, établir un rapprochement, une comparaison entre elles pour expliciter sa pensée, son jugement, son appréciation.

2) Selon le dictionnaire *Le Littré* en ligne [LIT 14], « comparer » veut dire :

- Examiner simultanément les ressemblances ou les différences.
- Égaler à.

3) Dans le grand dictionnaire terminologique de l'office québécois de la langue française [OFF 14], plusieurs définitions du mot « Comparaison » ont été présentées, nous notons ici une définition relative à l'informatique:

- La comparaison est la technique qui consiste à chercher les points de similitude entre les éléments de deux ensembles.

4) Selon le trésor de la langue française 2004 [TRE 04], « Comparer » veut dire :

- Rapprocher pour mettre en évidence des rapports de ressemblance ou de différence.
- Mettre en parallèle, égaler, faire rivaliser en intensité ou en valeur.

3.2.3. Différence entre l'évaluation de la binarisation et la comparaison des binarisations

Après avoir présenté les définitions des mots *évaluation* et *comparaison*, nous mettons le point sur la différence entre ces deux termes dans le domaine de la binarisation d'images.

Considérant la binarisation comme une segmentation de l'image en deux classes : *avant-plan* et *arrière-plan*, nous empruntons la déclaration de Zhang dans [ZHA 97] sur la différence entre l'évaluation de la segmentation et la comparaison des segmentations. L'évaluation et la comparaison sont étroitement liées. Elles sont en effet différentes.

- 1) **L'évaluation de la binarisation** est un processus intra-technique. L'objectif de l'évaluation pour un algorithme spécifique de binarisation est d'apprécier quantitativement son comportement lors du traitement de différentes images et / ou de l'aider à atteindre des performances meilleures en configurant ses paramètres.
- 2) **La comparaison des binarisations** est un processus inter-techniques. L'objectif de la comparaison de différents algorithmes est de rapprocher leurs performances, de les classer et de fournir un moyen de choisir les algorithmes les plus appropriés selon les applications visées. Elle permet aussi de promouvoir des nouveaux développements en combinant les points forts de plusieurs algorithmes.

3.3. Les différentes approches d'évaluation de la qualité de binarisation

Alors que le développement d'algorithmes de binarisation a attiré une attention considérable dans la littérature, relativement moins d'efforts ont été consacrés à leur évaluation. En outre, la plupart d'efforts sur l'évaluation ont été destinés à concevoir des nouvelles méthodes d'évaluation et seulement quelques auteurs ont essayé de caractériser et de catégoriser les techniques d'évaluation existantes [ZHA 92].

Nous présentons dans la suite de cette section, quelques travaux essentiels de catégorisation et différentes approches d'évaluation des méthodes de binarisation.

3.3.1. Catégorisation des approches d'évaluation

Ntirogiannis et al. [NTI 08] ont établi une classification des méthodologies d'évaluation des méthodes de binarisation en trois catégories. Dans la première catégorie, l'évaluation s'effectue par un ou plusieurs visionneurs humains. Dans la deuxième catégorie, le résultat de binarisation est soumis à un logiciel de reconnaissance optique de caractères (OCR) et la qualité de la binarisation est mesurée par la précision de la reconnaissance. La troisième catégorie combine l'évaluation humaine et la précision de l'OCR.

Les mêmes auteurs ont ajouté une quatrième catégorie de techniques d'évaluation dans [NTI 08b][NTI 13]. Les techniques d'évaluation de cette catégorie évaluent la qualité d'une méthode de binarisation en comparant son résultat avec une image de vérité terrain (image binaire décrivant le résultat optimal attendu de la binarisation) au niveau pixel.

Une autre catégorisation a été proposée par Kumar et al. [KUM 13] dans laquelle les stratégies d'évaluation sont regroupées en trois catégories : *supervisé-pixel*, *supervisé-composants* et *basée-reconnaissance*. La stratégie basée sur la reconnaissance estime la qualité des images binarisées en se basant sur les résultats de reconnaissance de ces images. Dans la classe des techniques *supervisées-pixel*, des informations statistiques de l'image

peuvent être utilisées. Dans la catégorie supervisée-composants, les composantes connexes de l'image de vérité terrain sont utilisées pour l'évaluation.

Considérant la binarisation comme une segmentation de l'image en deux régions : *avant-plan* et *arrière-plan*, les techniques d'évaluation de la qualité de segmentation ont été utilisées pour l'évaluation de la binarisation. Selon [ZHA 96], les algorithmes de segmentation peuvent être évalués analytiquement ou empiriquement, de sorte que les méthodes d'évaluation peuvent être subdivisées en deux catégories: les méthodes analytiques et les méthodes empiriques. Les méthodes analytiques examinent directement et évaluent les algorithmes de segmentation eux-mêmes en analysant leurs principes et propriétés. Les méthodes empiriques apprécient indirectement les algorithmes de segmentation en les appliquant sur des images de test et en mesurant la qualité de leurs résultats. Diverses méthodes d'évaluation empiriques ont été proposées. La plupart de ces méthodes peuvent encore être classées en deux groupes: *méthodes empiriques de bonté* (connues aussi sous l'appellation *méthodes non-supervisées*) et *méthodes empiriques de divergence* (connues aussi sous le nom de *méthodes supervisées*). Dans la première catégorie, certaines propriétés souhaitables des images segmentées, souvent établies en fonction de l'intuition humaine, sont mesurées par des paramètres de "bonté". Les performances des algorithmes de segmentation évalués sont jugées en fonction des valeurs des mesures de bonté. Dans la deuxième catégorie, quelques références représentant les résultats de segmentation idéales ou attendus sont d'abord établies. Les résultats de segmentation obtenus par l'application d'un algorithme de segmentation sont comparés aux données de référence, et la comparaison est quantifiée par des mesures de divergence. Les performances des algorithmes de segmentation sont ensuite jugées en fonction des valeurs des mesures de divergence obtenues. Cependant, cette approche est la plus utilisée pour l'évaluation objective [ZHA 03]. Après cette discussion, trois groupes de méthodes peuvent être distingués (figure 3.1).

Zhang et al. [ZHA 08] ont divisé les différentes méthodes d'évaluation de la segmentation en deux catégories principales: *subjective* et *objective*, considérant le fait qu'un évaluateur humain examine les images segmentées visuellement ou non. L'évaluation *objective* peut être décomposée en deux groupes : *évaluation au niveau système* et *évaluation directe*. L'évaluation au niveau système examine l'impact d'une méthode de segmentation sur un système/ application plus large utilisant cette méthode, tandis que l'*évaluation directe* étudie les méthodes de segmentation indépendamment. Dans l'*évaluation directe* on distingue deux groupes : *analytique* et *empirique* en se basant sur le fait que la méthode elle-même ou bien les résultats qu'elle génère sont examinés. Les méthodes empiriques sont finalement divisées en deux classes : méthodes *non-supervisées* dites aussi méthodes *empiriques de bonté*, et méthodes supervisées dites aussi méthodes *empiriques de divergence*, en reposant sur le fait que la méthode d'évaluation nécessite une image de vérité terrain (de référence) ou non. Il est nécessaire de noter que ces classes ne sont pas exclusivement mutuelles, et qu'une technique d'évaluation peut appartenir à plusieurs catégories en même temps. Pour plus d'éclaircissements, la figure 3.2 illustre cette hiérarchie des techniques d'évaluation.

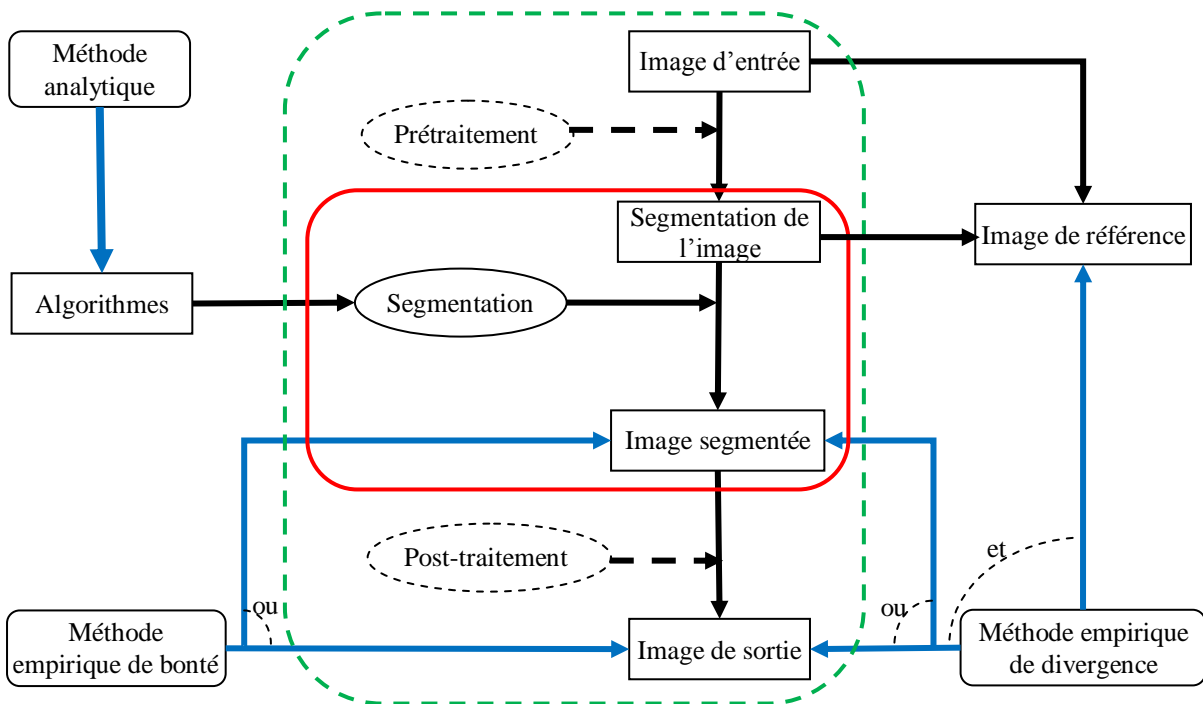


Figure 3. 1. Catégorisation des techniques d'évaluation des algorithmes de segmentation selon [ZHA 96]

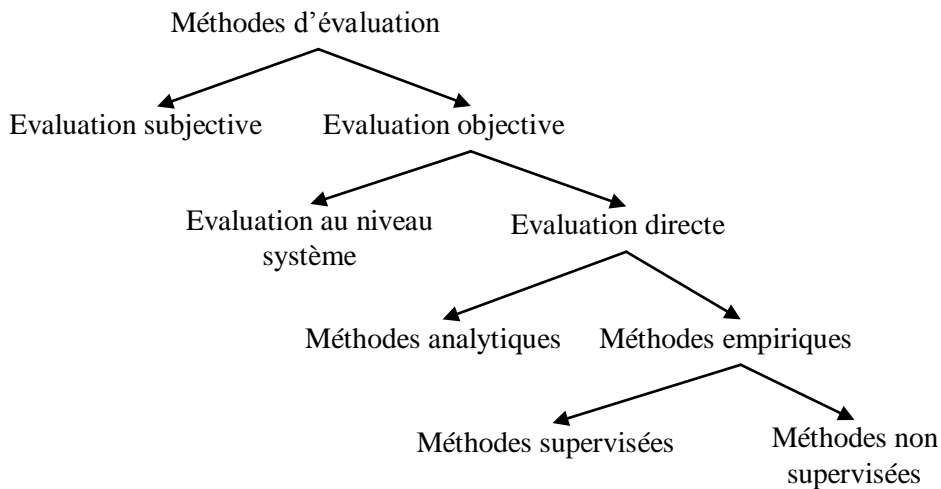


Figure 3. 2. Hiérarchie des méthodes d'évaluation des algorithmes de segmentation selon [ZHA 08]

3.3.2. Approches d'évaluation de la binarisation

Nous étudions dans cette section les différentes approches utilisées pour l'évaluation des performances des méthodes de binarisation. Nous essayons de catégoriser ces différentes approches d'évaluation en adoptant la taxonomie de Zhang et al. [ZHA 08] et nous mentionnons leurs utilisations, avantages et limites.

3.3.2.1. Évaluation subjective

L'évaluation subjective humaine est un aspect d'évaluation raisonnable et largement adoptif notamment pour des applications nécessitant un jugement humain [KAM 93].

L'évaluation subjective visuelle est l'une des techniques d'évaluation les plus fréquemment utilisées. Dans ce type d'évaluation, un ou plusieurs visionneurs humains choisissent pour chaque image considérée dans l'évaluation, la méthode de binarisation qui fournit le meilleur résultat. Ce choix peut être directement [PAL 86] [LEE 02][KEF 09][GOY 11][KHU 09], ou suivant des critères visuels bien déterminés. Dans ce dernier cas, le visionneur attribue un score à chaque image binarisée suivant chaque critère séparément.

Som et al. par exemple ont évalué les méthodes de binarisation en termes de *taux de symboles reconnais visuellement* par un humain [SOM 11]. Dans [TRI 95], les cinq critères suivants ont été utilisés : *Structures de ligne cassées (Broken line structures)*, *Symboles et texte cassés, (Broken symbols, text, etc.)*, *Confusion des lignes, symboles et texte (Blurring of lines, symbols and text)*, *Perte d'objets complets (Loss of complete objects)*, *Bruit dans les zones homogènes (Noise in homogeneous areas)*.

L'évaluation visuelle ou qualitative a l'avantage d'être guidée par l'objectif visé de l'évaluation, mais elle est très fastidieuse, couteuse en temps, manque la robustesse [NTI 08] et donne des résultats parfois non justifiés et peu précis. Un autre inconvénient est que les scores de l'évaluation subjective peuvent varier significativement d'un évaluateur à un autre, car chaque évaluateur possède ses propres standards pour l'estimation de la qualité des images binarisées [ZHA 08]. En outre, les résultats d'évaluation peuvent dépendre de l'ordre selon lequel les évaluateurs observent les résultats de binarisation, ainsi l'obtention d'une compréhension impartiale de l'efficacité d'un algorithme de binarisation est une tâche difficile. Elle nécessite une étude intensive de l'évaluation visuelle sur un large ensemble d'images par un groupe large d'évaluateurs.

3.3.2.2. Évaluation au niveau système

L'évaluation au niveau système examine l'impact d'une méthode de binarisation sur un système large employant cette méthode. L'évaluation au niveau système est une évaluation dirigée par le but ultime de l'application intégrant la méthode de binarisation. Cette approche permet aux chercheurs ou concepteurs d'affirmer qu'une méthode de binarisation est meilleure qu'une autre pour un système particulier en se basant sur les résultats finaux du système. Malheureusement, cette méthode d'évaluation est indirecte. Lorsque les étapes qui suivent la binarisation génèrent des bons résultats, cela ne signifie pas nécessairement que les résultats de binarisation étaient bons, et vice versa. Les résultats au niveau système des différentes méthodes de binarisation indiquent simplement que les caractéristiques des résultats ont été plus favorables pour ce système particulier [ZHA 08].

En effet, le système empirique le plus utilisé pour l'évaluation dans le cas des images de documents est l'OCR. Comme décrit [KAV 06], une approche d'évaluation objective doit impliquer des outils d'OCR pour tester la lisibilité des images de documents binarisées. Ainsi, plusieurs travaux d'évaluation ont utilisé l'OCR comme moyen d'évaluation quantitative, en calculant pour chaque image binarisée le taux ou la précision de reconnaissance des caractères/ mots détectés dans l'image. Cependant, l'utilisation de l'OCR comme moyen d'évaluation permet une décision quantitative claire mais il ne peut être appliqué que sur des documents imprimés modernes, supportés par des moteurs de reconnaissance contemporains [KAV 06]. Dans le cas des documents historiques par exemple, où la mauvaise qualité empêche la reconnaissance, cette manière d'évaluation ne peut être procédée.

Parmi les travaux d'évaluation utilisant l'OCR, nous citons [TRI 95b], [HE 05], [OGO 94], [LEE 03], [GUP 07], [KHU 09]. D'une manière analogue, Burgoyne et al. ont effectué une évaluation des méthodes de binarisation dirigée par la reconnaissance de la musique [BUR 07]. Ils ont utilisé pour cela un système de Reconnaissance Optique de la Musique (OMR).

Une évaluation dirigée par l'extraction des caractères présents dans les images binarisées a été procédée dans [KIM 02]. Les performances d'une méthode de binarisation ont été mesurées par le taux d'extraction des caractères défini comme le rapport du nombre de caractères extraits correctement sur le nombre total de caractères présents dans l'image.

3.3.2.3. Evaluation analytique

Les méthodes analytiques examinent et évaluent les algorithmes de binarisation indépendamment de leur résultat, en analysant leurs principes, exigences, propriétés, utilités, complexités d'exécution, etc. Comme a signalé [ZHA 96], l'utilisation des méthodes analytiques pour évaluer les algorithmes de binarisation permet d'éviter l'implémentation concrète de ces algorithmes. Cependant, pas toutes les propriétés des algorithmes de segmentation peuvent être obtenues par une étude analytique. Bien que certaines propriétés des algorithmes de binarisation puissent être obtenues facilement par analyse, d'autres propriétés ne peuvent pas être trouvées car aucun modèle formel n'existe.

De ce fait, les méthodes analytiques sont applicables seulement pour évaluer les propriétés algorithmiques ou d'implémentation des algorithmes de binarisation. Ces propriétés sont généralement indépendantes de la qualité des résultats de binarisation obtenus, ainsi ces propriétés ne sont pas efficaces pour l'estimation des performances réelles des algorithmes de binarisation [ZHA 08].

Parmi les travaux d'évaluation des algorithmes de binarisation appartenant à cette classe, nous citons celui de Kamel et Zhao [KAM 93] dans lequel les auteurs ont évalué les deux nouvelles méthodes qu'ils ont proposées avec d'autres méthodes de la littérature analytiquement suivant cinq aspects : *mémoire nécessaire, rapidité, restriction de la largeur de trait, nombre de paramètres, configuration des paramètres*. Pour avoir une idée plus claire sur ce type d'évaluation, nous présentons dans le tableau 3.1 les résultats d'évaluation publiés dans [KAM 93].

Tableau 3. 1. Evaluation analytique des méthodes de binarisation dans [KAM 93]

Technique	Evaluation subjective	Largeur	Temps CPU	Exigence mémoire	Nombre de paramètres	Réglage de paramètres
Seuillage global	Pire	Aucun	1.1	0	1	Facile
YDH	Lignes de blocs. Contours non souhaités	Partiellement	74.8	MN	5	Facile
Fonction non-linéaire	Pseudo-ombre Contours non souhaités	Partiellement	6.1	M	1276	Difficile
Fonction Intégrée	Bruit	Entièrement	14.5	$(2W+3)M$	2	Facile
Contraste local	Contours non souhaités	Partiellement	103.0	$9M$	4	Difficile
Niveau logique	Meilleure	Entièrement	7.7	$(4W-1)M$	2	Facile
Soustraction basée-masque	Mieux que d'autres	Entièrement	30.3	$2MN$	3	Facile

3.3.2.4. Evaluation empirique non-supervisée

Les méthodes non-supervisées connues aussi sous l'appellation de *méthodes empiriques de bonté* évaluent les performances des algorithmes de binarisation en estimant la qualité des images binarisées par l'utilisation de certaines mesures de la qualité qui s'accordent avec l'intuition humaine. Différents types de mesures de bonté, tels que *l'uniformité intra-région*,

le contraste inter-région, et la forme des régions, ont été proposées. L'évaluation non-supervisée est quantitative et objective. Son avantage principal est qu'elle évite la dépendance des images de références, ce que la rendre applicable dans plusieurs situations où les images de références n'existent pas. Selon [FUN 10], les expérimentations ont montré que les méthodes non-supervisées sont moins efficaces que les méthodes supervisées.

Cependant, l'évaluation non-supervisée en utilisant des mesures classiques de bonté a été précédée dans plusieurs travaux d'évaluation dans la littérature, citons par exemple [SAH 88], [SEZ 04], [RAM 10].

Kumar et al. ont proposé une nouvelle mesure pour l'évaluation des images de documents binarisées [KUM 13]. La mesure proposée est élaborée de l'analyse en composantes principales (ACP) et elle est basée sur la décomposition des valeurs propres. Pour chaque image binarisée, deux matrices de caractéristiques sont d'abord construites. Une pour l'avant-plan et l'autre pour l'arrière-plan. Les colonnes des matrices représentent : la valeur de gris, le numéro de ligne et le numéro de colonne d'un pixel. Le nombre de lignes dans chaque matrice égale au nombre de pixels appartenant à la classe associée à cette matrice. Après, les matrices de covariance sont estimées, à partir desquelles la décomposition des valeurs propres est effectuée. Ensuite, le produit des valeurs propres est calculé pour chaque image binarisée. L'image binarisée ayant la valeur de produit maximale est considérée comme la meilleure. Comme ont noté les auteurs, l'image binarisée sélectionnée peut ne pas être de bonne qualité pour la reconnaissance.

3.3.2.5. Evaluation empirique supervisée

Les méthodes d'évaluation *supervisées* aussi connues sous le nom de *méthodes empiriques de divergence*, comparent l'image binarisée avec une image de référence (de vérité terrain) préétablie et la différence entre les deux images détermine la qualité de la binarisation. Des mesures de divergence sont utilisées pour quantifier la différence entre les deux images. Un avantage potentiel des méthodes supervisées par rapport aux méthodes non supervisées est que la comparaison directe entre une image binarisée et une image de référence est censée fournir une résolution plus fine de l'évaluation. L'inconvénient majeur de ces méthodes est la nécessité de disposition des images de référence, ce qui est très difficile dans des applications réelles. Une solution possible à ce problème est d'utiliser des images synthétiques. Ainsi, on distingue deux catégories de techniques supervisées :

3.3.2.5.1. Evaluation supervisée-pixel (au niveau pixel)

Cette technique d'évaluation consiste à décider pour chaque pixel, s'il est bien classé ou non, et ce à travers la définition de plusieurs mesures quantitatives qui utilisent des informations statistiques de l'image [KAV 08]. Ainsi les mesures employées sont calculées en comparant les images binarisées avec des images de vérité terrain (de référence) préétablies, ce qui indique que cette évaluation est effectuée au niveau pixel et ne donne pas toujours des résultats précis pour des tâches comme la reconnaissance. Dans cette technique, soit du bruit est ajouté à l'image de vérité terrain afin d'obtenir une image en niveaux de gris synthétique, soit l'image en niveaux de gris originale est binarisée par un humain après plusieurs tests et nettoyages supplémentaires, pour produire une image de vérité terrain.

L'évaluation supervisée-pixel a été largement utilisée dans la littérature pour l'estimation objective de la qualité des images binarisées. Elle a été utilisée dans [ROS 03], [STA 08], [SEZ 04], etc. L'évaluation au niveau-pixel a été utilisée également dans le cadre des

compétitions internationales de binarisation DIBCO² 2009, H-DIBCO³ 2010, DIBCO 2011, H-DIBCO 2012 et DIBCO 2013.

Cependant, Lu et al. [LU 04] ont indiqué que les mesures empiriques supervisées ne s'accordent pas toujours avec l'évaluation subjective (visuelle humaine) et ils ont donné l'exemple du PSNR qui est l'une de mesures les plus connues. Pour les images de documents binaires, le PSNR ne s'accorde pas bien avec l'évaluation subjective, puisqu'il s'agit d'une mesure à base de points, et les relations mutuelles entre les pixels ne sont pas pris en compte.

3.3.2.5.2. Évaluation supervisée-composant

Dans l'évaluation supervisée-composant, les composantes connexes ou des caractéristiques extraites à partir de l'image de vérité terrain sont utilisées dans l'évaluation.

Nitrogiannis et.al par exemple ont évalué les performances des algorithmes de binarisation en comparant les squelettes détectées à partir des images binarisées avec des squelettes de vérité terrain préétablies semi-automatiquement [NTI 08].

3.4. Mesures d'évaluation empiriques

Les caractéristiques des méthodes d'évaluation sont principalement déterminées par les mesures d'évaluation employées. Cependant, les méthodes d'évaluation empiriques déterminent la qualité des images binarisées quantitativement en employant divers critères et mesures.

Nous citons dans cette section plusieurs mesures et critères empiriques largement utilisés pour des besoins d'évaluation. Les mesures d'évaluation empiriques sont regroupées en deux classes, *supervisées* et *non supervisées*, selon qu'elles imposent ou non la disponibilité d'images de vérité-terrain.

3.4.1. Mesures empiriques non-supervisées

Lorsque les images de vérité terrain ne sont pas disponibles, ce qui est le cas usuel, l'évaluation est dite non-supervisée et nous utilisons des mesures de la qualité de segmentation. De nombreux critères ont été proposés, cherchant à quantifier la qualité ou la lisibilité d'image. Selon Philipp-Foliguet et Guiges [PHI 06], ces critères peuvent être classés en deux grandes catégories : les critères de «contraste» et les critères d'«adéquation à un modèle». Les premiers recherchent une variabilité inter-région, alors que les seconds recherchent une uniformité en intensité ou en couleur à l'intérieur des régions. Zhang [ZHA 96] a séparé les mesures empiriques non-supervisées en trois classes: mesures basées sur l'uniformité intra-région, mesures basées sur le contraste inter-région, et mesures basées sur la forme des régions.

3.4.1.1. Mesure d'uniformité intra-région de Levine et Nazif (*GU*)

Levine and Nazif [LEV 85] ont admis qu'une segmentation parfaite produit des images avec des régions bien uniformes. Ainsi ils ont défini une mesure d'uniformité intra-région des niveaux de gris *GU* basée sur la somme des variances des régions comme suit :

² Document Image Binarization Contest

³ Handwritten Document Image Binarization Competition

$$GU = \sigma_B^2 + \sigma_F^2 = \sum_{i=1}^2 \sum_{(x,y) \in R_i} \left[I(x,y) - \frac{1}{|R_i|} \sum_{(x,y) \in R_i} I(x,y) \right]^2$$

Avec σ_B^2 et σ_F^2 dénotent les variances de l'arrière plan et de l'avant plan respectivement. I est l'image originale, R_i est la $i^{ème}$ région segmentée, $|R_i|$ le nombre de pixels dans R_i .

Une meilleure binarisation est obtenue avec une valeur réduite de GU .

Une mesure d'uniformité normalisée a été proposée et utilisée par Sahoo et al. [SAH 88]. Ainsi, la mesure d'uniformité normalisée est donnée par :

$$NGU = 1 - \frac{GU}{C}$$

Avec C un facteur de normalisation.

3.4.1.2. Mesure de Non-Uniformité de régions (NU)

La mesure de non-uniformité de régions juge la qualité des régions segmentées [SEZ 04]. Elle est définie par :

$$NU = \frac{|F|}{|F+B|} \frac{\sigma_F^2}{\sigma^2}$$

Tel que : σ^2 représente la variance de l'image entière, σ_F^2 la variance de l'avant plan, F et B dénotent l'avant plan et l'arrière plan de l'image respectivement, et $|R|$ la cardinalité (le nombre de pixels) de la région R .

Une image bien segmentée aura une mesure de non-uniformité proche de 0, et 1 dans le cas contraire.

3.4.1.3. Mesure de contraste inter-région (GC)

Levine and Navif ont admis aussi qu'une segmentation adéquate doit produire des images ayant des régions bien contrastées [LEV 85]. Notons μ_B et μ_F le niveau de gris moyen dans l'arrière plan et l'avant plan de l'image respectivement. La mesure de contraste inter-région (GC) utilisée dans [ZHA 96] est calculée par :

$$GC = \frac{|\mu_F - \mu_B|}{\mu_F + \mu_B}$$

Une image bien segmentée aura une valeur de GC proche de 1.

3.4.1.4. Contraste de Zeboudj (ZC)

Cette mesure prend en compte le contraste intérieur et le contraste extérieur aux régions, mesurés sur un voisinage $W(s)$ d'un pixel s . Ainsi le contraste de Zeboudj est défini par [COC 95]:

$$ZC = \frac{1}{A} \sum_{i=1,2} |R_i| \cdot C(R_i)$$

Tel que, A est le nombre de pixels dans l'image entière, R_i est la $i^{\text{ème}}$ région segmentée, $|R_i|$ le nombre de pixels dans R_i , $C(R_i)$ le contraste de R_i calculé par:

$$C(R_i) = \begin{cases} 1 - \frac{I_i}{E_i}, & \text{si } 0 < I_i < E_i \\ E_i, & \text{si } I_i = 0 \\ 0, & \text{sinon} \end{cases}$$

Avec, I_i et E_i dénotent le contraste intérieur et le contraste extérieur respectivement de la région R_i et ils sont donnés par :

$$I_i = \frac{1}{|R_i|} \sum_{s \in R_i} \max(c(s,t), t \in W(s) \cap R_i) \quad \text{et} \quad E_i = \frac{1}{l_i} \sum_{s \in F_i} \max(c(s,t), t \in W(s), t \notin R_i)$$

Tel que : $c(s,t)$ est le contraste entre deux pixels s et t défini par :

$$c(s,t) = \frac{|I(s) - I(t)|}{L - 1}$$

I représente l'intensité et L le maximum des intensités.

F_i est la frontière de R_i et l_i la longueur de F_i .

Cette mesure n'est pas adaptée aux images trop bruitées ou texturées d'après [PHI 06].

3.4.1.5. Mesure de forme des régions (SM)

La forme des régions segmentées peuvent être un critère discriminant dans l'évaluation de la qualité de segmentation. Sahoo et al. [SAH 88] ont proposé la mesure de forme (SM) pour l'évaluation des algorithmes de seuillage. Cette mesure est définie par :

$$SM = \frac{\sum_{(x,y)} Sgn(I(x,y) - \mu_{N(x,y)}) \Delta(x,y) Sgn(I(x,y) - T)}{C}$$

Avec $\mu_{N(x,y)}$ est le niveau de gris moyen du voisinage $N(x,y)$ d'un pixel de coordonnées (x,y) ayant un niveau de gris $I(x,y)$ et une valeur du gradient $\Delta(x,y)$. T est un seuil sélectionné pour la segmentation, C est un facteur de normalisation, et Sgn est la fonction seuil.

3.4.1.6. Variance pondérée (WV)

Otsu a proposé la variance pondérée WV comme suit [OTS 79]:

$$WV = \frac{|B| \cdot \sigma_B^2 + |F| \cdot \sigma_F^2}{|B + F|}$$

Tel que : σ_B^2 représente la variance de l'arrière plan, σ_F^2 la variance de l'avant plan, B et F dénotent l'arrière plan et l'avant plan de l'image respectivement, et $|R|$ la cardinalité (le nombre de pixels) de l'ensemble R .

Une valeur réduite de WV exprime une bonne binarisation.

3.4.1.7. Mesure de variance pondérée locale (LWV)

La mesure de variance pondérée locale (LWV) a été proposée par Ramirez-Ortegon et al. [RAM 10] et elle est dérivée de la variance pondérée d'Otsu (WV). Les auteurs ont affirmé que cette mesure est l'une des mesures les plus appropriées aux applications de l'OCR.

Notons F et B l'avant plan et l'arrière plan de l'image respectivement. On note $F_r(p)$ l'intersection entre F et le voisinage du pixel p dans une fenêtre de $(2r+1) \times (2r+1)$ centrée sur p . De même on note $B_r(p)$ l'intersection entre B et le voisinage du pixel p dans une fenêtre de $(2r+1) \times (2r+1)$ centrée sur p . LWV est donnée pour l'image entière par :

$$LWV = \frac{1}{|B+F|} \sum_p WV_r(p)$$

Avec $WV_r(p)$ est la mesure de variance pondérée locale à chaque pixel p en considérant son $(2r+1) \times (2r+1)$ voisinage.

$$WV_r(p) = \begin{cases} \frac{|B_r(p)| \cdot \tilde{\sigma}_{B_r(p)}^2 + |F_r(p)| \cdot \tilde{\sigma}_{F_r(p)}^2}{|B_r(p) + F_r(p)|}, & \text{si } |B_r(p)| \geq 2 \text{ et } |F_r(p)| \geq 2 \\ \tilde{\sigma}_{B_r(p)+F_r(p)}^2, & \text{sinon} \end{cases}$$

Avec $\tilde{\sigma}_A^2 = \ln\left(1 + \frac{\sigma_A^2}{[\mu_A]^2}\right)$, σ_A^2 dénote la variance de l'ensemble A et μ_A sa moyenne. $|A|$ la cardinalité de l'ensemble A .

Dans leurs tests, les auteurs utilisent $r=50$.

3.4.1.8. Mesure de variance uniforme locale (LUV)

La mesure de variance uniforme locale (LUV) est l'une des mesures les plus appropriées aux applications de l'OCR selon [RAM 10]. LUV de l'image entière est calculée par :

$$LUV = \frac{1}{|B+F|} \sum_p UV_r(p)$$

Avec $UV_r(p)$ est la mesure de variance uniforme locale à chaque pixel p en considérant son $(2r+1) \times (2r+1)$ voisinage.

$$UV_r(p) = \begin{cases} \frac{|B_r(p)| \cdot \tilde{\sigma}_{B_r(p)} + |F_r(p)| \cdot \tilde{\sigma}_{F_r(p)}}{|B_r(p) + F_r(p)|}, & \text{si } |B_r(p)| \geq 2 \text{ et } |F_r(p)| \geq 2 \\ \tilde{\sigma}_{B_r(p)+F_r(p)}, & \text{sinon} \end{cases}$$

Avec $\tilde{\sigma}_A = \ln\left(1 + \frac{\sigma_A}{[\mu_A]^2}\right)$, σ_A dénote l'écart type de l'ensemble A et μ_A sa moyenne. $|A|$ la cardinalité de l'ensemble A .

3.4.1.9. Mesure de dissimilarité de Liu et Yang

Ce critère est basé sur le nombre de régions, l'aire des régions et la couleur moyenne, dans l'espace RGB [LIU 94]. Il est défini par :

$$DM = \frac{1}{1000 \times A} \sqrt{2} \sum_{i=1}^2 \frac{e_i^2}{\sqrt{|R_i|}}$$

Tel que : A est le nombre de pixels dans l'image entière, e_i la somme des distances euclidiennes entre les vecteurs couleur des pixels de la région R_i et le vecteur couleur attribué à la région R_i dans l'image segmentée (en général la moyenne des couleurs de la région).

Le critère doit être faible pour une bonne segmentation.

3.4.2. Mesures empiriques supervisées

Ces mesures estiment l'efficacité des algorithmes de binarisation au niveau pixel. Elles sont calculées en comparant les résultats de seuillage avec des images 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 , FN le nombre de : vrais positifs, vrais négatifs, faux positifs, et faux négatives respectivement définis comme suit :

- 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} .

A partir des ces quatre valeurs, plusieurs mesures ont été définies.

3.4.2.1. Pourcentage de classification correcte (PCC)

C'est l'une des mesures les plus répandue dans la vision par ordinateur pour mesurer les performances de classification [ROS 03]. Elle exprime le pourcentage de pixels correctement classés.

$$PCC = \frac{TP + TN}{TP + FP + TN + FN}$$

3.4.2.2. Erreur de fausse-classification (ME)

Cette mesure a été utilisée aussi dans [KAV 08] [STA 08] sous l'appellation de *taux d'erreur de pixels* (*Pixel Error Rate* PEER). Elle exprime le pourcentage de pixels du fond incorrectement assignés à l'avant plan, et inversement [YAS 78]. L'erreur de fausse classification est donnée par :

$$ME = 1 - \frac{TP + TN}{TP + FP + TN + FN} = \frac{FP + FN}{TP + FP + TN + FN}$$

ME varie de 0 pour une image bien segmentée à 1 pour une image totalement mal binarisée.

3.4.2.3. Coefficient de Jaccard (JC)

Le coefficient de Jaccard est défini comme suit [ROS 03] :

$$JC = \frac{TP}{TP + FP + FN}$$

3.4.2.4. Coefficient de Yule (YC)

Le coefficient de Yule est défini comme suit [ROS 03] :

$$YC = \left| \frac{TP}{TP + FP} + \frac{TN}{TN + FN} - 1 \right|$$

3.4.2.5. F-mesure (FM)

F-mesure a été introduit la première fois par Chinchor dans [CHI 92]. A cause de sa simplicité, *F-mesure* est considéré 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-mesure* est défini par :

$$FM = \frac{2 \times Rappel \times Précision}{Rappel + Précision}$$

Avec : $Rappel = \frac{TP}{TP + FN}$ et $Précision = \frac{TP}{TP + FP}$

3.4.2.6. Pseudo F-mesure (p-FM)

Pseudo F-mesure a été introduit par Ntirogiannis et al. dans [NTI 08]. Il est motivé par le fait que chaque caractère a une silhouette unique qui peut être représentée par son squelette. *Pseudo F-mesure* est basée sur une mesure appelée pseudo-rappel (*p-Rappel*) calculée à partir de l'image de vérité terrain amincie (*SG*).

$$p - FM = \frac{2 \times p-Rappel \times Précision}{p-Rappel + Précision}$$

Avec : $p-Rappel = \frac{\sum_{x=1}^M \sum_{y=1}^N SG(x, y) \times I_b(x, y)}{\sum_{x=1}^M \sum_{y=1}^N SG(x, y)}$

$$SG(x, y) = \begin{cases} 0, & \text{fond} \\ 1, & \text{texte} \end{cases}$$

M et *N* dénotent la largeur et la hauteur de l'image respectivement.

3.4.2.7. Précision de la moyenne géométrique (GA)

Lorsque le nombre de pixels noirs soit très petit par rapport au nombre de pixels blancs, l'utilisation des mesures d'exactitude standards ne se montre pas un bon choix. Pour pallier à

ce problème, Paredes et Kavallieratou [PAR 10] ont utilisé une autre mesure : *la précision de la moyenne géométrique (GA)*. Cette dernière est donnée par :

$$GA = \sqrt{\frac{TP}{TP + FN} \times \frac{TN}{TN + FP}}$$

3.4.2.8. Erreur carrée (MSE)

C'est l'une des mesures statistiques traditionnelles de description de la qualité de l'image [KIT 00]. Pour une image de taille $M \times N$, elle est donnée par :

$$MSE = \frac{\sum_{x=1}^M \sum_{y=1}^N [I_{GT}(x, y) - I_b(x, y)]^2}{M \times N}$$

3.4.2.9. Rapport signal sur bruit (SNR)

SNR est défini comme le rapport du signal moyen carré sur le bruit moyen carré [KIT 00]. Il est donné par :

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

Plus la valeur de *SNR* est grande, mieux sera la qualité de binarisation.

3.4.2.10. 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 [KIT 00]. *PSNR* a été utilisée dans toutes les éditions des compétitions DIBCO/ H-DIBCO.

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

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

3.4.2.11. 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 [AGU 05]. Elle combine le taux de faux négatifs (NR_{FN}) et le taux de faux positifs (NR_{FP}) comme suit:

$$NRM = \frac{NR_{FN} + NR_{FP}}{2}$$

Avec: $NR_{FN} = \frac{FN}{FN + TP}$ et $NR_{FP} = \frac{FP}{FP + TN}$

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

3.4.2.12. Métrique de pénalité de fausse-classification (MPM)

La métrique de pénalité de mauvaise-classification évalue le résultat de binarisation par rapport à la vérité terrain sur une base objet par objet [AGU 05]. 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.

$$MPM = \frac{MP_{FN} + MP_{FP}}{2}$$

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

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 la segmentation 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.

3.4.2.13. Mesure de distorsion de distance réciproque (DRD)

La métrique de distorsion de distance réciproque (DRD) est une mesure objective de distorsion dans les images binaires, et elle a été proposée par Lu et al. dans [LU 04]. Après, elle a été utilisée pour l'évaluation des images binarisées. Cette mesure corrèle correctement avec la perception visuelle humaine et elle mesure la distorsion pour tous les S pixels renversés comme suit :

$$DRD = \frac{\sum_{k=1}^S DRD_k}{NUBN}$$

$NUBN$ est le nombre de blocs non-uniformes 8×8 dans l'image de vérité terrain.

DRD_k est la distorsion du $k^{\text{ème}}$ pixel renversé de coordonnées (x, y) et elle est calculée en utilisant une matrice pondérée normalisée (W_{Nm}) de taille 5×5 comme définie dans [LU 04]. DRD_k est donnée par :

$$DRD_k = \sum_{i=-2}^2 \sum_{j=-2}^2 |I_{GT}(x+i, y+j) - I_B(x, y)| \times W_{Nm}(i+2, j+2)$$

Lu et al. affirment que cette mesure surmonte $PSNR$ dans le sens où elle prend la perception visuelle humaine en compte et elle est bien corrélée avec l'évaluation subjective, qui est le juge ultime de la distorsion [LU 04].

3.4.2.14. Erreur de l'aire de l'avant plan relative (RAE)

La comparaison des propriétés de l'objet comme la forme et l'aire, obtenues à partir de l'image segmentée, en tenant compte de l'image de référence peut fournir une mesure utile d'évaluation. Ainsi, pour tester la précision de la segmentation, Zhang a proposé une mesure nommée « précision de mesure ultime relative » (*RUMA*) calculée en comparant une caractéristique obtenue de l'image de référence avec une caractéristique estimée de l'image segmentée [ZHA 96]. Sezgin et al. ont modifié cette mesure pour le caractéristique de l'aire A et ont proposé ainsi « l'erreur de l'aire de l'avant plan relative » définie comme suit [SEZ 04]:

$$RAE = \begin{cases} \frac{A_{GT} - A_b}{A_{GT}}, & \text{si } A_b < A_{GT} \\ \frac{A_b - A_{GT}}{A_b}, & \text{si } A_{GT} < A_b \end{cases}$$

Où A_{GT} est la surface de l'avant plan dans l'image de référence, et A_b la surface de l'avant plan dans l'image binarisée. Evidemment, pour des régions bien segmentées, RAE est zéro, dans le cas contraire la mesure est maximale.

3.4.2.15. Distance de Hausdorff modifiée (MHD)

La distance de Hausdorff HD peut être utilisée pour évaluer la similarité entre la forme des régions dans l'image binarisée et celle des régions dans l'image de vérité terrain. Sezgin et al. ont mesuré la déformation de la forme par la moyenne des distances de Hausdorff modifiées (*MHD*) sur tous les objets [SEZ 04]. Par exemple, les *MHDs* sont calculées pour chaque boîte de caractère de 19×19 pixels, puis moyennées sur tous les caractères du document. La distance de Hausdorff modifiée est défini par:

$$MHD(F_{GT}, F_b) = \frac{1}{|F_{GT}|} \sum_{f_{GT} \in F_{GT}} d(f_{GT}, F_b)$$

Où : $d(f_{GT}, F_b) = \min_{f_b \in F_b} ED(f_{GT}, f_b)$

Avec : F_{GT} et F_b dénotent l'avant plan de l'image de vérité terrain et de l'image binarisée respectivement, et $|R|$ la cardinalité de l'ensemble R .

$ED(f_{GT}, f_b)$ est la distance Euclidienne entre deux pixels de l'avant plan de l'image de vérité terrain et de l'image binarisée.

3.4.2.16. Probabilité de l'erreur : mesure de Lee et al.

Considérons la binarisation comme un processus de classification des pixels, le pourcentage de pixels mal-classés semble d'être une mesure empirique utile d'évaluation.

La probabilité d'erreur (PE) est une mesure proposée par Lee et al. dans [LEE 90] et elle est définie comme suit :

$$PE = P(F) \times P(B / F) + P(B) \times P(F / B)$$

Tel que: $P(B/F)$ est la probabilité de l'erreur de classification de l'avant plan en tant que arrière plan, $P(F/B)$ est la probabilité de l'erreur de classification de l'arrière plan en avant

plan, $P(F)$ et $P(B)$ sont les probabilités à priori de l'avant plan et de l'arrière plan respectivement dans les images.

3.4.2.17. Fragmentation

Cette mesure se base sur l'assumption que pour une segmentation parfaite, le nombre d'objets dans l'image segmentée doit être égal au nombre d'objets dans l'image de vérité terrain [ZHA 96]. Ainsi, cette mesure est définie comme suit [STR 91]:

$$Frag = \frac{1}{1 + p \times |T_N - A_N|^q}$$

Où : T_N est le nombre réel d'objets dans l'image de référence, A_N est le nombre d'objets dans l'image segmentée, p et q sont des paramètres de mise à l'échelle.

3.4.2.18. Position des pixels mal segmentés : mesure de Yasnoff et al.

Compter simplement le nombre de pixels mal segmentés est insuffisant, il est aussi important de tenir compte de la position de ces pixels en utilisant par exemple la distance entre un pixel mal segmenté et la région à laquelle il appartient dans la référence [PHI 06]. La mesure de Yasnoff et al. est définie par [YAS 78]:

$$YM = \frac{100}{A} \times \sqrt{\sum_{s=1}^S d^2(s)}$$

Où A est le nombre de pixels dans l'image entière, S le nombre de pixels mal segmentés, et $d(s)$ la distance du pixel mal segmenté s au pixel le plus proche de la région à laquelle il appartient.

3.4.2.19. Mesure de Strasters and Gerbrands

Une mesure bien connue dans l'évaluation de la détection de contours est le *FOM* (*mean-square distance figure of merit*) proposée par Pratt [PRA 78]. Strasters et Gerbrands ont utilisé *FOM* pour évaluer les résultats de segmentation. Ils ont ainsi défini une version modifiée de *FOM* calculé par [STR 91]:

$$FOM_e = \begin{cases} \frac{1}{S} \sum_{i=1}^S \frac{1}{1 + p \times d^2(i)}, & \text{si } S > 0 \\ 1, & \text{si } S = 0 \end{cases}$$

Tel que S est le nombre de pixels mal-classés, p est un paramètre de mise à l'échelle.

$d(i)$ dénote la distance entre le $i^{\text{ème}}$ pixel mal-classé et sa position correcte.

3.4.2.20. Inadéquation du contour

Cette mesure pénalise l'écart entre le contour de l'image en niveaux de gris et le contour de l'image binarisée. Notons CE le nombre de pixels de contour présents à la fois dans l'image de vérité terrain et dans l'image binarisée, EO l'ensemble de pixels de contour dans l'image de vérité terrain manquants dans l'image binarisée, et ET l'ensemble de pixels de contour dans l'image binarisées qui ne sont pas présents dans l'image de vérité terrain. L'inadéquation du contour est définie par [SEZ 01b]:

$$EMM = 1 - \frac{CE}{CE + a \left[\sum_{k \in \{EO\}} \delta(k) + b \sum_{k \in \{ET\}} \delta(k) \right]}$$

a et b sont des facteurs de pénalité associés aux pixels de contour mal-classés. $\delta(k)$ est donné comme suit :

$$\delta(k) = \begin{cases} |d_k| & \text{si } |d_k| < maxdist \\ D_{max} & \text{sinon} \end{cases}$$

d_k dénote la distance eclidienne du $k^{ème}$ pixel de contour mal-classé à un pixel de contour complémentaire dans une zone de recherche déterminée par le paramètre D_{max} .

[SEZ 01b] a suggéré les valeurs suivantes des paramètres: $maxdist = 0.025 \times N$, où N est la dimension de l'image, $D_{max} = 0.1 \times N$, $a = 10/N$, et $b = 2$.

3.4.2.21. Texte cassé (BT)

Cette mesure et les trois mesures qui suivent ont été proposées par Ntirogiannis et al. dans [NTI 08]. Le calcul de ces mesures est basé sur l'image de vérité terrain amincie SG et l'image étiquetée de composantes connexes correspondante SGC . Ces deux images sont définies comme suit :

$$SG(x, y) = \begin{cases} 0, & \text{fond} \\ 1, & \text{texte} \end{cases}$$

$$SGC(x, y) = \begin{cases} 0, & \text{si } SG(x, y) = 0 \\ i, & \text{sinon} \end{cases}$$

Avec $i \in \{1, 2, \dots, C_{GT}\}$, C_{GT} dénote le nombre de composantes connexes dans l'image de vérité terrain.

Le texte cassé (BT) est défini comme le pourcentage des parties de SG qui ne sont pas détectées dans l'image binarisée I_b , tout en appartenant à des composants partiellement détectés.

$$BT = \frac{\sum_{x=1}^M \sum_{y=1}^N f(SGC(x, y))(1 - I_b(x, y))}{\sum_{x=1}^M \sum_{y=1}^N SG(x, y)} \times 100\%$$

M et N dénotent la hauteur et la largeur de l'image respectivement. f est la fonction qui définit si une partie d'une composante de vérité terrain amincie est partiellement détectée dans l'image binaire I_b . Elle est donnée par :

$$f(i) = \begin{cases} 1, & \text{si } \sum_{x=1}^M \sum_{y=1}^N SG(x, y) I_b(x, y) > 0 \\ & \text{SGC}(x, y) = i \\ 0, & \text{sinon} \end{cases}$$

Avec $i \in K$

3.4.2.22. Texte manquant (MT)

Le Texte manquant (MT) est défini par le pourcentage des parties de SG qui ne sont pas détectées dans l'image binaire I_B tout en appartenant à des composants totalement détectés [NTI 08]. Il est donné par l'équation suivante :

$$MT = \frac{\sum_{x=1}^M \sum_{y=1}^N (1 - f(SGC(x, y))) \cdot SG(x, y)}{\sum_{x=1}^M \sum_{y=1}^N SG(x, y)} \cdot 100\%$$

3.4.2.23. Fausses alarmes (FA)

Les fausses alarmes réfèrent aux pixels de l'avant plan de l'image binaire I_b qui n'appartiennent pas à l'image de vérité terrain estimée. Elles sont définies par le pourcentage de tous les pixels des composants de l'image binaire I_b qui n'ont pas de pixels correspondants dans SG. Elles sont données par :

$$FA = \frac{\sum_{x=1}^M \sum_{y=1}^N h(BC(x, y)) \cdot I_b(x, y)}{\sum_{x=1}^M \sum_{y=1}^N I_b(x, y)} \cdot 100\%$$

Tel que : BC est l'image étiquetée de composants connexes de I_b définie par :

$$BC(x, y) = \begin{cases} 0, & \text{si } I_b(x, y) = 0 \\ j, & \text{sinon} \end{cases}$$

Avec $j \in L$ et $L = \{1, 2, \dots, C_B\}$, C_B dénote le nombre de composants connexes dans l'image binarisée I_b .

h est une fonction indiquant si un composant binaire n'est pas détecté dans SG.

$$h(i) = \begin{cases} 1, & \text{si } \sum_{x=1}^M \sum_{y=1}^N SG(x, y) \cdot I_b(x, y) = 0 \\ & BC(x, y) = i \\ 0, & \text{sinon} \end{cases}$$

Avec $i \in L$.

3.4.2.24. Déformations

Les déformations de l'image binaire sont définies par le pourcentage de tous les pixels de texte de l'image binaire I_b qui ne sont pas détectés dans l'image de vérité terrain estimée EG et ne font pas parti des fausses alarmes. EG est donnée par :

$$EG(x, y) = \begin{cases} 0, & \text{fond} \\ 1, & \text{texte} \end{cases}$$

Ainsi, nous distinguons la déformation des composantes (Def) de la déformation conduisant à la fusion de 2 composantes adjacentes ($MDef$). Elles sont définies comme suit :

$$Def = \frac{\sum_{x=1}^M \sum_{y=1}^N d(BC(x, y)) \cdot I_b(x, y) \cdot (1 - EG(x, y))}{\sum_{x=1}^M \sum_{y=1}^N I_b(x, y)} \cdot 100\%$$

$$MDef = \frac{\sum_{x=1}^M \sum_{y=1}^N m(BC(x, y)) \cdot I_b(x, y) \cdot (1 - EG(x, y))}{\sum_{x=1}^M \sum_{y=1}^N I_b(x, y)} \cdot 100\%$$

Où $d(i)$ et $m(i)$ sont des fonctions indiquant si une composante binaire correspond à un ou plusieurs composantes de vérité terrain respectivement. Elles sont définies comme suit:

$$d(i) = \begin{cases} 1, & \text{si } |SGC(x, y)| = 1 \\ & \forall (x, y): BC(x, y) = i \\ 0, & \text{sinon} \end{cases}$$

$$m(i) = \begin{cases} 1, & \text{si } |SGC(x, y)| > 1 \\ & \forall (x, y): BC(x, y) = i \\ 0, & \text{sinon} \end{cases}$$

3.5. Travaux antérieurs d'évaluation

Dans cette section, nous présentons quelques travaux antérieurs relatifs à l'évaluation et la comparaison des algorithmes de binarisation d'images. Nous essayons de donner un aperçu rapide de chaque travail en mentionnant : les différentes mesures d'évaluation utilisées, le type et le nombre d'images sur lesquelles les tests ont été effectués, et la méthode de binarisation gagnante.

Un des premiers travaux d'évaluation des techniques de binarisation est celui de Palumbo et al. [PAL 86]. Dans ce travail trois méthodes locales ont été comparées sur une collection de dix images de lettres postales bruitées. L'évaluation visuelle a favorisé la méthode de binarisation par mesure de contraste de Giuliano et al. [GIU 77].

Sahoo et al. ont présenté un survol sur plusieurs méthodes de seuillage globales et locales [SAH 88]. Parmi ces méthodes, neuf techniques globales ont été comparées sur trois images naturelles en utilisant deux mesures d'évaluation : la mesure d'uniformité (GU) et la mesure

de forme des régions (SM). Les méthodes les plus performantes selon cette étude comparative sont : la méthode de seuillage par préservation des moments [TSA 85], et la méthode d'Otsu [OTS 79].

Dans [OGO 94], l'auteur a comparé les performances de sept techniques de seuillage en les appliquant sur trente pages de journaux de mauvaise qualité. L'évaluation a été faite en calculant le taux de reconnaissance des images binarisées par un logiciel commercial de reconnaissance (*Calera model RS 9000*). Les meilleurs résultats ont été obtenus avec la méthode de préservation de connectivité [OGO 94b] et la méthode de seuillage par préservation des moments [TSA 85].

L'étude comparative de Trier et Taxt est considérée comme une des travaux d'évaluation les plus référencés. Dans ce travail, les auteurs ont évalué visuellement les performances d'onze algorithmes locaux de seuillage et quatre algorithmes globales sur un ensemble de sept images de cartes bruitées [TRI 95]. Un expert a estimé visuellement les résultats de binarisation et a attribué un score à chaque méthode suivant cinq critères : *Structures de ligne cassées (BLS)*, *Symboles et texte cassés, (BST)*, *Confusion des lignes, symboles et texte (BLST)*, *Perte d'objets complets (LCO)*, *Bruit dans les zones homogènes (NHA)*. Selon cette étude, la méthode de Niblack [NIB 86] avec une étape de post-traitement de Yanowitz et Bruckstein [YAN 86] semble la meilleure. Trier et Jane [TRI 95b] ont comparé les performances des 15 méthodes de binarisation précédentes objectivement. La comparaison a été effectuée en se basant sur le nombre de chiffres correctement reconnus à partir des images binarisées, et les résultats ont privilégié aussi la méthode de Niblack [NIB 86].

Dans [LEE 02], une comparaison de trois algorithmes de binarisation globale a été exposée dans le but de procéder un seuillage multi-niveaux des documents dégradés. L'évaluation visuelle sur un ensemble de 20 images de documents dégradés a annoncé que l'algorithme *QIR (Quadratic Integral Ratio)* [SOL 99] est le plus efficace pour la séparation texte/ fond. En 2003, les mêmes auteurs ont proposé deux nouvelles techniques et les ont comparées avec trois méthodes existantes [LEE 03]. L'évaluation a été effectuée sur 40 images de documents de 4 types différents (documents manuscrits historiques, chèques, formulaires, et pages de journaux) et elle a été exprimée en termes de *rappel* et *précision* de la reconnaissance. Selon cette étude, la nouvelle méthode proposée : *Seuillage par soustraction du fond* a donné les meilleures performances.

Dans [ROS 03], les auteurs ont développé un Framework pour l'évaluation quantitative des algorithmes de seuillage dans le cadre de la détection de changement dans les environnements de surveillance. Huit algorithmes ont été comparés sur une large collection contenant 4000 images pour lesquelles les images de vérité terrain ont été établies. L'évaluation a été procédée à 2 niveaux : niveau pixel en utilisant le *pourcentage de classification correcte (PCC)*, le *coefficient de Jaccard (JC)*, et le *coefficient de Yule (YC)*, et niveau dirigé-but pour la détection de changement. Cependant, quatre méthodes ont fourni des bons résultats avec la méthode de Kapur et al. [KAP 85] en première position.

Woo a comparé six méthodes de binarisation dans le but de l'analyse subséquente des empreintes de pieds numérisées (*foot-prints*) en utilisant un nouveau critère d'évaluation [WOO 04]. Le critère proposé est baptisé *MNFS (Nombre minimal de segments de l'avant-pant)* et il est relatif aux caractéristiques spécifiques aux images d'empreintes de pieds. Une collection de 71 images a été utilisée dans le test, sur lesquelles les 6 méthodes ont été appliquées. Les meilleurs résultats ont été obtenus avec la méthode de Abutaleb [ABU 89].

En 2004, Sezgin et Sankur ont publié un article de synthèse [SEZ 04] décrivant un des travaux les plus importants et les plus référencés dans le domaine de la binarisation d'images (cité 2545 fois selon *google scholar* jusqu'au 26/12/2014). Le travail exposé est une extension

d'une évaluation précédente des méthodes de binarisation sur des images naturelles (NDT) [SEZ 01]. L'importance de ce travail vient du fait qu'il introduit une nouvelle classification des algorithmes de binarisation basée sur le type d'informations exploitées, et exprime les formules des algorithmes présentés sous une notation uniforme. Quarante méthodes de binarisation de différentes catégories ont été testées sur deux types d'images : 40 images de documents et 40 images NDT. Pour une évaluation objective, les auteurs ont combiné cinq critères de la qualité de segmentation : *l'erreur de fausse-classification (ME)*, *l'inadéquation du contour (EMM)*, *l'erreur de l'aire de l'avant plan relative (RAE)*, *la distance de Hausdorff modifiée (MHD)* et *la non-uniformité de régions (NU)*. Les résultats finaux ont signalé que la méthode de *Kittler et Illingworth* [KIT 86] est la plus performante pour les 2 types d'images, suivie par les méthodes locales dans le cas des images de documents et par les méthodes globales entropiques dans le cas des images NDT.

Un autre travail de comparaison des techniques de binarisation pour les images de documents historiques a été effectué par He et al. [HE 05]. Dans cette étude comparative, cinq algorithmes de binarisation ont été évalués sur un ensemble composé de 4435 images de cartes d'archives contenant chacune un seul mot dactylographié. La qualité de binarisation a été mesurée par le taux de reconnaissance des images binarisées en utilisant le logiciel commercial de reconnaissance *Abbyy 6.0*. La méthode de *Niblack* [NIB 86] semble être la plus efficace ici et surmonte la méthode de *Sauvola* [SAU 00] malgré que cette dernière est une amélioration de la méthode de *Niblack*.

Vu l'influence du choix des valeurs des paramètres des algorithmes de binarisation sur le résultat final de la binarisation, *Badekas et al.* [BAD 05] ont proposé une nouvelle technique permettant l'estimation des meilleures valeurs de paramètres pour chaque méthode de binarisation et une évaluation automatique par la suite. Sept méthodes de binarisation ont été comparées sur un nombre important d'images de documents, et la méthode de *Sauvola* [SAU 00] est apparue la plus performante dans la plupart des cas.

Dans [GUP 07], les auteurs ont proposé deux nouvelles méthodes de seuillage et les ont comparé avec cinq méthodes de la littérature afin de tester leurs performances pour la recherche et l'indexation des documents latins imprimés historiques. L'évaluation a été faite sur 12 images de journaux et elle a été estimée en termes du taux de reconnaissance en utilisant un logiciel commercial (*ABBYY FineReader 7.1 SDK*). Selon les résultats obtenus, la méthode d'*Otsu* [OTS 79] et celles basées sur Otsu (les nouvelles méthodes proposées) se sont classées en premières places.

[BUR 07] a exposé une étude comparative de 25 méthodes de binarisation sur un ensemble de 100 images extraites de 5 anciens livres de musique. Dans les tests, le logiciel *Aruspix* de reconnaissance de musique (*OMR*) a été utilisé et l'évaluation a été quantifiée par le rappel et la précision de la reconnaissance. Selon cette étude comparative, les meilleurs résultats ont été obtenus avec la méthode de *Brink et Pendock* [BRI 96].

Dans le but de réduire l'intervention humaine dans le processus d'évaluation et comme l'évaluation par l'OCR n'est pas toujours pratique pour des images de documents manuscrits, *Ntirogiannis et al.* ont proposé une nouvelle méthodologie d'évaluation objective [NTI 08] [NTI 08b]. La méthodologie proposée est basée sur une procédure semi-automatique de construction des données de vérité terrain ainsi que sur 5 critères d'évaluation déduits du rappel et de précision: *F-mesure*, *Texte manquant (MT)*, *Texte cassé (BT)*, *Fausses alarmes (FA)*, et *Déformations (Def, MDef)*. Six algorithmes de seuillage ont été testés sur un nombre d'images de documents manuscrits et ont été évalués par la nouvelle méthodologie. Les résultats obtenus, indiquent que la méthode *adaptative de document dégradé* (ADD pour *Adaptive Degraded Document method*) [GAT 06] est la plus performante.

Kavalieratou [KAV 08] a conçu une technique d'appréciation des méthodes de binarisation pour les images de documents historiques. L'évaluation repose sur la comparaison des images binarisées avec celles de vérité terrain en calculant *le taux d'erreur*. Une comparaison entre trois algorithmes de seuillage a été décrite. Cette technique a été utilisée dans [STA 08][STA 08b] pour l'évaluation des performances de 30 méthodes de binarisation sur 150 images de documents synthétiques. Les performances ont été quantifiées par 4 mesures : *Taux d'erreur de pixels (PERR)*, *l'erreur moyenne carrée (MSE)*, *le rapport signal sur bruit (SNR)*, et *le rapport signal maximum sur bruit (PSNR)*. La conclusion tirée est que les méthodes locales sont légèrement mieux que celle globales, et que les méthodes globales basées sur l'histogramme ou la classification possèdent des performances proches des méthodes locales.

La technique d'évaluation précédente a été utilisée aussi dans la compétition organisée par ICFHR (*International Conference on Frontiers in Handwriting Recognition*) en 2010 pour l'évaluation quantitative de 6 méthodes de seuillage qui ont participé à la compétition [PAR 10]. Sur 150 images de documents historiques artificiels, les 6 méthodes ont été appliquées et leurs performances ont été estimées par *la précision de la moyenne géométrique (GA)*. Les meilleurs résultats ont été obtenus avec la méthode Synchron3 [MOG 09].

Kefali et al. [KEF 09] ont procédé une étude comparative visuelle de 12 méthodes de binarisation dans le but de choisir la plus adéquate pour leur système de recherche de documents arabes anciens. Les 12 méthodes ont été appliquées sur un ensemble d'anciens documents arabes et la méthode sélectionnée a été *la méthode NICK* [KHU 09].

[CHA 10] a décrit un Framework pour le choix automatique de l'algorithme de binarisation le plus approprié au traitement d'images d'anciens manuscrits. La méthodologie adoptée est basée sur l'apprentissage automatique en utilisant l'algorithme de rétro-propagation du gradient. Pour la validation, six algorithmes de binarisation ont été testés sur 380 images de d'anciens manuscrits. Les meilleurs résultats ont été obtenus avec *l'algorithme de Tsai* [TSA 85].

Dans [PAT 10], les auteurs ont comparé les performances de trois méthodes de binarisation dans le cadre de la reconnaissance de l'écriture Indienne sur une collection artificielle construite en ajoutant du bruit à 100 images de documents claires et nettoyées. L'évaluation a été faite à 2 niveaux : niveau pixel (en utilisant le critère *SNR*), et niveau dirigé-but en calculant le taux de reconnaissance. Les meilleures performances ont été obtenues avec *la méthode Adaptive* [PAT 10].

[FUN 10] a adressé un panorama des techniques d'évaluation des méthodes de binarisation pour la segmentation de caractères dans les images de manuscrits historiques. Une nouvelle technique a été proposée ici pour la sélection automatique de l'algorithme de binarisation le plus optimal en quatre étapes. La technique proposée est basée sur la détection de contour, l'extraction de caractéristiques et l'apprentissage automatique.

Ramirez-Ortegeon et Rojas [RAM 10] ont décrit une évaluation de l'efficacité de 6 mesures d'évaluation de la qualité de binarisation pour des applications de l'OCR. Quatre nouvelles mesures ont été aussi dérivées. Dans l'expérimentation, *la méthode de Sauvola* a été appliquée sur une base contenant 84 images et a été évaluée à la fois par les mesures précitées et par le taux de reconnaissance en utilisant le logiciel *TopOCR*. Les mesures les plus adéquates selon les auteurs pour des applications de l'OCR sont : *la variance pondérée (WV)* et *la variance uniforme (UV)* [RAM 10b].

[NIR 11] A donné un petit aperçu sur six techniques de binarisation en discutant leurs limites et avantages pour les images de documents dégradés. Le papier ne comporte aucune expérimentation, ni de comparaison. Un article similaire est celui d'Athimethphat [ATH 11].

Dans cet article l'auteur a présenté une vue globale de sept algorithmes de seuillage global pour l'amélioration de la qualité des images de documents historiques. Ainsi, l'auteur a mentionné la différence entre ces algorithmes, leurs avantages et leurs limites.

Intéressés à la préservation de documents Indiens historiques, Goyal et al. ont comparé visuellement les performances de quatre méthodes de seuillage sur différents types d'images de documents [GOY 11]. Des combinaisons d'algorithmes ont été testées et ont donné de bons résultats notamment *Kittler & Otsu*.

Une évaluation des techniques de binarisation pour des images naturelles a été décrite dans [SHA 12] dans le cadre de l'extraction du texte. Trois méthodes ont été testées sur 1027 images naturelles de dimensions réduites. L'analyse visuelle des résultats privilège la *méthode de Sauvola* [SAU 00].

Jing-Hao Xue et al., dans [XUE 12], ont fourni une clarification de la relation théorique entre les méthode de *Ridler et Calvard* [RID 78], *Kittler et Illingworth* [KIT 86], et *Otsu* [OTS 79]. Ils ont énoncé que la *méthode de Ridler et Calvard* est une version itérative de la *méthode d'Otsu* ; la *méthode d'Otsu* peut être considérée comme un cas spécial de la *méthode de Kittler et Illingworth*.

[NTI 13] a proposé une nouvelle méthodologie d'évaluation au niveau pixel destinée spécifiquement aux images de documents historiques. Deux nouvelles mesures ont été définies à savoir la *pseudo-rappel* et la *pseudo-précision*, à partir desquelles d'autres mesures ont été définies. Des expérimentations ont été effectuées pour tester la robustesse des mesures proposées. Ces dernières ont montré la meilleure corrélation avec l'OCR.

[JAC 14] a adressé un survol sur quelques méthodes de binarisation de documents historiques. Une petite discussion a été présentée et aucune évaluation n'a été réalisée. De la même manière, Kaur et Mahajan ont procédé une étude de quelques techniques de binarisation d'images de documents historiques dans le but de trouver les lacunes dans les techniques existantes [KAU 14]. Ainsi, l'étude a été soutenue par énoncer quelques limitations des travaux de recherche dans le domaine de la binarisation de documents historiques.

L'année 2009 a marqué le début des compétitions de binarisation DIBCO/ H-DIBCO. Ces compétitions s'intéressaient à l'évaluation quantitative des nouvelles méthodes proposées chaque année pour la binarisation d'images de documents. 43 algorithmes ont participé à la première édition de DIBCO en 2009 [GAT 09], et l'évaluation a été faite sur 10 images de documents historiques en utilisant les mesures suivantes : *F-measure*, *PSNR*, *NRM* (métrique de taux négatif), *MPM* (métrique de pénalité de fausse classification). La deuxième édition baptisée H-DIBCO 2010 [PRA 10] s'intéressait aux documents manuscrits historiques. Dans cette compétition, 17 algorithmes ont participé. L'évaluation a été faite sur 10 images de manuscrits historiques en utilisant les mesures suivantes : *F-measure*, *Pseudo F-measure*, *PSNR*, *NRM*, et *MPM*. Dans DIBCO 2011 [PRA 11], le nombre de méthodes participées était 18. Ces méthodes ont été évaluées sur un ensemble de 16 images de documents historiques. Les mesures d'évaluation utilisées ont été : *F-measure*, *PSNR*, *DRD* (métrique de distorsion de distance réciproque), et *MPM*. A la compétition H-DIBCO 2012 [PRA 12], 24 méthodes ont participé. Ces méthodes ont été comparées sur 14 images de manuscrits historiques en utilisant quatre mesures d'évaluation : *F-measure*, *Pseudo F-measure*, *PSNR*, et *DRD*. Durant la cinquième édition de la compétition (DIBCO 2013) les 23 méthodes participantes ont été évaluées sur 16 images de documents historiques (8 imprimés et 8 manuscrits) [PRA 13]. Finalement, H-DIBCO 2014 [NTI 14] a connu la participation de 8 méthodes de binarisation et l'évaluation a été effectuée sur 10 images de manuscrits historiques en termes de *F-measure*, *Pseudo F-measure*, *PSNR*, et *DRD*.

Le tableau 3.2 récapitule les travaux précédents selon l'ordre chronologique.

Tableau 3. 2. Travaux antérieurs sur l'évaluation des techniques de binarisation

Travaux	Nb de méthodes évaluées	Images de test	Type d'évaluation	Critères d'évaluation	Méthode suggérée
[PAL 86]	3 locales	10 images de lettres postales bruitées	Subjective	Visuel	binarisation par mesure de contraste [GIU 77]
[SAH 88]	9 globales	3 images naturelles	Objective non supervisée	- GU - SM	- Seuillage par préservation des moments [TSA 85] - Ostu [OTS 79].
[OGO 94]	7	30 pages de journaux	Objective au niveau-système	Taux de reconnaissance	- Seuillage par préservation de connectivité [OGO 94b] - Seuillage par préservation des moments [TSA 85]
[TRI 95]	11 locales 4 globales	7 images de cartes bruitées	Subjective	BLS, BST, BLST, LCO, NHA	Niblack [NIB 86] avec une étape de post-traitement de [YAN 86]
[TRI 95b]	11 locales 4 globales	Images de cartes	Objective au niveau-système	nombre de chiffres correctement reconnus	Niblack [NIB 86]
[LEE 02]	3 globales	20 images de documents dégradés	Subjective	Visuel	QIR (Quadratic Integral Ratio) [SOL 99]
[LEE 03]	1 globale 4 locales	40 images de documents	Objective au niveau-système	Rappel et précision des mots reconnus	Seuillage par soustraction du fond [LEE 03]
[ROS 03]	8	4000 images	Objectif -Supervisé - au niveau système	PCC, JC, YC détection de changement	Kapur et al. [KAP 85]
[WOO 04]	5 globales 1 locale	71 Images d'empreintes de pieds	Objective non supervisé	MNFS	Abutaleb [ABU 89]
[SEZ 04]	40 de différents types	40 images de documents 40 images NDT	Objective supervisée et non supervisée	ME, EMM, RAE, MHD, RU	Kittler et Illingworth [KIT 86]
[HE 05]	1 globale 4 locales	4435 images de cartes d'archives	Objective au niveau-système	Taux de reconnaissance	Niblack [Nib 86]
[GUP 07]	7	12 images de journaux	Objective au niveau-système	Taux de reconnaissance	Otsu [OTS 79]
[BUR 07]	25	100 images tirées de 5 anciens livres de musique	Objective au niveau-système	Rappel et précision de l'OMR	Brink et Pendock [BRI 96].
[NTI 08]	6	Images de documents manuscrits	Objective supervisée	F-mesure, MT, BT, FA, Def, MDef	ADD (Adaptive Degraded Document method) [GAT 06]
[STA 08] [STA 08b]	30	150 images de documents synthétiques	Objective supervisée	PERR, MSE, SNR, PSNR	
[PAR 10]	6	150 images de documents synthétiques	Objective supervisée	GA	Synchro3 [MOG 09]
[KEF 09]	6 globales 6 locales	Images d'anciens documents arabes	Subjective	Visuel	Nick [KHU 09]
[CHA 10]	6	380 images de d'anciens manuscrits	Objective supervisée	TP FP	Tsai [TSA 85].
[PAT 10]	3	100 images de documents	Objective -Supervisé - au niveau-système	SNR Taux d'OCR	Adaptative [PAT 10]

[GOY 11]	4	Images de documents	Subjective	Visuel	Combinaison de Kittler & Otsu
[SHA 12]	3	1027 images naturelles	Subjective	Visuel	Sauvola [SAU 00]
[GAT 09]	43	10 images de documents historiques	Objective supervisée	F-mesure, PSNR, NRM, MPM,	
[PRA 10]	17	10 images de manuscrits historiques	Objective supervisée	F-mesure, Pseudo F-mesure, PSNR, NRM, MPM	
[PRA 11]	18	16 images de documents historiques	Objective supervisée	F-mesure, PSNR, DRD,	
[PRA 12]	24	14 images de manuscrits historiques	Objective supervisée	F-mesure, Pseudo F-mesure, PSNR, DRD	
[PRA 13]	23	16 images de documents historiques	Objective supervisée	F-mesure, Pseudo F-mesure, PSNR, DRD	
[NTI 14]	8	10 images de manuscrits historiques	Objective supervisée	F-mesure, Pseudo F-mesure, PSNR, DRD	

3.6. Conclusion

L'évaluation des algorithmes de binarisation constitue un problème important et elle a suscité récemment une attention considérable, que ce soit pour : la sélection de l'algorithme le plus efficace pour une application particulière, la détermination des paramètres optimales d'un algorithme, ou la validation d'un nouvel algorithme en le comparant avec des méthodes déjà existantes. Plusieurs travaux ont été effectués dans la littérature pour l'évaluation et la comparaison des performances des techniques de binarisation. Plusieurs méthodes d'évaluation, mesures et critères de comparaison ont été proposés. L'étude des différentes techniques et mesures d'évaluation proposées dans la littérature nous semble un besoin essentiel. Le présent chapitre a eu comme objectif de présenter une vue globale sur cet axe important de recherche. Nous avons commencé le chapitre par quelques définitions de base pour souligner la différence entre l'évaluation et la comparaison. Ensuite, nous avons cité quelques classifications des approches d'évaluation existantes. Nous avons adopté par la suite une catégorisation des différentes techniques d'évaluation, et nous avons détaillé les caractéristiques de chaque catégorie de méthodes. Comme les différentes approches utilisent des mesures pour accomplir l'évaluation, il a été nécessaire aussi de présenter les mesures d'évaluation les plus connues. Finalement, nous avons listé et résumé certains travaux relatifs à ce domaine. L'étude présentée dans ce chapitre constitue la base théorique de nos deux premières contributions détaillées dans le chapitre suivant.

PARTIE II :
CONTRIBUTIONS

Chapitre 4 :
Contributions à
l'évaluation des
techniques de
binarisation

4.1. Introduction

Nous avons vu dans les chapitres précédents que la binarisation est une opération très importante de prétraitement, qui influence les étapes suivantes dans le processus d'analyse et de reconnaissance de documents, et par conséquent la sortie finale du système d'analyse. Considérant l'exemple d'un système de reconnaissance optique de caractères (OCR), une mauvaise binarisation produit souvent des taches de bruit et des ratures dans le document binarisé, affectant la tâche de segmentation. Elle peut également conduire à la fusion, la fracture, ainsi que d'autres déformations dans la forme de caractères présents dans le document binarisé. Tous ces problèmes entraînent la détérioration des performances de l'OCR. La figure 4.1 présente un exemple, où le document de la figure 4.1.a est binarisé par deux algorithmes différents. Le premier algorithme produit une image nette et claire (figure 4.1.b), tandis que le deuxième algorithme résulte une image bruitée contenant des taches, des caractères cassés, etc. (Figure 4.1.d). Les textes reconnus des deux images binarisées en utilisant un logiciel commercial de reconnaissance (Abby Fine Reader) sont affichés dans la figure 4.1.c et la figure 4.1.e respectivement.

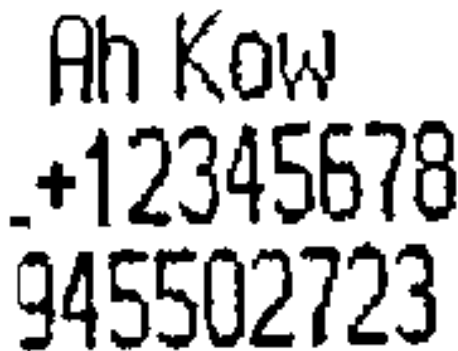
Nous trouvons dans la littérature un grand nombre de techniques de binarisation et ce nombre augmente continuellement chaque année. Avec ce nombre important de techniques proposées, nous nous trouvons face à une question importante qui nécessite d'être répondue: Quel algorithme de binarisation permet de donner le meilleur résultat pour une image ou pour une catégorie d'images? Afin de répondre à cette question, il est nécessaire de tester différents algorithmes de binarisation, d'évaluer leurs résultats et de comparer leurs performances pour chaque image séparément. La contribution à l'évaluation et à la comparaison des techniques de binarisation, que ce soit en proposant des nouvelles mesures, critères, ou méthodologies d'évaluation, ou plus simplement en procédant une étude comparative de certains nombre de techniques de binarisation sur un type particulier d'images est devenue depuis quelques années un axe actif de recherche dans le domaine de l'analyse et la reconnaissance de documents. Plusieurs travaux sur l'évaluation et la comparaison des performances des méthodes de binarisation ont été effectués, différentes approches ont été proposées, et divers critères et mesures ont été utilisés. C'est dans ce contexte que s'inscrivent nos deux premières contributions.

Dans ce chapitre nous présentons nos deux contributions à l'évaluation des techniques de binarisation d'images de documents historiques. La première contribution s'agit d'une étude comparative objective de plusieurs méthodes de seuillage en utilisant différents critères et mesures et critères d'évaluation, tandis que la deuxième contribution consiste à la proposition d'une nouvelle méthodologie d'évaluation des méthodes de binarisation pour la recherche dans les images de documents arabes dégradés.

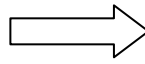
La suite de ce chapitre est subdivisée en deux grandes sections. La section 4.2 expose notre première contribution. Elle mentionne ainsi les objectifs visés, les méthodes de binarisation comparées, les mesures d'évaluations utilisées, les expérimentations effectués ainsi que les résultats obtenus. La section 4.3 détaille notre deuxième contribution. Elle présente ses objectifs, décrit la méthodologie d'évaluation proposée, et expose les tests réalisés et les résultats obtenus.



(a) Document dégradé

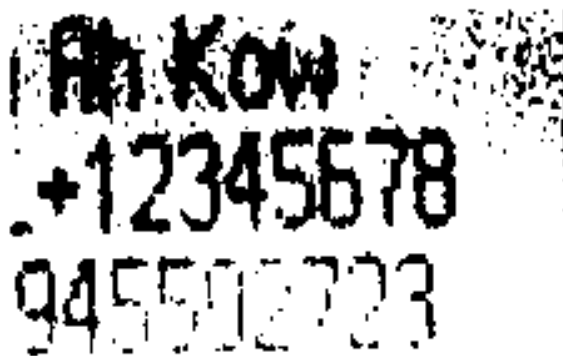


(b) Résultat d'une binarisation correcte

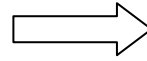


Ah Kow
t1 2345678
94 5S02723

(c) Texte reconnu correspondant



(d) Résultat d'une binarisation incorrecte



q4l r7' fl
— 4' .1 _I lbi. d •

(e) Texte reconnu correspondant

Figure 4. 1. Influence de la qualité de binarisation sur le résultat de l'OCR

4.2. Evaluation objective de plusieurs techniques de binarisation d'images de documents dégradés

4.2.1. Motivations et objectifs

Vu l'influence du résultat de la binarisation sur la sortie finale du système de traitement et d'analyse d'images, la détermination de la technique de binarisation la plus appropriée à un type particulier d'images est cruciale. Dans cette contribution, nous nous intéressons aux *images de documents historiques*.

Parmi les travaux d'évaluation des techniques de binarisation effectués dans la littérature, c'est le travail de Sezgin et Sankur publié en 2004 [SEZ 04] qui nous semble d'être le plus

important. Dans leur article intitulé « *Survey over image thresholding techniques and quantitative performance evaluation* », Sezgin et Sankur ont comparé les performances de 40 méthodes de binarisation sur deux types d'images : images de documents et images naturelles, en utilisant cinq mesures d'évaluations. L'article propose également une catégorisation des méthodes étudiées selon le type d'informations qu'elles exploitent en six groupes.

En effet, depuis 2004, un grand nombre de méthodes de binarisation de différentes catégories employant diverses techniques ont été proposées dans la littérature. Tant d'articles décrivant ces méthodes proposées ont été publiés. Dans chacun de ces articles, les auteurs présentent des tests effectués sur une base locale d'images et assument que leur méthode fonctionne parfaitement et qu'elle est la meilleure ; la décision qui nécessite d'être vérifiée par un tiers. Ce sont les travaux d'évaluation et de comparaison des techniques de binarisation qui peuvent accomplir cette tâche. Face au grand nombre de méthodes de binarisation proposées, les travaux sur l'évaluation et la comparaison des méthodes de binarisation ne sont pas de cette abondance, et une grande portion des méthodes proposées n'étaient incluses dans aucune étude comparative. En exception des compétitions internationales spécialisées DIBCO (2009, 2011, 2013) et H-DIBCO (2010, 2012, 2014), les travaux de comparaison procédés depuis 2004 sur les images de documents incluent un nombre restreint de techniques de binarisation, et les évaluent en utilisant quelques mesures, généralement supervisées, nécessitant la disponibilité d'images de vérité terrain, et qui ne s'accordent pas toujours avec l'évaluation visuelle humaine. Chaque étude comparative suggère une méthode de binarisation différente, et rarement que plusieurs évaluations s'accordent sur la même méthode.

Le besoin de sélectionner une méthode de binarisation adéquate aux images de documents dégradés, les lacunes dans les études comparatives précédentes et leur limitation à un nombre réduit de méthodes, souvent classiques, en ignorant des nouvelles techniques proposées en plus de leur désaccord sur la meilleure technique constituent nos principales motivations dans notre première contribution.

Le présent travail s'inscrit dans le cadre de l'évaluation et la comparaison des performances des méthodes de binarisation. Dans ce travail nous nous intéressons aux images de documents historiques connues par leur mauvaise qualité due aux différents types de dégradations et détériorations subies au cours de leur cycle de vie. Notre objectif essentiel est de procéder une étude comparative large afin de trouver un algorithme suffisamment efficace permettant de donner des résultats satisfaisants en matière de binarisation pour ce type particulier de documents. Afin d'atteindre cet objectif, nous incluons dans notre étude comparative un nombre important de techniques de binarisation. Pour une évaluation objective, nous utilisons une combinaison de plusieurs mesures quantitatives d'évaluation. Certaines mesures nécessitent la disponibilité d'images de référence (de vérité terrain) représentant la forme idéale de l'image binaire, et avec lesquelles on compare les images résultantes de la binarisation. Certaines autres mesures ne les nécessitent pas.

Nous nous intéressons dans cette étude aux méthodes de binarisation d'images en niveaux de gris car dans la plupart du temps, les documents en couleurs peuvent être transformés en niveaux de gris sans perte significative d'information, sauf quelques exceptions comme les documents publicitaires et certains styles de magazines. L'étude comparative présentée dans cette section est une extension d'un travail d'évaluation précédent de nous dont lequel nous avons effectué une comparaison visuelle entre douze méthodes de binarisation [KEF 09].

Le présent travail est bénéfique à plusieurs niveaux; (1) un nombre important de méthodes de différentes catégories sont implémentées et testées (51 méthodes) ce que surpasse les travaux d'évaluation effectués jusqu'à présent ; (2) une collection large et diversifiée d'images de

documents est construite et utilisée dans l'expérimentation; et (3) une combinaison de plusieurs mesures quantitatives supervisées et non supervisées sont utilisées pour l'évaluation

4.2.2. Base d'images utilisée

La base d'images que nous avons utilisée dans cette étude est composée de deux ensembles: un ensemble public d'images réelles, et un ensemble local d'images synthétiques. Notons que pour chaque image de la base, une image de vérité terrain représentant le résultat de binarisation idéal qu'on souhaiterait avoir doit être accompagnée car certaines mesures d'évaluation les nécessitent.

Nous décrivons dans cette section chacun des deux ensembles d'images.

4.2.2.1. Premier ensemble d'images réelles

Le premier ensemble d'images utilisé dans les expérimentations est constitué des cinq collections d'images proposées dans le cadre des compétitions internationales de binarisation DIBCO 2009⁴, H-DIBCO 2010⁵, DIBCO 2011⁶, H-DIBCO 2012⁷ et DIBCO 2013⁸. DIBCO (*Document Image Binarization Contest*) et H-DIBCO (*Handwritten Document Image Binarization Competition*) sont des compétitions internationales de binarisation d'images de documents émergées en 2009 dans le contexte de la conférence ICDAR 2009. Le but de ces compétitions est de suivre les avancées dans le domaine de la binarisation d'images de documents et d'évaluer les performances en utilisant des mesures quantitatives sur la même base. Chaque collection contient un certain nombre d'images de documents dégradées, et d'images de vérité terrain correspondantes créées semi-automatiquement en utilisant la procédure proposée dans [NTI 08]. Les images incluses dans les cinq collections DIBCO/H-DIBCO ont été sélectionnées attentivement par les organisateurs pour qu'elles couvrent la plupart des dégradations liées aux documents anciens (intensité variable dans l'arrière plan, faible contraste, ombre, taches d'encre et d'humidité, effet de transparence, plis et déchirures, etc.).

La collection de DIBCO 2009 contient 5 images de documents imprimés et 5 manuscrits proviennent de plusieurs bibliothèques (bibliothèque de l'état et de l'université Göttingen, bibliothèque de l'état de Bavière, bibliothèque britannique, et bibliothèque du Congrès), avec les images de vérité terrain associées. L'ensemble de test de H-DIBCO 2010 est composé de 10 images de documents manuscrits proviennent de la bibliothèque du Congrès, et des images de vérité terrain correspondantes. La collection de DIBCO 2011 comprends 8 images de documents imprimés et 8 manuscrits proviennent de: la bibliothèque de l'état et de l'université Göttingen, bibliothèque de l'état de Bavière, la bibliothèque britannique, et la bibliothèque du Congrès, accompagnées des images de référence. La collection de H-DIBCO 2012 contient 14 images de documents manuscrits, avec leur image de vérité terrain. L'ensemble d'images de DIBCO 2013 est composé de 8 images de documents imprimés et 8 manuscrits, et d'images de références associées.

Un total de 66 images est ainsi obtenu et elles sont réparties comme dans le tableau 4.1.

⁴ <http://users.iit.demokritos.gr/~bgat/DIBCO2009/benchmark/>

⁵ <http://www.iit.demokritos.gr/~bgat/H-DIBCO2010/benchmark>

⁶ <http://utopia.duth.gr/~ipratika/DIBCO2011/benchmark>

⁷ <http://www.iit.demokritos.gr/~bgat/H-DIBCO2010/benchmark>

⁸ <http://utopia.duth.gr/~ipratika/DIBCO2013/benchmark/>

Tableau 4. 1. Répartition d'images de documents dans les collections publiques de DIBCO et H-DIBCO

	Manuscrit	Imprimé	Somme
DIBCO 2009	5	5	10
H-DIBCO 2010	10	0	10
DIBCO 2011	8	8	16
H-DIBCO 2012	14	0	14
DIBCO 2013	8	8	16
Somme	45	21	66

Les figures 4.2 – 4.6 présentent quelques images extraites des collections DIBCO/ H-DIBCO avec les images de vérité terrain correspondantes.



Figure 4. 2. Image extraite de la collection DIBCO 2009 avec l'image de vérité terrain correspondante

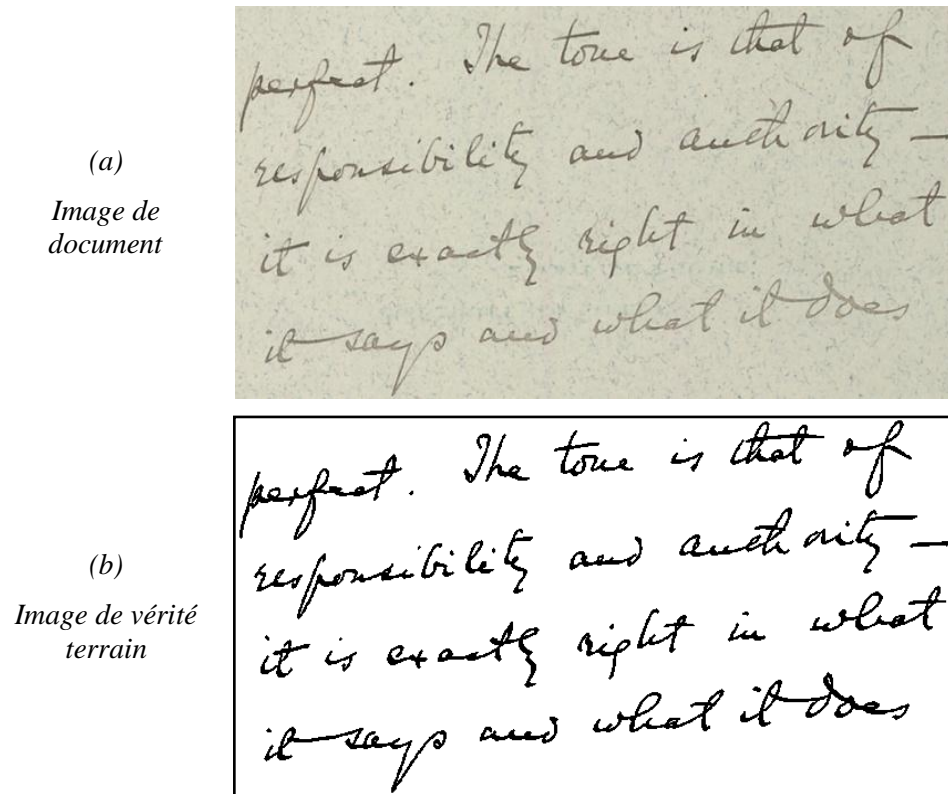


Figure 4. 3. Image extraite de la collection H-DIBCO 2010 avec l'image de vérité terrain correspondante

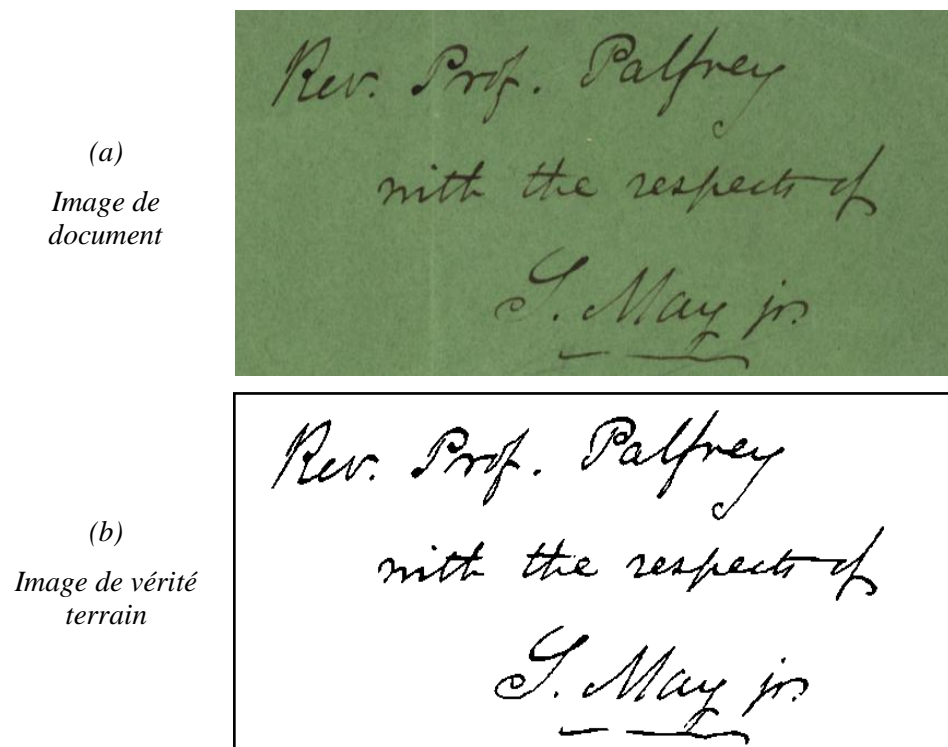


Figure 4. 4. Image extraite de la collection DIBCO 2011 avec l'image de vérité terrain correspondante

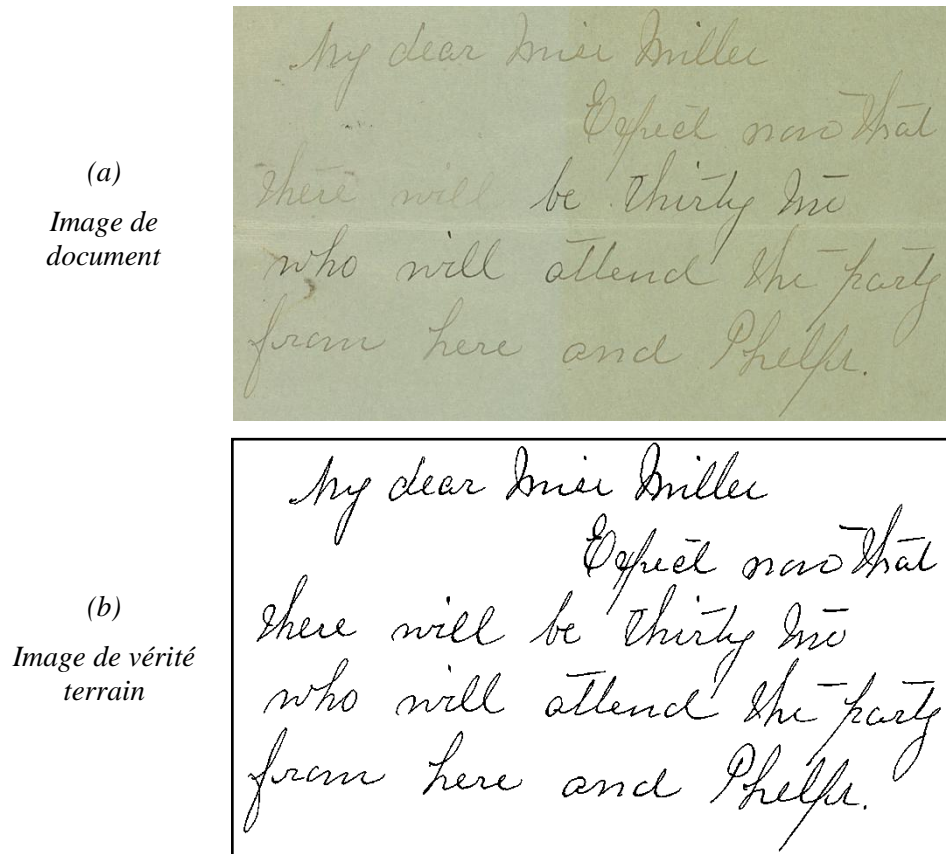


Figure 4. 5. Image extraite de la collection H-DIBCO 2012 avec l'image de vérité terrain correspondante

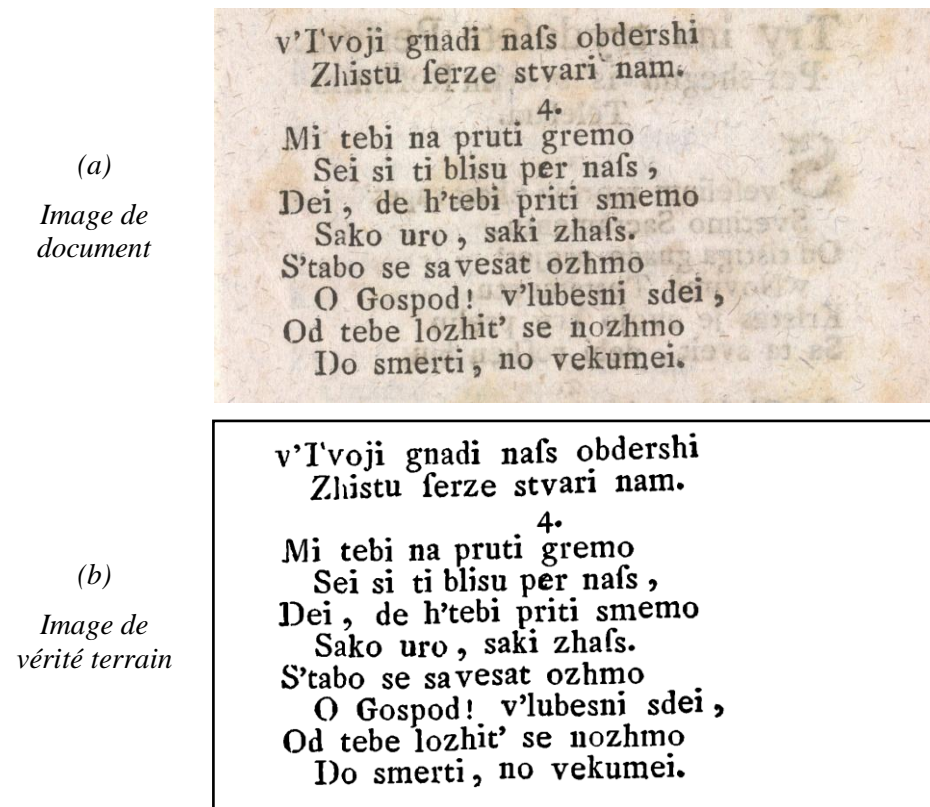


Figure 4. 6. Image extraite de la collection DIBCO 2013 avec l'image de vérité terrain correspondante

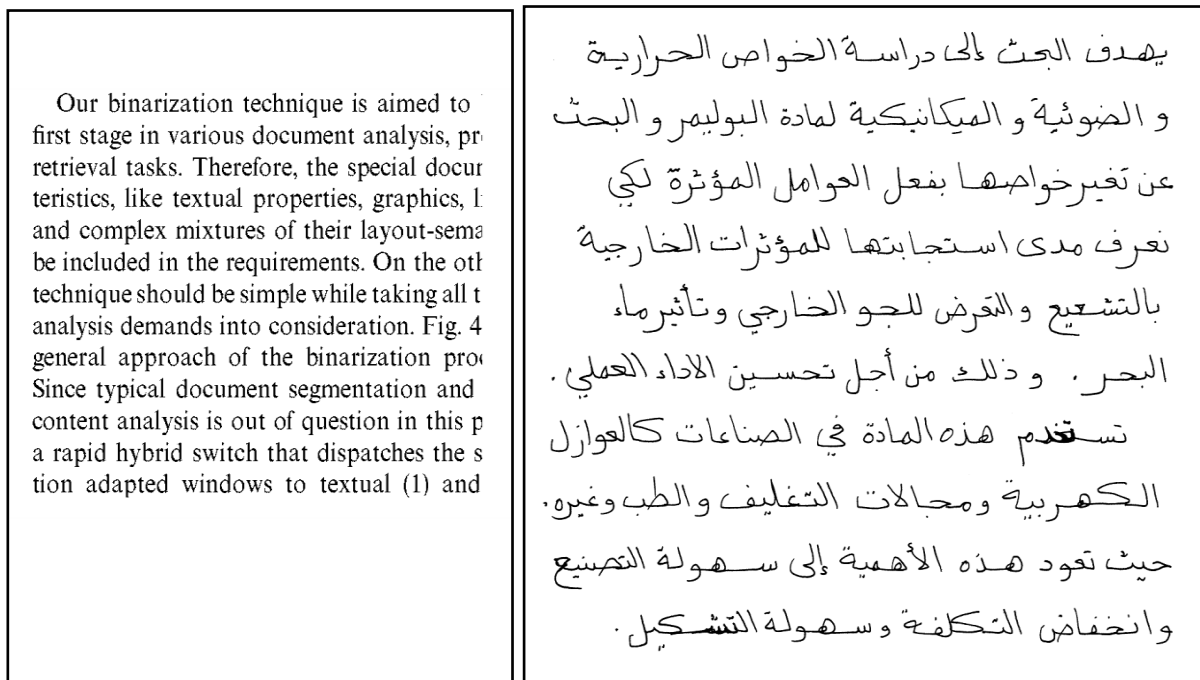
4.2.2.2. Deuxième ensemble d'images synthétiques

Pour que les résultats d'évaluation soient plus sûrs, les tests doivent être effectués sur une large collection de données. Ainsi, le premier ensemble d'images (images de DIBCO/H-DIBCO) nous semble insuffisant. Nous avons besoins donc d'images supplémentaires pour les utiliser dans les tests, tout en prenant en considération que chaque image doit être accompagnée de sa version binaire idéale (image de référence ou vérité terrain). Le deuxième ensemble d'images a été alors construit.

Cependant, il n'est possible pratiquement d'utiliser qu'un nombre réduit d'images réelles à cause de la difficulté de l'établissement manuel des images de vérité terrain, ce qui s'oppose à notre objectif. Pour cette raison, nous avons décidé d'utiliser des images synthétiques au lieu d'images réelles, et nous avons ainsi construit un ensemble d'images de documents historiques synthétiques. En effet, l'idée d'utiliser des images synthétiques dans les tests n'est pas d'aujourd'hui. Les coûts élevés des bases de données commerciales et des phases de collections de données ont motivé plusieurs groupes de recherche à développer des méthodes de création des bases de données synthétiques [ELA 08]. [MÄR 01] et [KIE 14] sont des exemples.

Dans notre cas, les images de documents dégradés synthétiques sont générées en partant d'un ensemble d'images de référence de départ (images en noir et blanc propres et épurées de bruit), et en les ajoutant du bruit relatif aux documents anciens (taches, transparence, trous, etc.). Les images de documents de référence sont prises de deux sources. Certaines images ont été tirées à partir des pages PDF contenant du texte imprimé et/ou manuscrit, avec ou sans figures, tableaux, colonnes, formules, etc. Les autres images ont été sélectionnées à partir de la base AHDB [ALM 02] et contenant des textes manuscrits arabes.

La figure 4.7 présente deux images différentes utilisées pour la génération d'images de documents synthétiques.



(a) Image prise d'une page PDF

(b) Image sélectionnée de la base AHDB

Figure 4.7. Exemples d'images de référence utilisées pour la génération d'images de documents synthétiques



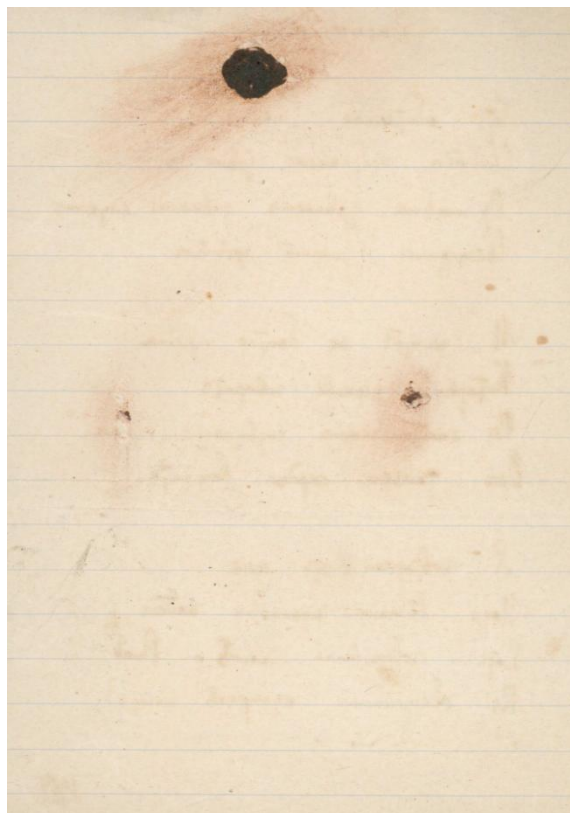
(a) Taches d'humidité



(b) Déchirure, Tampon, et bruit marginal



(c) Fond variable



(d) Trous et effets de transparence

Figure 4. 8. Quelques images de fonds anciens utilisées pour la génération d'images de documents synthétiques

Une image de document synthétique (artificielle) est obtenue par la superposition d'une image de référence et un masque. Les masques sont des images de fonds anciens vierges (ne contiennent aucun texte ou graphiques), collectées à partir du web et couvrent la plupart des dégradations rencontrés dans les documents anciens (présence des taches, illumination inégale, plis et déchirure, etc.). La Figure 4.8 illustre quelques exemples de fonds anciens de notre collection.

La combinaison est effectuée en utilisant la technique de *Mosaicing* d'images proposée par P. Stathis et al. dans [STA 08]. Ainsi dans [STA 08], deux techniques de combinaison ont été proposées : *l'intensité maximale*, et *l'intensité moyenne*. En utilisant la technique de *intensité maximale*, l'image résultante sera constituée en privilégiant les pixels les plus foncés, c.à.d. en attribuant à chaque pixel dans la nouvelle image, l'intensité du pixel le plus foncé entre les deux pixels correspondants des 2 images combinées (l'image de référence et l'image du fond). De telles façons, pour les pixels de l'avant plan, l'image du texte aura dans la plupart des cas l'avantage sur l'image du masque et pour les pixels de l'arrière plan, l'image du masque aura l'avantage car elle est généralement plus foncée. En utilisant la technique de *Mosaicing* par *intensité moyenne*, l'intensité de chaque pixel dans la nouvelle image égale à la moyenne des intensités des deux pixels correspondants dans les deux images combinées.

Alors que chacune des deux techniques précédentes permet de produire des images de documents synthétiques contenant du bruit et des dégradations, il nous semble que le mixage de ces deux techniques permettra de donner des résultats plus naturels et proches des images réelles. Nous avons adopté ainsi cette solution pour la construction de notre deuxième ensemble d'images de documents historiques synthétiques. Ainsi, les intensités des pixels dans l'image résultante sont attribuées en appliquant la technique de *l'intensité maximale* lorsque les pixels dans l'image du masque soient plus foncés que les pixels correspondants dans l'image du texte, et en appliquant la technique de *l'intensité moyenne* dans l'autre cas. En d'autres termes, un pixel dans l'image résultante aura l'intensité du pixel correspondant dans l'image du masque lorsque ce pixel est plus foncé que celui de l'image du texte, et aura la moyenne des intensités des deux pixels correspondants dans les deux images combinées dans le cas contraire.

Cette technique peut être résumée par le pseudo code suivant :

Algorithme de superposition par mosaicing

```

Entrées : GT: l'image du texte
           BG: l'image de vérité terrain
Sortie:   R: l'image résultat
Début
  Pour chaque pixel (i, j) de GT faire
    Si BG(i,j) est plus foncé que GT(i,j) alors  $R(i,j) \leftarrow BG(i,j)$ 
    Sinon  $R(i,j) \leftarrow (GT(i,j)+BG(i,j))/2$ 
  Fin Pour
Fin
    
```

La figure 4.9 présente un exemple d'une image synthétique obtenue en utilisant l'algorithme de combinaison précédent.

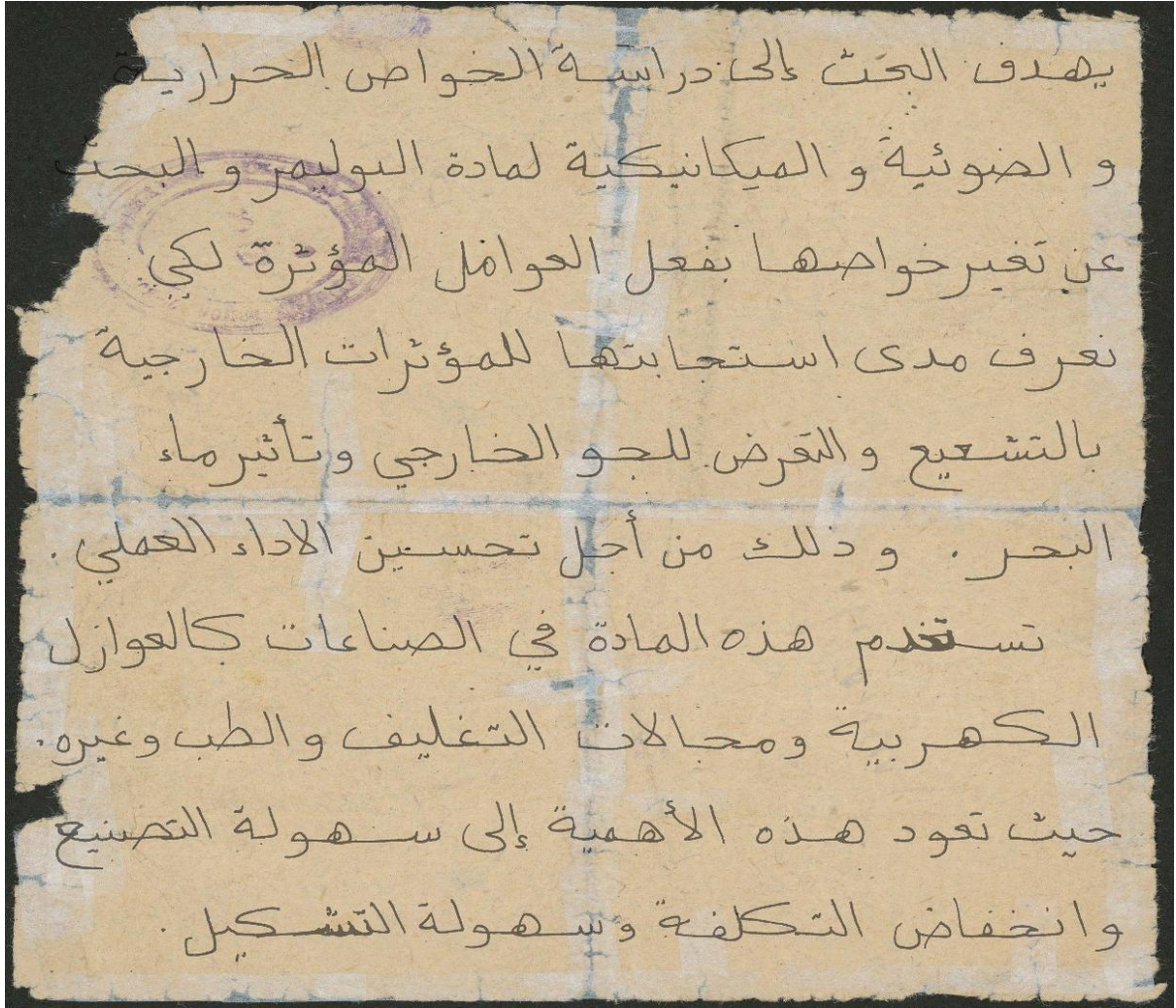


Figure 4. 9. Image de document synthétique obtenue par la combinaison de l'image de référence de la figure 4.7.b et de l'image du fond de la figure 4.8.b

En effet, la fusion en utilisant l'algorithme précédent est possible lorsque les deux images à fusionner ont la même taille ou lorsque l'image du masque est plus grande. Dans ce dernier cas, la fusion ne prend pas l'image du fond dans sa totalité mais une partie ayant la même taille que l'image du texte. Lorsque l'image du texte soit la plus grande, la fusion en utilisant l'algorithme précédent ne peut pas être effectuée correctement. La dernière partie de l'image du texte reste sans couverture par le masque. Pour surmonter ce problème, nous avons proposé d'étirer l'image du masque automatiquement pour qu'elle soit de la même taille que l'image du texte. De ce fait, aucun prétraitement manuel des images n'est nécessaire et aucune contrainte sur la taille des images du fond n'est imposée pour notre cas à la différence avec [STA 08] où les images du fond utilisées ont été redimensionnées manuellement pour avoir la même taille que les images du texte. La Figure 4.10 montre le résultat de la fusion par étirement du masque.

(a)
Image de
référence



(b)
Image du fond



(c)
Image
synthétique
obtenue



Figure 4. 10. Image de document synthétique obtenue de la fusion par étirement du masque

Notre ensemble d'images de documents historiques synthétiques est composé de 682 images générées à partir de 11 images de texte (de référence) et 62 images du fond.

4.2.3. Méthodes de binarisation incluses dans l'évaluation

Comme nous avons dit précédemment, notre but ultime est de trouver un algorithme de binarisation capable de donner des résultats satisfaisants pour les images de documents historiques dégradés, et ce en accomplissant une étude comparative large incluant un nombre important de méthodes de binarisation.

En fait, la présente étude comprend 51 méthodes de binarisation utilisant différentes techniques: 25 méthodes globales, 20 locales, et 6 hybrides. Certaines de ces méthodes ont été incluses dans des travaux d'évaluation précédents et ont montré leur efficacité. Certaines autres n'ont jamais été l'objet d'une étude comparative. Nous citons dans cette section les méthodes de binarisation testées. Les détails sur ces méthodes peuvent être trouvés dans le chapitre 2.

Les méthodes globales incluses sont :

- 1) Seuillage global fixe [GUP 07]
- 2) Seuillage percentile (méthode P-tile) [DOY 62]
- 3) Méthode de Khashman et Sekeroglu [KHA 07]

- 4) Méthode de Pun [PUN 80]
- 5) Méthode de Johannsen et Bille [JOH 82]
- 6) Méthode de Kapur et al. [KAP 85]
- 7) Méthode de seuillage par entropie croisée (Li et Lee) [LI 93]
- 8) Seuillage entropique flou (Shanhabag) [SHA 94]
- 9) Méthode de Yen et al. [YEN 95]
- 10) Méthode de Cheng et al. [CHE 98]
- 11) Première Méthode de Mello et Lins [MEL 08]
- 12) Méthode de Portes de Albuquerque et al. [ALB 04]
- 13) Méthode de Silva et al. [SIL 06]
- 14) Algorithme de seuillage basé sur l'entropie de Tsallis (deuxième méthode de Mello et al.) [MEL 08]
- 15) Seuillage par préservation des Moments [TSA 85]
- 16) Méthode de Huang et Wang [HUA 95]
- 17) Méthode de Fisher [FIS 58]
- 18) Méthode du Mode (Prewitt et Mendelsohn) [PRE 66]
- 19) Méthode du Triangle (Zack) [ZAC 77]
- 20) Méthode d'Otsu [OTS 79]
- 21) Méthode ISODATA [VEL 80]
- 22) Seuillage par erreur minimale (Kittler et Illingworth) [KIT 86]
- 23) Seuillage global itératif (IGT) [KAV 05]
- 24) Méthode de Boudra et al. [BOU 07]
- 25) Méthode de Xiao et al. [XIA 08]

Les méthodes locales comprises :

- 1) Méthode de Niblack [NIB 86]
- 2) Méthode de Sauvola et al. [SAU 00]
- 3) Méthode de Wolf et al. [WOL 02]
- 4) Méthode de Zhang et al. [ZHA 01]
- 5) Méthode de Bin Rais et al. [BIN 04]
- 6) Méthode de Feng et al. [FEN 04]
- 7) Seuillage par l'utilisation des moyennes mobiles [GON 07]
- 8) Méthode Nick [KHU 09]
- 9) Méthode de Singh et al. [SIN 11]
- 10) Méthode de Peng et al. [PEN 12]
- 11) Méthode de Laskov [LAS 12]
- 12) Méthode de Yasuda et al. [YAS 80]
- 13) Algorithme du seuil dynamique [WHI 83]
- 14) Méthode de Bernsen [BER 86]
- 15) Méthode de Giuliano et al. [GIU 77]
- 16) Technique du niveau logique [KAM 93]
- 17) Algorithme de fonction intégrée [WHI 83]
- 18) Technique de soustraction basée-masque [KAM 93]

- 19) Technique du moyen-gradient [LEE 03]
- 20) Segmentation hiérarchique floue [GAD 00]

Les méthodes hybrides sont :

- 1) Seuillage par soustraction du fond [LEE 03]
- 2) Méthode de Cavalcanti et al. [CAV 06]
- 3) Méthode de Tabatabaie et al. [TAB 10]
- 4) Méthode de Gangamma et al. [GAN 11]
- 5) Méthode IGT améliorée [KAV 06]
- 6) Binarisation en utilisant le minimum et le maximum locaux [SU 10]

4.2.4. Mesures d'évaluation utilisées

La binarisation peut introduire du bruit à l'avant plan et/ou à l'arrière plan des images de documents. Elle peut également provoquer des déformations altérant la forme des caractères, comme la coupure dans le trait des caractères, la fusion de plusieurs caractères, etc. Les pixels parasites et les déformations au niveau des caractères affectent considérablement les étapes suivantes dans le processus d'analyse et de reconnaissance de documents. Par conséquent, l'évaluation des algorithmes de binarisation doit prendre en considération à la fois le caractère bruyant de la segmentation et la déformation de la forme des caractères, spécialement dans les applications de traitement et d'analyse de documents.

Nous avons vu dans le chapitre 3 les différentes approches d'évaluation proposées dans la littérature : *subjective*, *objective*, *au niveau-système* (habituellement l'OCR), *analytique*, et *empirique*. Nous avons vu aussi que l'évaluation subjective est couteuse en temps et elle dépend du point de vue du visionneur. L'évaluation analytique examine les méthodes de binarisation indépendamment de leur résultat et elle se montre donc inappropriée pour l'estimation des performances réelles de ces méthodes. De même, l'évaluation en utilisant l'OCR ne peut pas être appliquée dans le cas des documents historiques à cause de leur mauvaise qualité qui entrave la reconnaissance.

Pour ces raisons, nous avons décidé de procéder une évaluation objective empirique des méthodes de binarisation en combinant plusieurs mesures quantitatives largement utilisées pour des besoins d'évaluation. Le choix d'employer plusieurs mesures est justifié par le fait que comme a noté [ROS 03], la quantification de la similarité entre deux images, qui semble d'être une tâche simple, est étonnamment délicate. Cependant une seule mesure peut, dans certains cas, donner d'estimations trompeuses, et l'utilisation de plusieurs mesures permet de surmonter ce problème. Nous avons ainsi utilisé un ensemble de mesures quantitatives de plusieurs points de vue pour mettre en évidence les différentes caractéristiques de performance des méthodes de binarisation étudiées.

Certaines des mesures utilisées nécessitent la disponibilité des images de vérité terrain représentant les images binaires idéales avec lesquelles les résultats doivent être comparés, tandis que les autres mesures ne les nécessitent pas. Ces mesures évaluent les méthodes de binarisation en estimant la qualité de la segmentation obtenue.

Nous citons dans cette section les mesures d'évaluation utilisées dans notre étude comparatives. Plus de détails sur ces mesures, peuvent être trouvés dans le chapitre 3.

Les mesures ne nécessitant pas d'images de références sont :

1. Mesure d'uniformité intra-région (*GU*)
2. Non-uniformité de régions (*NU*)
3. Mesure de contraste inter-région de niveaux de gris (*GC*)
4. Variance pondérée (*WV*)
5. Mesure de variance pondérée locale (*LWV*)
6. Mesure de variance uniforme locale (*LUV*)

Les mesures qui nécessitent d'images de références sont :

1. Pourcentage de classification correcte (*PCC*)
2. Erreur de fausse-classification (*ME*)
3. Coefficient de Jaccard (*JC*)
4. Coefficient de Yule (*YC*)
5. F-mesure (*FM*)
6. Précision de la moyenne géométrique (*GA*)
7. Erreur carrée (*MSE*)
8. Rapport signal à bruit (*SNR*)
9. Rapport maximum signal à bruit (*PSNR*)
10. Métrique de taux négatif (*NRM*)
11. Métrique de pénalité de fausse-classification (*MPM*)
12. Métrique de distorsion de distance réciproque (*DRD*)
13. Erreur de la zone de l'avant plan relative (*RAE*)
14. Distance de Hausdorff modifiée (*MHD*)
15. Texte cassé (*BT*)
16. Texte manquant (*MT*)
17. Fausses alarmes (*FA*)
18. Déformations (*D* et *MD*)

Le classement final des algorithmes de binarisation étudiés est effectué comme suit. D'abord, les performances de chaque algorithme, obtenues sur toutes les images sont moyennées. Ensuite, les méthodes de seuillage sont ordonnées de 1 à 51 suivant chaque mesure d'évaluation séparément, et la somme des rangs est calculée pour chaque méthode. La méthode la plus performante (m_{opt}) est celle dont la somme des rangs est minimale :

$$m_{opt} = \arg \min_{i=1..51} \left(\sum_{j=1..24} R_i^j \right)$$

Avec R_i^j le rang de la $i^{ème}$ méthode de binarisation suivant le $j^{ème}$ critère d'évaluation.

4.2.5. Expérimentations et résultats

Les méthodes de binarisation citées auparavant ont été appliquées sur l'ensemble d'images de test dont le but d'évaluer leurs performances. L'évaluation a été faite en mesurant la qualité des images binaires résultantes quantitativement par les différentes mesures et critères cités dans la section 4.2.4.

Avant d'évaluer les performances des différentes méthodes de binarisation, il est nécessaire de déterminer les valeurs les plus adéquates de leurs paramètres. Une première série d'expérimentations a été d'abord effectuée pour trouver ces valeurs. En effet, comme il est connu, les valeurs optimales des paramètres ne sont pas fixes et communes à toutes les images mais elles dépendent fortement des caractéristiques de l'image traitée [BEN 12]. Il n'y a ainsi aucune valeur de paramètre qui garantie le meilleur résultat de binarisation pour toutes les images. Les valeurs optimales d'un ensemble de paramètres sont appropriées à une image traitée et possible aux images similaires (pages du même livre par exemple), mais certainement non adéquates aux images différentes [BAD 09]. Cependant, nous avons choisi les valeurs de paramètres les plus adéquates à la majeure partie d'images de la base de test. Les valeurs des paramètres les plus appropriées sont résumées dans le tableau 4.2.

Tableau 4. 2. Valeurs optimales des paramètres des méthodes de binarisation incluses dans l'évaluation

Méthodes	Meilleures valeurs des paramètres
Portes de Albuquerque et al. [ALB 04]	$q=2$.
Seuillage global itératif [KAV 05]	$diffLim=0.004$
Boudra et al. [BOU 07]	$beta=5$
Xiao et al. [XIA 08]	$n=3$ et $eps=1$
Niblack [NIB 86]	$k=-0.2$ et $w=35$
Sauvola et al. [SAU 00]	$k=0.2$ et $w=35$
Wolf et al. [WOL 02]	$k=0.2$ et $w=15$
Zhang et al. [ZHA 01]	$k=-0.2$ et $w=35$
Bin Rais et al. [BIN 04]	$w=79$
Feng et al. [FEN 04]	$w1=15$ et $w2=45$
Seuillage par l'utilisation des moyennes mobiles [GON 07]	$n=15$ et $k=0.95$
Nike [KHU 09]	$w=19$ et $k=-0.1$
Singh et al. [SIN 11]	$w=35$ et $k=0.06$
Peng et al. [PEN 12]	$w=79$
Laskov [LAS 12]	$n=20$ et $k=0.2$
Yasuda et al. [YAS 80]	$w=3$
Algorithme du seuil dynamique [WHI 83]	$w=35$ et $b=1.1$
Bernsen [BER 86]	$w=31$ et $l=15$
Technique du niveau logique [KAM 93]	$w=31$ et $t0=20$
Algorithme de fonction intégrée [WHI 83]	$w=17$ et $t=40$

Technique de soustraction basée-masque [KAM 93]	$w=51$ et $w_{min}=9$ et $t=15$
Technique du moyen-gradient [LEE 03]	$w=11$ et $k=-0.06$ et $R=70$
Cavalcanti et al. [CAV 06]	$n=350$ et $d_{min}=40$
Méthode IGT améliorée [KAV 06]	$n=50$ et $k=3$
Binarisation en utilisant le minimum et le maximum locaux [SU 10]	$w=7$

Après avoir déterminé les valeurs optimales de paramètres, nous avons procédé à l'estimation des performances des différentes méthodes de binarisation examinées.

Les résultats d'évaluation moyens des algorithmes de binarisation (en considérant les valeurs optimales de paramètres) obtenus sur toutes les images sont présentés dans le tableau 4.3.

Le classement final des algorithmes de binarisation comparés est fourni dans le tableau 4.4, qui résume aussi les rangs partiels de chaque méthode selon chaque d'évaluation ainsi que la somme des rangs, la somme des 5 meilleurs rangs, et la somme des 20 meilleurs rangs. La figure 4.11 illustre graphiquement ces performances.

Tableau 4. 3. Résultats d'évaluation moyens des algorithmes de binarisation inclus dans l'étude comparative

Méthode	GU	NU	GC	WV	LWV	LUV	PCC	ME	JC	YC	FM	GA	MSE	SNR	PSNR	NRM	MPM	DRD	RAE	MHD	BT	MT	FA	D	MD
Seuillage global fixe	883.52	0.088	0.415	381.99	0.016	0.074	0.906	0.094	0.525	0.803	0.648	0.734	6138.34	13.29	13.71	0.204	42.80	9.53736	0.586	25.26	0.334	0.00607	0.026	0.026	0.139
Seuillage percentile	1349.46	0.636	0.116	673.54	0.029	0.088	0.590	0.410	0.182	0.181	0.297	0.741	26635.40	3.49	3.91	0.226	151.07	39.07817	0.182	1.28	0.003	0.00000	0.137	0.137	0.682
Méthode de Khashman et Sekeroglu	951.67	0.053	0.442	337.50	0.013	0.068	0.958	0.042	0.577	0.836	0.697	0.787	2756.20	14.27	14.69	0.168	5.64	2.88603	0.648	19.85	0.319	0.00285	0.040	0.040	0.093
Méthode de Pun	1016.81	0.146	0.490	782.12	0.022	0.086	0.860	0.140	0.325	0.776	0.433	0.507	9102.89	11.35	11.77	0.308	16.64	9.24605	0.350	2850.25	0.317	0.20177	0.029	0.029	0.144
Méthode de Johansen et Bille	953.40	0.115	0.445	464.32	0.015	0.071	0.905	0.095	0.577	0.830	0.679	0.748	6202.72	14.11	14.53	0.186	11.17	6.14525	0.624	21.04	0.247	0.04807	0.013	0.013	0.129
Méthode de Kapur et al.	1254.52	0.103	0.365	361.47	0.017	0.070	0.967	0.033	0.701	0.790	0.814	0.912	2171.47	15.29	15.71	0.079	4.65	2.01929	0.818	12.38	0.130	0.00531	0.069	0.069	0.128
Seuillage par entropie croisée	869.99	0.071	0.399	310.38	0.015	0.070	0.951	0.049	0.670	0.814	0.782	0.880	3154.91	14.86	15.28	0.111	20.42	5.36337	0.749	15.57	0.189	0.00040	0.046	0.046	0.122
Seuillage entropique flou	892.12	0.055	0.492	523.51	0.014	0.073	0.929	0.071	0.389	0.844	0.516	0.633	4646.88	12.29	12.71	0.271	14.34	5.77232	0.409	39.76	0.428	0.08652	0.054	0.054	0.055
Méthode Yen et al.	1325.40	0.120	0.350	376.96	0.018	0.071	0.961	0.039	0.676	0.763	0.795	0.911	2539.44	14.74	15.16	0.080	10.86	2.58554	0.787	11.39	0.123	0.00658	0.082	0.082	0.142
Méthode de Cheng et al.	1316.64	0.450	0.162	476.57	0.026	0.079	0.744	0.256	0.275	0.274	0.409	0.844	16654.49	5.83	6.25	0.144	154.68	25.54248	0.278	3.66	0.010	0.00000	0.211	0.211	0.513
1ère méthode de Mello et Lins	1086.61	0.110	0.450	596.08	0.016	0.075	0.888	0.112	0.382	0.779	0.527	0.633	7304.69	11.55	11.97	0.265	40.31	10.9206	0.408	2804.22	0.399	0.06416	0.040	0.040	0.132
Méthode de Portes de Albuquerque et al.	1325.40	0.120	0.350	376.96	0.018	0.071	0.961	0.039	0.676	0.763	0.795	0.911	2539.44	14.74	15.16	0.080	10.86	2.58554	0.787	11.39	0.123	0.00658	0.082	0.082	0.142
Méthode de Silva et al.	1177.72	0.318	0.290	414.27	0.020	0.079	0.759	0.241	0.502	0.568	0.618	0.740	15645.68	10.21	10.63	0.169	131.91	25.84213	0.574	13.34	0.093	0.00028	0.108	0.108	0.326
Algorithme de seuillage basé sur l'entropie de Tsallis	843.01	0.087	0.447	437.45	0.020	0.082	0.899	0.101	0.449	0.810	0.581	0.669	6552.96	12.73	13.15	0.243	20.16	7.39282	0.488	2807.70	0.341	0.07637	0.029	0.029	0.125
Seuillage par préservation des Moments	1175.75	0.124	0.322	292.87	0.017	0.071	0.949	0.051	0.629	0.690	0.755	0.923	3316.24	13.71	14.13	0.074	13.35	4.44538	0.704	12.25	0.102	0.00013	0.118	0.118	0.181
Méthode de Huang et Wang	1176.34	0.228	0.262	303.84	0.021	0.075	0.870	0.130	0.540	0.580	0.650	0.895	8467.03	11.47	11.89	0.098	98.58	15.95551	0.591	7.00	0.062	0.00004	0.102	0.102	0.311
Méthode de Fisher	1014.80	0.082	0.368	295.49	0.016	0.070	0.958	0.042	0.696	0.798	0.806	0.906	2745.72	15.30	15.72	0.088	4.73	3.03672	0.793	12.15	0.144	0.00020	0.040	0.040	0.147
Méthode du Mode	927.43	0.066	0.416	333.85	0.014	0.070	0.954	0.046	0.628	0.826	0.756	0.826	2998.72	14.91	15.33	0.137	4.42	3.20465	0.697	2792.35	0.215	0.03186	0.033	0.033	0.119
Méthode du Triangle	1355.69	0.295	0.254	366.12	0.021	0.074	0.849	0.151	0.502	0.513	0.631	0.882	9836.96	10.51	10.93	0.098	49.52	12.90432	0.541	4.66	0.034	0.00000	0.117	0.117	0.366

Chapitre 4 : Contributions à l'évaluation des techniques de binarisation

Méthode	GU	NU	GC	WV	LWV	LUV	PCC	ME	JC	YC	FM	GA	MSE	SNR	PSNR	NRM	MPM	DRD	RAE	MHD	BT	MT	FA	D	MD
Méthode d'Otsu	1014.80	0.082	0.368	295.49	0.016	0.070	0.958	0.042	0.696	0.798	0.806	0.906	2745.72	15.30	15.72	0.088	4.73	3.03672	0.793	12.15	0.144	0.00020	0.040	0.040	0.147
Méthode ISODATA	1010.35	0.080	0.370	296.17	0.016	0.070	0.958	0.042	0.697	0.802	0.807	0.905	2710.97	15.33	15.75	0.089	4.62	2.9978	0.794	12.38	0.147	0.00020	0.039	0.039	0.145
Seuillage par erreur minimale	1468.41	0.497	0.158	538.16	0.027	0.084	0.727	0.273	0.284	0.283	0.419	0.819	17737.71	5.91	6.33	0.152	143.06	27.26373	0.286	2.09	0.006	0.00000	0.198	0.198	0.518
Seuillage global itératif	1102.66	0.099	0.365	370.00	0.014	0.067	0.957	0.043	0.626	0.731	0.757	0.894	2804.32	13.85	14.27	0.099	15.88	3.55547	0.744	17.20	0.174	0.00021	0.111	0.111	0.135
Méthode de Booudra et al.	1070.83	0.303	0.384	405.85	0.022	0.081	0.750	0.250	0.433	0.678	0.549	0.664	16225.04	9.89	10.31	0.226	120.65	24.56378	0.467	22.50	0.169	0.02611	0.071	0.071	0.318
Méthode de Xiao et al.	1139.20	0.065	0.424	402.80	0.015	0.070	0.961	0.039	0.591	0.857	0.711	0.794	2540.10	14.58	15.00	0.162	3.78	2.27835	0.647	18.32	0.292	0.01862	0.034	0.034	0.081
Méthode de Niblack	2090.84	0.505	0.137	692.87	0.015	0.062	0.798	0.202	0.303	0.303	0.444	0.856	13151.41	6.73	7.15	0.140	163.63	21.0495	0.335	14.16	0.066	0.00005	0.603	0.603	0.085
Méthode de Sauvola et al.	1216.85	0.081	0.352	347.39	0.014	0.066	0.967	0.033	0.674	0.813	0.792	0.885	2173.98	15.07	15.49	0.103	7.59	2.29133	0.777	15.98	0.186	0.00032	0.107	0.107	0.062
Méthode de Wolf et al.	1229.50	0.076	0.362	385.28	0.018	0.072	0.965	0.035	0.660	0.818	0.786	0.871	2261.60	14.84	15.26	0.119	6.99	2.47231	0.771	22.11	0.222	0.00015	0.126	0.126	0.035
Méthode de Zhang et al.	1248.81	0.088	0.344	346.82	0.014	0.066	0.966	0.034	0.675	0.801	0.792	0.892	2235.41	15.00	15.42	0.097	8.35	2.40854	0.772	14.95	0.171	0.00026	0.113	0.113	0.070
Méthode de Bin Rais et al.	1864.83	0.435	0.164	571.63	0.019	0.068	0.824	0.176	0.347	0.348	0.492	0.888	11437.55	7.48	7.90	0.108	139.13	18.50232	0.362	3.98	0.027	0.00003	0.398	0.398	0.251
Méthode de Feng et al.	1542.76	0.144	0.295	407.58	0.019	0.071	0.950	0.050	0.628	0.705	0.753	0.905	3276.36	13.71	14.13	0.091	23.55	4.64211	0.763	18.28	0.141	0.00002	0.215	0.215	0.065
Seuillage par l'utilisation des moyennes mobiles	2140.59	0.300	0.193	612.29	0.024	0.080	0.876	0.124	0.409	0.443	0.559	0.840	8089.45	9.64	10.06	0.157	50.52	11.0944	0.577	76.90	0.197	0.00000	0.416	0.416	0.118
Méthode Nick	1512.86	0.131	0.306	403.90	0.017	0.070	0.951	0.049	0.619	0.717	0.747	0.888	3205.09	13.72	14.14	0.106	19.51	4.23232	0.763	24.10	0.175	0.00002	0.224	0.224	0.041
Méthode de Singh et al.	2148.19	0.276	0.241	475.19	0.014	0.065	0.934	0.066	0.520	0.561	0.667	0.897	4306.11	11.66	12.08	0.099	39.92	6.1017	0.650	31.00	0.139	0.00007	0.364	0.364	0.062
Méthode de Peng et al.	1867.78	0.459	0.156	591.23	0.019	0.068	0.811	0.189	0.331	0.331	0.473	0.878	12281.37	7.15	7.57	0.117	359.34	19.88697	0.347	5.05	0.032	0.00009	0.391	0.391	0.273
Méthode de Laskov	1280.53	0.079	0.357	382.09	0.016	0.070	0.965	0.035	0.657	0.822	0.780	0.865	2285.96	14.82	15.24	0.118	6.23	2.2736	0.773	19.43	0.219	0.00047	0.095	0.095	0.061
Méthode de Yasuda et al.	2145.10	0.186	0.250	622.90	0.023	0.082	0.940	0.060	0.563	0.680	0.704	0.854	3893.25	13.01	13.43	0.133	11.31	4.28496	0.720	27.53	0.225	0.00020	0.136	0.136	0.158
Algorithme du seuil dynamique	1512.02	0.155	0.286	382.82	0.015	0.065	0.946	0.054	0.621	0.678	0.747	0.917	3489.89	13.47	13.89	0.080	27.57	4.91496	0.723	16.57	0.112	0.00007	0.228	0.228	0.082
Méthode de Bernsen	2153.16	0.478	0.149	657.00	0.015	0.065	0.804	0.196	0.300	0.306	0.431	0.803	12753.97	7.14	7.56	0.193	163.72	20.84377	0.405	97.14	0.191	0.00029	0.636	0.636	0.037
Méthode de Giuliano et al.	1233.52	0.054	0.394	455.43	0.019	0.079	0.959	0.041	0.573	0.864	0.713	0.784	2660.33	14.07	14.49	0.180	3.25	2.37808	0.637	25.26	0.349	0.00183	0.061	0.061	0.040

Chapitre 4 : Contributions à l'évaluation des techniques de binarisation

Méthode	GU	NU	GC	WV	LWV	LUV	PCC	ME	JC	YC	FM	GA	MSE	SNR	PSNR	NRM	MPM	DRD	RAE	MHD	BT	MT	FA	D	MD
Technique du niveau logique	1406.31	0.090	0.348	412.23	0.012	0.064	0.967	0.033	0.673	0.815	0.795	0.880	2177.62	14.97	15.39	0.111	14.60	2.35772	0.821	19.47	0.203	0.00203	0.114	0.114	0.051
Algorithme de fonction intégrée	2063.43	0.637	0.067	958.16	0.029	0.090	0.604	0.396	0.154	0.122	0.260	0.672	25740.92	3.79	4.21	0.320	221.90	37.15169	0.211	129.97	0.228	0.00000	0.222	0.222	0.615
Technique de soustraction basée-masque	1652.98	0.072	0.353	589.76	0.016	0.075	0.943	0.057	0.438	0.751	0.579	0.690	3695.64	12.64	13.06	0.245	13.18	3.58055	0.613	33.51	0.470	0.00351	0.141	0.141	0.064
Technique du moyen-gradient	1971.43	0.164	0.285	499.11	0.022	0.078	0.950	0.050	0.600	0.709	0.735	0.870	3237.71	13.55	13.97	0.118	11.55	3.6074	0.768	130.90	0.196	0.00344	0.197	0.197	0.073
Segmentation hiérarchique floue	1454.72	0.732	0.097	873.29	0.030	0.089	0.506	0.494	0.165	0.163	0.276	0.668	32140.06	2.65	3.11	0.277	135.61	38.82209	0.166	3.13	0.007	0.00000	0.109	0.109	0.726
Seuillage par soustraction du fond	1441.97	0.069	0.391	639.02	0.018	0.076	0.939	0.061	0.516	0.824	0.662	0.745	3994.35	13.22	13.64	0.222	17.32	5.57858	0.591	51.45	0.423	0.00004	0.108	0.108	0.023
Méthode de Cavalcanti et al.	1136.65	0.060	0.511	581.83	0.019	0.075	0.938	0.062	0.399	0.870	0.558	0.567	4056.84	13.19	13.61	0.278	0.50	3.40741	0.450	11146.34	0.367	0.18195	0.012	0.012	0.073
Méthode de Tabatabaie et al.	1016.91	0.085	0.373	280.52	0.015	0.068	0.962	0.038	0.712	0.806	0.815	0.917	2474.94	15.63	16.05	0.080	10.83	3.43904	0.797	12.36	0.130	0.00017	0.062	0.062	0.120
Méthode de Gangamma et al.	1760.53	0.124	0.339	407.32	0.013	0.070	0.944	0.056	0.503	0.678	0.645	0.789	3631.07	12.71	13.13	0.185	30.30	4.90975	0.693	29.58	0.335	0.00430	0.237	0.237	0.055
Méthode IGT améliorée	1150.05	0.086	0.381	432.56	0.012	0.066	0.960	0.040	0.628	0.775	0.761	0.870	2624.13	14.17	14.59	0.120	13.72	3.01678	0.750	20.62	0.223	0.00062	0.109	0.109	0.088
Binarisation en utilisant le minimum et le maximum locaux	1318.06	0.115	0.372	402.63	0.018	0.075	0.951	0.049	0.687	0.807	0.800	0.885	3170.19	15.10	15.52	0.109	8.45	3.72932	0.822	18.89	0.182	0.00056	0.097	0.097	0.076

Tableau 4. 4. Classement final et partiel des algorithmes de binarisation inclus dans l'étude comparative

Rang	Méthode	GU	NU	GC	WV	LWV	LUV	PCC	ME	JC	YC	FM	GA	MSE	SNR	PSNR	NRM	MPM	DRD	RAE	MHD	BT	MT	FA	D	MD	somme	somme 5	somme 20
1	Méthode de Tabatabaie et al.	12	16	16	1	13	12	7	7	1	16	1	3	7	1	1	3	15	18	48	14	16	19	16	16	25	304	5	178
2	Méthode ISODATA	8	12	18	5	19	21	13	13	3	18	3	10	13	2	2	9	5	12	47	15	21	21	8	8	36	342	15	196
3	Méthode de Fisher	9	14	19	3	20	19	14	14	4	20	4	7	14	3	3	7	7	14	45	11	19	21	11	11	37	350	17	207
4	Méthode d'Otsu	9	14	19	3	20	19	14	14	5	20	4	8	14	3	3	7	7	14	45	11	19	21	11	11	37	352	18	209
5	Méthode de Kapur et al.	26	23	22	12	26	23	1	1	2	22	2	4	1	5	5	2	6	1	49	16	15	39	17	17	28	365	6	197

Chapitre 4 : Contributions à l'évaluation des techniques de binarisation

Rang	Méthode	GU	NU	GC	WV	LWV	LUV	PCC	ME	JC	YC	FM	GA	MSE	SNR	PSNR	NRM	MPM	DRD	RAE	MHD	BT	MT	FA	D	MD	somme	somme 5	somme 20
6	Méthode de Sauvola et al.	22	13	26	11	6	8	2	2	10	13	11	17	2	7	7	16	12	4	42	21	27	28	24	24	11	366	16	219
7	Méthode de Zhang et al.	25	19	30	10	8	6	4	4	9	19	10	14	4	8	8	11	13	7	40	19	23	25	30	30	14	390	25	235
8	Méthode de Xiao et al.	17	5	8	22	15	14	10	10	23	3	24	34	10	16	16	33	3	3	25	25	40	43	7	7	18	431	21	256
9	Seuillage par entropie croisée	2	8	11	7	14	15	19	19	12	12	13	21	19	11	11	21	31	29	34	20	28	29	13	13	26	438	39	287
10	Méthode du Mode	5	6	9	8	10	17	18	18	17	7	17	31	18	10	10	28	4	16	29	47	33	45	6	6	24	439	27	254
11	Méthode de Khashman et Sekeroglu	6	1	7	9	4	13	16	16	24	5	26	36	16	17	17	34	9	11	26	29	42	35	10	10	22	441	23	265
12	Méthode de Laskov	27	11	24	18	23	22	6	6	14	9	14	26	6	13	13	24	10	2	41	27	34	30	21	21	9	451	29	292
13	Technique du niveau logique	34	21	29	27	1	2	3	3	11	11	7	20	3	9	9	20	25	5	50	28	32	34	31	31	6	452	12	271
14	Méthode Yen et al.	30	27	27	15	29	25	8	8	7	26	8	5	8	14	14	5	16	9	43	9	13	41	19	19	33	458	33	282
15	Méthode de Portes de Albuquerque et al.	30	27	27	15	29	25	8	8	8	26	8	6	8	14	14	5	16	9	43	9	13	41	19	19	33	460	35	284
16	Méthode de Wolf et al.	23	10	23	20	28	29	5	5	13	10	12	23	5	12	12	25	11	8	39	32	35	18	34	34	2	468	25	294
17	Seuillage global itératif	15	22	21	14	5	9	17	17	19	29	16	13	17	21	21	15	26	19	33	23	24	24	29	29	31	509	56	358
	Méthode IGT améliorée	18	17	15	29	2	7	11	11	18	25	15	24	11	18	18	26	23	13	35	30	36	32	27	27	21	509	42	347
19	Binarisation en utilisant le minimum et le maximum locaux	29	25	17	21	31	35	20	20	6	15	6	18	20	6	6	19	14	22	51	26	26	31	22	22	17	525	38	348
20	Méthode de Giuliano et al.	24	2	12	31	37	40	12	12	26	2	23	37	12	20	20	36	2	6	24	36	46	33	15	15	4	527	16	331
21	Seuillage par préservation des Moments	19	29	32	2	27	24	24	24	15	33	18	1	24	24	24	1	22	25	30	13	11	17	33	33	40	545	28	374
22	Algorithme du seuil dynamique	38	34	36	19	16	3	25	25	20	36	21	2	25	26	26	4	33	28	32	22	12	15	44	44	19	605	36	407
23	Méthode de Johansen et Bille	7	26	6	32	11	28	34	34	25	6	27	38	34	19	19	38	18	33	23	31	39	46	2	2	29	607	23	412
24	Méthode Nick	39	31	33	23	25	18	21	21	21	30	20	15	21	22	22	17	29	23	36	34	25	8	43	43	5	625	63	430
25	Méthode de Feng et al.	40	32	34	26	34	27	23	23	16	32	19	9	23	23	23	10	32	26	37	24	18	8	41	41	13	634	56	441
26	Méthode de Cavalcanti et al.	16	4	1	39	33	34	30	30	39	1	38	50	30	29	29	49	1	17	13	51	47	50	1	1	16	649	5	402
27	Seuillage global fixe	3	20	10	17	18	32	33	33	29	17	31	42	33	27	27	40	37	36	19	35	43	40	3	3	32	660	36	458
28	Seuillage entropique flou	4	3	2	36	7	30	32	32	40	4	41	48	32	34	34	47	24	31	12	41	50	49	14	14	8	669	20	434

Chapitre 4 : Contributions à l'évaluation des techniques de binarisation

Rang	Méthode	GU	NU	GC	WV	LWV	LUV	PCC	ME	JC	YC	FM	GA	MSE	SNR	PSNR	NRM	MPM	DRD	RAE	MHD	BT	MT	FA	D	MD	somme	somme 5	somme 20
29	Seuillage par soustraction du fond	35	7	13	45	32	38	29	29	31	8	29	39	29	28	28	41	28	30	20	42	49	12	26	26	1	695	41	479
30	Méthode de Huang et Wang	20	37	38	6	41	36	38	38	28	38	30	12	38	37	37	13	40	40	21	8	8	11	23	23	43	704	45	502
31	Méthode de Singh et al.	50	38	41	33	9	5	31	31	30	40	28	11	31	35	35	14	35	32	27	39	17	14	46	46	10	728	49	505
32	Algorithme de seuillage basé sur l'entropie de Tsallis	1	18	5	30	38	46	35	35	35	14	35	45	35	31	31	44	30	34	15	49	45	48	4	4	27	734	28	501
33	Méthode du Triangle	33	39	39	13	40	31	40	40	34	41	33	19	40	39	39	12	38	39	16	6	7	1	32	32	46	749	39	542
34	Technique du moyen-gradient	45	35	37	35	42	39	22	22	22	31	22	25	22	25	25	23	20	21	38	46	30	36	38	38	15	754	100	544
35	Méthode de Gangamma et al.	42	30	31	25	3	16	26	26	32	35	32	35	26	32	32	37	34	27	28	38	44	38	45	45	7	766	77	552
36	Technique de soustraction basée-masque	41	9	25	40	22	33	27	27	36	28	36	43	27	33	33	45	21	20	22	40	51	37	37	37	12	782	84	562
37	1 ^{ère} méthode de Mello et Lins	14	24	4	42	24	37	36	36	41	23	40	49	36	36	36	46	36	37	11	48	48	47	9	9	30	799	47	561
38	Méthode de Yasuda et al.	49	36	40	44	45	45	28	28	27	34	25	28	28	30	30	27	19	24	31	37	37	20	35	35	39	821	115	598
39	Méthode de Bin Rais et al.	43	43	43	38	36	10	41	41	42	43	42	16	41	43	43	18	44	41	9	5	5	10	48	48	41	834	39	608
40	Méthode de Pun	11	33	3	49	43	48	39	39	44	24	45	51	39	38	38	50	27	35	8	50	41	51	5	5	35	851	32	600
41	Méthode de Silva et al.	21	42	35	28	39	41	45	45	33	39	34	41	45	40	40	35	42	47	17	17	10	26	25	25	45	857	90	630
42	Méthode de Peng et al.	44	45	46	41	35	11	42	42	43	44	43	22	42	44	44	22	51	42	7	7	6	16	47	47	42	875	47	639
43	Méthode de Boudra et al.	13	41	14	24	44	44	46	46	37	37	39	47	46	41	41	42	41	45	14	33	22	44	18	18	44	881	77	651
44	Méthode de Niblack	47	48	48	48	17	1	44	44	45	46	44	27	44	46	46	29	48	44	6	18	9	13	50	50	20	882	46	638
45	Méthode de Cheng et al.	28	44	44	34	47	42	47	47	48	48	48	29	47	48	48	30	47	46	4	4	4	1	40	40	47	912	41	672
46	Seuillage par l'utilisation des moyennes mobiles	48	40	42	43	46	43	37	37	38	42	37	30	37	42	42	32	39	38	18	43	31	1	49	49	23	927	103	692
47	Seuillage par erreur minimale	37	47	45	37	48	47	48	48	47	47	47	32	48	47	47	31	45	48	5	2	2	1	39	39	48	932	41	692
48	Méthode de Bernsen	51	46	47	46	12	4	43	43	46	45	46	33	43	45	45	39	49	43	10	44	29	27	51	51	3	941	56	692
49	Seuillage percentile	32	49	49	47	50	49	50	50	49	49	49	40	50	50	50	43	46	51	2	1	1	1	36	36	50	980	37	729
50	Segmentation hiérarchique floue	36	51	50	50	51	50	51	51	50	50	50	46	51	51	51	48	43	50	1	3	3	1	28	28	51	995	36	740
51	Algorithme de fonction intégrée	46	50	51	51	49	51	49	49	51	51	51	44	49	49	49	51	50	49	3	45	38	7	42	42	49	1116	132	861

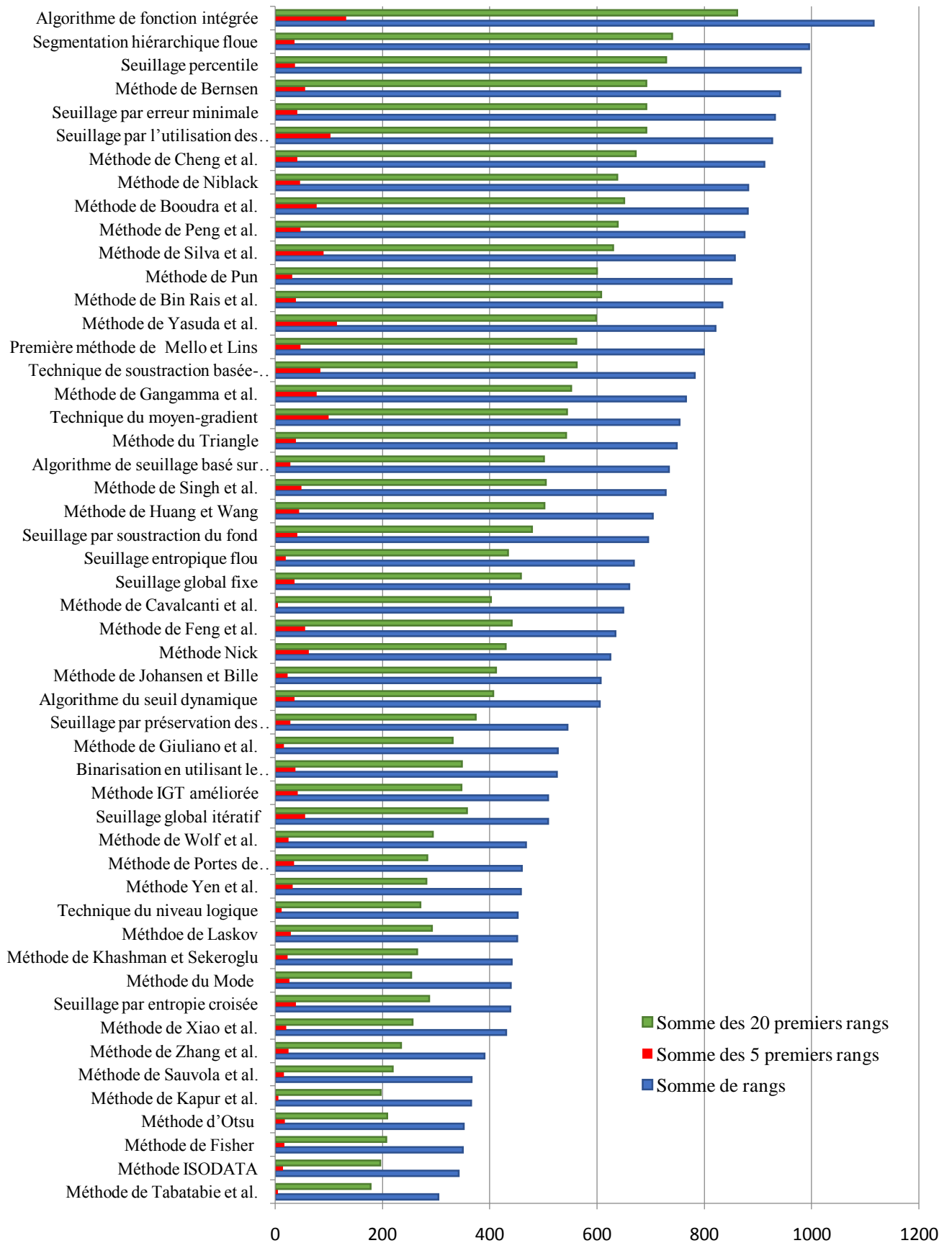


Figure 4. 11. Performances des méthodes de binarisation incluses dans l'étude comparative

En analysant les résultats obtenus pour chaque image et les performances moyennes résumées dans les tableaux 4.3 et 4.4 précédents ainsi que dans le graphe de la figure 4.11, nous pouvons tirer les remarques suivantes :

- Evidemment, aucun algorithme de binarisation ne paraît le plus performant et le plus adéquat pour toutes les images de la base de test. Un algorithme peut être performant et fournir de bons résultats en qualité de binarisation pour une image ou un ensemble d'images (usuellement ayant des caractéristiques similaires), mais se montre inefficace pour d'autres images. La figure 4.12 présente un exemple de binarisation des images des figures 4.2.a, 4.3.a, 4.4.a, 4.5.a, 4.6.a par la méthode de Khashman et Sekeroglu [KHA 07]. Comme nous pouvons remarquer visuellement, cette méthode a réussi à binariser les images des figures 4.2.a, 4.3.a, et 4.6.a, et a fourni des résultats proches des images de vérité terrain. Par contre les résultats de binarisation des images des figures 4.4.a et 4.5.a en utilisant le même algorithme sont loin d'être acceptables. Le tableau 4.5 expose les performances de la méthode de Khashman et Sekeroglu [KHA 07] pour les 5 images.

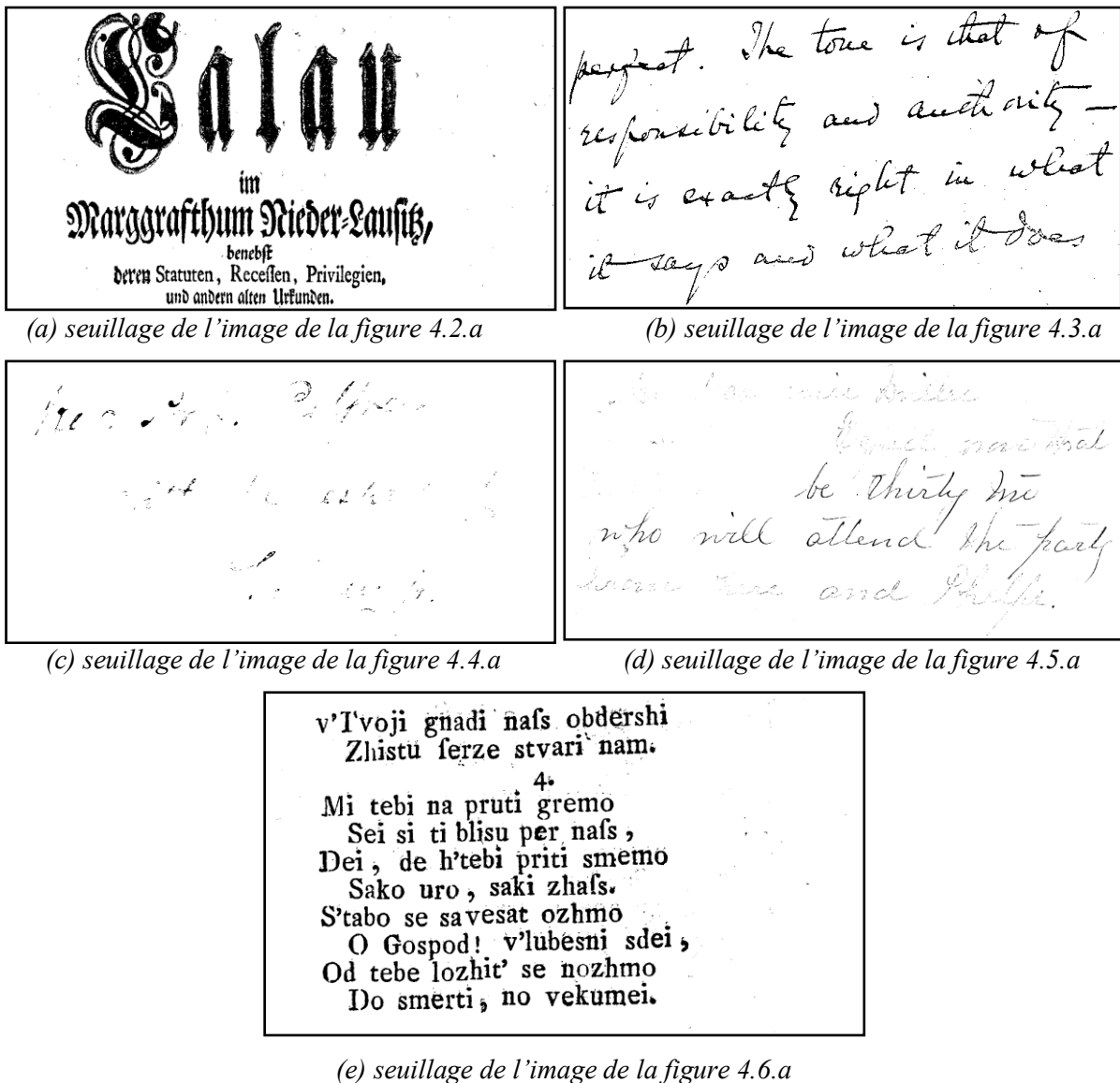


Figure 4. 12. Résultats de binarisation de différentes images par la méthode de Khashman et Sekeroglu [KHA 07]

Tableau 4. 5. Performances de la méthode de Khashman et Sekeroglu [KHA 07] pour différentes images

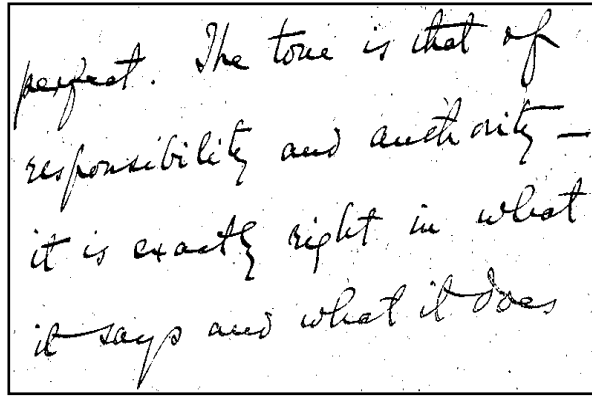
<i>Image</i> <i>Mesure</i>	Figure 4.2.a	Figure 4.3.a	Figure 4.4.a	Figure 4.5.a	Figure 4.6.a
GU	688.24	473.68	127.44	321.15	680.67
NU	0.00952	0.05253	0.00002	0.01053	0.02069
GC	0.5924	0.2715	0.6000	0.3406	0.4228
WV	465.67	75.32	122.89	142.78	326.57
LWV	0.0214	0.0025	0.0128	0.0042	0.0071
LUV	0.1063	0.0358	0.0782	0.0512	0.0638
PCC	0.9654	0.9739	0.9531	0.9644	0.9885
ME	0.0346	0.0261	0.0469	0.0356	0.0115
JC	0.7979	0.6326	0.0130	0.2654	0.8732
YC	0.9581	0.9673	0.9531	0.9546	0.9690
FM	0.8876	0.7749	0.0256	0.4195	0.9323
GA	0.8939	0.7967	0.1139	0.5158	0.9421
MSE	2248.88	1698.78	3049.43	2315.89	749.67
SNR	13.80	15.51	13.08	14.27	18.98
PSNR	14.61	15.83	13.29	14.48	19.38
NRM	0.1005	0.1827	0.4935	0.3670	0.0563
MPM	0.0694	0.1900	0.4908	0.4029	0.0994
DRD	1837.52	1215.70	3850.16	2646.35	534.40
RAE	0.8011	0.6384	0.0130	0.2686	0.9076
MHD	15.05	17.94	41.64	34.93	5.31
BT	0.2006	0.3652	0.9511	0.7323	0.1108
MT	0.0000	0.0000	0.0360	0.0016	0.0000
FA	0.00176	0.00180	0.00000	0.00795	0.00777
D	0.00176	0.00180	0.00000	0.00795	0.00777
MD	4.6E-04	3.9E-03	0.0E+00	1.4E-03	1.2E-02

- Aucun algorithme de binarisation n'a obtenu les meilleures performances moyennes pour toutes les mesures d'évaluation. Un algorithme de binarisation peut être le plus performant suivant une mesure, mais le pire suivant une ou plusieurs autres mesures. L'algorithme de *segmentation hiérarchique floue* [GAD 00] est un exemple. Cet algorithme est classé en première position (rang 1) avec les mesures *RAE* et *MT*, et en même temps il a eu des mauvaises performances avec d'autres mesures à savoir *NU*, *LWV*, *PCC*, *ME*, *MSE*, *SNR*, *PSNR* et *MD*.

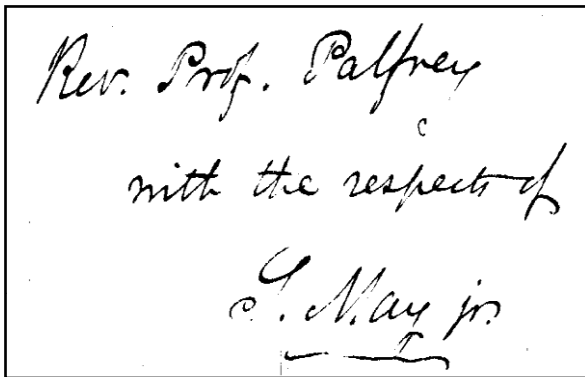
- La méthode de binarisation classée en première place est celle de Tabatabaie et al. [TAB 10], qui est une méthode hybride non paramétrique basée sur la morphologie. Des exemples de binarisation en utilisant cette méthode sont illustrés par la figure 4.13. La méthode de Tabatabaie et al. s'est montrée la plus performante et elle a eu la somme des rangs minimale, malgré qu'elle n'a obtenu la meilleure performance que pour cinq mesures : *WV*, *JC*, *FM*, *SNR*, *PSNR*.
- Étonnamment, les quatre méthodes mieux classées qui suivent (classées de 2 à 5) sont des méthodes globales simples, par opposition à plusieurs études comparatives antérieures où les méthodes locales souvent réussissent.
- Néanmoins, entre ces quatre méthodes, trois n'ont eu des plus fortes performances pour aucune mesure d'évaluation, et malgré ça elles ont été placées en premières positions avec des sommes de rangs minimales. Ces méthodes ont été arrangées même avant d'autres techniques obtenant les plus fortes performances suivant plusieurs mesures d'évaluation.
- Parmi les méthodes globales, les méthodes basées sur le groupage (clustering), notamment la méthode ISODATA [VEL 80] et la méthode d'Otsu [OTS 79] ont été les plus performantes suivies par les méthodes entropiques.
- La méthode de Kapur et al. [KAP 85] a paru la méthode entropique la plus puissante et elle a pris la cinquième position.
- Comme dans plusieurs travaux d'évaluation antérieurs, la méthode de Sauvola et al. [SAU 00] a montré son efficacité dans notre étude comparative et elle a surmonté les autres méthodes locales. Elle a paru ainsi la méthode locale la plus puissante (classée en sixième place), en considérant les valeurs des paramètres les plus adéquates.
- Prodigieusement, les méthodes locales de binarisation ont présenté des faibles performances. Ceci peut être justifié par le fait que les méthodes locales sont généralement sensibles aux valeurs de paramètres (comme la taille de fenêtre glissante, etc.) qui nécessitent d'être fixées attentivement et qui ne sont pas identiques pour toutes les images de la base de test.
- Le seuillage global fixe, qui est l'algorithme de binarisation le plus simple, a été rangé le 27^{ème}, avant plusieurs autres méthodes récentes et complexes, ce qui a été inattendu aussi.
- Certaines méthodes suggérées par des travaux d'évaluation précédents dans la littérature n'ont pas montré leur efficacité dans la présente étude. Nous pouvons citer par exemple la méthode de Giuliano [GIU 77], préservation des moments [TSA 85], Niblack [NIB 86], Nick [KHU 09], etc.
- Deux méthodes ont eu les meilleures performances pour le plus grand nombre de mesures simultanément ; la méthode de Tabatabaie [TAB 10] et celle de Cavalcanti et al. [CAV 06]. Quoique les deux méthodes ont été les meilleures suivant cinq mesures d'évaluation en même temps, elles ont été arrangées dans deux positions fortement éloignées (1 pour la méthode de Tabatabaie et al., et 26 pour la méthode de Cavalcanti et al).



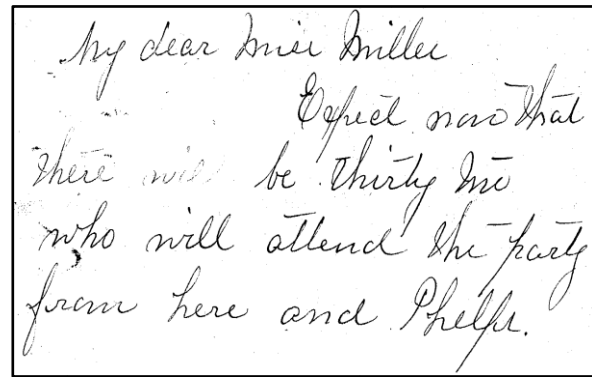
(a) Seuillage de l'image de la figure 4.2.a



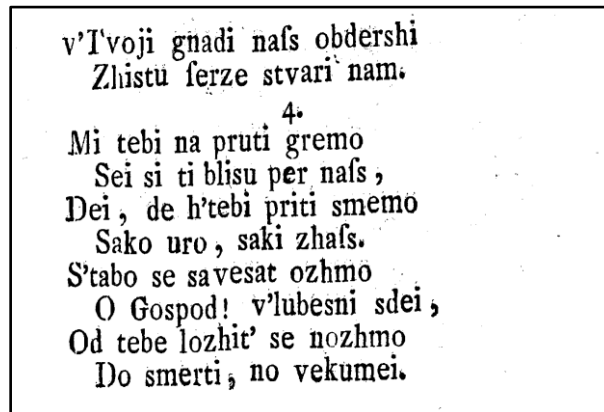
(b) Seuillage de l'image de la figure 4.3.a



(c) Seuillage de l'image de la figure 4.4.a



(d) Seuillage de l'image de la figure 4.5.a



(e) Seuillage de l'image de la figure 4.6.a

Figure 4. 13. Résultats de binarisation de différentes images en utilisant la méthode de Tabatabai et al. [TAB 10]

- Certaines méthodes comme la *segmentation hiérarchique floue* [GAD 00] et *l'algorithme de fonction intégrée* [WHI 83] ont produit des résultats trop bruités et loin d'être satisfaisants, ce qui indique que ces méthodes sont inadéquates pour la binarisation d'images de documents dégradés.
- La dernière méthode classée est *l'algorithme de fonction intégrée* de White et Rohrer [WHI 83]. C'est une méthode locale qui utilise des informations de voisinage.

- Ce qui a attiré également notre attention, c'est le fait que certaines méthodes conçues spécifiquement pour les images de documents historiques (IGT [KAV 05], IGT améliorée [KAV 06], Nick [KHU 09], la méthode de binarisation en utilisant le maximum et le minimum locaux [SU 10], etc.) ont été classées après des méthodes à usage général.
- Il est évident que les algorithmes récents ne sont pas toujours plus performants que les anciens. Nous pouvons remarquer que l'algorithme simple et très ancien de Fisher [FIS 58] a été le troisième meilleur.

4.2.6. Conclusion

Dans cette première partie du chapitre, nous avons présenté notre première contribution à l'évaluation des techniques de binarisation pour les images de documents historiques, le type de documents le plus difficile à traiter automatiquement à cause de leur mauvaise état. Dans cette contribution, nous avons été intéressés à la comparaison et à l'estimation des performances d'un nombre important d'algorithmes de seuillage de différentes catégories (globales, locales, et hybrides) dans le but de déterminer l'algorithme le plus adéquat aux documents historiques. Bien que plusieurs travaux d'évaluation des techniques de binarisation ont été mis au point dans la littérature, la diversité des objectifs et la variété des résultats de ces travaux ont constitué une de nos principales motivations.

Cinquante-et-un méthodes de binarisation de différentes catégories ont été sélectionnées et testées sur une large base d'images constituée de deux ensembles : le premier est composé des collections proposées dans les compétitions spécialisées de binarisation DIBCO/ H-DIBCO et le deuxième a été créé par nous d'une manière automatique. La littérature révèle diverses mesures d'évaluation proposées par différents chercheurs pour différents buts et applications. Nous avons choisi 25 mesures quantitatives pour avoir différents aspects d'évaluation. Les résultats obtenus confirment le fait qu'aucun algorithme de seuillage ne peut surpasser tous les autres pour toutes les images. Ceci peut être justifié par le fait que les dégradations (ou le bruit en général) que souffrent les documents historiques sont arbitraires et ne suivent aucun modèle statistique connu (voir [BAI 93]). En considérant les performances globales moyennes, la méthode qui a eu le meilleur score global et qui a paru la plus puissante est celle de Tabatabaie et al. [TAB 10]. Le présent travail d'évaluation est le premier qui regroupe ce grand nombre de méthodes de binarisation et en même temps ce nombre considérable de métriques.

4.3. Une nouvelle méthode d'évaluation des techniques de binarisation pour la recherche dans les images de documents arabes dégradés

4.3.1. Motivations et objectifs

La binarisation d'images de documents a attiré une attention considérable des chercheurs depuis longtemps en raison de son importance et de son influence sur la sortie finale des systèmes d'analyse et de reconnaissance de documents. Pour un système de reconnaissance de l'écriture par exemple, si une erreur est introduite lors de l'étape de binarisation, nous observons une détérioration dans les performances de l'OCR (Figure 4.1).

Un nombre important d'algorithmes de binarisation employant diverses techniques ont été proposés dans la littérature dont le but de s'approcher au maximum de la binarisation idéale. Dans le cas des images de documents dégradés, pour qu'une méthode de binarisation arrive à

atteindre ce but, elle doit conserver la forme des caractères et en plus elle doit lutter contre le bruit et les dégradations subis par ces documents. Avec le déploiement de ce nombre immense d'algorithmes de binarisation, un nouveau problème a été posé ; celui de savoir quelle image binarisée est la meilleure. *L'évaluation des performances des méthodes de binarisation et l'estimation de la qualité des images binarisées* a été ainsi apparu comme champ actif de recherche dans le domaine de l'analyse et la reconnaissance de documents ayant comme but de répondre à la question précédente. A cause de la diversité d'objectifs pour lesquels l'évaluation est procédée, différentes approches ont été proposées pour l'estimation de la qualité des images binarisées, et une multitude de mesures ont été utilisées pour la quantification de cette qualité. Les différentes approches d'estimation, les diverses mesures d'évaluation, leurs avantages et limites ont été abordées en détails dans le chapitre 3.

Cependant, une grande partie des travaux d'évaluation antérieurs ont été basées sur la comparaison avec des images de vérité terrain. Cette évaluation est une évaluation empirique supervisée qui détermine pour chaque pixel s'il est bien classé ou non. Deux causes compliquent l'évaluation supervisée au niveau des pixels ; la première cause est la difficulté de l'obtention des images de référence, qui se fait souvent manuellement ou semi-automatiquement; la deuxième est que ce type d'évaluation ne s'accorde pas toujours avec l'évaluation subjective humaine. En effet, comme nous avons déjà mentionné, la binarisation n'est pas une finalité en elle-même ; elle constitue une étape dans une séquence de traitements (système) ayant un objectif bien déterminé. Plusieurs travaux d'évaluation essayent d'examiner les performances des méthodes de binarisation en fonction de leur impact sur le système entier, ce qui classe les algorithmes de binarisation selon leur capacité d'améliorer le résultat final du système. C'est une évaluation au niveau-système dirigée par le but ultime de l'application. L'exemple de l'évaluation au niveau-système la plus répandue est l'OCR.

Dans le présent travail [SAR 15], nous considérons la binarisation comme une étape de prétraitement dans le processus de recherche dans les images documents arabes dégradés. Nous nous intéressons à choisir la meilleure méthode de binarisation pour l'employer dans notre système de recherche de documents arabes dégradés. Comme il n'y a pas jusqu'à présent un logiciel efficace de reconnaissance de l'écriture arabe manuscrite pour pouvoir l'utiliser dans l'évaluation, et à cause des limites de l'évaluation supervisée-pixel, tout en gardant en esprit notre objectif du présent travail, la deuxième contribution s'agit d'une nouvelle méthodologie d'évaluation qui évite la dépendance du logiciel de reconnaissance et qui réduit l'intervention humaine dans la préparation des données de vérité terrain. Dans cette méthodologie, l'intérêt d'une méthode de binarisation est jugé par la performance induite sur une étape après-binarisation (qui peut être la phase finale) dans le processus de recherche. La meilleure technique de binarisation sera donc celle avec laquelle les performances du système de recherche soient maximales. La méthode d'évaluation proposée appartient à la catégorie de techniques d'évaluation objective au niveau système, et en même temps à la catégorie de techniques d'évaluation empirique supervisée-composants.

4.3.2. Description de l'approche proposée

Nous décrivons dans cette section la méthodologie que nous avons proposée pour l'évaluation objective des algorithmes de binarisation. Comme nous avons dit précédemment, l'évaluation n'est pas conçue pour trouver une méthode générique pour toutes les applications de traitement d'images mais pour répondre à un objectif bien précis qui est dans notre cas la recherche de mots dans les images de documents arabes dégradés. L'efficacité d'une technique de binarisation est alors déterminée en fonction de sa capacité d'améliorer les performances du système de recherche.

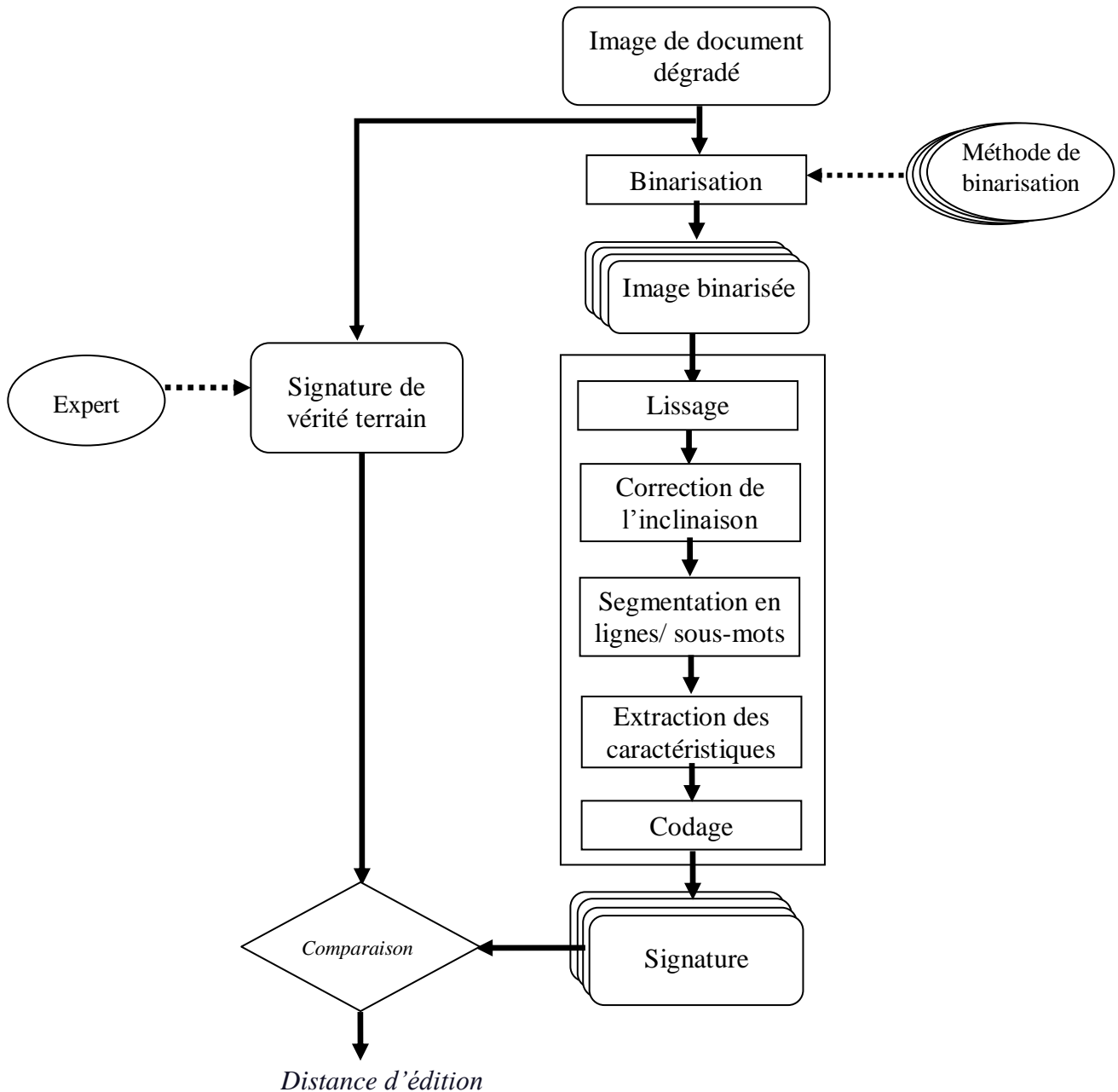


Figure 4. 14. Schéma général de la méthodologie d'évaluation proposée

L'idée de base de notre proposition est d'évaluer les performances d'une méthode de binarisation en comparant les caractéristiques extraites des images de documents binarisées avec des caractéristiques de vérité terrain (de référence) au lieu de comparer les images elles-mêmes. La qualité d'une image binarisée est mesurée par la distance entre les caractéristiques extraites d'elle et les caractéristiques de vérité terrain. La meilleure méthode est celle avec laquelle la distance soit minimale. Nous pouvons justifier ce choix par le fait que d'une part les caractéristiques extraites à partir d'une image de document décrivent le texte présent dans ce document, qui est l'objet principal du système de recherche, et de l'autre part la préparation des caractéristiques de vérité terrain (comme nous allons décrire dans les sections suivantes) est beaucoup plus facile et moins coûteuse que la préparation des images de vérité terrain. Le système que nous avons utilisé dans l'évaluation est notre système de recherche de documents arabes dégradés proposé dans [KEF 10][SAR 11].

Le principe est comme suit : nous représentons tout d'abord chaque image de document de notre collection de test par une signature constituée de l'ensemble d'informations pertinentes qu'on peut extraire du document. Cette signature contient les caractéristiques de l'écriture et elle est codée en ASCII. La signature obtenue est considérée comme une signature de vérité terrain. Après, nous binarisons chaque image de test par plusieurs méthodes de binarisation et sur chaque image binaire résultante nous appliquons une série de traitements, intégrés dans le système de recherche, et qui visent à extraire les caractéristiques de l'écriture. Ces dernières sont représentées sous forme d'une signature textuelle correspondante à l'image binarisée. La signature textuelle ainsi obtenue est comparée avec la signature de vérité terrain et la qualité de l'image binarisée est mesurée par la distance d'édition entre les deux signatures. Le schéma de la figure 4.14 résume les étapes impliquées dans le processus d'évaluation.

4.3.2.1. Préparation des signatures de vérité terrain

La première chose à faire est de préparer pour chaque image de document considérée dans l'évaluation la signature de vérité terrain correspondante. Cette dernière est une description textuelle des caractéristiques de l'écriture présente dans le document. En effet, pour que la signature de vérité terrain pourra être comparée avec les signatures extraites automatiquement à partir des images de documents binarisées, elle doit être composée des mêmes caractéristiques utilisées dans le système de recherche pour décrire les documents, à savoir, les caractéristiques structurelles et plus précisément :

- **Les hampes** : correspondent aux extensions hautes pouvant exister dans les caractères ;
- **Les jambages** : correspondent aux extensions basses pouvant exister dans les caractères ;
- **Les points diacritiques** : correspondent aux parties secondaires des caractères. Des points simples ou multiples pouvant être écrits au dessus ou au dessous du corps principal des caractères ;
- **Les boucles** : correspondent aux occlusions pouvant exister dans les caractères.

La figure 4.15 présente un exemple de telles caractéristiques.

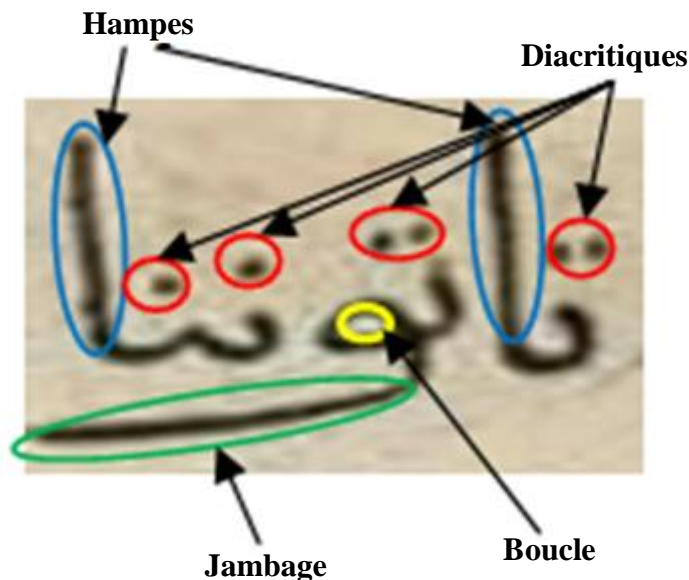


Figure 4.15. Caractéristiques structurelles présentes dans un mot arabe

Afin d'établir la signature de vérité terrain, un expert analyse visuellement les images de documents dans le but d'extraire les différentes caractéristiques structurelles existantes. A chaque caractéristique on attribue un code précis :

- *H* pour les hampes,
- *J* pour les jambages,
- *B* pour les boucles,
- *P* pour les points diacritiques hauts,
- *Q* pour les points diacritiques bas.

Ainsi, la signature de vérité terrain d'un document est obtenue par la concaténation des codes des différentes caractéristiques structurelles détectées, tout en séparant les sous-mots⁹ par un point virgule “ ; ”.

En tenant compte que l'écriture arabe s'écrit cursivement de la droite vers la gauche, le mot présenté dans la figure 4.15 est composé de trois sous-mots :

- « تا », caractérisé par la présence de deux points diacritiques hauts et une hampe respectivement. Il peut être codé par la séquence : “**PPH**”
- « تو », caractérisé dans l'ordre par deux points diacritiques hauts, une boucle et un jambage. La séquence de code correspondante est : “**PPBJ**”
- « ننا », contient deux points diacritiques hauts et une hampe et il est donc codé par la séquence : “**PPH**”

La signature de vérité terrain de l'image du mot précédente est alors obtenue par la concaténation des séquences de codes des ses trois sous-mots comme suit:

“ **PPH;PPBJ;PPH** ”

Afin de minimiser le temps de préparation manuelle des signatures de vérité terrain, nous pouvons avoir recours aux images de documents arabes synthétiques au lieu des images réelles. La minimisation du temps de préparation des signatures se fait en automatisant le processus d'extraction des caractéristiques et en limitant l'intervention humaine uniquement dans la correction d'éventuelles erreurs qui peuvent se produire lors de l'extraction automatique des caractéristiques.

Les images de documents synthétiques sont obtenues chacune par la superposition d'une image binaire nette et épurée du bruit d'un texte arabe imprimé, considérée comme image de vérité terrain avec un masque de fond ancien. La superposition est effectuée en utilisant la technique de *Mosaicing* d'images proposée par P. Stathis et al. dans [STA 08] comme il est détaillé dans la section 4.2.2.2.

Sur chacune des images de vérité terrain on applique une série de traitements soutenus par la représentation de l'image par une signature textuelle exprimant leurs caractéristiques structurelles. Les caractéristiques extraites sont codées comme auparavant en faisant correspondre à chaque caractéristique un code précis (*H* pour les hampes, *J* pour les jambages, *P* pour les points hauts et *Q* pour les points bas) ce qui nous fournit une signature de vérité terrain. Un expert peut intervenir à chaque niveau de traitement pour corriger les erreurs éventuelles.

Le schéma général de l'évaluation en utilisant des images de synthèse est montré dans la figure 4.16.

⁹ Un sous-mot est une partie de mot composée d'une ou plusieurs lettres connectées

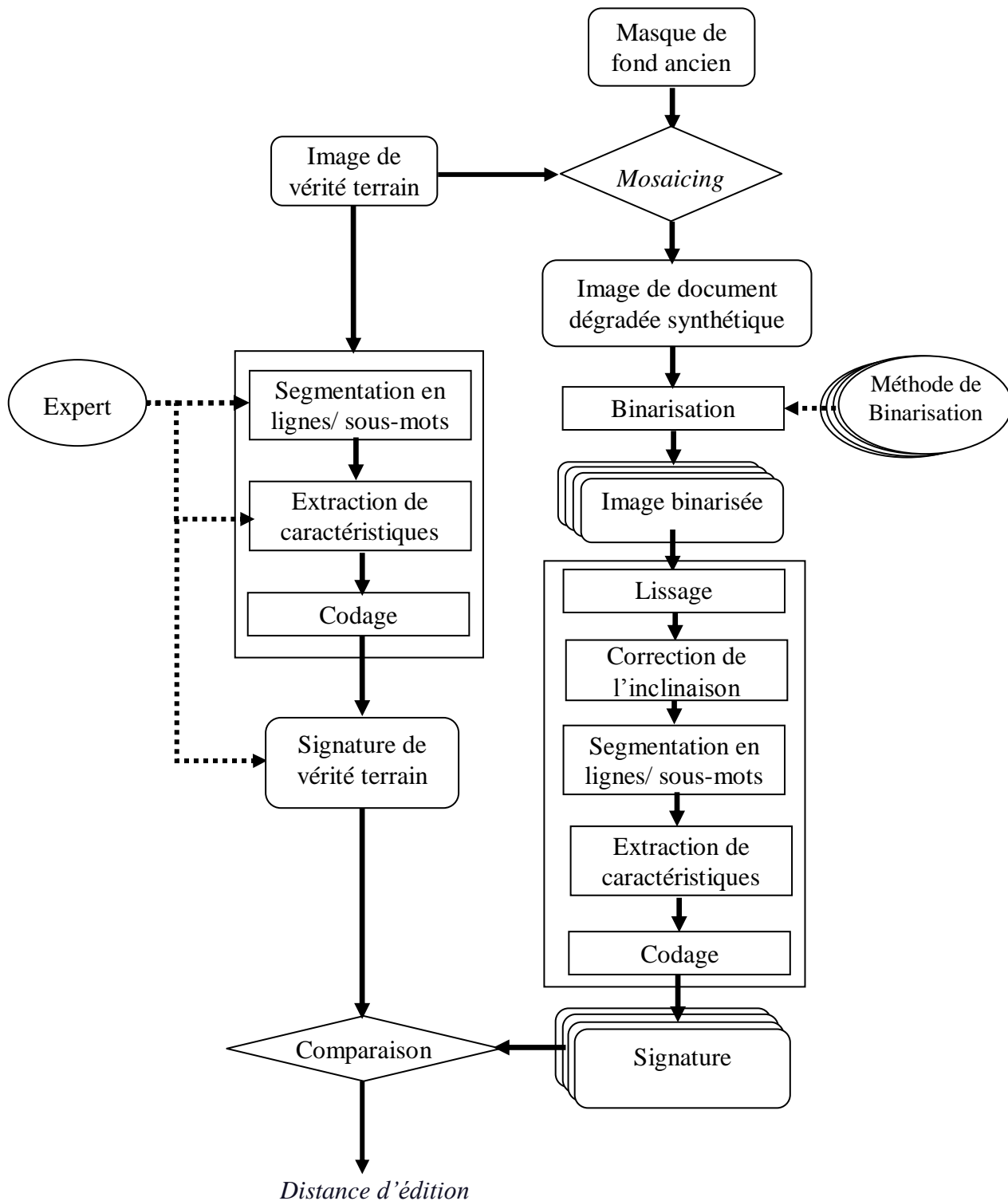


Figure 4. 16. Schéma général de la méthodologie d'évaluation proposée en utilisant des images synthétiques

4.3.2.2. Traitement des images de documents binarisées

Sur chaque image (réelle ou synthétique) de la collection de test nous appliquons les différentes méthodes de seuillage considérées dans l'évaluation, et les images binarisées résultantes sont soumises à un certain nombre de traitements automatiques visant à représenter chaque image par une signature textuelle décrivant les caractéristiques

structurelles (hampes, jambages, boucles, et points diacritiques) extraites à partir des sous-mots présents dans le document. Les étapes de traitement sont les suivantes :

4.3.2.2.1. Lissage

La binarisation peut introduire du bruit dans l'image, qui se traduit en particulier par la présence d'irrégularités le long des tracés des caractères. Pour pallier à ce problème, nous appliquons un lissage en utilisant l'algorithme de [MAH 94] qui réduit le bruit d'une image binaire en éliminant les pixels isolés d'une part et en bouchant les trous vides de l'autre part. Cette technique simple et efficace est basée sur une décision statistique. En effet, la nouvelle valeur de chaque pixel dans l'image binarisée est calculée en fonction de sa valeur initiale et celles des 8 pixels voisins (Figure 4.17).

P ₁	P ₂	P ₃
P ₄	P₀	P ₅
P ₆	P ₇	P ₈

Figure 4. 17. Pixel courant P₀ et ses voisins

La figure 4.18 montre le résultat d'application de cet algorithme sur une image de mot.

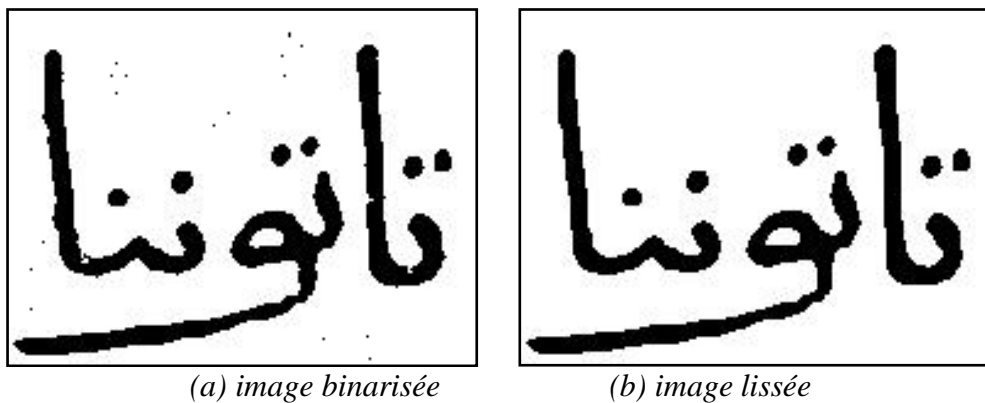


Figure 4. 18. Lissage d'une image de mot binarisée

4.3.2.2.2. Correction de l'inclinaison (Redressement de l'image)

Les anciens documents composant notre collection sont parfois inclinés. La présence des inclinaisons influe considérablement sur les étapes de segmentation et d'extraction des caractéristiques et par conséquent sur le résultat final du système. Une étape de redressement est alors nécessaire. Pour cela, nous avons utilisé une méthode de correction de l'inclinaison basée sur les projections partielles. La méthode procède en cinq étapes comme suit (voir [KEF 10][SAR 11] pour plus de détails):

- 1- Division de l'image en colonnes de taille fixe (taille d'un mot environ)
- 2- Calcul des histogrammes des projections horizontales de chaque colonne
- 3- Extraction des lignes de base correspondant aux pics des histogrammes calculés auparavant.
- 4- Calcul de l'angle d'inclinaison du document θ comme la moyenne de tous les angles formés par deux lignes de base appartenant à deux colonnes successives.
- 5- Rotation de l'image d'angle θ

La figure 4.19 illustre le résultat d'application de cette méthode de redressement sur une

image de document incliné.

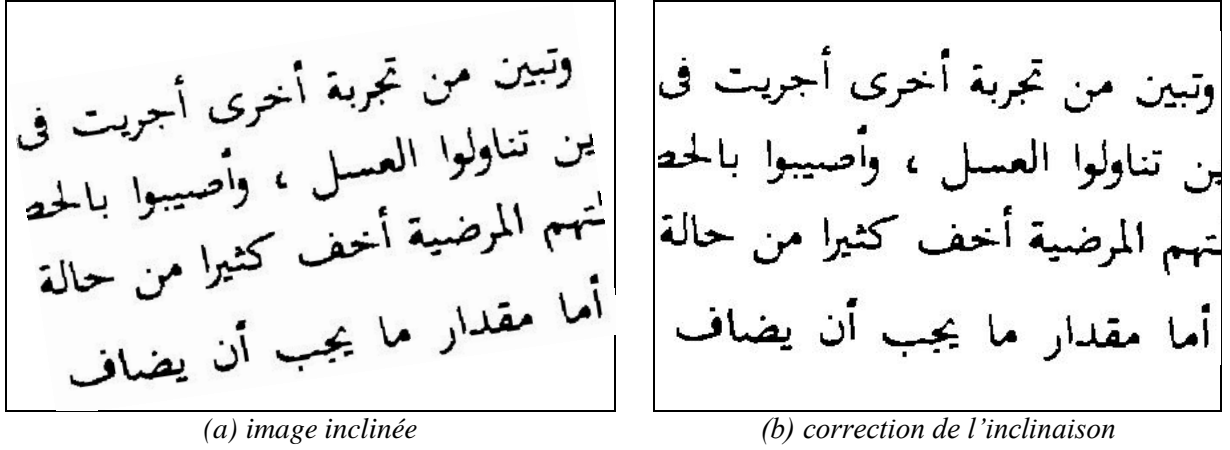


Figure 4.19. Résultat de redressement d'une image de document incliné

4.3.2.2.3. Segmentation en lignes/ sous-mots

Après le lissage et le redressement, vient l'étape de segmentation. Le but de cette étape est d'extraire les différents sous-mots du texte. Le choix de travailler avec les sous-mots est justifié par le fait qu'ils nous semblent offrir un meilleur compromis entre la complexité de la segmentation des mots arabes manuscrits en lettres et l'approche globale qui emploie des dictionnaires de petites tailles. En effet, avec un petit nombre de sous-mots on peut construire un grand nombre de mots et leur extraction est plus facile [KEF 10][SAR 11].

Pour extraire les sous-mots, nous devons tout d'abord segmenter le texte en lignes, et par la suite segmenter chaque ligne en sous-mots.

Les lignes du texte sont séparées en utilisant des mesures de densité des plages blanches à partir des projections horizontales. Les vallées de l'histogramme correspondent aux zones de séparation entre les lignes. L'algorithme de séparation en lignes procède comme suit (voir [KEF 10][SAR 11] pour plus de détails):

- 1- Calcul de l'histogramme des projections horizontales de l'image.
- 2- Extraction des minima locaux correspondant aux zones de séparation entre les lignes.
- 3- Filtrage des minima locaux obtenus auparavant.
- 4- Résolution des conflits en attribuant les pixels noirs existants dans les zones séparatrices à la ligne du texte la plus proche par analyse de proximité.

Le résultat de segmentation en lignes d'une image de document est donné par la figure 4.20.

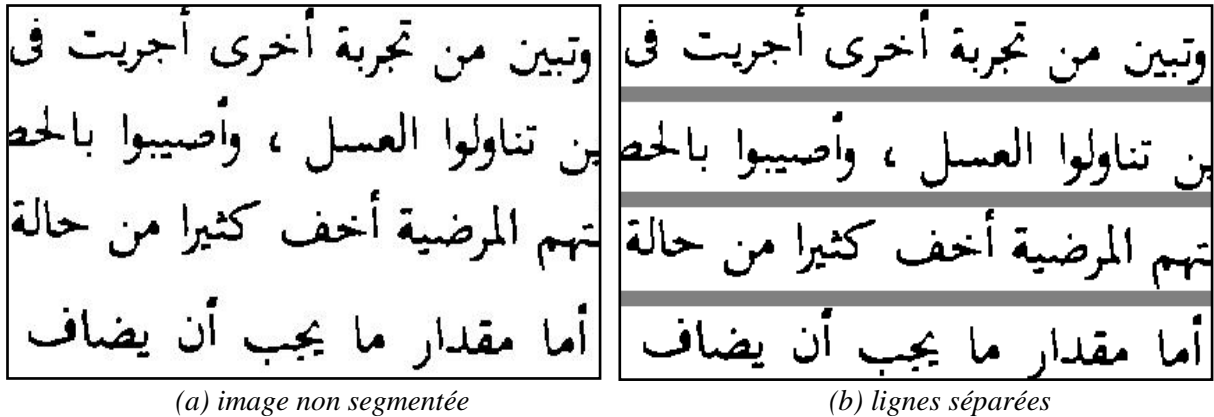


Figure 4.20. Résultat de segmentation en lignes d'une image de document

Après, les sous-mots sont extraits à partir de chaque ligne du texte. Dans l'écriture arabe, le sous-mot correspond à une composante connexe principale (corps du sous-mot) et des composantes connexes auxiliaires représentant les signes diacritiques. L'extraction des sous-mots consiste à étiqueter les différentes composantes connexes de l'image et ensuite d'attribuer chaque point diacritique à son propre corps du sous-mot. L'extraction des composantes connexes se fait en regroupant les pixels noirs voisins dans une unité distincte (Figure 4.21), et on utilise pour cela la méthode d'agrégation des pixels. Cette méthode peut être résumée par les étapes suivantes :

- 1- Trouver un premier pixel noir non visité.
- 2- Chercher tous ses voisins : si l'un des voisins est un pixel noir on le regroupe avec le premier et on réitère récursivement l'opération pour tous les voisins.
- 3- On s'arrête lorsque tous les pixels noirs soient visités.
- 4- Revenir à l'étape 1.



Figure 4. 21. Ligne de texte segmentée en composantes connexes

La décision qu'une composante connexe est un point diacritique ou non est basée sur une heuristique : les points diacritiques sont généralement plus petits que les corps des lettres. Les composantes connexes, dont la taille est inférieure à un certain seuil sont donc considérées comme des points diacritiques. Le seuil est choisi relatif à la taille moyenne des composantes connexes : $seuil = taille\ moyenne / n$. Avec n est fixé par expérimentations. Finalement, les points diacritiques sont attribués au corps du sous-mot le plus proche par analyse de proximité.

4.3.2.2.4. Extraction des caractéristiques

L'extraction des caractéristiques a comme but d'identifier les propriétés les plus importantes pour la discrimination de classes de formes. Cette étape doit être effectuée le plus fidèlement possible, car les caractéristiques extraites composent la signature qui représente le document. Un des problèmes fondamentaux dans l'analyse d'images est de déterminer quelles caractéristiques à employer pour avoir de bons résultats. Les primitives structurales issues de la perception humaine qui sont reliées à la forme de l'écriture sont considérées comme des caractéristiques pertinentes pour la discrimination des caractères manuscrits [CHE 98b]. Nous avons choisi d'extraire à partir de chaque sous mot, quatre caractéristiques structurales : les hampes, les jambages, les boucles et les points diacritiques.

Notons que les caractéristiques structurales précédentes sont extraites à partir des contours des sous-mots, et en s'aidant de la connaissance a priori que les extensions hautes et basses des caractères sont toujours en dehors d'une zone médiane. Afin d'extraire donc ces caractéristiques, nous procédons comme suit :

a) Détection de la ligne de base

La ligne de base est la ligne sur laquelle repose les caractères qui ne possèdent pas des descendants. Dans les textes arabes, la ligne de base porte une information assez importante sur l'orientation des textes et la position des points diacritiques. La méthode la plus répandue pour détecter les lignes de base est la projection horizontale de l'image [KEF 10]. La ligne de base correspond à la ligne dont la projection contient le plus grand nombre de pixels noirs (la

ligne rouge dans la Figure 4.22).

b) Localisation de la zone médiane

Le corps des mots arabes, apparaît dans une zone appelée zone médiane, c'est pour cette raison que la localisation de cette zone est importante pour l'extraction des caractéristiques. Nous détectons cette zone en traçons 2 frontières, haute et basse par rapport à la ligne de base (les 2 lignes bleues dans la Figure 4.22). La zone médiane sera donc l'espace compris entre ces deux frontières (Voir [KEF 10][SAR 11] pour plus de détails).

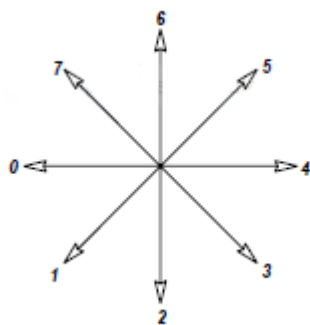


Figure 4. 22. Ligne de base et zone médiane détectées sur une ligne de texte

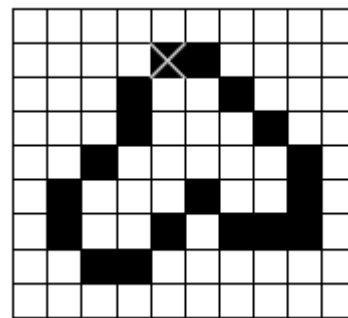
c) Suivre de contours

Le suivi de contours est communément utilisé pour l'extraction des caractéristiques structurelles des caractères. Le contour d'un sous-mot est l'ensemble de points délimitant le tracé, et pour le codage du contour, le code généré est la chaîne de Freeman, qui représente les directions des pixels du contour à 8 voisins. Le code de Freeman est une technique très utile pour représenter le contour d'un objet sur l'image. Elle consiste à décrire le contour d'une forme en spécifiant un point de départ et en indiquant les pixels voisins immédiats par l'une des 8 directions possibles. Les 8 directions indiquent les 8 voisins immédiats possibles du pixel analysé.

Ainsi l'objet de la figure 4.23.b possède le code de Freeman suivant, le point d'origine étant celui identifié par X se trouvant en (4, 1) :1-2-1-1-2-3-4-5-5-3-4-4-6-6-7-7-7-0.



(a) Directions du contour



(b) contour d'un objet

Figure 4. 23. Suivre de contours d'un objet

L'analyse de ces contours, leur type (interne ou externe) et leur position nous permet de distinguer la présence des boucles, des points diacritiques, et des extensions hautes et basses.

d) Extraction des Diacritique

Une composante connexe est considérée comme un point diacritique si sa taille est inférieure à un certain seuil.

e) Extraction des Jambages

Un jambage correspond à une descente qui se termine en dehors de la zone médiane c'est-à-dire située dans la zone inférieure.

f) Extraction des Hampes

Une hampe correspond à une montée qui dépasse la zone médiane (au dessus de la frontière haute).

g) Extraction des Boucles

Les boucles sont utiles pour l'identification de certains caractères. Pour détecter les boucles, on regarde si toutes les coordonnées d'un contour englobent toutes les coordonnées d'un autre, alors on a une boucle.

Chaque caractéristique est représentée par une couleur dans la Figure 4.24.



Figure 4. 24. Caractéristiques extraites sur une ligne de texte

4.3.2.2.5. Codage

Après l'extraction des caractéristiques, une signature textuelle est attribuée au document comme précédée lors de la préparation des signatures de vérité terrain, c'est à dire en faisant correspondre à chaque caractéristique un code en ASCII (les hampes en *H*, les jambages en *J*, les boucles en *B*, les points diacritiques hauts en *P*, et les points diacritiques bas en *Q*). Un autre caractère est ajouté (le point virgule '*;*') qui sert de séparateur inter-sous-mots. Par exemple, les caractéristiques détectées de l'image de la figure 4.24 sont codées par la signature :

J;PQQJP;JP;PHQJ;QP;PH;PJ;J;PH;QJ;QP;PH;QH;H;H;H;HBPJ

4.3.2.3. Comparaison

Finalement, pour chaque image de la collection de test, les signatures extraites à partir des images résultantes des différentes méthodes de binarisation sont comparées avec la signature de vérité terrain préétablie. La comparaison est effectuée en utilisant une célèbre mesure de distance entre les chaînes, à savoir la distance d'édition. Cette dernière est parfois référencée comme la distance de Levenshtein en reconnaissance de l'article [LEV 66] par Vladimir Levenshtein où la distance d'édition a été certainement discutée pour la première fois. La distance d'édition entre deux chaînes est définie comme le nombre minimal d'opérations d'édition (insertion, suppression, substitution) nécessaires pour transformer la première chaîne en la seconde. En effet, Levenshtein a introduit la définition de la distance d'édition mais il n'a jamais décrit l'algorithme permettant de trouver la distance d'édition entre deux chaînes [PEV 06]. Un algorithme de calcul de la distance d'édition entre les chaîne X et Y a été proposé par Wagner et Fischer dans [WAG 74]. Ainsi pour comparer une signature extraite après le traitement avec une signature de vérité terrain en utilisant l'algorithme de Wagner et Fischer, nous considérons les deux signatures comme deux suites $X = (x_1 \dots x_n)$ et $Y = (y_1 \dots y_m)$.

L'algorithme de Wagner et Fischer est un algorithme de programmation dynamique, qui utilise une matrice D de dimension $(n+1) \times (m+1)$ où n et m sont les longueurs des deux chaînes de caractères. L'approche consiste à remplir la table D tel que la valeur d'une case dépend des valeurs précédemment obtenues. La première ligne de la table $D(0, i)$ est initialisée à i et la première colonne $D(j, 0)$ est initialisée à j . L'algorithme peut être résumé par le pseudo-code suivant.

```

Pré-requis :  $n$  = longueur de la première chaîne  $X$ 
Pré-requis :  $m$  = longueur de la deuxième chaîne  $Y$ 
Pré-requis :  $C(a, b)$  = coût de la transformation pour passer de  $a$  à  $b$ 
Pré-requis :  $\lambda$  = vide
Assure :  $D[m, n]$  = distance d'édition de  $X$  à  $Y$ 
Pour  $i = 0 \dots n$  Faire  $D[0, i] = i$ 
Pour  $j = 0 \dots m$  Faire  $D[j, 0] = j$ 
Pour  $i = 1 \dots m$  Faire
    Pour  $j = 1 \dots n$  Faire
         $D[i, j] = \text{Min} \{$ 
             $D[i-1, j-1] + C(X_i, Y_j), // \text{ substitution (0 si les caractères } X_i \text{ et } Y_j \text{ sont égaux)}$ 
             $D[i-1, j] + C(X_i, \lambda), // \text{ effacement}$ 
             $D[i, j-1] + C(\lambda, Y_j) // \text{ insertion}$ 
         $\}$ 
    Fin Pour
Fin Pour

```

La distance entre les deux signatures X et Y est contenue dans la dernière case de la matrice D .

La meilleure méthode de binarisation pour chaque image de la collection de test est finalement celle avec laquelle les caractéristiques extraites sont les plus proches des caractéristiques de vérité terrain, c'est-à-dire dont la distance entre la signature extraite et la signature de vérité terrain est minimale.

4.3.3. Expérimentations et résultats

La méthodologie d'évaluation proposée a été utilisée pour évaluer les performances de différents algorithmes de binarisation pour l'objectif de la recherche d'images de documents arabes dégradées.

4.3.3.1. Méthodes de binarisation évaluées

Dans nos expérimentations, nous avons considéré huit algorithmes de binarisation appartenant à différentes catégories :

- 1- Méthode d'Otsu [OTS 79] : une des méthodes globales les plus connues et utilisées.
- 2- Méthode de Kapur et al. [KAP 85] : représentant des méthodes entropiques.
- 3- Méthode de Bernsen [BER 86] : méthode basée sur l'intervalle de gris local.
- 4- Méthode de Niblack [NIB 86] : méthode classique basée sur la variance locale.
- 5- Méthode de Sauvola et al. [SAU 00] : célèbre méthode basée sur la variance locale.
- 6- Méthode de Wolf et al. [WOL 02] : méthode basée sur la variance locale.
- 7- Méthode Nick [KHU 09] : méthode basée sur la variance locale.

- 8- Méthode de Binarisation en utilisant le minimum et le maximum locaux [SU 10] : la méthode gagnante dans la compétition H-DIBCO 2010 [PRA 10].

4.3.3.2. Images de test

Deux ensembles d'images ont été utilisées dans l'expérimentation.

4.3.3.2.1. Premier ensemble

Le premier ensemble contient 20 images de documents arabes dégradés collectées à partir du web, et donc elles ne possèdent pas les mêmes caractéristiques. Elles sont de natures très différentes, provenant de plusieurs sources, couvrant plusieurs domaines, et présentant différents types de dégradations liées aux documents anciens. La figure 4.25 illustre 2 images de cet ensemble.



Figure 4. 25. Exemples d'images de documents historiques réelles utilisées pour tester la méthodologie d'évaluation proposée

4.3.3.2.2. Deuxième ensemble

Le deuxième ensemble est composé d'images synthétiques. En effet, le coût bas de l'obtention des signatures de vérité terrain à partir des images synthétiques comparativement aux images réelles, rend possible l'utilisation dans les tests d'un nombre considérable d'images. Ces images ont été créées en utilisant la technique de fusion par *Mosaicing* proposée par P. Stathis et al. [STA 08] comme il a été détaillé dans la section 4.2.2.2. Ainsi, 100 images de documents ont été obtenues par la fusion de 10 images propres (images en noir et blanc sélectionnées à partir des documents PDF) contiennent du texte imprimé, et 10 images de fonds anciens vierges (ne contiennent aucun texte ou graphiques). La figure 4.26 présente 2 images synthétiques de cet ensemble.

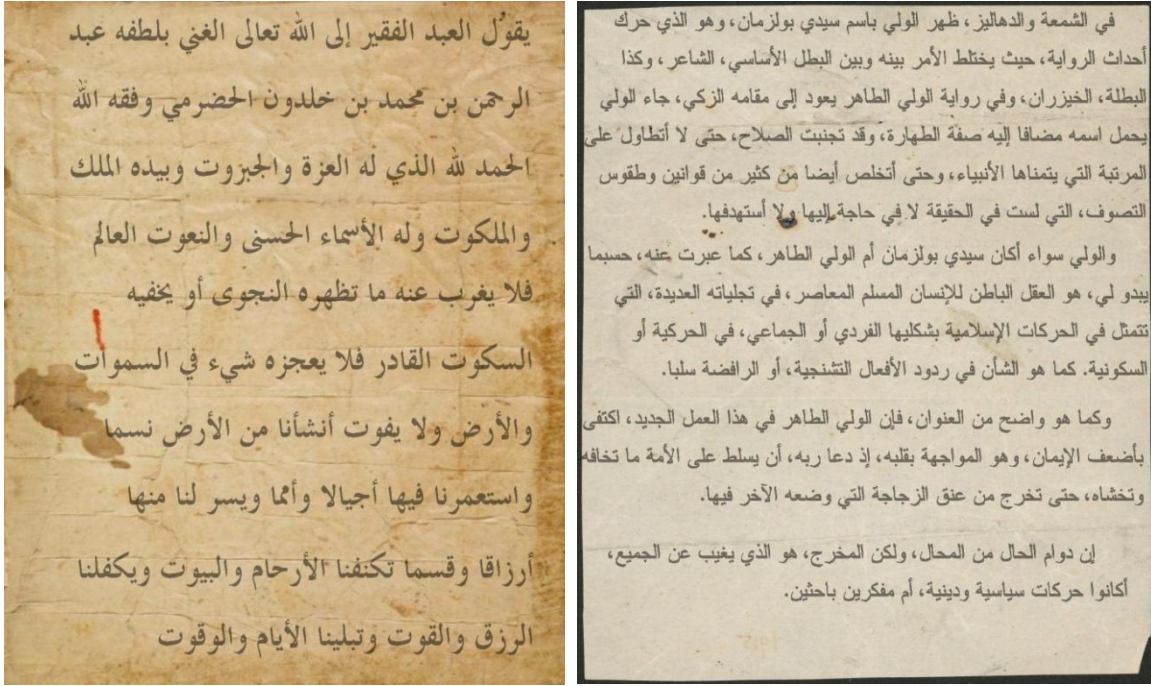


Figure 4. 26. Exemples d'images de documents historiques synthétiques utilisées pour tester la méthodologie d'évaluation proposée

4.3.3.3. Procédure d'évaluation

L'objectif des expérimentations est non seulement d'utiliser la nouvelle méthodologie proposée pour estimer les performances de différents algorithmes de binarisation. Il est également de vérifier sa robustesse, de tester sa concordance avec d'autres approches d'évaluation, et de trouver l'algorithme le plus adéquat aux applications de recherche dans les images de documents arabes dégradés.

Cependant, et conjointement avec l'évaluation dirigée-but (au niveau-système) en employant notre méthodologie, nous avons procédé une évaluation supervisée-pixel dans le but d'avoir une vision globale sur les performances des techniques de binarisation étudiées selon deux points de vue différents. Nous avons ainsi employé quatre mesures empiriques très connues et qui ont été utilisées dans le cadre des compétitions DIBCO et H-DIBCO. Ces mesures sont :

- 1- Fmeasure (FM)
- 2- Rapport maximum signal à bruit ($PSNR$)
- 3- Métrique de taux négatif (NRM)
- 4- Métrique de pénalité de fausse-classification (MPM)

Comme l'humain est toujours le visionneur idéal et l'ultime exploitant des systèmes de traitement et d'analyse de documents, et en vu d'estimer la robustesse de la méthodologie d'évaluation proposée sur un niveau élevé (niveau compréhension), nous avons procédé conjointement avec l'évaluation dirigée-but et l'évaluation supervisée-pixel, une évaluation subjective visuelle. Cinq visionneurs ont été interrogés pour mettre leurs observations indépendamment pour chaque image. Ainsi, pour toute image de la collection de test, chaque visionneur examine les résultats de binarisation de cette image en utilisant les différentes méthodes testées, et classe ces méthodes selon leur résultat de la meilleure à la mauvaise (de 1 à 8). Finalement, nous calculons pour chaque méthode la moyenne des scores attribués.

Sauvola et al.		QH H H H B; H; H H H B; H; H J; B P; H; H; Q J; J; Q; P P Q P; H; H; H H B; J; Q; H; H H; H Q P; J; B H Q; H; H H B; H Q; Q; P H; B; J; H J Q
Wolf et al.		QH H H H B; H; H H H B; H; H J; B P; H; H; Q J; J; Q; P P Q P; H; H; H H B; J; Q; H; H H; H Q P; B H Q; H; H H B; H Q; Q; P H; B; J; H J Q
Nick		QH H P P P B; H; H H H B; Q H; H J; B P; H; H; Q J; J; Q; P P Q P; P; H; H; H H B; J; P Q; H; H H; H Q P; B H Q; H; H H B; H Q; Q; P H; B; J; H J Q
Binarisation en utilisant le minimum et le maximum locaux		J P; Q H P P P H P Q P Q B Q; P; P; Q; P H; H P H P P H B; H; Q H Q P Q J; Q; P Q Q B P J P P Q; Q H; H P J; Q; P Q Q Q P P P P Q J; Q; B Q J; Q P P; P Q Q Q P P P Q P P P; P; H; H Q P; Q H H P P B P; P; J; Q Q; H; P; H B H; Q P H Q P P; Q Q J; B P Q P H Q Q P; H; H H B; H Q P; Q Q Q Q P; P P H; P B P; P; Q; Q J Q; Q P P P H Q J P Q Q Q

Tableau 4. 7. Résultats d'évaluation (visuelle et en termes de distance d'édition) des méthodes de binarisation sur une image exemple

Critère d'évaluation	Visuelle						Distance d'édition
	V1	V2	V3	V4	V5	Moyenne	
Méthode de binarisation							
Méthode d'Otsu	2	2	2	1	1	1.6	19
Méthode de Kapur et al.	5	5	5	5	5	5	37
Méthode de Bernsen	7	7	7	7	8	7.2	107
Méthode de Niblack	6	6	6	6	6	6	75
Méthode de Sauvola et al.	1	1	1	2	2	1.4	18
Méthode de Wolf et al.	3	3	3	3	3	3	21
Méthode Nick	4	4	4	4	4	4	25
Binarisation en utilisant le minimum et le maximum locaux	8	8	8	8	7	7.8	111

Comme nous pouvons remarquer des deux tableaux précédents, les deux approches d'évaluation procédées fournissent le même classement des méthodes de binarisation pour l'image précédente : 1- Sauvola et al., 2- Otsu, 3- Wolf et al., 4- Nick, 5- Kapur et al., 6- Niblack, 7- Binarisation en utilisant le minimum et le maximum locaux, 8- Bernsen. Cette constatation montre que l'évaluation au niveau système en utilisant notre approche concorde totalement avec l'évaluation visuelle de l'humain considéré comme l'ultime visionneur. Notons que le classement obtenu est valide pour l'image précédente mais il pourra être totalement différent pour d'autres images.

Les performances moyennes de chaque méthode de binarisation sont calculées sur chaque ensemble de test séparément, et le classement global des méthodes est effectué en analysant leurs performances moyennes. Les résultats d'évaluation moyens des méthodes de binarisation sur le premier et le deuxième ensemble d'images sont résumés dans les tableaux 4.8 et 4.9 respectivement. Les résultats sont classés suivant la distance d'édition moyenne mesurée entre les signatures extraites des images binarisées et les signatures de vérité terrain.

Tableau 4. 8. Résultats d'évaluation (visuelle et en termes de distance d'édition) des algorithmes de binarisation sur le premier ensemble de test

Critère d'évaluation Méthode de binarisation	Rang visuel moyen	Distance d'édition moyenne
Méthode de Sauvola et al.	2	38.33333
Binarisation en utilisant le minimum et le maximum locaux	2.6	44.26
Méthode Nick	2.9	46
Méthode d'Otsu	3.3	50.943
Méthode de Wolf et al.	3.6	62.111
Méthode de Kapur et al.	5.4	70.9
Méthode de Bernsen	7.4	143
Méthode de Niblack	8.8	158.66667

Tableau 4. 9. Résultats d'évaluation (visuelle, supervisé-pixel, et en termes de distance d'édition) des algorithmes de binarisation sur le deuxième ensemble de test

Critère d'évaluation Méthode de binarisation	Rang visuel moyen	Distance d'édition moyenne	FM	PSNR	NRM ($\times 10^{-2}$)	MPM ($\times 10^{-3}$)
Méthode de Sauvola et al.	2.3	28.666667	0.812	16.207	9.7	6.66
Binarisation en utilisant le minimum et le maximum locaux	2.5	30.9413333	0.789	14.53	11.9	8.87
Méthode Nick	3	31.337	0.751	14.84	11.5	14.5
Méthode d'Otsu	3.2	38.2	0.796	15.33	10.83	7.22
Méthode de Wolf et al.	3.7	53.166667	0.773	14.82	12.9	7.8
Méthode de Kapur et al.	5.1	64	0.792	15.31	10.81	8.04
Méthode de Bernsen	7.6	132.66667	0.57	8.26	19.8	141
Méthode de Niblack	8.6	146	0.545	7.95	16.1	156.13

La figure 4.27 et 4.28 illustrent graphiquement les performances des méthodes de binarisation sur le premier et le deuxième ensemble d'images respectivement.

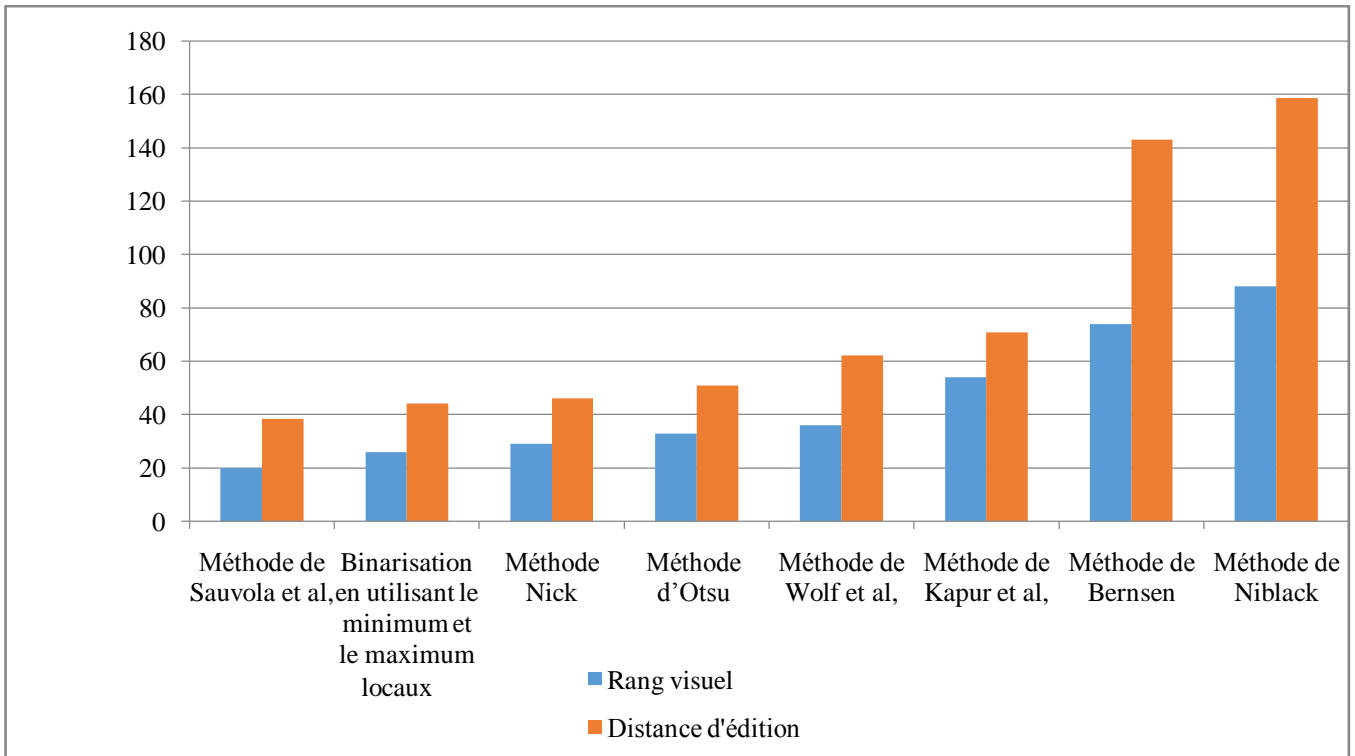


Figure 4. 27. Performances des méthodes de binarisation sur le premier ensemble

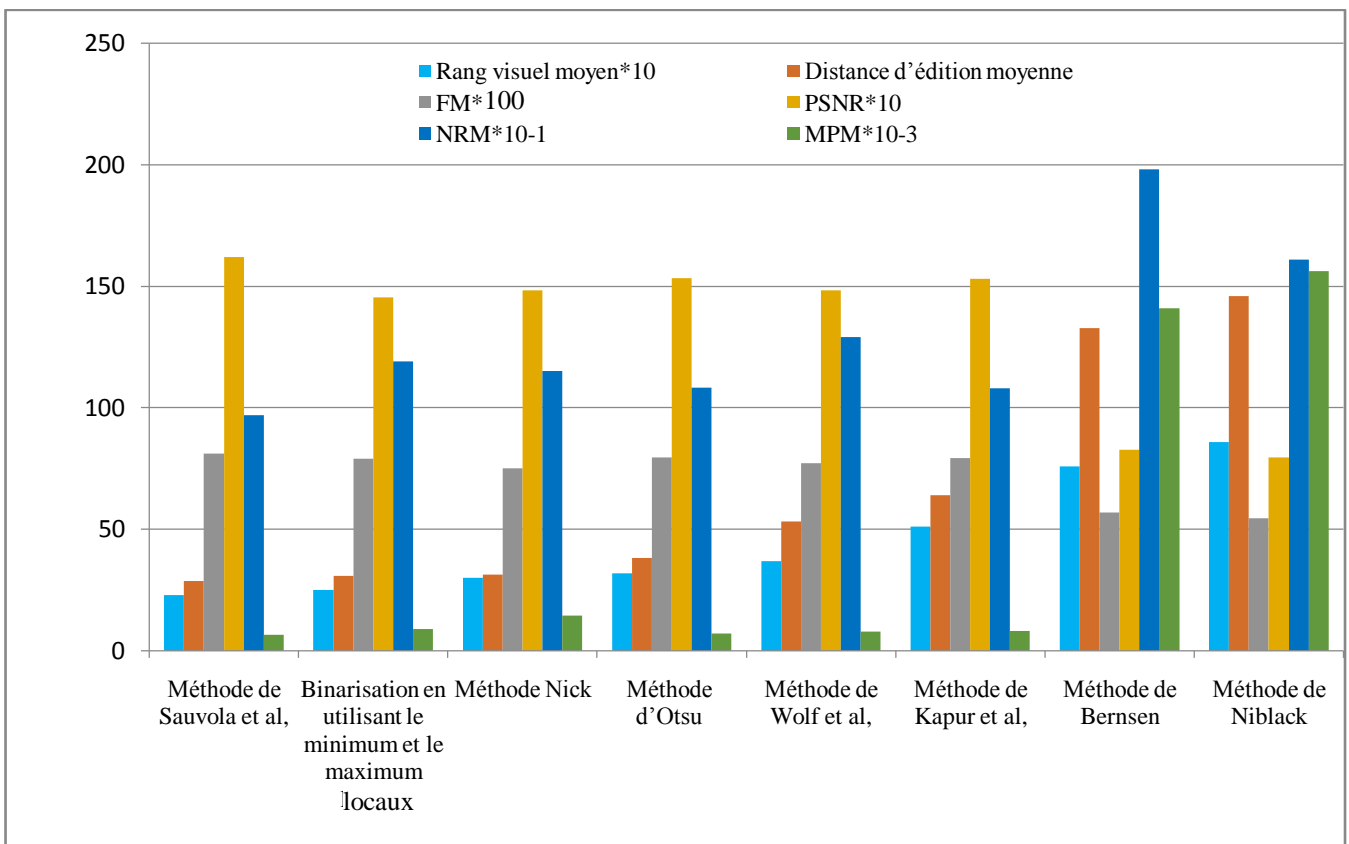


Figure 4. 28. Performances des méthodes de binarisation sur le deuxième ensemble

Notons qu'aucun algorithme de binarisation n'a surpassé les autres pour toutes les images et que le classement précédent a été effectué en se basant sur les résultats moyens.

A partir des tableaux 4.8 et 4.9 précédents et des figures 4.27 et 4.28, nous pouvons remarquer que les résultats d'évaluation obtenus en utilisant la méthodologie d'évaluation proposée classent la méthode de Sauvola et al. [SAU 00] en première place avec une distance d'édition moyenne minimale, et elle a présenté des performances légèrement meilleures que la méthode de binarisation en utilisant le minimum et le maximum locaux [SU 10]. Ce classement est identique pour les deux ensembles d'images, mais on observe une petite détérioration dans performances (en termes de distance d'édition) sur le premier ensemble comparativement au deuxième. Cette détérioration est raisonnable tenant compte de la variabilité d'écritures manuscrites existantes dans les documents réels par opposition aux documents synthétiques créés par la fusion de documents imprimés en noir et blanc avec des images de fond ancien.

La présence de la méthode de Sauvola et al. en tête du classement obtenu par notre approche montre que cette méthode a réussi à binariser les images des documents des deux ensembles de test en préservant la majorité de leurs caractéristiques structurelles. Ces caractéristiques ont été extraites avec une bonne précision par le système de recherche employé ce qui a été traduit par une distance d'édition minimale entre les signatures associées aux images binarisées décrivant les caractéristiques extraites et les signatures de vérité terrain décrivant les caractéristiques réelles. La méthode de Sauvola et al. se montre donc la plus appropriée pour notre système de recherche d'images de documents arabes dégradés.

Comme nous avons dit précédemment, l'évaluation visuelle a été procédée pour tester la robustesse de l'approche proposée en examinant ses résultats selon un point de vue visuel humain. En comparant le classement obtenu par la méthodologie d'évaluation proposée avec le classement de l'évaluation visuelle, nous trouvons que les deux approches qui ont deux points de vue différents sont parfaitement conformes. Le même classement a été obtenu sur les deux ensembles de test. La concordance de notre approche avec l'évaluation visuelle constitue un point fort de la méthodologie proposée et montre sa robustesse car l'évaluation visuelle reste la plus convaincante puisqu'elle s'effectue par l'humain, le décideur idéal. Ce dernier dispose d'une faculté de compréhension à l'aide de laquelle il peut juger la qualité d'une image binarisée en tenant compte divers critères : la clarté du document, la lisibilité du texte, l'absence du bruit, etc.

Concernant la relation avec l'évaluation supervisée-pixel, nous remarquons que le classement obtenu en utilisant notre méthodologie d'évaluation ne s'accorde pas bien avec le classement obtenu de l'évaluation supervisée-pixel (en termes des 4 mesures : Fmeasure, PSNR, NRM et MPM). Cette dernière a fourni l'ordre suivant : 1- Sauvola et al., 2- Otsu, 3- Kapur et al., 4- Wolf et al., 5- Nick, 5- Binarisation en utilisant le minimum et le maximum locaux, 7- Bernsen, 8- Niblack. En effet, les deux approches d'évaluation ont deux points de vue totalement différents. Cependant, dans notre méthodologie d'évaluation, la pertinence d'un algorithme de binarisation est mesurée par sa capacité à préserver toutes les caractéristiques structurelles de l'écriture présente dans le document. Elle s'oppose à l'évaluation supervisée-pixel qui utilise des mesures statistiques basées sur le nombre de pixels bien classés sans tenir compte de leur position, le critère important dans le cas des images de document où le texte constitue l'essentiel de l'image. Ainsi, en utilisant les mesures supervisées-pixel précédentes, l'amincissement ou l'épaississement des caractères peuvent engendrer une détérioration considérable dans les performances de la méthode de binarisation appliquée. Pour mieux comprendre, la figure 4.29 présente une image de vérité terrain (de la lettre ρ) ainsi que deux autres images résultantes de sa dilatation et son érosion respectivement. Nous considérons ces

deux dernières images comme les résultats de binarisation en utilisant deux algorithmes différents.

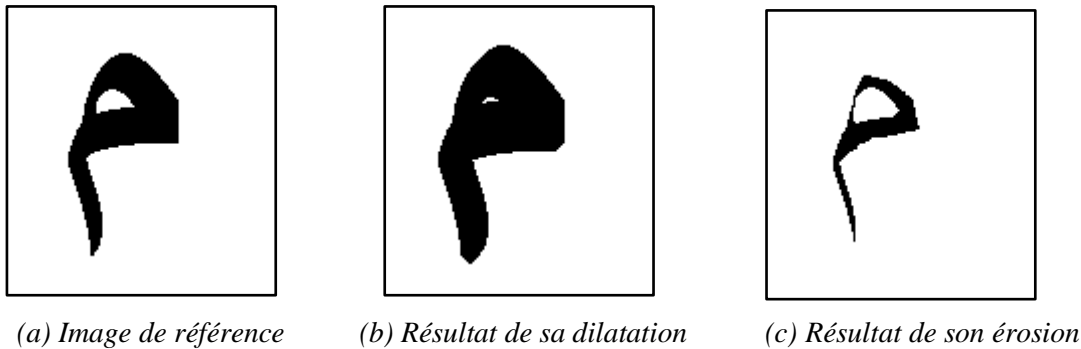


Figure 4. 29. Exemple d'une image de référence, de sa dilatation et de son érosion

L'estimation de la qualité des deux images binarisées à la fois par les mesures supervisées-pixel et en utilisant notre approche est donnée par le tableau 4.10.

Tableau 4. 10. Résultats d'évaluation (supervisée-pixel et en termes de distance d'édition) de la qualité des deux images de la figure 4.29

Image \ Critère	Distance d'édition	FM	PSNR	NRM ($\times 10^{-2}$)	MPM ($\times 10^{-3}$)
Image de la figure 4.29.b	0	0.7863	12.12117	3.458	0.1815
Image de la figure 4.29.c	0	0.5424	11.4949	31.392	0.23002

Comme nous pouvons remarquer du tableau 4.10, l'évaluation supervisée-pixel de la qualité des deux images binarisées a montré des performances faibles si on considère que les deux images sont lisibles, acceptables visuellement, et décrivent l'essentiel des informations contenues dans l'image de référence (le caractère م).

En utilisant notre approche, on parvient à extraire une boucle et un jambage des deux images binarisées et les assigne ainsi la signature **BJ** qui est identique à la signature de vérité terrain. Cependant malgré les déformations dues à l'épaississement et à l'amincissement, les deux images conservent la forme globale initiale du caractère et par conséquent ses caractéristiques structurales, les seules informations importantes dans la méthodologie proposée. Comme résultat, notre méthodologie considère que les images binarisées précédentes sont de bonne qualité avec une distance d'édition nulle.

4.3.4. Conclusion

Dans la deuxième partie du chapitre, nous avons présenté notre deuxième contribution à l'évaluation des techniques de binarisation. Cette deuxième contribution consiste à la proposition d'une nouvelle méthodologie d'évaluation des techniques de binarisation pour la recherche dans les images de documents arabes dégradés. La méthodologie d'évaluation proposée appartient à la classe de techniques d'évaluation au niveau-système qui estime la puissance d'un algorithme de binarisation pour une application particulière en analysant ses résultats finaux. La méthodologie proposée est basée sur la comparaison des caractéristiques structurales extraites des images binarisées avec des caractéristiques de vérité terrain préétablies, à la différence avec l'évaluation supervisée-pixel où la comparaison est effectuée entre l'image binarisée et l'image de vérité terrain au niveau pixel. Nous avons présenté et

détaillé le principe de la méthodologie proposée en illustrant les différentes étapes de traitement par des figures. L'approche proposée conjointement avec une évaluation visuelle et une évaluation supervisée-pixel a été utilisée pour estimer les performances de huit algorithmes de binarisation sur un ensemble d'images réelles et un ensemble d'images synthétiques. Les résultats obtenus montrent la robustesse de notre proposition et sa concordance totale avec l'évaluation visuelle humaine, et par contre, son incohérence avec l'évaluation supervisée-pixel à cause de la limitation de cette dernière au niveau bas d'évaluation (niveau pixel). Finalement, la méthode la plus performante selon notre méthodologie est la méthode de Sauvola et al. [SAU 00].

4.4. Conclusion du chapitre

Dans ce chapitre nous avons décrit nos deux contributions à l'évaluation des performances de méthodes de binarisation d'images de documents arabes dégradés. Une étude comparative large de 51 méthodes de binarisation évaluées en utilisant une combinaison de mesures quantitatives sur un ensemble d'images de documents historiques réels et synthétiques a été d'abord exposée. Après, nous avons présenté une nouvelle méthodologie d'évaluation des techniques de binarisation pour l'objectif général de la recherche dans les images de documents arabes dégradées. Plusieurs constatations et remarques ont été tirées des deux travaux précédents et les résultats obtenus sont promoteurs.

Chapitre 5 :
Contributions à la
binarisation d'images
de documents dégradés

5.1. Introduction

Dans les images de documents, le texte et les éléments de l'avant-plan constituent la part essentielle de l'image. Le texte et les autres éléments de l'avant-plan sont ainsi le cœur des systèmes d'analyse et de reconnaissance de documents. Pour cette raison, les applications de traitement et d'analyse d'images de documents s'occupent dans la plupart de temps de la séparation entre l'avant-plan et l'arrière-plan de l'image (*Foreground/ Background Separation*). L'objectif d'un algorithme de séparation avant-plan/ arrière-plan est d'extraire les informations pertinentes (texte, figures, tableaux, etc.), ou « l'avant-plan », à partir de l'image en éliminant le fond et le bruit considérés comme informations inutiles. La séparation avant-plan/ arrière-plan est communément appelée « binarisation » lorsque l'image résultante de la séparation soit composée uniquement de deux couleurs ; noir décrivant l'avant-plan, et blanc représentant le reste de l'image.

En fait, 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 lui-même, 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, l'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 nous présentons nos deux contributions à la séparation avant-plan/ arrière-plan dans les images de documents historiques. Dans chacune des deux contributions, une nouvelle méthode de binarisation est proposée. La première méthode s'agit d'une technique de binarisation basée sur la classification des pixels en « avant-plan » ou « arrière-plan » en utilisant un réseau de neurones de type MLP (Perceptron multi couches). La deuxième méthode est une technique de seuillage hybride qui combine plusieurs méthodes de seuillage dont le but d'avoir un résultat plus sûr. La suite de ce chapitre est consacrée à la présentation des deux méthodes proposées.

5.2. Séparation avant-plan/ arrière-plan dans les images de documents dégradés en utilisant un réseau de neurones artificiel

5.2.1. Motivations et objectifs

Nous avons vu dans le chapitre 2 que la binarisation est généralement effectuée à l'aide d'une valeur de seuil et que la littérature est riche en méthodes de binarisation basées sur le seuillage [PRA 10] [GAT 09] [SEZ 04] [KEF 09]. Ces méthodes utilisent différentes techniques pour calculer le seuil et elles sont regroupés en trois classes principales ; les méthodes globales lorsqu'un seuil unique est utilisé dans l'image entière, les méthodes locales lorsque un seuil est calculé pour une partie de l'image, et les méthodes hybrides qui combinent des informations globales et locales afin de trouver les valeurs optimales des seuils. Les méthodes précédentes sont dites méthodes non-supervisées [KUM 13]. D'autres méthodes de binarisation dites supervisées ont été proposées mais en nombre réduit comparativement aux méthodes non supervisées. Dans ces méthodes, la binarisation est effectuée à l'aide d'un

algorithme d'apprentissage automatique entraîné en se basant sur un ensemble d'apprentissage.

Parmi les techniques d'apprentissage automatique les plus connues, nous trouvons les réseaux de neurones artificiels (ANNs pour *Artificiel Neural Network*). Les ANNs ont contribué depuis leur introduction dans les années cinquante à résoudre des problèmes complexes dans de nombreux domaines y compris la classification, la prédiction, l'approximation de fonction, la reconnaissance des formes, la simulation de mémoires, etc. Le ANN est capable de généraliser, à partir d'un ensemble limité de données dites « d'apprentissage », son comportement pour des nouvelles données n'ont pas été vues auparavant. En dépit du succès des ANNs dans différentes applications de traitement d'images et de reconnaissance de formes, leur contribution dans des tâches plus antérieures comme la binarisation est encore limitée et n'a pas généré beaucoup de recherches [SEZ 04] [EGM 02].

L'objectif du présent travail est de tester la puissance des ANNs pour la séparation avant-plan/ arrière-plan dans les images de documents dégradés. Dans ce travail [KEF 14], nous exploitons la capacité de généralisation des ANNs pour concevoir une nouvelle méthode de binarisation. La méthode proposée ne repose pas sur le seuillage (comme la grande partie des méthodes de la littérature), mais sur la classification en utilisant un ANN de type *Perceptron multicouches* (MLP pour *Multilayer Perceptron*). Cette proposition peut être justifiée par le fait que la binarisation est en fait un processus de classification où les pixels de l'image sont séparés en deux classes «Noirs» et «Blancs», la tâche qui peut être accomplie efficacement par les ANNs.

L'avantage d'utiliser un réseau de neurones supervisé pour la séparation avant-plan/ arrière-plan est la réduction de la complexité des méthodes classiques de seuillage et de la disponibilité des bases de données importantes pour l'apprentissage et la validation.

5.2.2. Description de la technique proposée

Dans cette section, nous décrivons la nouvelle méthode de binarisation proposée. La méthode proposée peut être placée dans la classe des méthodes de binarisation supervisées qui utilisent des informations globales et locales. Dans cette méthode, la séparation des pixels de l'image en «noirs» ou «blancs» est assurée par un Perceptron multicouches (MLP) entraîné par l'algorithme de rétro-propagation du gradient (voir [KOT 07] pour plus de détails). Dans cette approche, le MLP ne calcule ou apprend aucun seuil mais effectue une binarisation directe en classifiant les pixels de l'image en deux classes. Comme nous l'avons décrit ci-dessus la binarisation est une sorte de problème de classification en deux classes.

Le MLP est un ANN avec un apprentissage supervisé, à savoir, l'algorithme de rétro propagation du gradient dédié spécialement pour des fins de classification. C'est le type de ANNs le plus utilisé et le plus productif au niveau d'application. Il a la possibilité de séparer des classes de formes non linéairement séparables. Nous avons choisi d'utiliser le MLP du fait que les niveaux de gris des pixels de l'avant-plan et de l'arrière-plan peuvent se chevaucher dans certains cas. Même si ce n'est pas toujours vrai, le MLP va souvent plus loin que beaucoup d'autres classificateurs [SMI 05].

(a)

collection de volumes dans les
mêmes conditions de
conservation



(b)

Livres historiques



(c)

Différentes pages du même
volume avec des dégradations
similaires

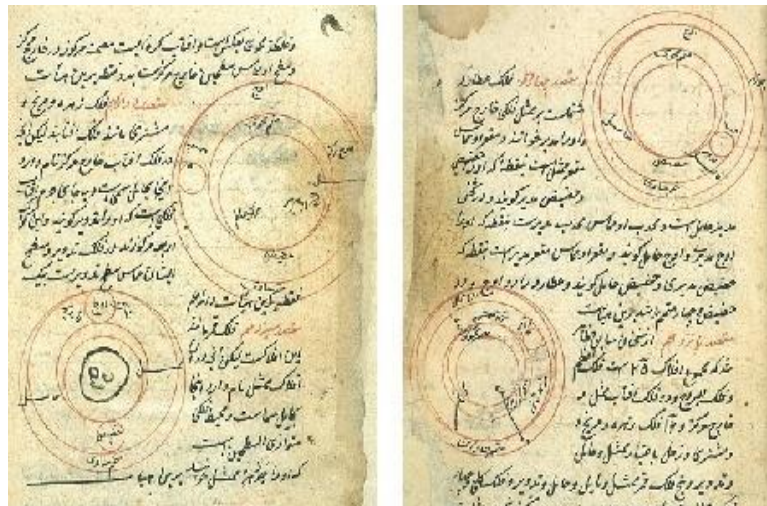


Figure 5.1. Livres et pages historiques

La motivation d'adopter une méthodologie basée sur un apprentissage supervisé est justifiée par le fait que, comme le montre la figure 5.1, les pages d'un même volume possèdent des défauts semblables car elles sont généralement rédigées par le même auteur, sur le même type de papier, dans la même période, et elles sont conservées dans les mêmes conditions. Comme résultat, un algorithme de binarisation aura les mêmes performances pour toutes les pages d'un même livre. Par conséquent, nous pouvons entraîner l'algorithme sur un nombre limité de pages, ou certaines zones spécifiques de différentes pages, et l'algorithme peut fonctionner correctement sur les autres pages. Ceci est effectué en apprenant la relation entre les exemples

appris et les nouvelles données, la tâche qui peut être effectuée avec excellence par le MLP. Pour les autres volumes, nous pouvons ré-entraîner le classificateur afin de l'adapter aux nouvelles données. Le même schéma est également possible pour une seule page; nous pouvons entraîner l'algorithme sur les données de certaines zones de la page et l'algorithme généralise son comportement sur le reste de la page.

Pour chaque pixel p_i appartenant à une image I_n de l'échantillon d'apprentissage, des informations globales et locales de son voisinage sont utilisées pour entraîner le MLP. Les informations locales sont la valeur de gris du pixel p_i avec celles de ses voisins dans une fenêtre $w \times w$ centrée sur p_i . Les informations globales sont la moyenne globale (M_n) et l'écart type global (S_n) de l'image I_n . En effet, comme il est noté dans [CHE 93], ces deux derniers et d'autres mesures statistiques de l'image sont des paramètres largement utilisés par les méthodes de binarisation, [OTS 79] et [KIT 86], par exemple, pour calculer les seuils. Le MLP devra ensuite sortir 0 ou 1 tels que 0 dénote le noir et 1 dénote le blanc. Voir la figure 5.2.

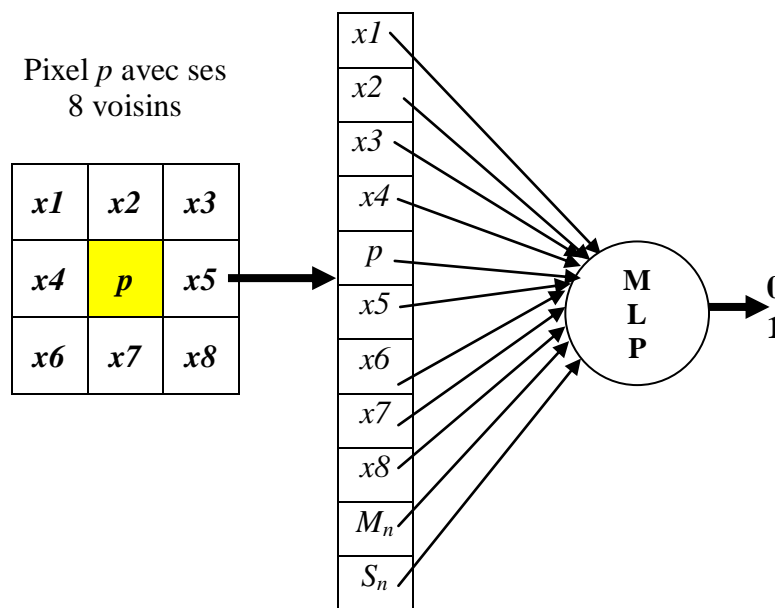


Figure 5. 2. MLP pour classifier un pixel p dans une fenêtre de 3×3 pixels

La méthode proposée se compose de quatre étapes: prétraitement de l'image, définition du MLP, apprentissage du MLP, et enfin binarisation. Le schéma fonctionnel de la méthode proposée est illustré dans la figure 5.3.

5.2.2.1. Prétraitement de l'image

Cette phase a comme objectif la préparation et l'extraction des données d'apprentissage (et de validation) à partir d'un échantillon d'images d'apprentissage et elle s'effectue hors ligne.

L'échantillon d'images d'apprentissage est composé d'un certain nombre d'images de documents dégradés et des images de vérité terrain associées.

5.2.2.1.1. Données d'apprentissage et de validation

Afin de garantir une forte précision, le MLP doit être entraîné sur un ensemble large de modèles dans le but de déterminer avec une grande exactitude ses paramètres libres (les poids des liens). C'est la propriété la plus intéressante des MLPs : la capacité de résoudre des problèmes très difficiles à expliciter ou à formaliser à condition de la disposition d'un ensemble d'exemples dit « ensemble d'apprentissage ».

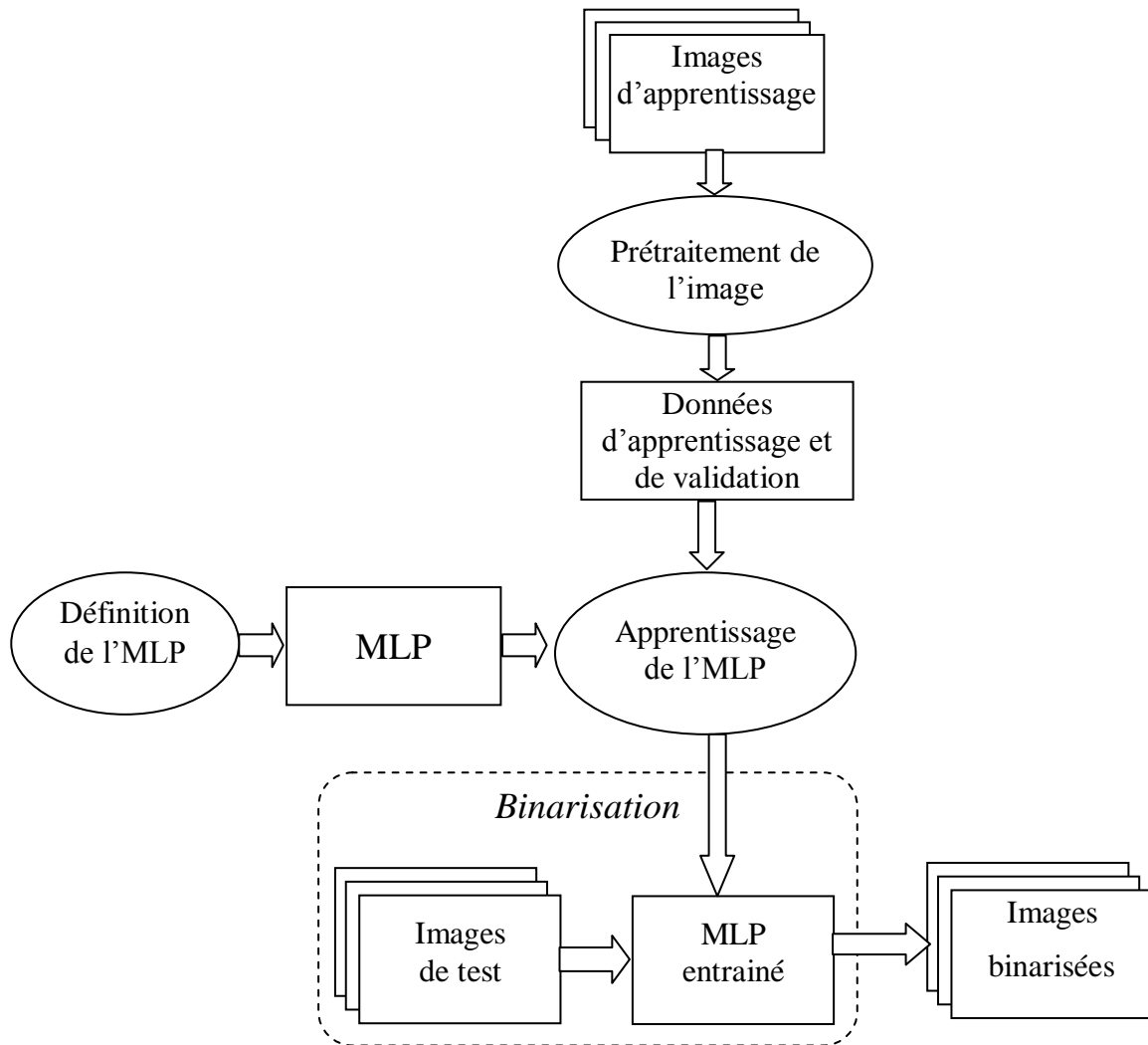


Figure 5. 3. Schéma général de notre première méthode de binarisation proposée

Pratiquement, l'utilisation d'un ensemble d'apprentissage seulement peut entraîner un phénomène dit de *sur-apprentissage* ; où le MLP apprend par cœur les données d'apprentissage mais il ne sera pas capable de généraliser le comportement pour d'autres données non observées durant la phase d'apprentissage. Pour surmonter ce problème, nous utilisons un autre ensemble de données dit « *ensemble de validation* » et durant l'apprentissage nous vérifions à chaque époque en plus de l'erreur d'apprentissage, l'erreur de validation et nous la comparons à la valeur précédente pour déterminer le moment de détérioration des performances.

En effet, le MLP est un ANN supervisé où le système est fourni avec des exemples d'entrées, et des valeurs désirées de sortie. Les données d'apprentissage et de validation doivent être donc composées de paires (X_i, Y_i) de vecteurs d'entrées X_i et des vecteurs de sorties désirées correspondantes Y_i , avec $1 \leq i \leq P$, et P dénote le nombre de pixels dans l'ensemble d'apprentissage.

Chaque vecteur d'entrées X_i représente le niveau de gris d'un pixel p_i appartenant à une image I_n avec les niveaux de gris de ses voisins dans une fenêtre $w \times w$ centrée sur p_i , en plus de la moyenne globale (M_n) et l'écart-type global (S_n) de l'image I_n . Le vecteur de sorties désirées

correspondant Y_i contient la valeur binaire b_i du pixel p_i dans l'image de référence. Ainsi le vecteur d'entrées d'un pixel p_i est de la forme suivante :

$X_i = \{g_i^1, g_i^2, \dots, g_i^{w \times w}, M_n, S_n\}$, avec g_i^j la valeur de gris du $j^{\text{ème}}$ pixel dans la fenêtre $w \times w$ centrée sur le pixel p_i . La figure 5.4 présente un exemple de tel vecteur pour un pixel (de niveau de gris 173) dans une fenêtre 3×3 .

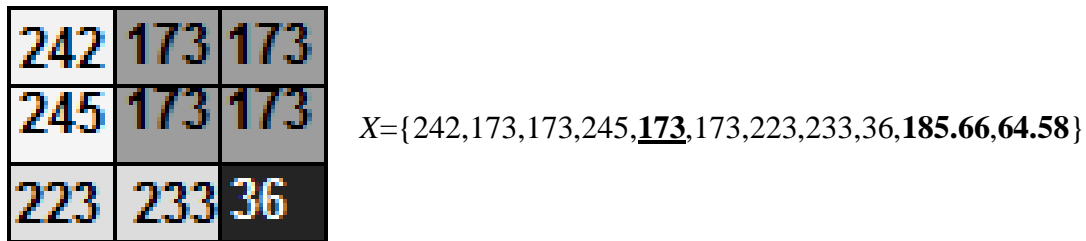


Figure 5. 4. Vecteur d'entrées pour un pixel dans une fenêtre 3×3

De même, le vecteur de sortie désirée associée est de la forme $Y_i = \{b_i\}$, avec $b_i = 0$ ou 1 .

L'étape de préparation des données d'apprentissage est schématisée par la partie ① de la figure 5.5.

Comme les niveaux de gris sont compris entre 0 et 255, les données d'apprentissage doivent être normalisées à l'intervalle $[0, 1]$ (0 correspond au noir et 1 au blanc).

5.2.2.1.2. Images d'apprentissage et de validation

Nous avons utilisé deux ensembles d'images pour la construction des données d'apprentissage et de validation. Le premier est un ensemble public composé d'images de documents provenant des quatre collections proposées dans le cadre des compétitions DIBCO 2009, H-DIBCO 2010, DIBCO 2011 et H-DIBCO 2012. Ces quatre collections contiennent un total de 50 images de documents réels (37 manuscrits et 13 imprimés) provenant des collections de plusieurs bibliothèques, avec les images de vérité terrain correspondantes. Toutes ces images contiennent des dégradations représentatives qui apparaissent fréquemment dans les documents historiques (taches d'encre et d'humidité, transparence, plis et déchirures, etc.). Des exemples d'images de ces collections sont montrés dans les figures 4.2 – 4.5 (Chapitre 4).

Le deuxième ensemble d'images est une collection synthétique préparée en vue d'inclure la plupart des dégradations qui peuvent être rencontrées dans les documents anciens. Elle a été créée en appliquant la technique de *Mosaicing* d'images par superposition proposée par P. Stathis et al. dans [STA 08]. L'idée est la suivante, nous commençons avec quelques images de documents en noirs et blancs, qui représentent les images de vérité terrain, et avec des fonds extraits des documents anciens et nous appliquons une procédure de fusion pour obtenir autant d'images différentes de documents historiques synthétiques. Pour plus de détails sur la procédure de fusion, veuillez voir la section 4.2.2.2 du chapitre 4. Un exemple d'une image de document historique synthétique est illustré par la figure 4.9.

Nous notons ici que nous avons utilisé un grand nombre de fonds avec différentes variétés de dégradations (effets de transparence, trous, taches, etc.) pour permettre au MLP d'apprendre à partir de la vaste diversité de cas possibles.

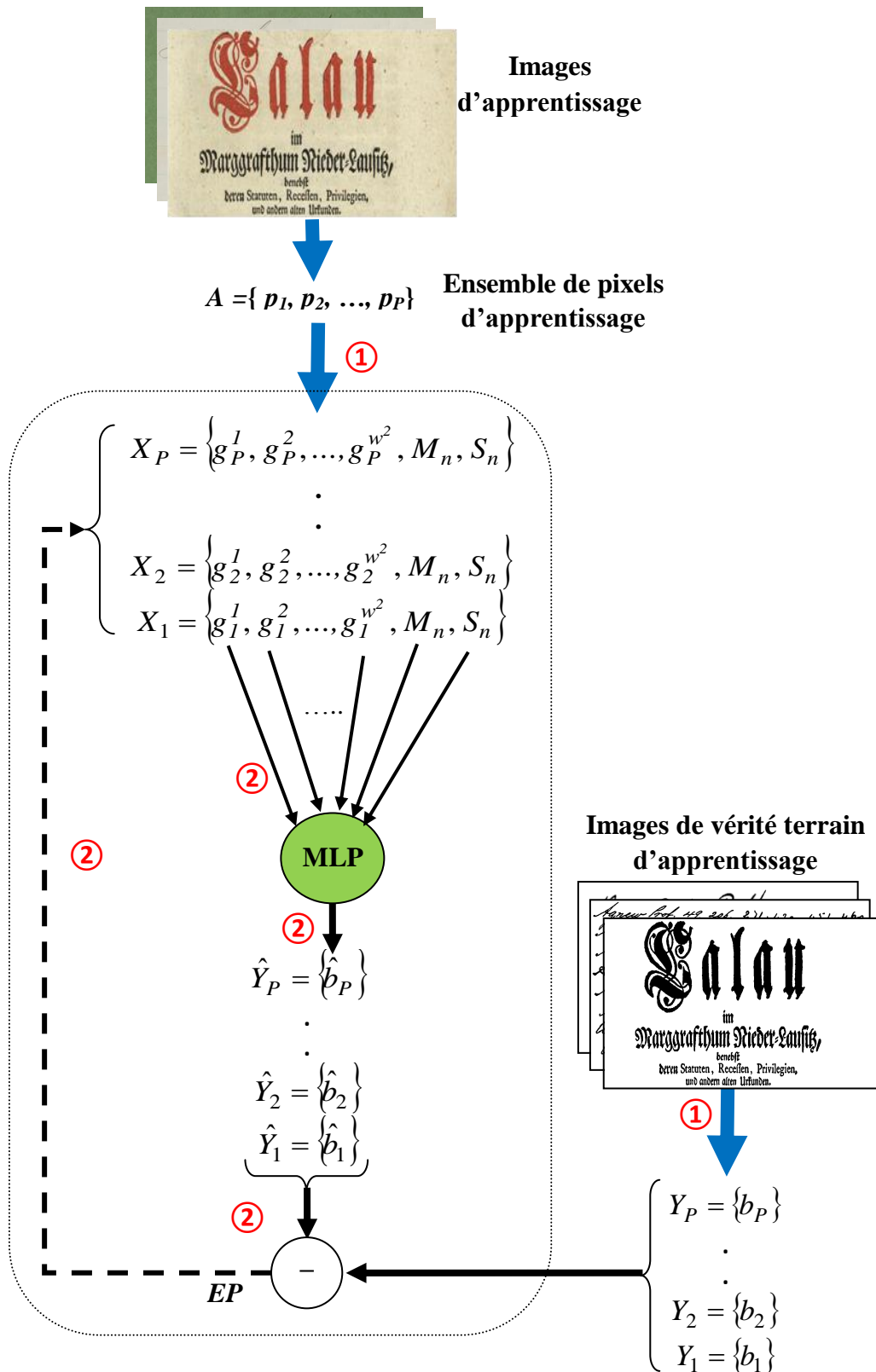


Figure 5. 5. Prétraitement et apprentissage du MLP

5.2.2.2. Définition du MLP

Après avoir préparé l'ensemble d'apprentissage et de validation, il est nécessaire de fixer la structure "adéquate" du ANN (le nombre de couches, le nombre de neurones dans chaque couche, le nombre d'entrées et de sorties, et la fonction d'activation de chaque neurone).

Comme nous avons dit, nous avons choisi d'utiliser un réseau de type MLP qui a montré sa fiabilité dans la résolution de plusieurs problèmes. Malgré l'importance de la détermination de la topologie optimale d'un MLP pour un problème donné, il n'est pas toujours facile de la concevoir et même pas toujours justifiée.

Kolmogorov [KAL 57] a montré que toute fonction de n variables possède une représentation exacte en termes de superpositions et compositions finies d'un petit nombre de fonctions d'un seul variable [CYB 89]. Cela signifie que toute fonction continue de plusieurs variables peut être simulée par un ANN avec deux couches cachées [GIR 89]. En plus, selon G.V. Cybenko dans [CYB 89] « *un réseau MLP qui possède uniquement une couche cachée est capable d'approximer presque tout types de fonctions non linéaires* ».

Ainsi, nous avons choisi l'architecture optimale de notre MLP par expérimentations (par essai - erreur). Nous avons conçu un MLP avec un seul neurone dans la couche de sortie correspondant à la valeur binaire du pixel (0 pour noir ou 1 pour blanc). Le nombre d'entrées est le nombre de pixels dans la fenêtre (9 pour une fenêtre 3×3, 25 pour 5×5, etc.), en plus de la moyenne (M) et l'écart type (S) de l'image entière. La fonction d'activation choisie est la fonction Sigmoidé définie par :

$$f(x) = \frac{1}{1 + e^{-ax}}$$

Cependant, une question reste : quelle taille de fenêtre va donner les meilleurs résultats ?

Pour répondre à cette question, nous avons testé plusieurs tailles de fenêtre avec différents ANNs et différents nombres de couches, et neurones dans chaque couche ; entraîné toutes ces configurations et évalué pour chaque configuration ses performances sur la base de test, en utilisant plusieurs mesures d'évaluation. Les résultats obtenus montrent que la configuration optimale est :

- Taille de fenêtre égale à 3×3 et donc le nombre d'entrées est 11,
- Une seule couche cachée de 11 neurones,
- Un seul neurone comme sortie.

5.2.2.3. Apprentissage du MLP

D'après [DRE 02], on appelle « apprentissage » des réseaux de neurones la procédure qui consiste à estimer les paramètres des neurones du réseau, à fin que celui-ci remplisse au mieux la tâche qui lui est assignée.

Dans cette étape, le MLP est entraîné en utilisant les données d'apprentissage et de validation préparées précédemment. L'algorithme utilisé pour l'apprentissage du MLP est l'algorithme de *rétro-propagation du gradient*. La mise à jours des poids se fait dans le but de minimiser l'erreur entre les sorties du MLP (sorties estimées) et les données réelles (sorties désirées).

Le processus d'apprentissage se fait de la manière suivante (partie ② de la figure 5.5): le MLP reçoit tous les vecteurs d'apprentissage comme entrées et calcule l'erreur d'apprentissage (EP) entre les sorties du MLP (sorties estimées correspondant aux valeurs binaires estimées des pixels) et les sorties désirées (valeurs binaires réelles). Une mise à jour

de poids est ensuite effectuée pour tous les neurones du MLP. Après, le MLP prend comme entrées les vecteurs de validation et calcule l'erreur de validation (EV) comme précédemment, mais sans mettre à jour les poids des neurones. Le processus est répété avec les nouveaux poids en vue de minimiser à la fois l'erreur d'apprentissage et l'erreur de validation. L'apprentissage est arrêté lorsque l'erreur de validation commence à accroître. Le MLP obtenu avec les poids finaux est enregistré afin d'être utilisé dans la dernière phase d'exploitation.

Après l'apprentissage, le MLP apprend à fournir la valeur correcte de sortie quand il reçoit les valeurs d'entrée seulement.

5.2.2.4. Binarisation

Une fois l'apprentissage est terminé, la dernière phase de la méthode proposée est l'utilisation du MLP entraîné pour binariser d'autres images de documents différentes de celles utilisées durant l'apprentissage ou la validation. La binarisation d'une image I contenant R pixels se fait en faisant exécuter le MLP entraîné avec une seule passe en avant pour chaque pixel de l'image.

Ainsi, pour chaque pixel p_i de l'image I , avec $1 \leq i \leq R$, nous fournissons au MLP entraîné le vecteur de caractéristiques de p_i comme dans la phase d'apprentissage et le MLP devra faire sortir une valeur numérique correspondant à la valeur binaire du pixel : $\hat{Y}_i = \{\hat{b}_i\}$. Voir la Figure 5.6.

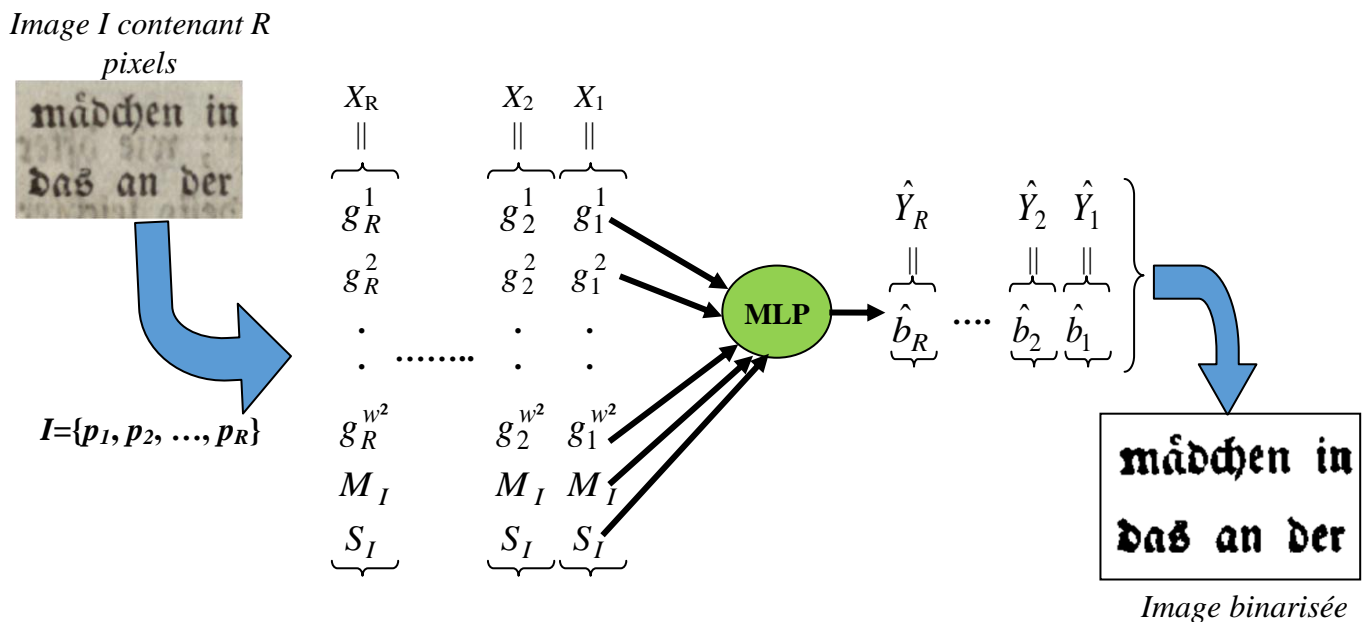


Figure 5. 6. Utilisation du MLP entraîné pour binariser une nouvelle image

5.2.3. Expérimentations et résultats

Notre application a été développée en langage de programmation Java. Pour la manipulation des réseaux de neurones nous avons utilisé la bibliothèque libre et open source *Neuroph*¹⁰. Cette bibliothèque contient la majorité des architectures des réseaux de neurones et des algorithmes d'apprentissage standards.

¹⁰ <http://neuroph.sourceforge.net>

Comme nous avons dit précédemment, nous avons utilisé deux ensembles d'images pour l'apprentissage et la validation. Le premier ensemble contient 50 images de documents réels prises des collections publiques utilisées dans le contexte des compétitions internationales de binarisation DIBCO 2009, H-DIBCO 2010, DIBCO 2011 et H-DIBCO 2012. Le second ensemble est composé de 240 images de documents synthétiques construites par notre algorithme de fusion détaillé dans la section 4.2.2.2 à partir de 20 images de fonds et 12 images binaires de textes.

Une série d'expérimentations a été réalisée. Dans la première expérimentation, seules les images des quatre collections publiques (DIBCO 2009, H-DIBCO 2010, DIBCO 2011, H-DIBCO 2012) ont été utilisées pour l'apprentissage, la validation, et le test. Les résultats obtenus ont été comparés uniquement avec les résultats publiés des compétitions DIBCO 2009 [GAT 09], H-DIBCO 2010 [PRA 10], DIBCO 2011 [PRA 11], et H-DIBCO 2012 [PRA 12] afin d'estimer les performances de notre méthode vis-à-vis des méthodes participantes aux compétitions.

Malheureusement, le premier ensemble (composé d'images de DIBCO/H-DIBCO) n'est pas suffisamment large pour permettre une estimation réelle des performances. Afin d'avoir une vue plus claire, nous avons mené une deuxième expérimentation sur l'ensemble plus large d'images synthétiques en plus de l'ensemble d'images publiques. Comme nous pouvons observer clairement, les algorithmes participants aux compagnes DIBCO/H-DIBCO sont loin d'être les plus populaires dans le domaine, plusieurs autres techniques existent [SEZ 04]. Ainsi, nous avons sélectionné un nombre de méthodes bien connues de la littérature pour la comparaison.

L'évaluation a été effectuée quantitativement en utilisant des mesures standards largement utilisées, en particulier dans les compétitions DIBCO et H-DIBCO. Ces mesures sont: FMeasure (*FM*), PSNR, Pseudo FMeasure (*p-FM*), NRM, MPM, et DRD.

Comme l'évaluation quantitative supervisée-pixel ne reflète pas toujours la qualité réelle des images binarisées et leur appropriation pour une application particulière, la troisième expérimentation a été mise en place. Elle a eu comme objectif de vérifier l'adéquation de la méthode proposée aux applications de recherche dans les images de documents arabes dégradés. Nous avons utilisé pour cela notre méthodologie d'évaluation présentée dans le chapitre 4, section 4.3.

5.2.3.1. Première expérimentation

Dans cette expérimentation, nous avons utilisé les collections publiques de DIBCO et H-DIBCO pour l'apprentissage, la validation et le test.

Comme premier essai, nous avons essayé d'entraîner le MLP en utilisant toutes les 36 images des collections DIBCO 2009, H-DIBCO 2010, et DIBCO 2011, et de tester la capacité de généralisation du MLP sur les images de la collection H-DIBCO 2012 uniquement. Nous avons ainsi utilisé 26 images pour l'apprentissage et 10 pour la validation. Nous avons également utilisé tous les pixels de toutes les images d'apprentissage et de validation. Cela n'était pas pratique car les images sont de grandes tailles (**651479** pixels en moyenne) ce qui résulte un ensemble massif d'apprentissage et de validation d'à peu près **16938454** et **6514790** vecteurs respectivement.

Dans la deuxième tentative, nous avons utilisé toutes les 36 images des 3 collections précédentes à la fois pour l'apprentissage et la validation, mais nous avons opté la sélection d'une partie des vecteurs uniquement. Ainsi, nous avons sélectionné aléatoirement **1000**

vecteurs à partir de chaque image pour l'apprentissage et **500** autres vecteurs pour la validation résultant **36000** vecteurs d'apprentissage et **18000** vecteurs de validation.

Après, nous avons lancé l'apprentissage avec ces données en calculant à chaque époque (Une fois que chaque exemple de l'ensemble d'apprentissage est présenté au MLP), l'erreur d'apprentissage et l'erreur de validation. Comme il est connu, l'erreur d'apprentissage continue à se décroître au fur et à mesure de l'apprentissage mais l'erreur de validation décroît au début puis va commencer à se détériorer ce qui signifie qu'un sur-apprentissage a eu lieu et nous devons arrêter l'apprentissage. Pour notre cas, cela a été produit à l'époque **22714** (figure 5.7).

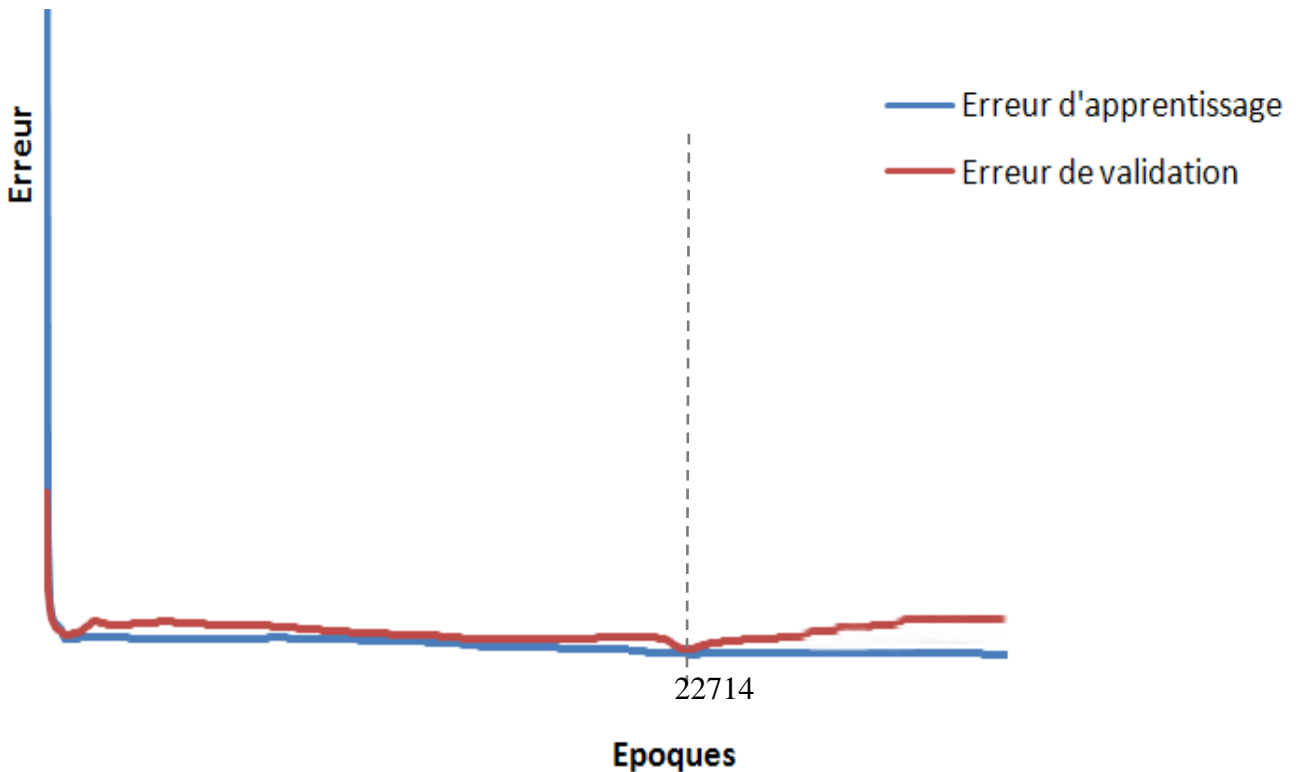


Figure 5. 7. Erreurs d'apprentissage et de validation du MLP

Après l'apprentissage, nous avons testé notre MLP sur les 36 images des collections DIBCO 2009, H-DIBCO 2010, et DIBCO 2011 utilisées dans l'apprentissage et nous avons comparé les images binaires résultantes avec les images de référence associées en utilisant les mesures d'évaluation précitées. Comme uniquement 1000 pixels de chaque image ont été appris, cela nous permet d'évaluer la capacité de généralisation du MLP sur le reste de l'image (pour les pixels non appris). Nous avons également testé le MLP sur les 14 images de la collection H-DIBCO 2012 qui n'ont pas été présentées au MLP auparavant. Les résultats moyens obtenus en appliquant notre méthode sur chacune des quatre collections sont résumés dans le tableau 5.1.

En comparant les résultats obtenus avec ceux publiés par les quatre compétitions dans [GAT 09], [PRA 10], [PRA 11], et [PRA 12] respectivement, nous pouvons déduire que notre méthode a été classée la troisième sur la collection DIBCO 2009, la deuxième sur la collection H-DIBCO 2010, la deuxième sur la collection DIBCO 2011, et la quatrième sur la collection H-DIBCO 2012. Même si notre méthode n'a pas fourni les meilleures performances, nous admettons que les résultats obtenus sont prometteurs. Ces résultats montrent que le MLP a bien appris la relation entre les caractéristiques décrivant chaque pixel et la valeur binaire

optimale de ce denier. Par conséquent, le MLP a réussi à généraliser son comportement (la détermination de la valeur binaire d'un pixel) pour d'autres pixels à la fois à l'intérieur des images utilisées dans l'apprentissage et pour d'autres images n'ont pas été apprises, comme elles ont approximativement les mêmes caractéristiques que les premières.

Tableau 5. 1. Résultats moyens obtenus du 1^{er} MLP sur les quatre collections publiques d'images

Mesure Collection	FM	p-FM	PSNR	NRM ($\times 10^{-2}$)	MPM ($\times 10^{-3}$)	DRD
DIBCO 2009	0.8967	0.9266	18.0218	4.98	0.7712	1.3256
H-DIBCO 2010	0.9084	0.9482	18.9795	4.13	0.5592	2.1570
DIBCO 2011	0.8468	0.8804	17.7204	8.48	0.6017	3.6680
H-DIBCO 2012	0.9078	0.9377	20.0168	5.01	0.5877	2.9802

5.2.3.2. Deuxième expérimentation

Pour une évaluation plus robuste de notre méthode, nous avons testé le MLP (sans ré-apprentissage) sur l'ensemble plus large d'images synthétiques composé de 240 images de documents. Les résultats moyens obtenus sont affichés dans le tableau 5.2.

Tableau 5. 2. Résultats moyens obtenus par le 1^{er} MLP sur la collection d'images synthétiques

	FM	p-FM	PSNR	NRM ($\times 10^{-2}$)	MPM ($\times 10^{-3}$)	DRD
Collection synthétique	0.5871	0.6365	11.1901	22.03	5.9437	8.7390

En effet, il n'est pas surprenant que les performances ont été diminuées considérablement comparativement à celles obtenues sur les collections publiques. Cela est dû aux différences remarquables entre les dégradations existantes dans les images synthétiques et celles dans les images réelles composant l'ensemble d'apprentissage ce qui a influencé la capacité de généralisation du MLP.

Pour pallier ce problème, nous avons décidé d'inclure plus de diversité de pixels dans la phase d'apprentissage. Cependant, nous avons utilisé dans la deuxième expérimentation les deux ensembles d'images : public de DIBCO/H-DIBCO et synthétique. Les images ont été réparties comme suit : 15 images des collections publiques et 70 images de la collection synthétique pour l'apprentissage, 10 images des collections publiques et 30 images de la collection synthétique pour la validation. Le reste est laissé pour le test. Comme il n'est pas pratique d'intégrer tous les pixels de toutes les images, nous avons considéré seulement une portion des pixels pour les utiliser dans l'apprentissage et la validation. Ainsi, nous avons sélectionné au hasard **500** pixels de chaque image d'apprentissage et **500** pixels de chaque image de validation ce qui nous a fournit un total de **42500** vecteurs d'apprentissage et **20000** vecteurs de validation.

Après que l'apprentissage a été terminé à l'époque **25382**, nous avons évalué les performances du nouveau MLP entraîné, à la fois sur les quatre collections publiques et sur la collection synthétique afin de les comparer avec les performances obtenues par l'ancien MLP. Les résultats obtenus sont résumés dans le tableau 5.3.

Tableau 5. 3. Résultats moyens obtenus par le 2^{ème} MLP sur les collections publiques et synthétique

Mesure Collection	FM	p-FM	PSNR	NRM ($\times 10^{-2}$)	MPM ($\times 10^{-3}$)	DRD
DIBCO 2009	0.8715	0.8966	16.1254	5.25	0.799	1.7203
H-DIBCO 2010	0.8798	0.9081	17.0018	4.727	0.6508	2.9851
DIBCO 2011	0.8403	0.8704	16.3280	7.16	1.0695	3.842
H-DIBCO 2012	0.8824	0.9013	18.9844	5.34	1.7456	2.6640
Collection Synthétique	0.7751	0.8418	16.0757	10.39	1.1084	3.7845

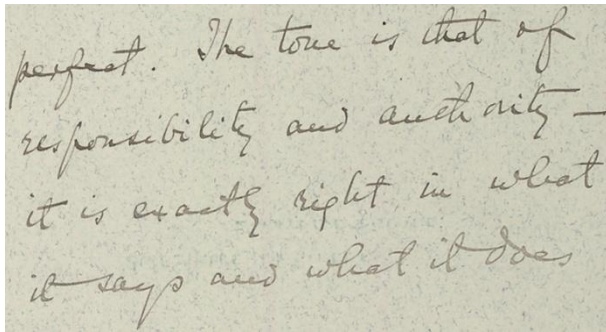
En comparant ces résultats avec ceux présentés dans les tableaux 5.1 et 5.2, nous trouvons que les performances obtenues par le nouveau MLP ont été légèrement diminuées par rapport aux performances du premier MLP sur les collections publiques. La diminution des performances se rapporte à une réduction de la capacité de généralisation du nouveau MLP comparativement à l'ancien. Pour la collection synthétique, les performances obtenues avec le nouveau MLP se sont augmenté considérablement comparativement à celles du premier MLP. Cela est dû principalement à la richesse de l'ensemble d'apprentissage utilisée pour l'entraînement du nouveau MLP et qui inclue une variété de pixels appartenant aux images des collections publiques et aux images synthétiques ayant des caractéristiques différentes. De ce fait la probabilité qu'un nouveau pixel à binariser ressemble à d'autres pixels vus durant l'apprentissage a été augmentée, et par conséquent le MLP arrive à déterminer avec plus de précision la valeur binaire correspondante. Ces résultats prouvent l'efficacité de la méthode proposée et confirment notre supposition.

Les résultats visuels de binarisation de trois images des collections publiques en utilisant le nouveau MLP sont illustrés dans la figure 5.8.

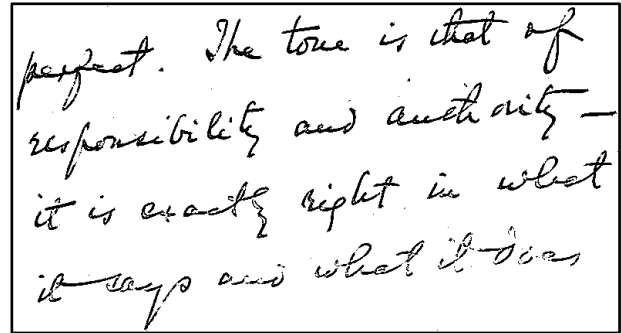
Nous avons également comparé notre méthode avec plusieurs techniques bien connues de la littérature pour la binarisation d'images de documents [SEZ 04] [STA 08b] [KEF 09].

Les résultats moyens obtenus avec toutes les méthodes comparées sur le premier et le deuxième ensemble de test sont résumés dans les tableaux 5.4 et 5.5 respectivement. Les résultats moyens entre les deux ensembles sont présentés dans le tableau 5.6.

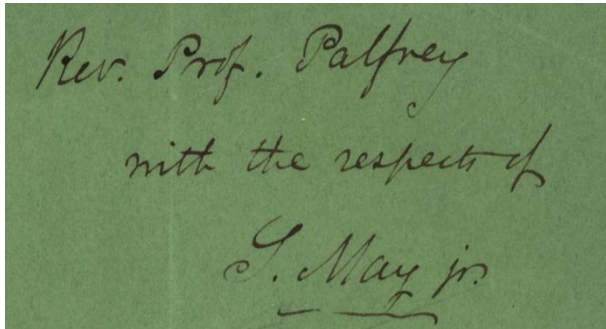
Le classement final des méthodes comparées est affiché dans le tableau 5.7, qui récapitule également les rangs partiels de chaque méthode selon chaque mesure d'évaluation et la somme des rangs. Ensuite, nous fournissons dans les figures 5.9 – 5.13 des graphes représentant les performances moyennes des méthodes comparées en termes de chaque mesure séparément.



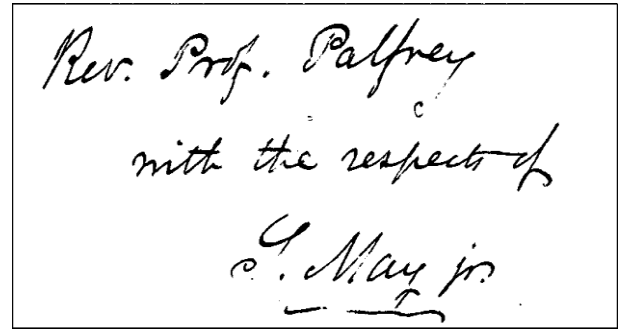
(a) Image de H-DIBCO 2010



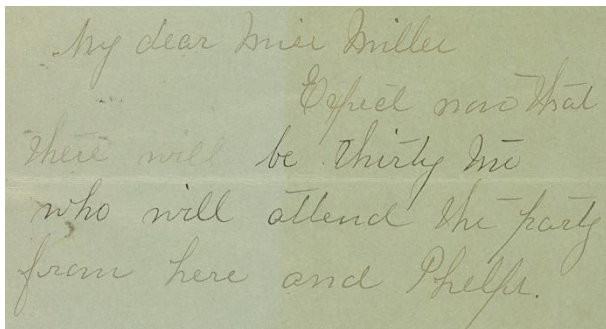
(b) Image binarisée correspondante



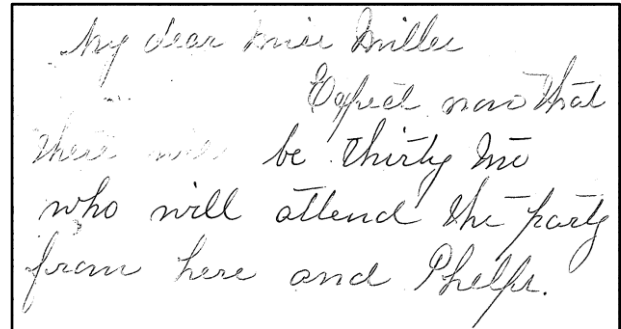
(c) Image de DIBCO 2011



(d) Image binarisée correspondante



(e) Image de H-DIBCO 2012



(f) Image binarisée correspondante

Figure 5. 8. Résultats de binarisation de quelques images publiques en utilisant le 2^{ème} MLP

Tableau 5. 4. Résultats moyens des différentes méthodes de binarisation sur l'ensemble d'images publiques

Méthode	Mesure	FM	PSNR	NRM ($\times 10^{-2}$)	MPM ($\times 10^{-3}$)	DRD
Otsu [OTS 79]		0.8266	15.886	6.28	1.915	1.755
ISODATA [VEL 80]		0.8214	15.734	6.32	1.938	1.833
Seuillage par erreur minimale [KIT 86]		0.4673	8.3621	16.3	97.698	26.13
Kapur et al. [KAP 85]		0.8202	15.093	5.09	2.673	2.161
1 ^{ère} méthode de Mello et Lins [MEL 00]		0.6496	13.713	20.55	3.465	3.051
2 ^{ème} méthode de Mello et al. [MEL 08]		0.5203	13.092	27.12	0.901	3.092
Seuillage global itératif (IGT) [KAV 05]		0.7622	13.683	11.62	17.347	4.215
Niblack [NIB 86]		0.5321	7.7542	12.59	83.005	17.85
Sauvola et Pietikainen [SAU 00]		0.8360	16.387	7.38	2.486	1.634

Nick [KHU 09]	0.7764	15.477	9.77	15.542	4.116
Feng et Tan [FEN 04]	0.7798	15.148	8.2	22.608	5.016
Bernsen [BER 86]	0.5150	8.1056	18.07	109.36	19.62
Technique du moyen-gradient [LEE 03]	0.7914	13.872	10.09	2.631	2.834
Cavalcanti et al. [CAV 06]	0.5926	12.928	31.17	0.807	4.255
Tabatabaie et Bohlool [TAB 10]	0.8002	15.660	6.71	24.013	5.309
Gangamma et Srikanta Murthy [GAN 11]	0.6714	14.317	17.65	41.804	6.549
IGT améliorée [KAV 06]	0.7530	13.695	13.64	13.784	3.751
Binarisation en utilisant le minimum et le maximum locaux [SU 10]	0.7891	14.342	9.71	22.913	5.439
Méthode proposée	0.8362	16.328	7.19	7.248	1.970

Tableau 5. 5. Résultats moyens des différentes méthodes de binarisation sur l'ensemble d'images synthétiques

Méthode	Mesure	FM	PSNR	NRM ($\times 10^{-2}$)	MPM ($\times 10^{-3}$)	DRD
Otsu [OTS 79]		0.6245	15.101	15.8	18.783	4.6768
ISODATA [VEL 80]		0.6153	15.059	16.66	19.31	4.7320
Seuillage par erreur minimale [KIT 86]		0.4239	6.2878	21.63	177.27	28.682
Kapur et al. [KAP 85]		0.6218	15.094	15.96	23.474	7.6726
1 ^{ère} méthode de Mello et Lins [MEL 00]		0.3574	12.708	33.14	7.512	4.3837
2 ^{ème} méthode de Mello et al. [MEL 08]		0.3367	11.532	45.47	3.561	3.9472
Seuillage global itératif (IGT) [KAV 05]		0.6049	13.637	17.11	23.889	4.0455
Niblack [NIB 86]		0.3996	7.5398	19.37	94.157	17.059
Sauvola et Pietikainen [SAU 00]		0.6818	16.130	15.35	7.63	2.2288
Nick [KHU 09]		0.6778	13.882	16.56	9.653	2.5658
Feng et Tan [FEN 04]		0.6744	13.826	15.76	10.867	2.7634
Bernsen [BER 86]		0.4081	8.1610	20.69	87.728	14.517
Technique du moyen-gradient [LEE 03]		0.6023	12.832	18.58	12.63	3.7058
Cavalcanti et al. [CAV 06]		0.3443	12.815	34.54	8.949	4.0059
Tabatabaie et Bohlool [TAB 10]		0.6377	14.744	15.72	22.36	3.6973
Gangamma et Srikanta Murthy [GAN 11]		0.6037	12.879	21.79	10.963	2.4942
IGT améliorée [KAV 06]		0.5980	13.486	19.21	18.568	3.7286
Binarisation en utilisant le minimum et le maximum locaux [SU 10]		0.5713	11.622	18.35	13.603	4.5757
Méthode proposée		0.7208	15.396	14.96	5.333	2.9860

Tableau 5. 6. Résultats moyens des différentes méthodes de binarisation entre les deux ensembles

Méthode	Mesure	FM	PSNR	NRM ($\times 10^{-2}$)	MPM ($\times 10^{-3}$)	DRD
Otsu [OTS 79]		0.72555	15.4935	11.04	10.349	3.2159
ISODATA [VEL 80]		0.71835	15.3965	11.49	10.624	3.2825
Seuillage par erreur minimale [KIT 86]		0.4456	7.32495	18.965	137.484	27.406
Kapur et al. [KAP 85]		0.721	15.0935	10.525	13.0735	4.9168
1 ^{ère} méthode de Mello et Lins [MEL 00]		0.5035	13.2105	26.845	5.4885	3.71735
2 ^{ème} méthode de Mello et al. [MEL 08]		0.4285	12.312	36.295	2.231	3.5196
Seuillage global itératif (IGT) [KAV 05]		0.68355	13.66	14.365	20.618	4.13025
Niblack [NIB 86]		0.46585	7.647	15.98	88.581	17.4545
Sauvola et Pietikainen [SAU 00]		0.7589	16.2585	11.365	5.058	1.9314
Nick [KHU 09]		0.7271	14.6795	13.165	12.5975	3.3409
Feng et Tan [FEN 04]		0.7271	14.487	11.98	16.7375	3.8897
Bernsen [BER 86]		0.46155	8.1333	19.38	98.544	17.0685
Technique du moyen-gradient [LEE 03]		0.69685	13.352	14.335	7.6305	3.2699
Cavalcanti et al. [CAV 06]		0.46845	12.8715	32.855	4.878	4.13045
Tabatabaie et Bohlool [TAB 10]		0.71895	15.202	11.215	23.1865	4.50315
Gangamma et Srikanta Murthy [GAN 11]		0.63755	13.598	19.72	26.3835	4.5216
IGT améliorée [KAV 06]		0.6755	13.5905	16.425	16.176	3.7398
Binarisation en utilisant le minimum et le maximum locaux [SU 10]		0.6802	12.982	14.03	18.258	5.00735
Méthode proposée		0.7785	15.862	11.075	6.2905	2.478

Tableau 5. 7. Classement final des méthodes comparées sur les deux ensembles de test

Rang	Méthode	FM	PSNR	NRM	MPM	DRD	Somme des rangs
1	Sauvola et Pietikainen [SAU 00]	2	1	5	3	1	12
2	Méthode proposée	1	2	3	5	2	13
3	Otsu [OTS 79]	5	3	2	7	3	20
4	ISODATA [VEL 80]	8	4	6	8	5	31
5	Nick [KHU 09]	3	7	8	9	6	34
6	Kapur et al. [KAP 85]	6	6	1	10	15	38
7	Feng et Tan [FEN 04]	3	8	7	12	10	40
8	Technique du moyen-gradient [LEE 03]	9	12	10	6	4	41

9	Tabatabaie et Bohlool [TAB 10]	7	5	4	15	13	44
10	Seuillage global itératif (IGT) [KAV 05]	10	9	11	14	11	55
11	1 ^{ère} méthode de Mello et Lins [MEL 00]	14	13	17	4	8	56
12	IGT améliorée [KAV 06]	12	11	13	11	9	56
13	2 ^{ème} méthode de Mello et al. [MEL 08]	19	16	19	1	7	62
14	Cavalcanti et al. [CAV 06]	15	15	18	2	12	62
15	Binarisation en utilisant le minimum et le maximum locaux [SU 10]	11	14	9	13	16	63
16	Gangamma et Srikanta Murthy [GAN 11]	13	10	16	16	14	69
17	Niblack [NIB 86]	16	18	12	17	18	81
18	Bernsen [BER 86]	17	17	15	18	17	84
19	Seuillage par erreur minimale [KIT 86]	18	19	14	19	19	89

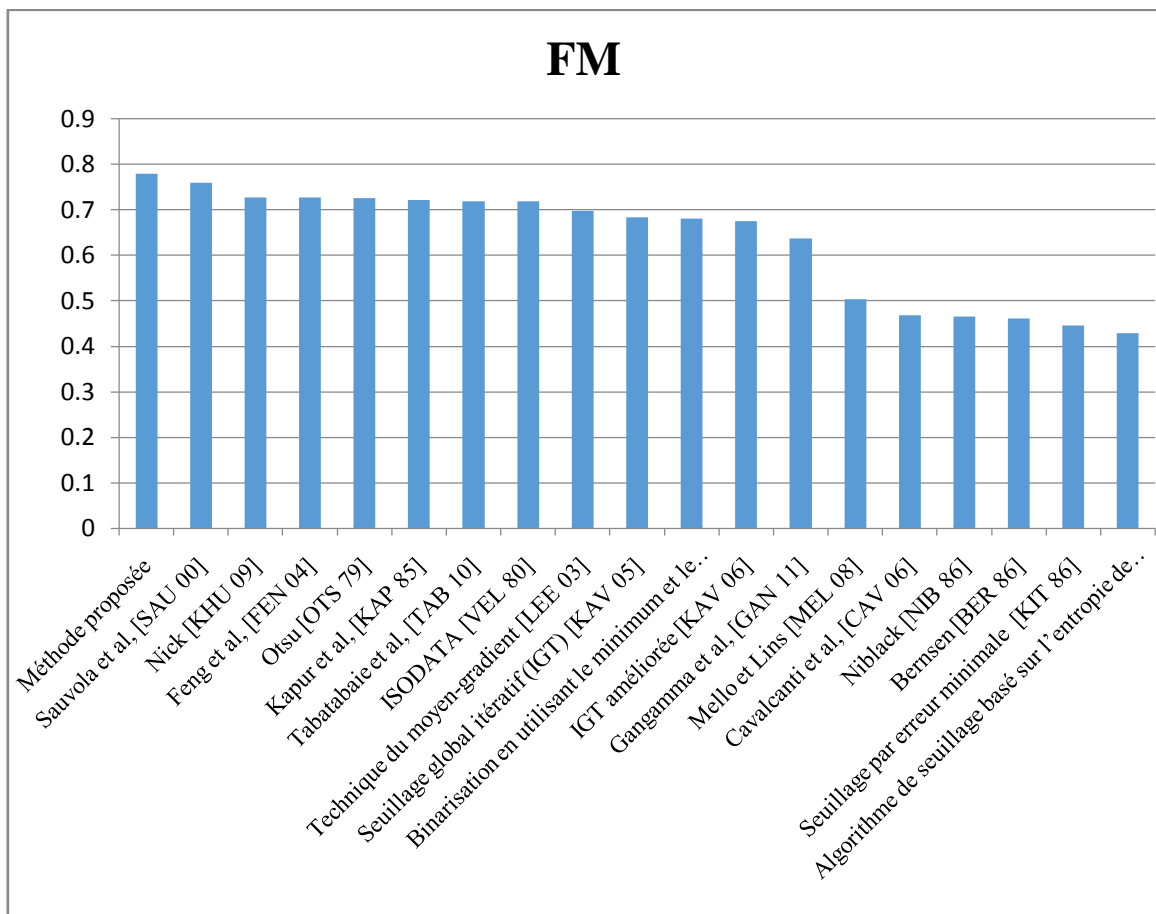


Figure 5. 9. Performances des méthodes de binarisation comparées en termes de FMeasure

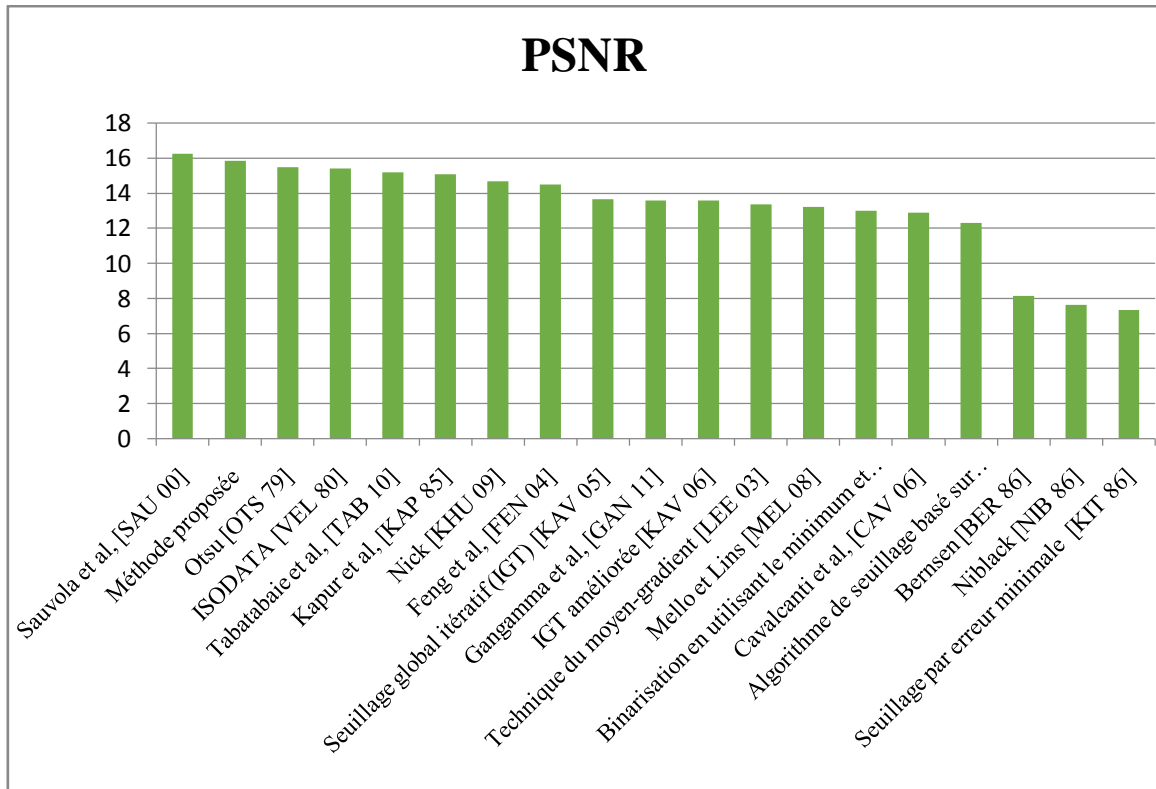


Figure 5. 10. Performances des méthodes de binarisation comparées en termes de PSNR

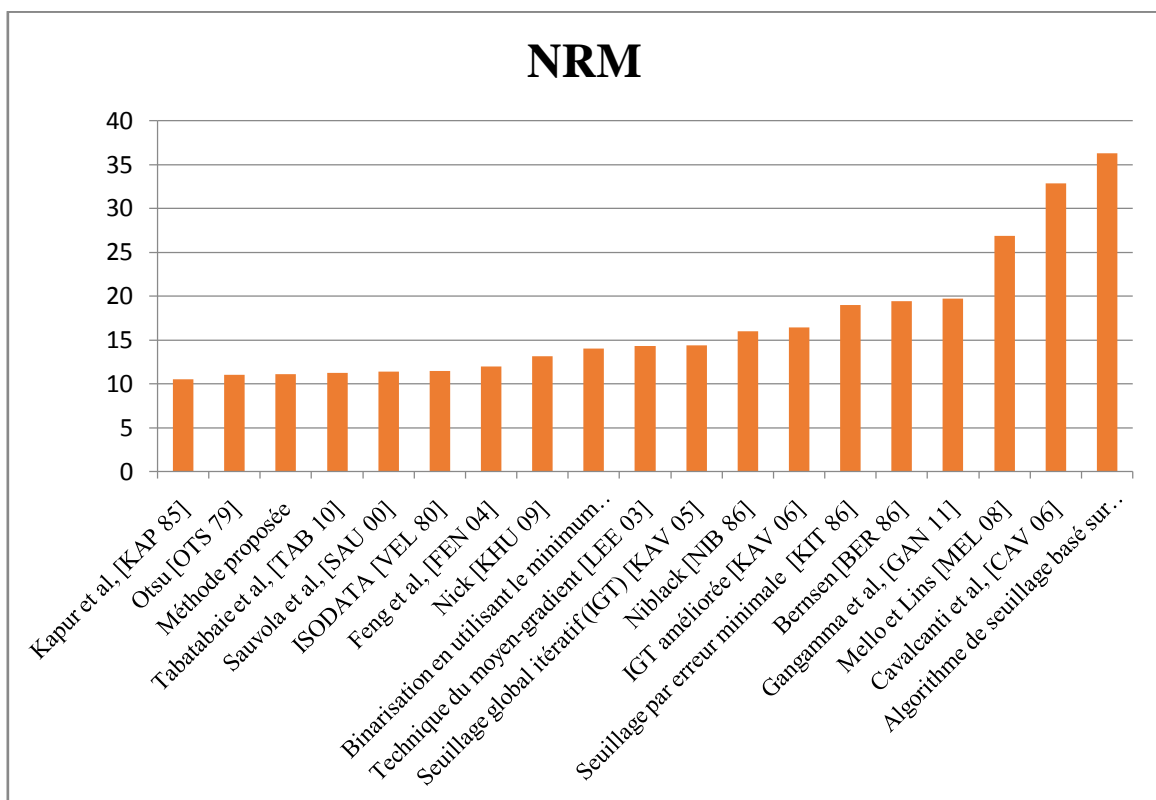


Figure 5. 11. Performances des méthodes de binarisation comparées en termes de NRM

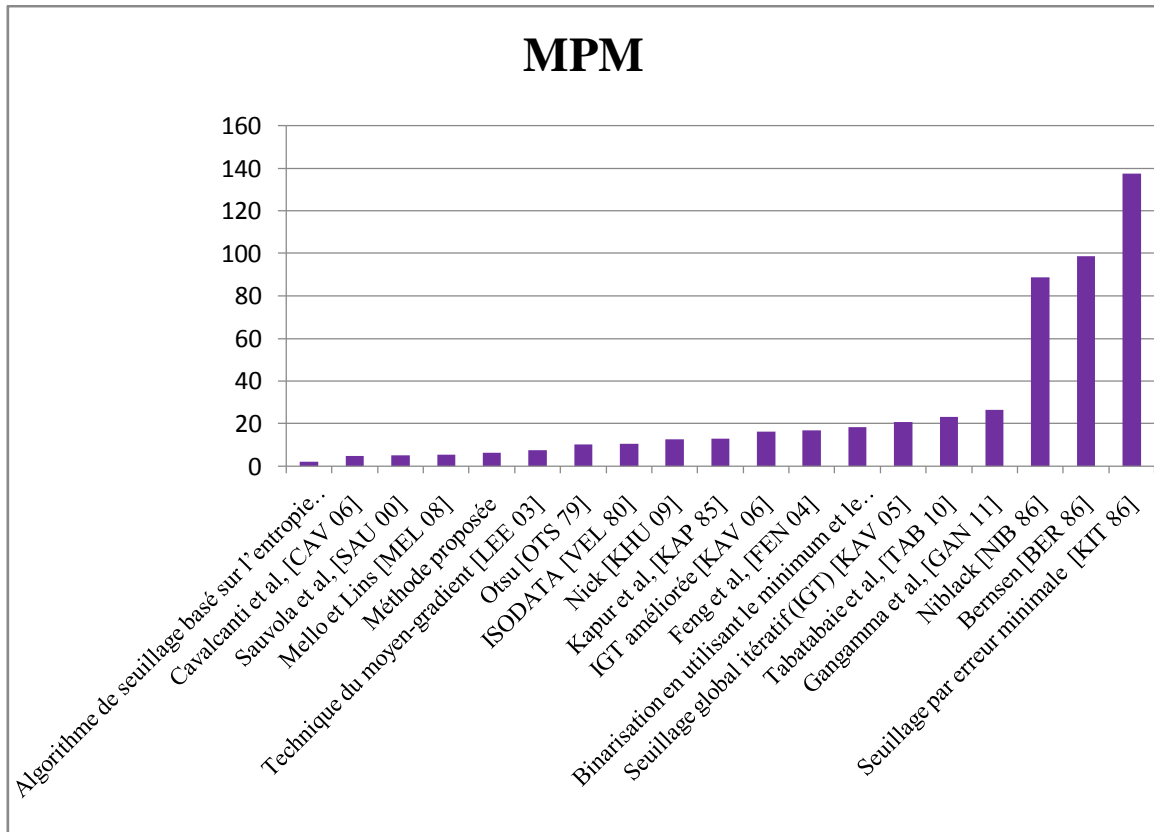


Figure 5. 12. Performances des méthodes de binarisation comparées en termes de MPM

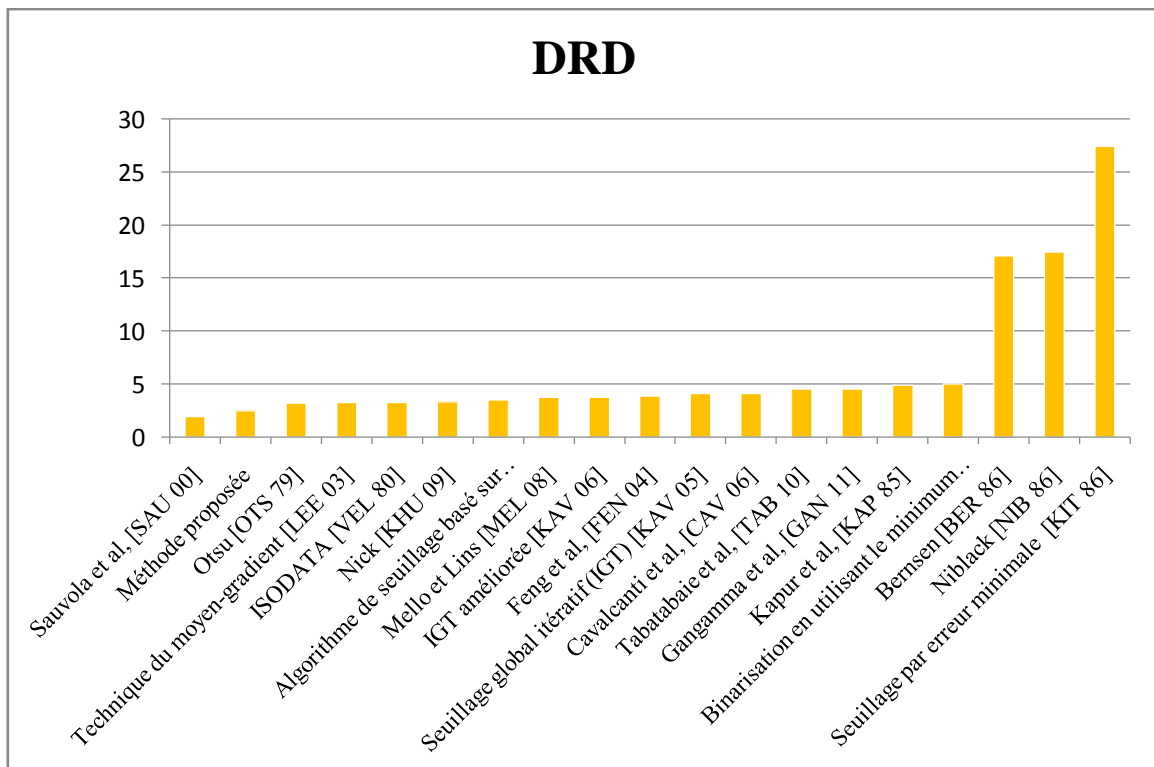


Figure 5. 13. Graphe montrant les performances des méthodes de binarisation comparées en termes de DRD

En analysant les résultats moyens obtenus sur les deux ensembles de test résumés dans le tableau 5.6, et le classement final donné par le tableau 5.7, on trouve que notre méthode de binarisation a été classée en deuxième position juste après la méthode de Sauvola et Pietikainen. Ainsi notre méthode a eu une somme de rangs égale à 13 qui est légèrement plus grande que celle obtenue par la méthode de Sauvola et Pietikainen (égale à 12). Les figures 5.9 – 5.12 précédentes nous indiquent aussi que la méthode proposée a présenté des bonnes performances suivant chaque mesure d'évaluation séparément. Elle se montre la meilleure méthode selon $F_{measure}$, la deuxième meilleure suivant $PSNR$, la deuxième meilleure en termes de NRM , la cinquième meilleure selon MPM , et la deuxième en termes de DRD .

A notre opinion, ces résultats sont encourageants en considérant le principe simple de la méthode proposée qui se base sur un classement direct des pixels de l'image en noirs ou blancs, par le biais d'un algorithme d'apprentissage automatique : le MLP.

5.2.3.3. Troisième expérimentation

Dans cette expérimentation, nous avons été intéressés à l'évaluation des performances de la méthode proposée pour la recherche dans les images de documents arabes dégradés. L'évaluation a été effectuée en utilisant la méthodologie que nous avons proposée dans le chapitre 4, section 4.3. L'évaluation en utilisant la méthodologie pré-mentionnée juge la qualité d'une image binarisée en fonction de la distance d'édition calculée entre la signature de vérité terrain associée à l'image originale et la signature extraite automatiquement de l'image binarisée.

Pour les tests, nous avons utilisé un autre ensemble composé de 20 images de documents dégradés réelles en langue arabe. Cet ensemble d'images est le même utilisé dans les expérimentations du chapitre 4 afin d'examiner la méthodologie d'évaluation proposée (voir section 4.3.3.2.1). Deux exemples d'images de cet ensemble sont illustrés dans la figure 4.25.

Cependant, les 20 images de l'ensemble de test ont été binarisées en utilisant les différentes méthodes comparées et la qualité de chaque image binarisée est mesurée en termes de distance d'édition. Les résultats d'évaluation moyens obtenus par les différentes méthodes de binarisation sont résumés dans le tableau 5.8.

Tableau 5. 8. Performances moyennes obtenues par les différentes méthodes de binarisation en utilisant notre méthodologie d'évaluation

Méthode	Distance d'édition moyenne
Méthode proposée	34.7666
Sauvola et Pietikainen [SAU 00]	38.33333
Binarisation en utilisant le minimum et le maximum locaux [SU 10]	44.26
Nick [KHU 09]	46
Otsu [OTS 79]	50.943
ISODATA [VEL 80]	51.017
Technique du moyen-gradient [LEE 03]	57.33333
Tabatabaie et Bohlool [TAB 10]	57.823
Feng et Tan [FEN 04]	60.25

IGT améliorée [KAV 06]	62.47
1 ^{ère} méthode de Mello et Lins [MEL 00]	68.163
Kapur et al. [KAP 85]	70.9
Cavalcanti et al. [CAV 06]	71.45
2 ^{ème} méthode de Mello et al. [MEL 08]	71.638
Seuillage global itératif (IGT) [KAV 05]	79
Gangamma et Srikanta Murthy [GAN 11]	83.3333
Bernsen [BER 86]	143
Niblack [NIB 86]	158.66667
Seuillage par erreur minimale [KIT 86]	169.6666

Comme nous pouvons remarquer du tableau 5.8, notre méthode a présenté les meilleures performances selon la méthodologie d'évaluation au niveau-système proposée dans le chapitre précédent. Ainsi notre méthode a été classée en premier rang avec la distance d'édition minimale, et même avant la méthode de Sauvola et Pietikainen qui a paru la plus performante selon l'évaluation supervisée-pixel précédente.

Les résultats obtenus par notre méthode de binarisation montrent qu'elle a réussi à binariser les images de documents de l'ensemble de test en conservant leurs caractéristiques structurales. Ces dernières ont été extraites avec une bonne précision par le système de recherche employé dans l'évaluation, ce qui a été traduit par une distance d'édition minimale entre les signatures associées aux images binarisées décrivant les caractéristiques extraites et les signatures de vérité terrain décrivant les caractéristiques réelles.

5.2.4. Conclusion

Dans cette première partie du 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 la classification directe des pixels en une des deux classes « noir » et « blanc » et non pas sur le seuillage comme la plupart des techniques de la littérature. Ainsi, la classification a été effectuée à l'aide d'un ANN de type MLP entraîné en utilisant l'algorithme de rétro-propagation du gradient. L'objectif de l'utilisation des ANNs, et plus précisément le MLP, dans la binarisation d'images est de combler le manque de l'utilisation des techniques d'apprentissage automatique dans tel problème, et de profiter de la capacité de généralisation du MLP.

Plusieurs expérimentations ont été effectuées afin d'évaluer les performances de la méthode proposée selon différents points de vue. Une évaluation supervisée-pixel a été d'abord procédée dans laquelle notre méthode a été classée la deuxième après la méthode de Sauvola et Pietikainen. Ensuite, nous avons évalué les performances de la méthode proposée pour la recherche d'images de documents arabes dégradés par l'utilisation de la méthodologie d'évaluation proposée dans le chapitre 4. Suivant cette évaluation, notre méthode se montre la plus performante. Les résultats obtenus sont, à notre opinion, encourageants en considérant le principe simple de la méthode proposée.

En effet, comme il a été confirmé par les résultats expérimentaux, le MLP 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.

De nombreuses extensions sont possibles et nous continuerons à améliorer la méthode proposée. Nous croyons que les résultats obtenus peuvent être améliorés en augmentant la taille de la base d'apprentissage et en lui enrichissant par plus de variétés d'exemples. Ils peuvent être améliorés aussi en modifiant la procédure du choix des pixels d'apprentissage pour qu'elle ne soit pas aléatoire mais par sélection.

5.3. Une nouvelle méthode hybride de seuillage pour la séparation texte/ fond dans les images de documents dégradés

5.3.1. Motivations et objectifs

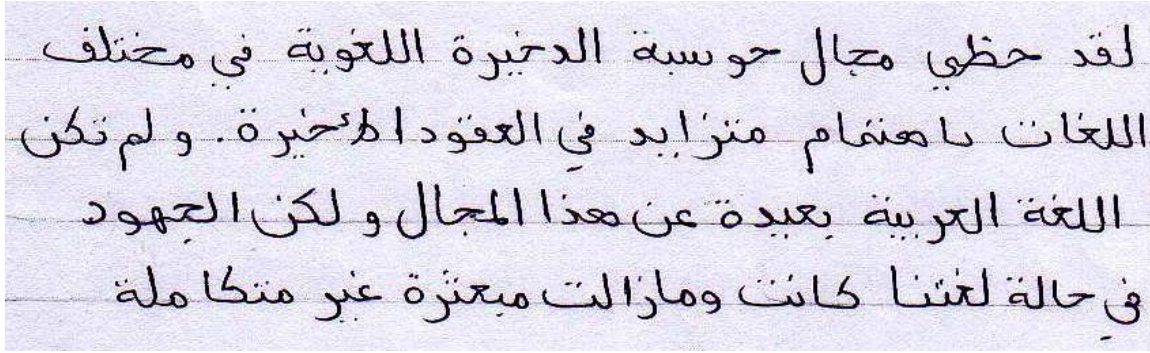
Comme nous avons vu dans le chapitre 2, les techniques de binarisation par seuillage peuvent être regroupées en trois grandes classes : seuillage global, seuillage local, et seuillage hybride. Les méthodes de seuillage global sont largement utilisées dans de nombreuses applications d'analyse d'images à cause de leur simplicité, rapidité, et efficacité [SU 11]. 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 et l'arrière-plan. La figure 5.14 présente un exemple de binarisation d'une image de document de bonne qualité ayant un histogramme bimodal par une méthode de seuillage global (la méthode d'Otsu [OTS 79]).

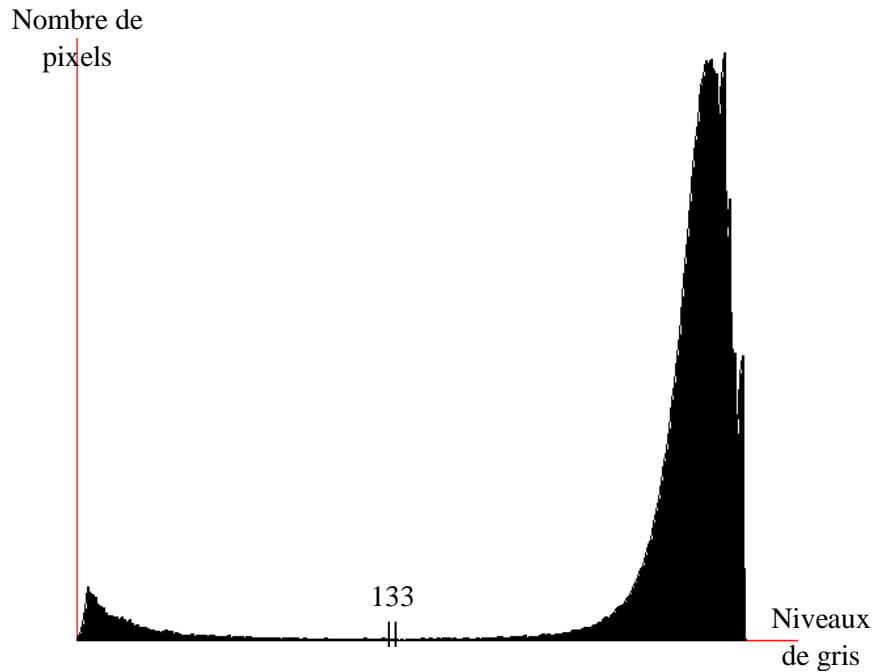
Pour les images de documents historiques qui sont généralement de mauvaise qualité, contenant différents types de détériorations et dégradations, avec un fond texturé et illumination inégale, ou lorsque les niveaux de gris des pixels de l'avant-plan et de l'arrière-plan se chevauchent, les méthodes de 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. Le résultat de binarisation d'une image de document de mauvaise qualité par l'algorithme d'Otsu [OTS 79] est illustré par la figure 5.15.

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 très coûteuses en temps d'exécution 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 lent chaque fois où on augmente la taille du voisinage considérée. Les méthodes hybrides, en revanche, combinent des informations globales et locales pour la segmentation de l'image.

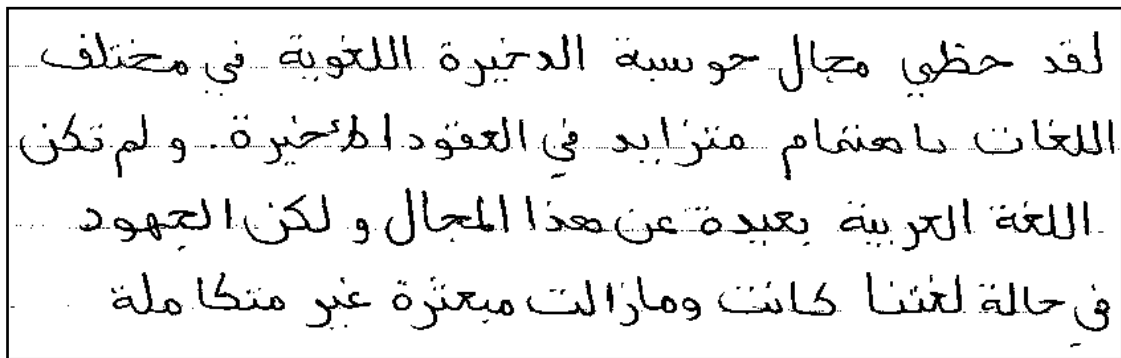
Le but du présent travail est de parvenir à une technique de binarisation rapide et efficace en même temps qui permet une séparation satisfaisante entre l'avant-plan et l'arrière-plan dans les images de documents historiques au bout d'un temps raisonnable. Notre contribution dans ce travail consiste à proposer une nouvelle technique de binarisation basée sur le seuillage hybride. La technique proposée utilise des informations globales et locales pour atteindre la séparation avant-plan / arrière-plan et elle combine ainsi les avantages des deux familles de techniques, à savoir l'efficacité et la rapidité d'exécution.



(a) Image en niveaux de gris de bonne qualité,

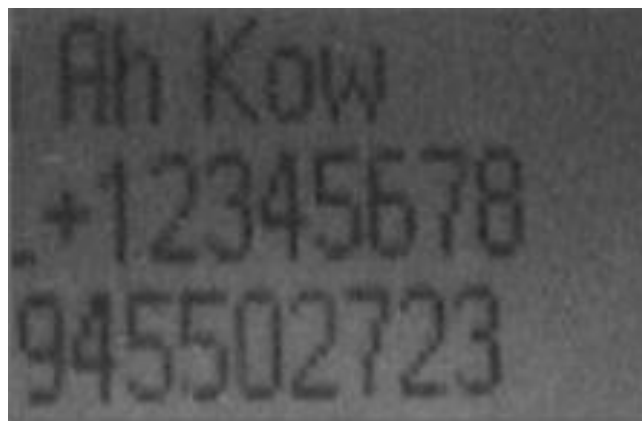


(b) Histogramme de niveaux de gris correspondant

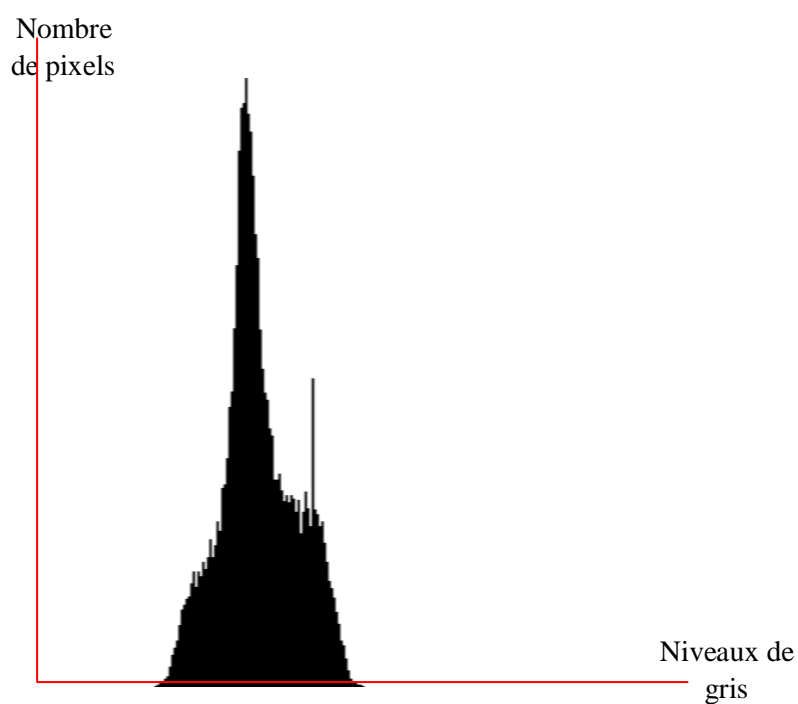


(c) Résultat de seuillage par l'algorithme d'Otsu

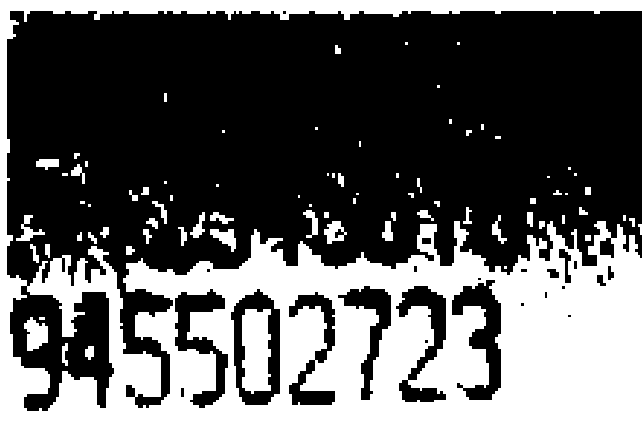
Figure 5. 14. Seuillage global d'une image bimodale



(a) Image en niveaux de gris de mauvaise qualité



(b) Histogramme de niveaux de gris correspondant



(c) Résultat de seuillage par l'algorithme d'Otsu

Figure 5. 15. Seuillage global d'une image de mauvaise qualité

5.3.2. Description de la technique proposée

Dans cette section, nous présentons notre deuxième technique proposée pour la binarisation d'images de documents historiques [KEF 14b] [SAR 14]. Comme nous avons déjà dit, la technique proposée appartient à la classe des méthodes de binarisation basées sur le seuillage, et plus précisément le seuillage hybride. Ces dernières peuvent être rapides et efficaces en même temps.

La technique proposée utilise deux seuils globaux différents T_1 et T_2 calculés en se basant sur des informations statistiques globales de l'image et elle s'exécute en deux passes. Dans la première passe, un seuillage global est appliqué sur l'image dans le but de classer le maximum de ses pixels en *avant-plan* ou *arrière-plan*. Dans la deuxième passe, les pixels incertains sont assignés à l'une des deux classes : *avant-plan* ou *arrière-plan* en se basant sur une analyse locale. Pour ce faire, plusieurs méthodes bien connues de seuillage local sont appelées pour déterminer séparément les valeurs binaires des pixels restants. La valeur binaire finale de chaque pixel incertain est sélectionnée comme la valeur la plus probable entre les sorties des différentes méthodes locales interrogées.

Nous détaillons dans la suite de cette section les étapes incluses dans la technique proposée.

5.3.2.1. Seuillage global

Durant cette première étape, la majeure partie de l'image est binarisée en appliquant un seuillage global sur l'image entière. Cependant, les pixels sont classés en trois groupes: *avant-plan*, *arrière-plan*, et *incertains* en se basant sur deux seuils globaux différents. Comme le but de la binarisation est de segmenter l'image en deux classes: *avant-plan* et *arrière-plan*, et comme un seuil unique n'est pas capable d'accomplir efficacement cette tâche, l'utilisation de deux seuils nous semble être une solution.

La première chose à faire dans le processus de binarisation est la détermination des deux seuils globaux T_1 et T_2 . Ces deux seuils sont estimés à partir de l'histogramme de niveaux de gris de l'image originale et ils représentent respectivement l'intensité moyenne de l'avant-plan et l'intensité moyenne de l'arrière-plan. L'avant-plan et l'arrière-plan sont séparés par un seuil global T calculé en utilisant un algorithme de seuillage global. Les étapes comprises dans cette phase sont les suivantes:

5.3.2.1.1. Calculer un seuil global T

Nous commençons par l'application d'un algorithme de seuillage global sur l'image traitée dans le but de calculer un seuil global T à partir duquel les deux seuils T_1 et T_2 sont déterminés. Ainsi, le seuil T sépare l'histogramme de niveaux de gris de l'image en deux classes : *avant-plan* et *arrière-plan*. Dans notre approche, nous avons choisi d'utiliser l'algorithme d'Otsu [OTS 79], car ce dernier a montré son efficacité et a surmonté les autres algorithmes de seuillage global dans plusieurs études comparatives précédentes [SAH 88] [RAM 10] [TRI 95], et dans notre étude comparative décrite dans le chapitre 4.

5.3.2.1.2. Déterminer les deux seuils T_1 et T_2

Les deux seuils T_1 et T_2 sont estimés en se basant sur le seuil global T calculé dans l'étape précédente.

Notons d_{min} la distance minimale entre l'intensité moyenne de l'avant-plan représenté par le premier moitié de l'histogramme et l'intensité moyenne de l'arrière-plan représenté par le deuxième moitié. T_1 et T_2 sont calculés comme suit :

$$T_1 = T - \frac{d_{\min}}{2}$$

$$T_2 = T + \frac{d_{\min}}{2}$$

La valeur de d_{\min} est choisie par expérimentations. Après plusieurs tests, nous avons adopté la valeur $d_{\min} = 40$.

5.3.2.1.3. Seuillage

Après que les valeurs des deux seuils T_1 et T_2 sont fixées, nous appliquons un seuillage à l'image entière afin de classifier la plupart des pixels en arrière-plan ou en avant-plan. La classification est effectuée ainsi comme suit. Tous les pixels ayant un niveau de gris inférieur au premier seuil T_1 sont affectés à la classe *avant-plan* et par conséquent, ils sont colorés en noir et gardés. De la même manière, tous les pixels ayant un niveau de gris supérieur au deuxième seuil T_2 sont attribués à la classe *arrière-plan*, et donc ils sont éliminés (mis en blanc). Les pixels restants (dont le niveau de gris est entre T_1 et T_2) sont considérés comme des pixels incertains et ils sont laissés à la deuxième phase.

Notons I l'image originale, et I_b l'image binarisée avec 0 dénote le noir et 255 dénote le blanc, le seuillage en utilisant les deux seuils T_1 et T_2 est résumé par l'équation suivante:

$$I_b(x, y) = \begin{cases} 0, & \text{si } I(x, y) < T_1 \\ 255, & \text{si } I(x, y) > T_2 \\ I(x, y), & \text{Sinon} \end{cases}$$

La majeure partie de l'image résultante est en noir et blanc mais elle contient encore des pixels en niveaux de gris correspondants aux pixels incertains laissés à la prochaine étape. Ainsi, l'image résultante contient du bruit mais tous les pixels de l'avant-plan sont préservés. La figure 5.16 présente un exemple d'application de cette étape sur l'image de la figure 5.15.a.

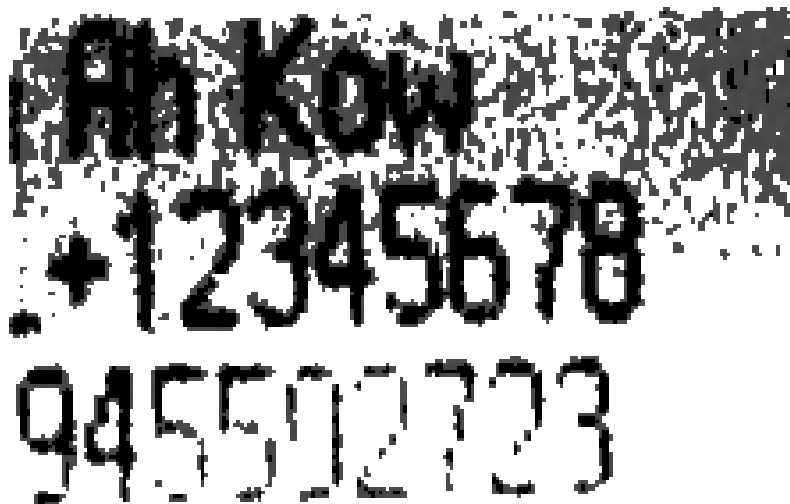


Figure 5.16. Seuillage global en utilisant les deux seuils T_1 et T_2 de l'image de la figure 5.15.a

5.3.2.2. Seuillage local des pixels restants

Les pixels non traités dans la phase précédente (ceux ayant un niveau de gris compris entre T_1 et T_2) peuvent appartenir à l'avant-plan et donc ils doivent être préservés, comme ils peuvent être des pixels du fond ou du bruit et dans ces deux cas ils doivent être éliminés. La décision d'attribution des pixels incertains à l'une des deux classes : avant-plan ou arrière-plan s'effectue par un processus local en examinant leur voisinage. Des algorithmes de seuillage local doivent être donc appliqués dans cette étape.

Pour assurer une classification plus ou moins précise, nous proposons d'appliquer plusieurs méthodes de seuillage local, et de choisir la valeur binaire (noir ou blanc) la plus probable pour chaque pixel. Cependant, pour chaque pixel de coordonnées (x, y) non encore classifié, nous calculons localement ses valeurs binaires (0 pour noir ou 255 pour blanc) obtenues par l'application des différentes méthodes de seuillage local intégrées.

Notons $b_1(x,y), b_2(x,y), \dots, b_n(x,y)$ ces valeurs avec n est le nombre de méthodes locales appliquées. La valeur binaire $b_i(x,y)$ du pixel (x,y) attribuée par la $i^{\text{ème}}$ méthode locale est calculée comme suit :

$$b_i(x, y) = \begin{cases} 0, & \text{si } I(x, y) < TL_i(x, y) \\ 255, & \text{sinon} \end{cases}$$

Tel que $TL_i(x,y)$ représente le seuil local au pixel (x,y) calculé par la $i^{\text{ème}}$ méthode locale.

La valeur binaire finale $b(x,y)$ du pixel (x,y) est celle attribuée par plus de la moitié des méthodes de binarisation locale. Formellement, la valeur finale $b(x,y)$ est donnée comme suit:

$$b(x, y) = \left(\text{round} \left(\frac{\sum_{i=1}^n b_i(x, y)}{n \times 255} \right) \right) \times 255$$

Round retourne l'entier le plus proche d'un nombre réel.

Le résultat final de binarisation de l'image de la figure 5.15.a en utilisant notre technique est illustré par la Figure 5.17.

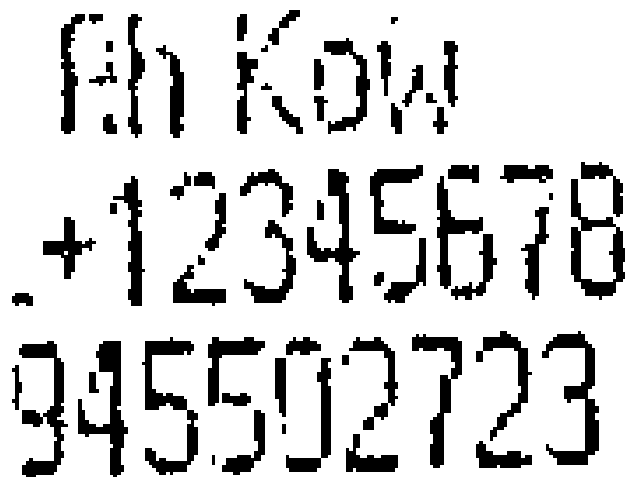


Figure 5. 17. Résultat final de binarisation de l'image de la figure 5.15.a en utilisant notre technique de seuillage hybride proposée

5.3.3. Expérimentations et résultats

Nous avons expérimenté deux versions de la méthode proposée. La première version intègre trois méthodes de binarisation locales : Niblack [NIB 86], Sauvola et Pietikainen [SAU 00], Nick [KHU 09]. Ces méthodes ont été souvent citées dans la littérature et elles ont montré des bonnes performances dans plusieurs études comparatives précédentes [KHU 09] [TRI 95][KEF 09]. Dans la deuxième version, et dans le but d'augmenter la précision de la classification, nous avons ajouté aux trois méthodes précédentes deux autres techniques: la méthode de binarisation en utilisant le minimum et le maximum locaux [SU 10] qui a été classée première dans la compétition H-DIBCO 2010, et notre méthode basée sur les ANNs proposée dans la première partie de ce chapitre. L'expérimentation des deux versions de la méthode proposée a eu comme objectif de tester l'influence du nombre de techniques de binarisation locale intégrées sur les performances de la méthode, et par suite de sélectionner la version qui se montre la plus efficace. Cependant, en comparant les performances obtenues par les deux versions, nous avons considéré la deuxième qui intègre cinq techniques de seuillage local.

Plusieurs tests ont été menés afin d'évaluer les performances de la technique proposée selon trois points de vue : la qualité de binarisation, le temps d'exécution, et l'adéquation aux applications de recherche dans les images de documents arabes dégradés. Les tests effectués ont été regroupés en deux expérimentations. Les performances obtenues sur une large collection de test ont été également comparées avec celles obtenues par des méthodes bien connues de la littérature, y compris des méthodes globales, locales et hybrides.

5.3.3.1. Base de test

Trois ensembles d'images ont été utilisées pour l'évaluation des performances de la méthode proposée comparativement aux méthodes de la littérature. Le premier est un ensemble public composé d'images de documents provenant des quatre collections proposées dans le cadre des compétitions DIBCO 2009, H-DIBCO 2010, DIBCO 2011 et H-DIBCO 2012. Ces quatre collections contiennent un total de 50 images de documents réels avec les images de vérité terrain correspondantes. Des exemples d'images de ces collections sont montrés dans les figures 4.2 – 4.5.

Le deuxième ensemble d'images est une collection synthétique composée de 150 images de documents artificielles construites par notre algorithme de fusion détaillé dans la section 4.2.2.2 à partir de 15 images de fonds et 10 images binaires de texte. Cet ensemble d'images a été construit pour pouvoir estimer les performances de notre approche sur une large base de test.

Le troisième ensemble de test est composé de 20 images de documents dégradés réelles en langue arabe. Cet ensemble d'images est le même utilisé pour examiner la méthodologie d'évaluation proposée dans le chapitre 4 (voir section 4.3.3.2.1), et dans la troisième expérimentation de notre méthode de binarisation basée sur les ANNs décrite dans la première partie du présent chapitre (section 5.2.3.3). Deux exemples d'images de cet ensemble sont illustrés dans la figure 4.25.

Les deux premiers ensembles ont été utilisés pour estimer les performances des méthodes comparées en jugeant la qualité de binarisation au niveau-pixel. Le troisième ensemble a été utilisé pour examiner l'efficacité de la méthode proposée pour des applications de recherche dans les images de documents arabes dégradés.

5.3.3.2. Première expérimentation

La première expérimentation a eu pour objectif d'estimer les performances de la méthode proposée (qui intègre cinq techniques locales) selon deux aspects : le temps d'exécution et la qualité de binarisation. Ainsi, la qualité de binarisation est mesurée quantitativement au niveau-pixel à l'aide des cinq mesures standards d'évaluation : *F-Mesure (FM)*, *PSNR*, *NRM*, *MPM*, et *DRD*. Pour plus de détails sur ces mesures, veuillez voir le chapitre 3.

Cependant, en exception du temps d'exécution, les mesures utilisées dans cette expérimentation estiment la qualité d'une image binarisée en la comparant avec l'image de référence associée. De ce fait, l'évaluation supervisée-pixel procédée dans cette expérimentation a été effectuée uniquement sur les deux premiers ensembles d'images mentionnés dans la section précédente.

Pour avoir une idée plus claire sur les performances de notre méthode, nous l'avons comparé avec plusieurs techniques de seuillage de la littérature. Trois méthodes globales ont été incluses dans la comparaison (Otsu [OTS 79], Kapur et al. [KAP 85], IGT [KAV 05]), trois méthodes locales (Niblack [NIB 86], Sauvola et Pietikainen [SAU 00], Nick [KHU 09]) et quatre méthodes hybrides (IGT améliorée [KAV 06], Gangamma et Srikanta Murthy [GAN 11], Tabatabaie et Bohlool [TAB 10], Seuillage par soustraction du fond [LEE 03]). Nous avons également comparé notre méthode avec la technique de binarisation basée sur les ANNs que nous l'avons proposée dans la première partie de ce chapitre. Pour les méthodes paramétrables, les valeurs optimales de paramètres utilisées sont celles déterminées dans le chapitre 4, section 4.2.5.

Les différentes méthodes de binarisation considérées ont été appliquées sur les images des deux premiers ensembles de test (l'ensemble public et l'ensemble synthétique), et leurs résultats ont été comparés avec les images de vérité terrain en utilisant les cinq mesures d'évaluation mentionnées ci-dessus. Les performances moyennes (en termes des cinq mesures précédentes et du temps d'exécution) obtenues par les différentes méthodes sur le premier et le deuxième ensemble de test sont récapitulées dans les tableaux 5.9 et 5.10 respectivement. Les résultats moyens entre les deux ensembles sont résumés dans le tableau 5.11. Le classement final des méthodes comparées est affiché dans le tableau 5.12, qui récapitule également les rangs partiels de chaque méthode selon chaque mesure d'évaluation ainsi que la somme des rangs. Ensuite, nous fournissons dans les figures 5.18–5.23 des graphes représentant les performances moyennes des méthodes comparées en termes de chaque mesure séparément.

Tableau 5.9. Résultats moyens des différentes méthodes de binarisation sur l'ensemble d'images publiques

Méthode	Mesure	Temps d'exécution (ms)	FM	PSNR	NRM ($\times 10^{-2}$)	MPM ($\times 10^{-3}$)	DRD
Otsu		402.78	0.8262	15.12	5	2.14	2.206
Kapur et al.		459.123	0.8128	15.06	5.1	2.87	2.7112
IGT		249.147	0.772	14.398	9.33	16.35	3.129
Niblack		69732.14	0.6361	8.64	11.66	61.78	13.1601
Sauvola et Pietikainen		65932.23	0.8509	18.56	6.033	1.2288	1.41245
Nick		57120.32	0.78848	16.268	8.23	15.658	3.104
IGT améliorée		410.96	0.79146	16.091	9.26	12.832	3.002

Gangamma et Srikanta Murthy	64892.314	0.7134	14.378	11.93	37.074	4.549
Seuillage par soustraction du fond	869.89	0.712	14.64	18.6	14.71	4.07456
Tabatabaie et Bohlool	60715.145	0.8215	16.99	5.81	13.54	2.366
Notre méthode basée ANNs	81054.31	0.8685	17.109	5.619	1.0662	2,80285
Technique de seuillage hybride proposée	1223.12	0.9125	19.076	4.83	0.9640	1.1459

Tableau 5. 10. Résultats moyens des différentes méthodes de binarisation sur l'ensemble d'images synthétiques

Méthode	Mesure	Temps d'exécution (ms)	FM	PSNR	NRM ($\times 10^{-2}$)	MPM ($\times 10^{-3}$)	DRD
Otsu		514.72	0.7851	15.107	10.66	6.0437	3.9325
Kapur et al.		582.533	0.7816	15.002	11.118	6.9706	3.128
IGT		369.939	0.73966	14.216	11.89	13.621	4.166
Niblack		87276.06	0.543	7.92	16.47	103	18.962
Sauvola et Pietikainen		81896.17	0.8044	16.886	8.9	4.2792	2.7540
Nick		87841.48	0.73592	14.826	12.39	12.622	3.952
IGT améliorée		564.926	0.7487	14.865	11.84	12.608	4.388
Gangamma et Srikanta Murthy		83106	0.6303	13.56	18.49	48.526	5.071
Seuillage par soustraction du fond		1374.224	0.63614	13.224	22.08	19.29	5.86544
Tabatabaie et Bohlool		91034.855	0.7949	16.39	11.59	10.12	3.512
Notre méthode basée ANNs		10117.9	0.7666	16.129	10.07	4.756	2.913
Technique de seuillage hybride proposée		1612.6	0.8553	18.226	8.19	5.79	2.477

Tableau 5. 11. Résultats moyens des différentes méthodes de binarisation entre les deux ensembles

Méthode	Mesure	Temps d'exécution (ms)	FM	PSNR	NRM ($\times 10^{-2}$)	MPM ($\times 10^{-3}$)	DRD
Otsu		458.75	0.80565	15.1135	7.83	4.09185	3.06925
Kapur et al.		520.828	0.7972	15.031	8.109	4.9203	2.9196
IGT		309.543	0.75583	14.307	10.61	14.9855	3.6475
Niblack		78504.1	0.58955	8.28	14.065	82.39	16.06105
Sauvola et Pietikainen		73914.2	0.82765	17.723	7.4665	2.754	2.083225
Nick		72480.9	0.7622	15.547	10.31	14.14	3.528
IGT améliorée		487.943	0.77008	15.478	10.55	12.72	3.695
Gangamma et Srikanta Murthy		73999.157	0.67185	13.969	15.21	42.8	4.81
Seuillage par soustraction du fond		1122.057	0.67407	13.932	20.34	17	4.97
Tabatabaie et Bohlool		75875	0.8082	16.69	8.7	11.83	2.939

Notre méthode basée ANNs	45586.105	0.81755	16.619	7.8445	2.9111	2.857925
Technique de seuillage hybride proposée	1417.86	0.8839	18.651	6.51	3.377	1.81145

Tableau 5. 12. Classement final des méthodes comparées sur les deux ensembles de test

Rang	Méthode	Temps d'exécution (ms)	FM	PSNR	NRM	MPM	DRD	Somme des rangs
1	Technique de seuillage hybride proposée	6	1	1	1	3	1	13
2	Sauvola et Pietikainen	9	2	2	2	1	2	18
3	Notre méthode basée-ANNs	7	3	4	4	2	3	23
4	Otsu	2	5	7	3	4	6	27
5	Kapur et al.	4	6	8	5	5	4	32
6	Tabatabaie et Bohlool	11	4	3	6	6	5	35
7	IGT améliorée	3	7	6	8	7	9	40
8	Nick	8	8	5	7	8	7	43
9	IGT	1	9	9	9	9	8	45
10	Seuillage par soustraction du fond	5	10	11	12	10	11	59
11	Gangamma et Srikanta Murthy	10	11	10	11	11	10	63
12	Niblack	12	12	12	10	12	12	70

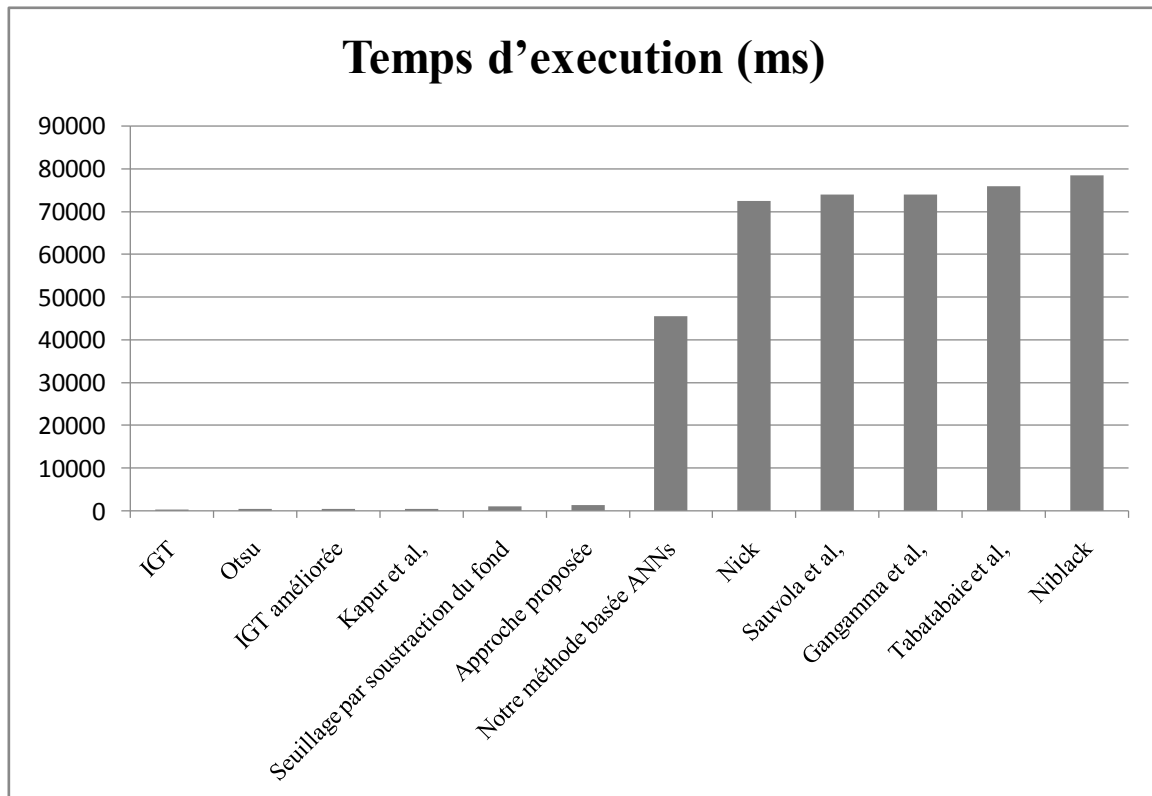


Figure 5. 18. Performances des méthodes de binarisation comparées en termes du temps d'exécution

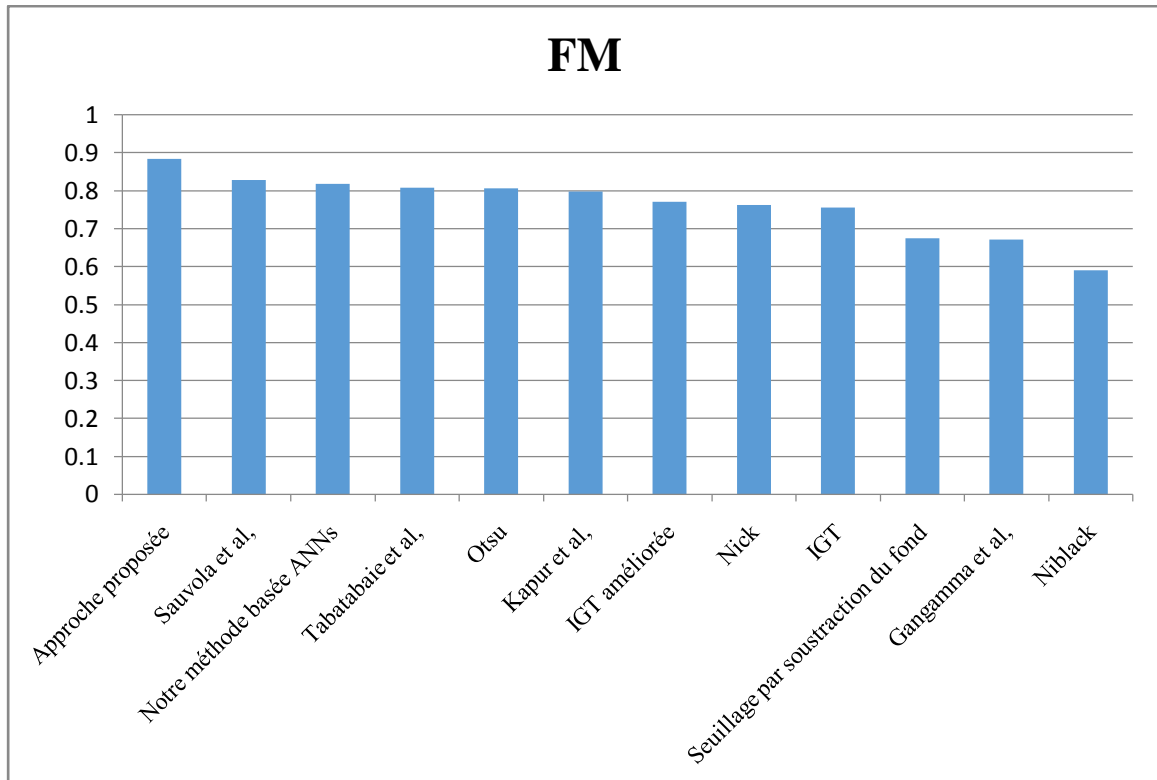


Figure 5. 19. Performances des méthodes de binarisation comparées en termes de Fmeasure.

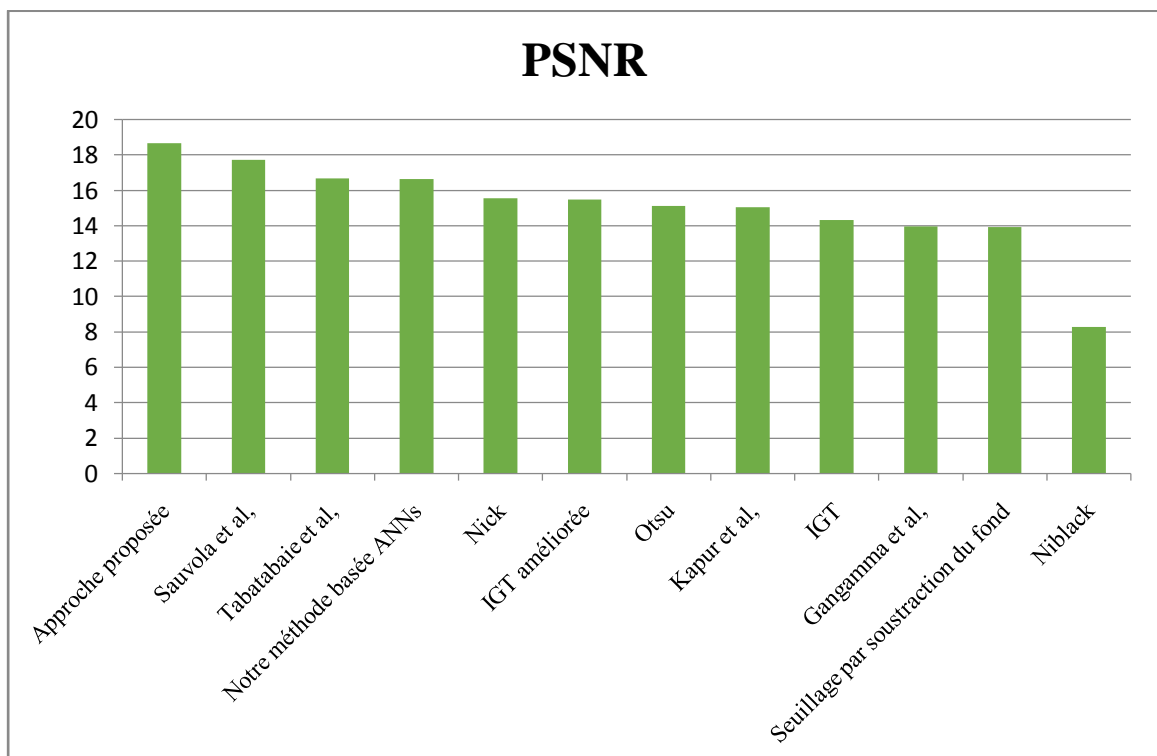


Figure 5. 20. Performances des méthodes de binarisation comparées en termes de PSNR

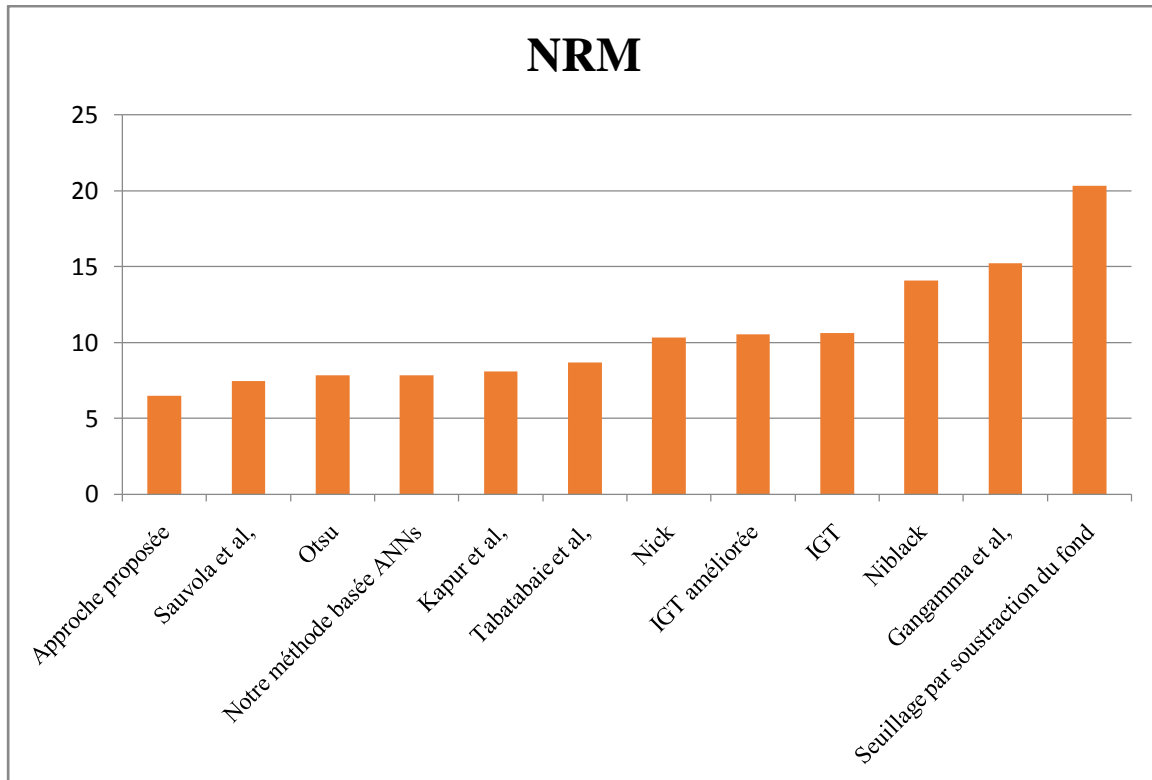


Figure 5. 21. Performances des méthodes de binarisation comparées en termes de NRM

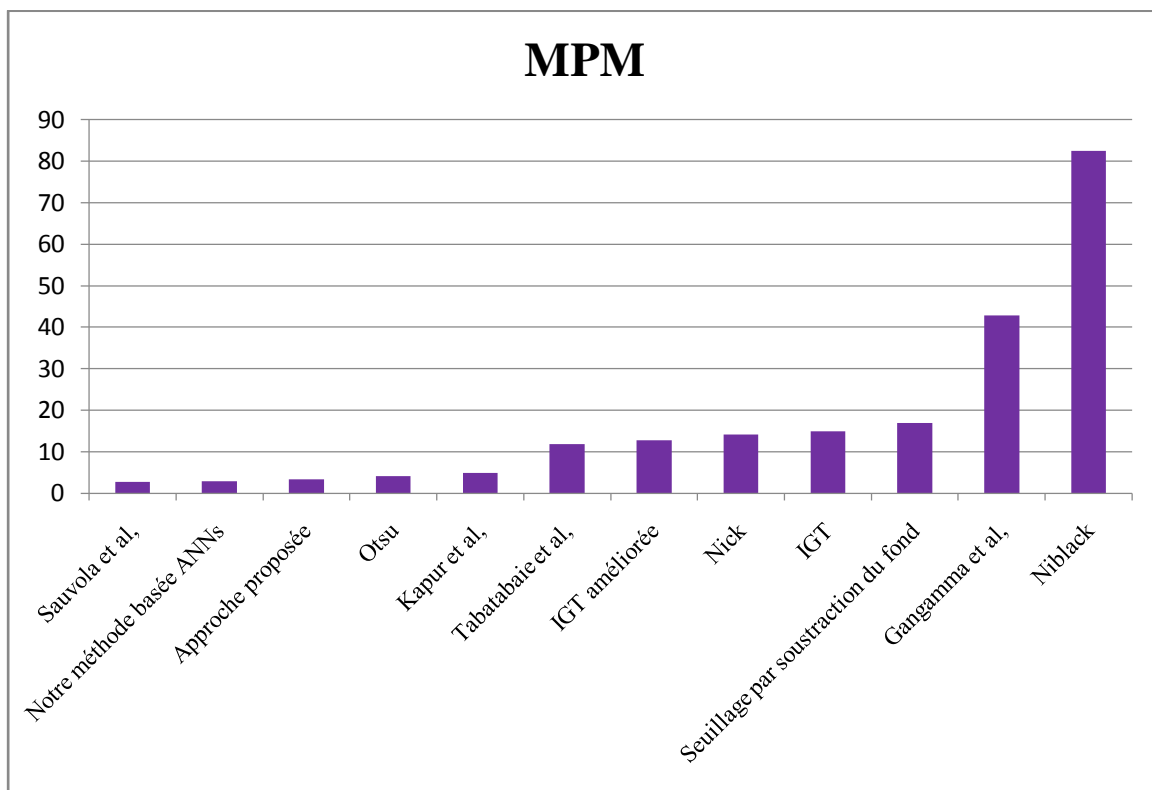


Figure 5. 22. Performances des méthodes de binarisation comparées en termes de MPM

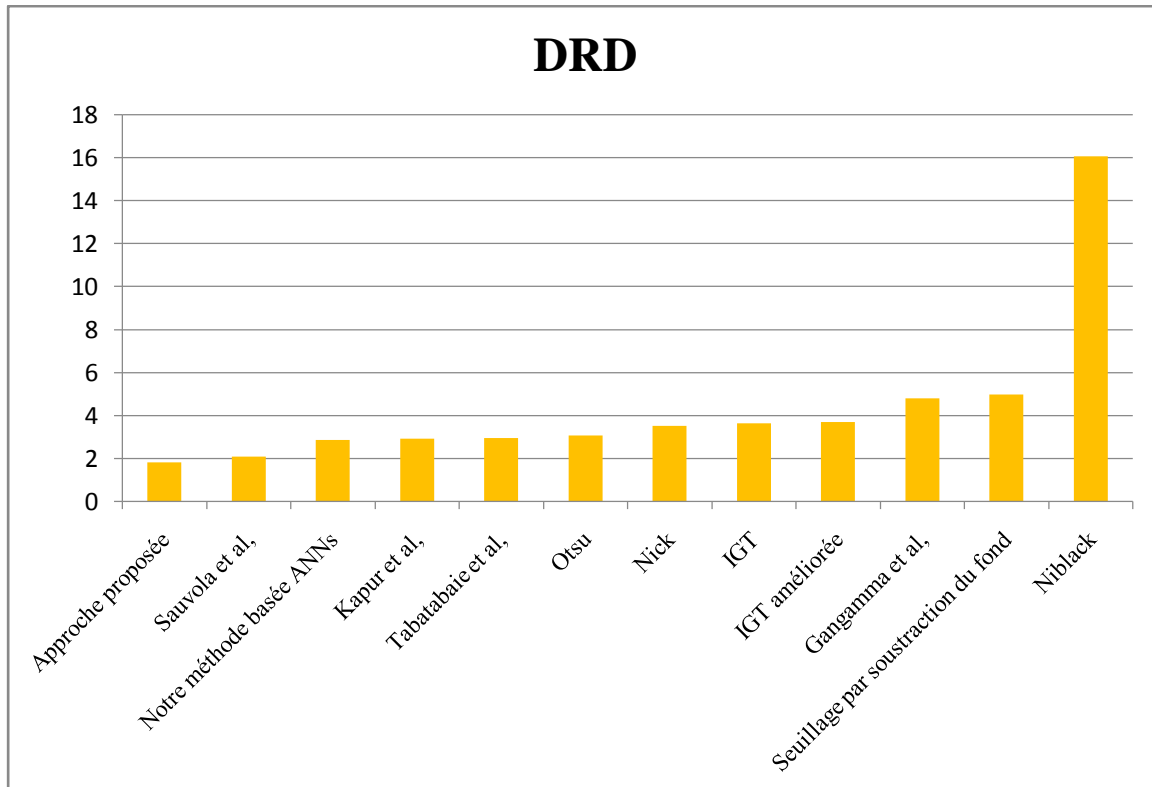


Figure 5. 23. Graphe montrant les performances des méthodes de binarisation comparées en termes de DRD

D'après les tableaux 5.10 et 5.11 qui résument les résultats moyens d'évaluation des méthodes de binarisation comparées, et le tableau 5.12 récapitulant le classement final de ces méthodes, il est clair que notre méthode de binarisation hybride proposée a été classée en première position avec une somme de rangs minimale égale à 18. Selon les résultats moyens obtenus en termes des mesures empiriques de la qualité de binarisation ($F_{measure}$, $PSNR$, NRM , MPM , DRD), notre méthode a paru la plus efficace et elle a présenté les meilleures performances globales. Ces performances signifient que dans la totalité, les résultats de binarisation de notre méthode sont les plus proches des images de référence. Cependant, la combinaison d'une technique de seuillage global et cinq techniques de seuillage local a permis une détermination plus crédible de la valeur binaire de chaque pixel. Cette dernière a été attribuée en prenant la valeur la plus probable statistiquement entre les sorties des différentes techniques intégrées. Notre méthode a surpassé les autres techniques hybrides et locales incluses dans la comparaison comme la méthode de Sauvola et Pietikainen considérée comme l'une des méthodes les plus efficaces dans plusieurs études antérieures. En analysant les graphes des figures 5.19 – 5.23, nous trouvons que la méthode proposée a montré des bonnes performances suivant chaque mesure d'évaluation séparément. Elle a été classée la première en termes de $F_{measure}$, $PSNR$, NRM , et DRD , et la troisième suivant MPM .

En ce qui concerne le temps d'exécution considéré comme le deuxième aspect d'évaluation dans cette expérimentation, notre méthode a eu un temps d'exécution moyen égal à 1612.6 ms. De ce fait, elle se révèle très rapide comparativement aux méthodes locales (environ 50 fois plus rapide que la méthode de Sauvola et Pietikainen), ce qui nous a permis de gagner une portion considérable du temps d'exécution. Cependant, ce résultat est raisonnable puisque seulement une partie des pixels (ayant un niveau de gris compris entre les deux seuils T_1 et T_2) a été traitée localement.

On peut conclure que la méthode proposée a montré des bonnes performances et en même temps elle a accompli la tâche au bout d'un temps très raisonnable comparativement aux méthodes locales, et de ce fait nous pouvons affirmer que l'objectif de notre présent travail a été atteint.

5.3.3.3. Deuxième expérimentation

Puisque l'évaluation quantitative supervisée-pixel ne reflète pas toujours la qualité réelle des images binarisées et leur appropriation pour une application particulière, la deuxième expérimentation a été mise en place. Elle a eu comme objectif de vérifier l'adéquation de la méthode proposée aux applications de recherche dans les images de documents arabes dégradées. Nous avons utilisé pour cela notre méthodologie d'évaluation décrite dans le chapitre 4, section 4.3. L'évaluation en utilisant cette méthodologie juge la qualité d'une image binarisée en fonction de la distance d'édition calculée entre la signature de vérité terrain associée à l'image originale et la signature extraite automatiquement de l'image binarisée.

Nous avons utilisé pour les tests le troisième ensemble d'images pré-mentionné. Cependant, les images de l'ensemble de test ont été binarisées par les différentes méthodes comparées et la qualité de chaque image binarisée est mesurée en termes de distance d'édition. Les résultats d'évaluation moyens obtenus par les différentes méthodes de binarisation sur l'ensemble de test sont résumés dans le tableau 5.13.

Tableau 5. 13. Performances moyennes obtenues par les différentes méthodes de binarisation en utilisant notre méthodologie d'évaluation

Méthode	Distance d'édition moyenne
Technique de seuillage hybride proposée	29.1635
Notre méthode basée sur les ANNs	34.7666
Sauvola et Pietikainen	38.33333
Nick	46
Otsu	50.943
Tabatabaie et Bohlool	57.823
IGT améliorée	62.47
Kapur et al.	70.9
Seuillage global itératif (IGT)	79
Gangamma et Srikanta Murthy	83.3333
Seuillage par soustraction du fond	96.752
Niblack	158.66667

Comme nous pouvons remarquer du tableau 5.13, notre nouvelle technique de binarisation a présenté les meilleures performances selon la méthodologie d'évaluation au niveau système proposée dans le chapitre précédent. Ainsi notre méthode de binarisation a été classée en premier rang avec la distance d'édition minimale.

Les résultats obtenus par la nouvelle technique de seuillage montre que cette technique a réussi à binariser les images de documents de l'ensemble de test en conservant leurs caractéristiques structurales. Ces dernières ont été extraites avec une bonne précision par le système de recherche employé dans l'évaluation ce qui a été traduit par une distance d'édition minimale entre les signatures associées aux images binarisées décrivant les caractéristiques extraites et les signatures de vérité terrain décrivant les caractéristiques réelles.


5.3.4. Conclusion

Dans la deuxième partie du chapitre, nous avons présenté notre deuxième contribution à la binarisation d'images de documents dégradés. Dans cette contribution, nous avons proposé une technique de binarisation basée sur le seuillage hybride dans le but de combiner les avantages des techniques globales et locales, à savoir la rapidité et l'efficacité. La méthode commence par appliquer un seuillage global sur l'image entière en vue de la segmenter en trois ensembles de pixels: *avant-plan*, *arrière-plan*, et *incertains*. Après, les pixels incertains sont assignés à une des deux classes *avant-plan* ou *arrière-plan* en se basant sur une analyse locale. Cinq techniques de seuillage local ont été combinées dans cette étape pour assurer une classification plus précise. Comme le nombre de pixels traités localement est faible par rapport au nombre total de pixels, le temps d'exécution de notre méthode reste raisonnable et proche de celui des méthodes globales, sans réduction des performances. Pour valider notre approche, nous l'avons comparé avec certaines techniques de la littérature et avec notre technique basée sur les ANNs proposée dans la première partie du chapitre et les résultats obtenus sur une collection d'images publiques et synthétiques sont encourageants et confirment notre proposition.

5.4. Conclusion du chapitre

Dans ce chapitre nous avons décrit nos deux contributions à la séparation avant-plan/ arrière-plan dans les images de documents historiques. Chaque contribution s'agit d'une nouvelle méthode de binarisation utilisant un principe différent pour extraire le texte du fond. La première méthode de binarisation utilise un ANN de type MLP pour classifier directement les pixels de l'image en *avant-plan* ou *arrière-plan* en se basant sur des informations globales et locales. Cependant, comme nous avons signalé précédemment, malgré le succès des ANNs dans plusieurs applications de traitement d'images, leur contribution dans le prétraitement et la binarisation plus spécifiquement est encore limitée. Ainsi, dans le but de combler le manque de l'utilisation des ANNs pour la binarisation et d'examiner leur applicabilité dans cette tâche, nous avons parti de l'hypothèse que la binarisation n'est qu'une opération de classification en deux classes. De ce fait, nous avons proposé d'exploiter la capacité de généralisation des ANNs pour concevoir une méthode de binarisation basée sur la classification. Dans la deuxième contribution, nous avons été intéressé à concevoir une méthode de binarisation efficace et rapide en même temps. Nous avons ainsi proposée une méthode de seuillage hybride afin de minimiser le temps d'exécution. Plusieurs techniques de seuillage ont été combinées pour augmenter la précision de la méthode proposée.

Différentes expérimentations ont été menés pour évaluer les performances de nos deux méthodes comparativement aux techniques de la littérature, et les résultats obtenus par les deux méthodes sont promoteurs et confirment nos propositions.



Conclusion
générale et
perspectives

Les bibliothèques, les musées et autres institutions à caractères pédagogiques ou sociopolitiques contiennent des collections considérables de documents, le plus souvent manuscrits. Les documents historiques des civilisations anciennes et les archives nationales sont l'exemple typique de telles richesses qui représentent le patrimoine et l'histoire des nations. En effet, ces documents encourent une dégradation progressive parce qu'ils ne sont pas conservés dans de bonnes conditions, et par suite, ils sont menacés d'un réel danger de disparition. Les bibliothèques et les services d'archives ont deux missions essentielles. La première est de conserver les documents afin que le patrimoine puisse être transmis aux générations futures, car l'avenir d'une nation, d'un peuple ou d'une société ne saurait s'envisager sans la connaissance de son passé. La deuxième est de les rendre accessibles au grand public afin que le patrimoine reste vivant et puisse faire l'objet de recherche et d'enrichissement. Ces deux missions semblent être contradictoires, car comme il est connu une des principaux facteurs de détériorations des documents est la manipulation physique. Comment donc arriver à conserver en communiquant ? et comment communiquer sans endommager ? En essayant de préserver les documents originaux, les bibliothèques ont élaboré des politiques de préservation à long terme qui tend à améliorer les conditions de conservation, à limiter la consultation des documents fragiles à quelques spécialistes. Cependant, cette restauration physique permet de minimiser les dégradations et de ralentir le vieillissement des documents, mais elle ne permet pas d'augmenter leur accessibilité.

Afin de surmonter ces problèmes et en vue de répondre aux deux questions précédentes, les bibliothèques et les conservateurs de documents ont commencé depuis plusieurs années à numériser les collections de documents historiques qui ont tendance à se détériorer. De nombreuses campagnes de numérisation ont été lancées et de nombreuses bibliothèques numériques ont vu le jour depuis les années soixante-dix. Ainsi, la numérisation permet une préservation numérique des documents, un partage des images, un accès simultané et un feuilletage virtuel, mais cette simple numérisation n'est pas suffisante pour répondre aux besoins des utilisateurs des bibliothèques numériques. Les documents numérisés, ne sont pas destinés à un simple archivage ; ils peuvent être sujets à divers traitements d'analyse du contenu (OCR, indexation, etc.). Néanmoins, les images de documents historiques numérisés sont souvent en mauvaise qualité comportant divers défauts et détériorations provenant du processus même de numérisation ou du mauvais état des documents originaux (vieillesse, modifications apportées par l'humain, etc.) compliquant l'analyse et la reconnaissance de ces documents à plusieurs niveaux. Afin d'atténuer l'effet des dégradations sur les traitements inclus dans le processus d'analyse et de reconnaissance de documents, l'acquisition doit être suivie par des opérations de bas niveau appliquées sur les images reçues du capteur communément appelés des *prétraitements*.

En réalité, le prétraitement d'images de documents peut être une finalité en lui-même en cherchant à augmenter la visibilité d'images, ou bien une étape dans une chaîne et dans ce cas l'objectif du prétraitement est de préparer le terrain aux traitements ultérieurs dans la chaîne. Tenant compte des différentes finalités des prétraitements, de la variété de défauts et de dégradations intervenant sur l'image, et de la diversité des applications d'analyse et de reconnaissance de documents, plusieurs opérations peuvent être considérées comme des prétraitements.

Bilan

Le travail adressé dans ce mémoire s'inscrit dans la démarche de sauvegarde et de valorisation de données patrimoniales dont la communauté internationale a pris conscience de l'intérêt. Nous étions intéressés dans ce mémoire au prétraitement d'images de documents

historiques connues par leur mauvaise qualité à cause des dégradations et détériorations présentes dans ces documents et qui compliquent leur traitement à plusieurs niveaux. Nous avons considéré dans ce travail uniquement une des opérations de prétraitement les plus importantes, à savoir, la séparation entre l'avant plan et l'arrière plan.

Ce manuscrit de thèse a permis tout d'abord de présenter l'étude que nous avons menée, et qui constitue la base théorique de nos contributions. Cette étude théorique a été répartie en trois chapitres. Le premier chapitre a survolé les différentes dégradations qui peuvent intervenir sur les images de documents numérisés, et les principales opérations de prétraitement fréquemment appliquées dans les systèmes d'analyse et de reconnaissance de documents. L'application des prétraitements est parfois précédée par une étape d'identification du bruit. Le chapitre expose ainsi à chaque niveau de prétraitement, les différentes approches, méthodes et techniques proposées dans la littérature. Dans le deuxième chapitre, un état de l'art approfondi sur la séparation avant-plan/ arrière-plan a été présenté. Nous avons concentré dans ce chapitre sur la binarisation, la technique de séparation avant plan/ arrière plan la plus répandue, en décrivant les différentes approches et méthodes existantes. Le troisième chapitre a abordé la problématique de l'évaluation des performances des techniques de binarisation. Dans ce chapitre, nous avons exposé les principales approches d'évaluation et de comparaison proposées dans la littérature, ainsi que les différentes mesures et critères utilisés par les chercheurs pour accomplir l'évaluation. Des trois chapitres précédents, nous avons constitué une idée sur les axes de recherche dans le champ de la séparation avant-plan/ arrière-plan et de binarisation plus précisément.

La deuxième partie de ce manuscrit décrit les quatre contributions que nous avons élaboré dans le cadre de la séparation avant-plan / arrière-plan dans les images de documents historiques. Les quatre contributions ont été regroupées en deux chapitres (chapitre 4 et 5) selon deux axes principaux: évaluation des techniques de binarisation, et proposition des nouvelles méthodes de binarisation d'images de documents.

Ainsi, le chapitre 4 détaille les deux contributions à l'évaluation des techniques de binarisation. La première est une étude comparative de 51 méthodes de binarisation appartenant à différentes catégories (globales, locales, et hybrides). Certaines de ces méthodes sont des techniques bien connues de la littérature qui ont montré leur efficacité dans des travaux antérieurs d'évaluations. D'autres sont des techniques non-connues et n'ont été incluses dans aucune étude comparative. La comparaison entre les différentes méthodes rapprochées a été effectuée quantitativement en utilisant plusieurs mesures empiriques supervisées (nécessitent d'images de référence) et non-supervisées (ne nécessitent pas d'images de référence). Dans les expérimentations, une collection large composée d'un ensemble d'images publiques (DIBCO/ H-DIBCO) et un ensemble d'images synthétiques a été utilisée. Les résultats obtenus nous ont permis de tirer plusieurs remarques. Ainsi, les résultats ont confirmé le fait qu'aucune méthode de binarisation ne se montre la plus puissante pour toutes les images. Nous avons remarqué aussi que des techniques simples ont fourni des résultats meilleurs que ceux obtenus par des méthodes désignées spécifiquement pour les images de documents historiques. La méthode qui a fourni les meilleures performances en moyenne est celle de Tabatabaie et al. [TAB 10] qui est une technique hybride basée sur la morphologie. Nous pouvons confirmer que notre étude comparative est la plus large comparativement aux études de la littérature, à cause du nombre important de techniques de binarisation comparées, de la taille de l'ensemble de test utilisé, et du nombre de critères d'évaluation combinés. La deuxième contribution exposée dans le chapitre 4 s'agit d'une nouvelle méthodologie d'évaluation des techniques de binarisation, proposée dans le but de sélectionner la technique la plus performante pour la recherche dans les images de documents arabes dégradées. L'approche proposée appartient à la classe de techniques d'évaluation au

niveau système, qui juge la qualité d'une binarisation en fonction du résultat final d'un système d'analyse. Cependant, La méthodologie proposée est basée sur la comparaison des caractéristiques structurelles extraites des images binarisées avec des caractéristiques de vérité terrain préétablies. Pour valider notre approche, nous l'avons comparé à l'évaluation visuelle humaine et l'évaluation supervisée-pixel, sur un ensemble d'images de documents arabes réels et synthétiques. Les résultats obtenus ont montré la concordance de l'approche proposée avec l'évaluation subjective humaine et par contre, son incohérence avec l'évaluation supervisée-pixel.

Concernant la contribution à la binarisation d'images de documents dégradées, deux nouvelles méthodes ont été proposées dans ce travail de thèse, et détaillées dans le chapitre 5. La première méthode effectue une classification directe des pixels de l'image en « avant-plan » ou « arrière-plan » en utilisant un réseau de neurones artificiel de type perceptron multicouches (MLP). Ce dernier est connu par sa capacité de généralisation à partir d'un ensemble d'apprentissage, le comportement pour d'autres données non vues auparavant. De ce fait, le MLP a été entraîné avec un certains nombre de pixels sélectionnés aléatoirement à partir de différentes images de documents dégradées. Le MLP entraîné est utilisé par la suite pour déterminer la classe d'appartenance de chaque pixel. Plusieurs expérimentations ont été menées afin de déterminer la topologie optimale du MLP, et d'évaluer les performances de la méthode proposée comparativement aux méthodes de la littérature. L'évaluation a été effectuée sur un ensemble d'images réelles et synthétiques, selon plusieurs points de vue et les résultats obtenus sont encourageants. La deuxième méthode s'agit d'une technique simple de binarisation basée sur le seuillage hybride. Cette technique a été proposée dans le but de surmonter les problèmes causés par les méthodes globales et locales, en combinant les avantages des deux familles de méthodes. La méthode proposée procède en deux étapes ; d'abord un premier seuillage global en utilisant l'algorithme d'Otsu est appliqué sur l'image entière dans le but de binariser la majorité des pixels. Ensuite, cinq techniques de seuillage locales sont combinées pour déterminer la valeur binaire des pixels restants de l'étape précédente. Malgré son principe simple, cette technique a montré des bonnes performances sur une collection de test composée d'images réelles et synthétique suivant différents aspects.

Perspectives

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

Comme le nombre de techniques de binarisation proposées augmente continuellement, la première des perspectives que nous nous fixons est d'élargir l'étude comparative effectuée pour qu'elle intègre d'autres méthodes de binarisation qui ne sont pas basées uniquement sur le seuillage. Il est ainsi important d'étudier les méthodes basées sur la classification et de rapprocher leurs performances pour déterminer quel classifieur permettra de produire les meilleurs résultats.

L'investigation sur les méthodes de seuillage local nous a permis de percevoir l'impact du choix des valeurs de paramètres sur les résultats de la binarisation. Une mauvaise initialisation de ces paramètres engendre des résultats de mauvaise qualité. Usuellement, les paramètres d'une technique de seuillage sont estimés à l'aide d'un humain qui détermine visuellement la valeur la plus adéquate pour chaque paramètre par essai-erreur. Ainsi, la conception d'un outil permettant le paramétrage automatique des techniques de binarisation constitue notre deuxième perspective.

Quant à la méthodologie d'évaluation que nous avons proposée, il est nécessaire de la tester sur une collection plus large d'images réelles pour confirmer son accord avec l'évaluation visuelle, et afin de s'assurer de ses résultats. Il nous semble intéressant aussi de développer le système de recherche intégré pour qu'il soit plus robuste contre les dégradations d'un côté et applicable sur des documents non-arabes de l'autre côté. Cependant le développement du système de recherche se fait en améliorant les différents traitements inclus dans le système, et chacun de ces traitements mérite une étude complète et approfondie.

La quatrième perspective de ce travail est d'améliorer les deux méthodes de binarisation proposées. L'amélioration de la méthode de seuillage hybride revient à pondérer la participation des différentes techniques de binarisation intégrées, et à réviser le calcul des seuils. Pour la méthode de binarisation basée sur les ANNs, l'amélioration repose principalement sur le développement de la procédure de sélection des pixels les plus représentatifs pour l'apprentissage. Nous pensons également à expérimenter d'autres classifieurs (HMM ou SVM par exemple) ou encore à combiner plusieurs classifieurs pour la binarisation d'images de documents.

Un autre travail important à accomplir est la construction et la publication d'une base d'images de documents historiques. Nous envisageons à construire une base de documents synthétiques en se basant sur les travaux de Baird [BAI 92], Rabeux [RAB 13], et Stathis et al. [STA 08]. La base de documents à construire doit contenir un grand nombre de documents avec les images de vérité terrain correspondantes, et des fichiers d'annotations décrivant des informations sur ces documents (nombre de mots, position des lignes, etc.).

Finalement la perspective générale que nous allons travailler à atteindre est de concevoir et d'implémenter une réelle plateforme de prétraitement et d'analyse de documents historiques qui regroupe plusieurs traitements, diverses techniques, et différentes mesures d'évaluation, et de le rendre accessible au grand public.

Bibliographie

- [ABD 06] A. Abdulkader, “A two-tier Arabic offline handwriting recognition based on conditional joining rules”, *Proceedings of the Conference on Arabic and Chinese Handwriting Recognition*, Maryland-USA, pp. 70-81, 2006.
- [ABU 89] A.S. Abutaleb, “Automatic thresholding of gray-level pictures using two-dimensional entropy”, *Computer Vision, Graphics, and Image Processing*, vol. 47, No. 1, pp. 22–32, 1989.
- [ADI 13] Y. Adibe, “Evaluation des acquis d’apprenants du Français Langue Etrangère en module de Technique d’Expression Ecrite et Orale en 1^{ère} année Universitaire”, *Thèse de doctorat en Didactique*, Université d’Oran – Algérie, 2012/2013.
- [AGR 09] M. Agrawal, D. Doermann, “Clutter noise removal in binary document images”, *Proceedings of the 10th International Conference on Document Analysis and Recognition (ICDAR)*, pp. 556-560, 2009.
- [AGU 05] J. Aguilera, H. Wildenauer, M. Kampel, M. Borg, D. Thirde, J. Ferryman, “Evaluation of motion segmentation quality for aircraft activity surveillance”, *Proceedings of the Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS)*, pp. 317-324, 2005.
- [ALB 04] M.P. de Albuquerque, I.A. Esquef, A.G. Mello, “Image thresholding using Tsallis entropy”, *Pattern Recognition Letters*, vol. 25, No. 9, pp. 1059-1065, July 2004.
- [ALB 06] A. Albaati, M. Kherallah, A.M. Alimi, A. Ennaji, “De l’hors-ligne vers un système de reconnaissance en-ligne: application à la modélisation de l’écriture Arabe manuscrite ancienne”, *Semaine du document numérique- Proceedings d’Atelier sur la Numérisation de l’Ecrit Ancien et des Grandes Masses de données (ANAGRAM)*, 2006.
- [ALL 03] B. Allier, “Contribution à la numérisation des collections : Apports des Contours Actifs”, *Thèse de doctorat*, Institut National des Sciences Appliquées de Lyon, 28/11/2003.
- [ALH 09] R. Al-Hajj, L. Likforman-Sulem, C. Mokbel, “Combining slanted-frame classifiers for improved HMM-based Arabic handwriting recognition”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, No. 7, pp. 1165-1177, 2009.
- [ALK 08] J.H. Alkhateeb, J. Ren, S.S. Ipson, J. Jiang, “Knowledge-based baseline detection and optimal thresholding for word segmentation in efficient preprocessing of handwritten Arabic text”, *Proceedings of the 5th International Conference on Information Technology: New Generation*, Las Vegas-Nevada, USA, pp. 1158-1159, 2008.
- [ALM 02] S. Al-Ma’adeed, D. Elliman, CA. Higgins, “A database for Arabic handwritten text recognition research”, *Proceedings of the International Workshop on Frontiers in Handwriting Recognition (IWFHR)*, pp. 485-489, 2002.
- [AMI 00] A. Amin, S. Fisher. “A document skew detection method using the Hough transform”, *Pattern Analysis & Applications*, vol. 3, No. 3, pp. 243–253, 2000.
- [AND 03] T. Andersen, W. Zhang, “Features for neural net based region identification of newspaper documents”, *Proceedings of the 7th International Conference on Document Analysis and Recognition (ICDAR)*, Washington, DC, USA, pp. 403-407, 2003.
- [ANT 05] A. Antonacopoulos, D. Karatzas, “Semantics-based content extraction in typewritten historical documents”, *Proceedings of the 8th International Conference on Document Analysis and Recognition (ICDAR)*, pp. 48–53, 2005.
- [ANT 04] A. Antonacopoulos, D. Karatzas, “Document image analysis for World War II personal records”, *Proceedings of the 1st International Workshop on Document Image Analysis for Libraries*, pp. 336–341, 2004.
- [ARA 06] H. K. G. Manjunath Aradhya, P. Shivakumara, “Skew Detection Technique for Binary Document Images based on Hough Transform”, *International Journal of Information Technology*, vol. 3, No. 3, 2006.

- [ATH 11] M. Athimethphat, “A Review on Global Binarization Algorithms for Degraded Document Images”, *AU Journal of Technology*, vol. 14, No. 3, pp. 188-195, January 2011.
- [AVI 04b] B.T. Avila, R.D. Lins, “Efficient Removal of Noisy Borders from Monochromatic Documents”, *Image Analysis and Recognition*, Springer Berlin Heidelberg, vol. 3212, pp. 249-256, 2004.
- [BAD 05] E. Badekas, N. Papamarkos, “Automatic Evaluation of Document Binarization Results”, *Progress in pattern recognition, image analysis and applications*, Springer Berlin Heidelberg, vol. 3773, pp. 1005-1014, 2005.
- [BAD 06] E. Badekas, N. Nikolaou, N. Papamarkos, “Text binarization in color documents”, *International Journal of Imaging Systems and Technology*, vol. 16, No. 6, pp. 262–274, 2006.
- [BAD 07] E. Badekas, N. Papamarkos, “Optimal Combination of Document Binarization Techniques Using a Self Organizing Map Neural Network”, *Engineering Applications of Artificial Intelligence*, vol. 20, No. 1, pp. 11-24, 2007.
- [BAD 09] E. Badekas, N. Papamarkos, “Estimation of Appropriate Parameter values for Document Binarization Techniques”, *International Journal of Robotics and Automation*, vol. 24, No. 1, pp. 66-78, 2009.
- [BAI 87] H.S. Baird, “The skew angle of printed documents”, *Proceedings of SPSE the 40th Symposium on Hybrid Imaging Systems*, Rochester, NY, pp. 739-743, 1987.
- [BAI 91] H.S. Baird, H. Bunke, K. Yamamoto, “Structured document image analysis”, *Springer Publishing Company, Incorporated*, 2011.
- [BAI 92] H.S. Baird, “Document Image Defect Models”, *Structured Document Image Analysis*, Springer Berlin Heidelberg, pp. 546-556, 1992.
- [BAI 93] H.S. Baird, “Document image defect models and their uses”, *Proceedings of the Second International Conference on Document Analysis and Recognition (ICDAR)*, pp. 62-67, Tsukuba-Japan, 1993.
- [BAP 98] F. Bapst, “Reconnaissance de documents assistée: architecture logicielle et intégration de savoir-faire”, *Thèse de doctorat*, Université de Fribourg - Suisse, 1998.
- [BAU 05] E. Baudrier, “Comparaison d’images binaires reposant sur une mesure locale des dissimilarités - Application à la classification”, *Thèse de doctorat*, Université de Reims Champagne-Ardenne, 9/12/2005.
- [BEL 92b] A. Belaid, Y. Belaid, “Reconnaissance des formes : méthodes et application”, *InterEdition* 1992.
- [BEL 01] A. Belaïd, “Reconnaissance automatique de l’écriture et du document”, *Pour la science (Article dans une revue de vulgarisation)*, 2001.
- [BEN 99] A. Bennisri, A. Zahour, B. Taconet, “Extraction des lignes d’un texte manuscrit arabe”, *Vision interface*, pp. 42-48, 1999.
- [BEN 08] A. Benouareth, E. Ennaji, and M. Sellami, “Semi-continuous HMM with explicit state duration for unconstrained Arabic word modeling and recognition”, *Pattern Recognition Letters*, vol. 29, No. 12, pp. 1742-1752, 2008.
- [BEN 11] I. Ben Messaoud, H. El Abed, H. Amiri, “New Method for the Selection of Binarization Parameters Based on Noise Features of Historical Documents”, *Proceedings of the 2011 Joint Workshop on Multilingual OCR and Analytics for Noisy Unstructured Text Data*, pp. 1, 2011.
- [BEN 12] I. Ben Messaoud, H. Amiri, H. El Abed, V. Margner, “Document Preprocessing System – Automatic Selection of Binarization”, *Proceedings of the 10th IAPR International Workshop on Document Analysis Systems (DAS)*, pp. 85-89, 2012.

- [BER 86] J. Bernsen, “Dynamic thresholding of grey-level images”, *Proceedings of the 8th International Conference on Pattern Recognition (ICPR)*, pp. 1251-1255, Paris - France, 1986.
- [BIB 11] Koninklijke Bibliotheek, Digitisation by Proquest of early printed books in KB collection, <http://www.kb.nl/en>. Visité le 17/03/2015.
- [BIN 04] N. Bin Rais, M.S. Hanif, I.A. Taj, “Adaptive Thresholding Technique for Document Image Analysis”, *Proceedings of the 8th International Multitopic Conference (INMIC)*, pp.61-66, 2004.
- [BLS 10] British Library Search Beta, http://searchbeta.bl.uk/primo_library/libweb/action/search.do. Visité le 17/03/2015.
- [BOS 08] D. Bossy, “Le traitement d’image dans l’analyse de documents anciens”, *Séminaire de recherche du groupe DIVA*, Université de Fribourg, 29 mai 2008.
- [BOT 98] L. Bottou, P. Haffner, P.G. Howard, P. Simard, Y. Bengio, Y. Lecun, “High quality document image compression with DjVu”, *Journal of Electronic Imaging*, vol. 7, No. 3, pp. 410-425, 1998.
- [BOU 97] F.L. Bourgeois, “Robust multifont OCR system from gray level images”, *Proceedings of the 4th International Conference on Document Analysis and Recognition (ICDAR)*, pp. 1–5, 1997.
- [BOU 06] W. Boussellaa, A. Zahour, B. Taconet, A. Benabdelhafid, A. Alimi, “Segmentation texte /graphique : Application au manuscrits Arabes Anciens”, *Proceedings of the 9^{ème} Colloque International Francophone sur l’Ecrit et le Document (CIFED)*, pp 139 – 144, 2006.
- [BOU 07] A.O. Boudraa, E.H.S. Diop, F. Salzenstein, A. Bouchikhi, “Seuillage d’images basé sur l’opérateur de Teager-Kaiser”, *Proceedings of the 21^{ème} Colloque du Groupe d’Etudes du Traitement du Signal et des Images (GRETSI)*, pp. 885-888, Troyes, 2007.
- [BOU 08] H. Boulehmi, B. Seddik, A. Kricha, N.E.B. Amara, “Prétraitement de documents anciens”, *Proceedings of the Colloque International Francophone sur l’Ecrit et le Document (CIFED)*, Groupe de Recherche en Communication Ecrite, pp. 209-210, 2008.
- [BOU 08b] M. Boughanem, “Recherche d’information : de la RI orientée système vers la RI orientée contexte”, *Conférence invitée, Colloque International Francophone sur l’Ecrit et le Document*, Rouen - France, 2008.
- [BRE 01] L. Breiman, “Random Forests”, *Machine Learning*, vol. 45, No. 1, pp. 5-32, 2001.
- [BRI 96] A.D. Brink, N.E. Pendock, “Minimum cross-entropy threshold selection”, *Pattern Recognition*, vol. 29, No. 1, pp. 179-188, 1996.
- [BUR 07] J.A. Burgoyne, L. Pugin, G. Eustace, I. Fujinaga, “A comparative survey of image binarisation algorithms for optical recognition on degraded musical sources”, *Proceedings of the International Society for Music Information Retrieval (ISMIR)*, pp. 509-512, Vienna-Austria, 2007.
- [CAO 03] H. Cao, X. Ding, C. Liu. “Rectifying the Bound Document Image Captured by the Camera: A Model Based Approach”, *Proceedings of the International Conference on Document Analysis and Recognition (ICDAR)*, pp. 71-75, Edinburgh - Scotland, 2003.
- [CAV 06] G.D.C. Cavalcanti, E.F.A. Silva, C. Zanchettin, B.L.D. Bezerra, R.C. Doria, J.C.B. Rabelo, “A Heuristic Binarization Algorithm for Documents with Complex Background”, *Proceedings of the International Conference on Image Processing (ICIP)*, pp. 389-392, 2006.
- [CHA 90] H. Charles, “L’évaluation, règles du jeu, des intentions aux outils”, *ESF éditeur*, Paris – France, 1990.
- [CHA 97] A. Chaudhuri, S. Chaudhuri, “Robust detection of skew in document images”, *IEEE Transactions on Image Processing*, vol. 6, No. 2, pp. 344-349, 1997.
- [CHA 06] C. Chatelain, “Extraction de séquences numériques dans des documents manuscrits quelconques”, *Thèse de doctorat*, Université de Rouen - France, 5/12/2006.

- [CHA 10] R. Chamchong, C.C. Fung, “Optimal Selection of Binarization Techniques for the Processing of Ancient Palm Leaf Manuscripts”, *Proceedings of the IEEE International Conference on Systems Man and Cybernetics (SMC)*, pp. 3796 - 3800, 2010.
- [CHA 12] M. Charfi, M. Kherallah, A. El Baati, A.M. Alimi, “A new approach for Arabic handwritten postal addresses recognition”, *The International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 3, No. 3, pp. 1 - 6, 2012.
- [CHE 93] T. Chen, M. Takagi, “Image Binarization by Back Propagation Algorithm”, *International Archives of Photogrammetry and Remote Sensing*, vol. 29, pp. 345-345, 1993.
- [CHE 95] S. Chen, R.M. Haralick, I.T. Phillips, “Automatic text skew estimation in document images”, *Proceedings of the 3rd International Conference on Document Analysis and Recognition (ICDAR)*, pp. 1153 - 1156, 1995.
- [CHE 98] H.D. Cheng, J. R. Chen, J. Li, “Threshold selection based on fuzzy c-partition entropy approach”, *Pattern Recognition*, vol. 31, No. 7, pp. 857-870, 1998.
- [CHE 98b] M. Cheriet, H. Miled, C. Olivier, Y. Lecourtier, “Visual Aspect of Cursive Arabic Handwriting Recognition”, *Proceedings of Vision Interface (VI)*, vol. 99, pp. 263-270, 1998.
- [CHE 99] M. Chen, X. Ding, “A robust skew detection algorithm for grayscale document image”, *Proceedings of the 5th International Conference on Document Analysis and Recognition (ICDAR)*, pp. 617-620, 1999.
- [CHE 13] L. Chergui, “Combinaison de classifieurs pour la reconnaissance de mots arabes manuscrits”, *Thèse de doctorat*, Université de Mentouri - Constantine, 2013.
- [CHI 92] N. Chinchor, “MUC-4 Evaluation Metrics”, *Proceedings of the 4th Message Understanding Conference*, pp. 22–29, 1992.
- [CHO 10] C.H. Chou, W.H. Lin, F. Chang, “A binarization method with learning-built rules for document images produced by cameras”, *Pattern Recognition*, vol. 43, No. 4, pp. 1518-1530, 2010.
- [CHR 01] A.M. Christin, “Histoire de l’écriture : De l’idéogramme au multimédia”, *Ouvrage collectif, Flammarion (Ed.)*, 405 p, 2001.
- [CIN 02] L. Cinque, S. Levialdi, L. Lombardi, S. Tanimoto, “Segmentation of page images having artifacts of photocopying and scanning”, *Pattern Recognition*, vol. 35, No. 5, pp. 1167–1177, 2002.
- [COC 95] J.P. Cocquerez, S.Philipp-Foliguet, “Analyse d’images: filtrage et segmentation”, *Masson Ed.*, Paris-France, 1995.
- [CYB 89] G. Cybenko, “Approximation by superposition of a sigmoidal function”, *Mathematics of control, Signals, and Systems*, vol. 2, No. 4, pp. 303-314, 1989.
- [DEF 95] O. Deforges, P. Piquin, C. Viard-Gaudin, D. Barba, “Segmentation d’images de documents par une approche multirésolution. Extraction précise des lignes de texte”, *Traitement du Signal*, vol. 12, No. 6, pp. 527-539, 1995.
- [DIN 08] X. Ding, H. Liu, “Segmentation-driven offline handwritten Chinese and Arabic script recognition”, *Proceedings of the Conference on Arabic and Chinese Handwriting Recognition (SACH)*, pp. 196-217, Berlin-Germany, 2008.
- [DOY 62] W. Doyle, “Operation useful for similarity-invariant pattern recognition”, *Journal of the Association for Computing Machinery (JACM)*, vol. 9, No. 2, pp. 259-267, 1962.
- [DRE 02] G. Dreyfus, J.M. Martinez, M. Samuelides, M.B. Gordon, F. Badran, S. Thiria, L. Héroult, “Réseaux de neurones : méthodologies et applications”, *Editions Eyrolles*, Paris - France, 2002.
- [DRI 07] F. Drira, “Contribution à la Restauration des Images de Documents Anciens”, *Thèse de doctorat*, Institut National des Sciences Appliquées de Lyon, 14/12/2007.
- [DU 09] X. Du, W. Pan, T.D. Bui, “Text line segmentation in handwritten documents using Mumford-Shah model”, *Pattern Recognition*, vol. 42, No. 12, pp. 3136–3145, 2009.

- [DUB 01] E. Dubois, A. Pathak. “Reduction of bleed-through in scanned manuscripts documents”, *Proceeding of the Image processing, Image Quality, Image capture Systems conference (PICS)*, pp. 177-180, Montreal-Canada, 2001.
- [DUO 05] J. Duong, “Etude des documents imprimés : approche statistique et contribution méthodologique”, *Thèse de doctorat*, Institut National des Sciences Appliquées de Lyon, 30/06/2005.
- [EGM 02] M. Egmont-Petersen, D. de Ridder, H. Handels, “Image processing with neural networks - a review”, *Pattern Recognition*, vol. 35, No. 10, pp. 2279-2301, 2002.
- [ELA 08] H. El Abed, V. Märgner, “Base de données et compétitions-Outils de développement et d'évaluation de systèmes de reconnaissance de mots manuscrits arabes”, *Proceedings of the 10^{ème} Colloque International Francophone sur l'Ecrit et le Document (CIFED)*, pp. 103-108, 2008.
- [ELL 90] D.G. Elliman, “A review of segmentation and contextual analysis techniques for text recognition”, *Pattern Recognition*, vol. 23, No. 3, pp. 316-337, 1990.
- [ESI 10] G.P. E-Silva, R.D. Lins, J.M. Silva, “HistDoc-a toolbox for processing images of historical documents”, *Image Analysis and Recognition*, Springer Berlin Heidelberg, pp. 409-419, 2010.
- [EUR 08] Europeana, <http://www.europeana.eu/portal/>. Visité le 17/03/2015.
- [FAN 02] C. Fan, Y.K. Wang, T.R. Lay, “Marginal noise removal of document images”, *Pattern Recognition*, vol. 35, No. 11, pp. 2593-2611, 2002.
- [FAR 13] A. Farahmand, A. Sarrafzadeh, J. Shanbehzadeh, “Document Image Noises and Removal Methods”, *Proceedings of the International MultiConference of Engineers and Computer Scientists*, vol.1, pp. 436 – 440, 2013.
- [FEL 01] M. Feldbach, K.D. Tönnies, “Line detection and segmentation in historical church registers”, *Proceedings of the 6th International Conference on Document Analysis and Recognition (ICDAR)*, pp. 743–748, Washington-USA, 2001.
- [FEN 04] M.L. Feng, Y.P. Tan, “Adaptive binarization method for document image analysis”, *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME)*, pp. 339-342, 2004.
- [FER 10] B. Fernando, S. Karaoglu, “Extreme Value Theory Based Text Binarization in Documents and Natural Scenes”, *Proceedings of the 3rd International Conference on Machine Vision (ICMV)*, pp. 144-151, 2010.
- [FIS 58] W.D. Fisher, “On grouping for maximum homogeneity”, *Journal of the American Statistical Association*, vol. 53, No. 284, pp. 789-798, 1958.
- [FUN 10] C.C. Fung, R. Chamchong, “A Review of Evaluation of Optimal Binarization Technique for Character Segmentation in Historical Manuscripts”, *Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining*, pp. 236- 240, 2010.
- [GAD 00] T. Gadi, R. Benslimane, “Segmentation hiérarchique floue”, *Traitement du signal*, vol. 7, No. 1, 2000.
- [GAL 97] Bibliothèque numérique Gallica, <http://gallica.bnf.fr/>. Visité le 17/03/2015.
- [GAN 11] B. Gangamma, K. Srikanta Murthy, “Enhancement of Degraded Historical Kannada Documents”, *International Journal of Computer Applications*, vol. 29, No. 11, pp. 1-6, 2011.
- [GAP 06] U. Garain, T. Paquet, L. Heutte, “On Foreground-Background Separation in Low Quality Document Images”, *International Journal of Document Analysis and Recognition (IJ DAR)*, vol. 8, No. 1, pp. 47-63, 2006.
- [GAR 04] C. Garcia, M. Delakis, “Convolutional face finder: A neural architecture for fast and robust face detection”, *IEEE Transactions on pattern analysis and machine intelligence*, vol. 26, No. 11, pp. 1408-1423, 2004.

- [GAR 06] U. Garain, T. Paquet, L. Heutte, “On foreground-background separation in low quality document images”, *International Journal of Document Analysis and Recognition (IJ DAR)*, vol. 8, No. 1, pp. 47-63, 2006.
- [GAT 06] B. Gatos, I. Pratikakis, S.J. Perantonis, “Adaptive Degraded Document Image Binarization”, *Pattern Recognition*, vol. 39, No. 3, pp. 317-327, 2006.
- [GAT 09] B. Gatos, K. Ntirogiannis, I. Pratikakis, “ICDAR 2009 Document Image Binarization Contest (DIBCO 2009)”, *Proceedings of the International Conference on Document Analysis and Recognition (ICDAR)*, pp. 1375-1382, 2011.
- [GIR 89] F. Girosi, T. Poggio, “Representation properties of networks: Kolmogorov’s theorem is irrelevant”, *Neural Computation*, vol. 2, No. 4, pp. 465-469, 1989.
- [GIU 77] E. Giuliano, O. Paitra, L. Stringer, “Electronic character reading system”, *U.S. Patent*, No. 4.042.15, 1977.
- [GON 07] R. Gonzalez, R. Woods, “Digital Image Processing”, (3rd Edition), *Prentice-Hall*, Upper Saddle River, NJ, USA, 2007.
- [GOO 04] Google Livres, <http://books.google.com/>. Visité le 17/03/2015.
- [GOY 11] R. Goyal, A. Kaur, “A Review of Optimal Binarization Techniques on Documents with Damaged Background”, *International Journal of Computer Science and technology (IJCST)*, vol. 2, No. 2, pp.237-239, 2011.
- [GRA 00] I. Granado, M. Mengucci, F. Muge, “Extraction de textes et de figures dans les livres anciens à l’aide de la morphologie mathématique”, *Proceedings of the Colloque International Francophone sur l’Ecrit et le Document (CIFED)*, pp. 81-90, 2000.
- [GRA 07] H. Grayloo, H. Kabir, “Fixing undesirable fragmentations in Persian printed documents using morphological dilation”, *Proceedings of the 12th International Conference of Computer Society of Iran*, 2007.
- [GRA 09] A. Graves, J. Schmidhuber, “Offline Handwriting recognition with multidimensional recurrent neural networks”, *Proceedings of the Advances in Neural Information Processing Systems conference*, pp. 545-552, 2009.
- [GUP 07] M.R. Gupta, N.P. Jacobson, E.K. Garcia, “OCR binarization and image pre-processing for searching historical documents”, *Pattern Recognition*, vol. 40, No. 2, pp. 389–397, 2007.
- [GUS 99] Gusnard de Ventadert (nom collectif), “Les documents anciens”, *Document numérique*, Hermès, vol. 3, No 1-2, pp. 57-73, 1999.
- [HAD 06] K. Hadjar, “Une étude de l’évolutivité des modèles pour la reconnaissance de documents arabes dans un contexte interactif”, *Thèse de doctorat*, Université de Fribourg - Suisse, 2006.
- [HAM 05] H. Hamza, E. Smigiel, A. Belaid, “Neural based binarization techniques”, *Proceedings of the 8th International Conference on Document Analysis and Recognition (ICDAR)*, pp. 317-321, 2005.
- [HAR 72] L.D. Harmon, “Automatic recognition of print and scripts”, *Proceedings of the IEEE*, vol. 60, No. 10, pp. 1165–1177, 1972.
- [HE 05] J. He, Q.D.M. Do, A.C. Downton, J.H. Kim, “A Comparison of Binarization Methods for Historical Archive Documents”, *Proceedings of the 8th International Conference on Document Analysis and Recognition (ICDAR)*, pp. 538-542, 2005.
- [HEN 68] R.B. Hennis, “The IBM 1975 optical page reader, system design”, *IBM Journal of Research and Development*, vol. 12, No. 5, pp. 346–353, 1968.
- [HER 88] L. Hertz, R.W. Schafer, “Multilevel thresholding using edge matching”, *Computer Vision, Graphics, and Image Processing*, vol. 44, No. 3, pp. 279–295, 1988.

- [HIR 01] H. Takahashi, H. Hata, K. Noguchi, "Image processing device for providing correction of a curvature distortion of a bound portion of a spread book", Brevet US6330050 B1, Ricoh Company, LTD, Washington, 2001.
- [HOB 97] J.D. Hobby, T.K. Ho, "Enhancing Degraded Document Images via Bitmap Clustering and Averaging", *Proceeding of the 4th International Conference on Document Analysis and Recognition (ICDAR)*, pp. 394-400, 1997.
- [HON 02] X. Hong-bo, T. Yan, "Skew Detection for Binary Document Images Using Mathematical Morphology", *Wuhan University Journal of Natural Sciences*, vol. 7 No. 3, pp. 338-340, 2002.
- [HUA 95] L.K. Huang, M.J.J. Wang, "Image thresholding by minimizing the measures of fuzziness", *Pattern Recognition*, vol. 28, No. 1, pp. 41-51, 1995.
- [HUL 98] J.J. Hull, "Document Image Skew Detection: Survey and Annotated Bibliography", *Series in Machine Perception and Artificial Intelligence*, vol. 29, pp. 40-66, 1998.
- [HYV 99] A. Hyvarinen, E. Oja, "A fast fixed-point algorithm for independent component analysis", *Neural Computation*, vol. 9, No. 7, pp. 1483-1492, 1997.
- [IDP] The International Dunhuang Project (IDP), <http://idp.bl.uk/>. Visité le 17/03/2015.
- [IMP] Improving Access to Text (IMPACT), <http://www.impact-project.eu/>. Visité le 17/03/2015.
- [JAC 14] B. Jacob, S.B. Waykar, "A Survey on Binarization of Historical Degraded Documents", *International Journal of Informative & Futuristic Research*, vol. 2, No. 2, pp. 400-405, 2014.
- [JAW 97] C.V. Jawahar, P.K. Biswas, A.K. Ray, "Investigations on fuzzy thresholding based on fuzzy clustering", *Pattern Recognition*, vol. 30, No. 10, pp. 1605-1613, 1997.
- [JIA 99] X. Jiang, H. Bunke, D. Widmer-Kljajic, "Skew Detection of Document Images by Focused Nearest-Neighbor Clustering", *Proceeding of the 5th International Conference on Document Analysis and Recognition (ICDAR)*, pp. 629-632, 1999.
- [JOH 82] G. Johannsen, J. Bille, "A threshold selection method using information measures", *Proceedings of the 6th International Conference on Pattern Recognition (ICPR)*, pp. 140-143, Munich-Germany, 1982.
- [JOU 06] N. Journet, "Analyse d'images de documents anciens : une approche texture", *Thèse de doctorat*, Université de La Rochelle - France, 1/12/2006.
- [KAL 57] A.N. Kalmogorov, "On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition", *Dokl. Acad. Nauk SSSR*, vol. 114, pp. 953-956, 1957.
- [KAM 93] M. Kamel, A. Zhao, "Extraction of binary character/graphics images from grayscale document images", *CVGIP: Graphical Models and Image Processing*, vol. 55, No. 3, pp. 203-217, 1993.
- [KAM 99] H. Kamada, K. Fujimoto, "High-speed, high-accuracy binarization method for recognizing text in images of low spatial resolutions", *Proceedings of the 5th International Conference on Document Analysis and Recognition (ICDAR)*, pp.139-142, 1999.
- [KAN 93] T. Kanungo, R.M. Haralick, I. Phillips. "Global and Local Document Degradation Models", *Proceedings of the 2nd International Conference on Document Analysis and Recognition (ICDAR)*, pp. 730-734, Tsukuba-Japan, 1993.
- [KAN 00] T. Kanungo, R. M. Haralick, H. S. Baird, W. Stuezel, D. Madigan, "A Statistical, Nonparametric Methodology for Document Degradation Model Validation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, No. 11, pp. 1209-1223, 2000.
- [KAP 85] J.N. Kapur, P.K. Sahoo, A.K.C. Wong, "A New method for gray-level picture threshold using the entropy of the histogram", *Computer Vision, Graphics, and Image Processing*, vol.29, No. 3, 1985.

- [KAU 14] D. Jagroop Kaur, R. Mahajan, “A Review of Degraded Document Image Binarization Techniques”, *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 3, No. 5, pp. 6581-6586, May 2014.
- [KAV 05] E. Kavallieratou, “A Binarization Algorithm Specialized on Document images and Photos”, *Proceedings of the 8th International Conference on Document Analysis and Recognition (ICDAR)*, pp. 463-467, 2005.
- [KAV 06] E. Kavallieratou, S. Stathis, “Adaptive Binarization of Historical Document Images”, *Proceedings of the 18th International Conference on Pattern Recognition (ICPR)*, vol. 3, pp. 742-745, 2006.
- [KAV 08] E. Kavallieratou, “An objective way to evaluate and compare binarization algorithms”, *Proceedings of the 2008 ACM symposium on Applied computing (SAC)*, pp. 397-401, 2008.
- [KEF 09] A. Kefali, T. Sari, M. Sellami, “Evaluation de plusieurs techniques de seuillage d’images de documents arabes anciens”, *Proceedings of the 5^{ème} Symposium International Images Multimédias Applications Graphiques et Environnements (IMAGE)*, pp.123-134, Biskra-Algeria, 2009.
- [KEF 10] A. Kefali, “Concurrence entre Concurrence entre recherche approximative et classification pour la recherche d’images de documents Arabes dégradés”, *Mémoire de Magistère*, Université Badji Mokhtar-Annaba, 20 Juillet 2010.
- [KEF 14] A. Kefali, T. Sari, H. Bahi, “Foreground-Background Separation by Feed-forward Neural Networks in Old Manuscripts”, *Informatica*, vol. 38, No. 4, pp. 329-338, 2014.
- [KEF 14b] A. Kefali, T. Sari, H. Bahi, “Text/ Background separation in the degraded document images by combining several thresholding techniques”, *WSEAS Transactions on Signal Processing*, Vol. 10, pp. 436-443, 2014.
- [KET 10] D. Ketata, M. Khemakhem, “Un survol sur l’analyse et la reconnaissance de documents : imprimé, ancien et manuscrit”, *Proceedings of the Colloque International Francophone sur l’Écrit et le Document (CIFED)*, 12 pages, Sousse-Tunisie, 2010.
- [KHA 07] A. Khashman , B. Sekeroglu, “A Novel Thresholding Method for Text Separation and Document Enhancement”, *Proceedings of the 11th Panhellenic Conference in Informatics (PCI)*, pp. 323-330, 2007.
- [KHA 07b] A. Khashman, B. Sekeroglu, “Global Binarization of Document Images Using a Neural Network”, *Proceedings of the 3rd International IEEE Conference on Signal-Image Technologies and Internet-Based System (SITIS)*, pp. 665-672, Shanghai, 2007.
- [KHU 09] K. Khurshid, I. Siddiqi, C. Faure, N. Vincent, “Comparison of Niblack inspired Binarization methods for ancient documents”, *Proceedings of the 16th Document Recognition and Retrieval Conference (DRR)*, USA, 2009.
- [KHU 09b] K. Khurshid, C. Faure, N. Vincent, “Recherche de mots dans des images de documents par appariement de caractères”, *Proceedings of the 10^{ème} Colloque International Francophone sur l’Écrit et le Document (CIFED)*, pp. 91-96, 2008.
- [KIE 14] V.C. Kieu, M. Mehri, V. Rabeux, N. Journet, M. Visani, “Génération d’images semi-synthétiques de documents anciens à des fins d’évaluation de performances et d’apprentissage”, *Proceedings of the Colloque International Francophone sur l’Écrit et le Document (CIFED)*, Tours-France, 2014.
- [KIM 02] I.K. Kim, D.W. Jung, R.H. Park, “Document image binarization based on topographic analysis using a water flow model”, *Pattern Recognition*, vol. 35, No. 1, pp. 265-277, 2002
- [KIS 98] K. Kise, A. Sato, M. Iwata, “Segmentation of page images using the area Voronoi diagram”, *Computer Vision and Image Understanding*, vol. 70, No. 3, pp. 370-382, 1998.
- [KIT 86] J. Kittler, J. Illingworth, “Minimum error thresholding”, *Pattern Recognition*, vol. 19, No. 1, pp. 41-47, 1986.

- [KIT 00] N. Damera-Venkata, T.D. Kite, W.S. Geisler, B.L. Evans, A.C. Bovik, "Image quality assessment based on a degradation model", *IEEE Transactions on Image Processing*, vol. 9, No. 4, pp. 636-650, 2000.
- [KIT 10] K. Kita, T. Wakahara, "Binarization of Color Characters in Scene Images Using k-means Clustering and Support Vector Machines", *Proceedings of the 20th International Conference on Pattern Recognition (ICPR)*, pp. 3183-3186, 2010.
- [KOT 07] S.B. Kotsiantis, "Supervised Machine Learning: A Review of Classification Techniques", *Informatica*, vol.31, No. 3, pp. 249-268, 2007.
- [KUM 13] D. Kumar, M.N Anil Prasad, A.G. Ramakrishnan, "Evaluation of document binarization using eigen value decomposition", *Proceedings of the Document Recognition and Retrieval Conference (DRR)*, Burlingame, California-USA, 2013.
- [KUN 07] A. Kundu, T. Hines, J. Philips, B.D. Hucky, L.C. Van Guilder, "Arabic handwriting recognition using variable duration HMM", *Proceedings of the 9th International Conference on Document Analysis and Recognition (ICDAR)*, vol. 2, pp. 644-648, Curitiba-Brazil, 2007.
- [LAR 14] Dictionnaire français « *Larousse* » (en ligne), <http://www.larousse.fr/dictionnaires/francais/>. Visité le 17/03/2015.
- [LAS 12] L. Laskov, "Adaptive document image binarization with application in processing astronomical logbooks", *Serdica Journal of Computing*, vol.6, No. 1, pp. 111-120, 2012.
- [LAZ 10] J. Lazaro, J.L Martin, J. Arias, A. Astarloa, C. Cuadrado, "Neuro Semantic Thresholding for High Precision OCR Applications", *Image and Vision Computing*, vol. 28, No. 4, pp. 571-578, 2010.
- [LE 96] D.X. Le, G.R. Thoma and H. Wechsler, "Automated borders detection and adaptive segmentation for binary document images", *Proceedings of the International Conference on Pattern Recognition (ICPR)*, pp. 737-741, Vienna-Austria, 1996.
- [LEE 90] S.U. Lee, S.Y. Chung, R.H. Park, "A comparative performance study of several global thresholding techniques for segmentation", *Computer Vision, Graphics, and Image Processing*, vol. 52, No. 2, pp. 171-190, 1990.
- [LEE 01] S.W. Lee, D.S. Ryu, "Parameter-Free Geometric Document Layout Analysis", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, No. 11, pp. 1240-1256, 2001.
- [LEE 02] G. Leedham, S. Varma, A. Patankar, V. Govindaraju, "Separating text and background in degraded documents images - a comparison of global thresholding techniques for multi-stage thresholding", *Proceedings of the 8th International Workshop on Frontiers in Handwriting Recognition*, pp. 244-249, 2002.
- [LEE 03] G. Leedham, C. Yan, K. Takru, J.H.N. Tan, L. Mian, "Comparison of Some Thresholding Algorithms for Text/Background Segmentation in Difficult Document Images", *Proceedings of the 12th International Conference on Document Analysis and Recognition (ICDAR)*, vol.2, pp. 859-864, 2003.
- [LEV 66] A. Levenstein, "Binary Codes Capable of Correcting Deletions, Insertions and Reversals", *Soviet Physics Doklady*, vol. 10, No. 8, pp. 707-710, 1966.
- [LEV 85] M.D. Levine, A.M. Nazif, "Dynamic measurement of computer generated image segmentations", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 7, No. 25, pp. 155-164, 1985.
- [LEY 04] Y. Leydier, F.L. Bourgeois, H. Emptoz, "Serialized k-Means for Adaptive Color Image Segmentation Application to Document Images and Others", *Document Analysis Systems VI*, Springer Berlin Heidelberg, pp. 252-263, 2004.
- [LEY 06] Y. Leydier, "Numérisation et exploitation des manuscrits médiévaux", *Thèse de doctorat*, Institut National des Sciences Appliquées de Lyon, 2006.

- [LI 93] C.H. Li, C.K. Lee, “Minimum Cross Entropy Thresholding”, *Pattern Recognition*, vol. 26, No. 4, pp. 617-625, 1993.
- [LI 08] Y. Li, Y. Zheng, D. Doermann, S. Jaeger, “Script independent text line segmentation in free style handwritten documents”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, No. 8, pp. 1313-1329, 2008.
- [LIK 94] L. Likforman-Sulem, C. Faure, “Extracting lines on handwritten documents by perceptual grouping”, *Advances in Handwriting and drawing : a multidisciplinary approach*, Europia press, Paris, , pp. 117-135, 1994.
- [LIK 95] L. Likforman-Sulem, A. Hanimyan, C. Faure, “A Hough Based Algorithm for Extracting Text Lines in Handwritten Documents”, *Proceedings of the 3rd International Conference on Document Analysis and Recognition (ICDAR)*, pp. 774-777, 1995.
- [LIK 03] L. Likforman-Sulem, “Apport du traitement des images à la numérisation des documents manuscrits anciens”, *Document Numérique*, vol. 7, No. 3, pp. 13-26, 2003.
- [LIK 08] L. Likforman-Sulem, “Analyse d’images de documents non contraints: de la structuration à la reconnaissance”, *Thèse d’habilitation à Diriger des Recherches*, Université Pierre et Marie Curie (Paris VI), 16/07/2008.
- [LIN 09] R.D. Lins, “A Taxonomy for Noise Detection in Images of Paper Documents-The Physical Noises”, *Proceedings of the 6th International Conference on Image Analysis and Recognition (ICIAR)*, pp. 844-854, Halifax-Canada, 2009.
- [LIN 10] R.D. Lins, S. Banergee, M. Thielo, “Automatically detecting and classifying noises in document images”, *Proceedings of the 2010 ACM Symposium on Applied Computing*, pp. 33-39, 2010.
- [LIN 14] Dictionnaire de la langue française « *Linternaute* » (en ligne), <http://www.linternaute.com/dictionnaire/fr/definition/>. Visité le 17/03/2015.
- [LIT 14] Dictionnaire de français « *Littré* » (en ligne), <http://litre.reverso.net/dictionnaire-francais/>. Visité le 17/03/2015.
- [LIU 94] J. Liu, Y.H. Yang, “Multiresolution color image segmentation”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, No. 7, pp. 689-700, 1994.
- [LLO 85] D.E. Lloyd, “Automatic Target Classification Using Moment Invariant of Image Shapes”, Technical Report RAE IDN AW126, Farnborough-UK, December 1985.
- [LOU 09] G. Louloudis, B. Gatos, I. Pratikakis, C. Halatsis, “Text line and word segmentation of handwritten documents”, *Pattern Recognition*, vol. 42, No. 12, pp. 3169-3183, 2009.
- [LU 04] H. Lu, A.C. Kot, Y.Q. Shi, “Distance-Reciprocal Distortion Measure for Binary Document Images”, *IEEE Signal Processing Letters*, vol. 11, No. 2, pp. 228-231, 2004.
- [LU 04b] Y. Lu, Z. Wang, C.L. Tan, “Word grouping in document images based on Voronoi tessellation”, *Document Analysis Systems VI*, Springer Berlin Heidelberg, pp. 147-157, 2004.
- [MAH 94] A.S. Mahmoud, “Arabic Character Recognition Using Fourier Descriptors and Character Contour Encoding”, *Pattern Recognition*, vol. 27, No. 6, pp. 815-824, 1994.
- [MAL 09] V. Malleron, V. Eglin, H. Emptoz, S. Dord-Crouslé, P. Régner, “Text lines and snippets extraction for 19th century handwriting documents layout analysis”, *Proceedings of the International Conference on Document Analysis and Recognition (ICDAR)*, pp. 1001-1005, 2009.
- [MAR 01] V. Märgner, M. Pechwitz, “Synthetic data for Arabic OCR system development”, *Proceedings of the 6th International Conference on Document Analysis and Recognition (ICDAR)*, pp. 1159-1163, 2001.
- [MAR 08] S. Marinai, “Introduction to document analysis and recognition”, *Machine learning in document analysis and recognition*, Springer Berlin Heidelberg, pp. 1-20, 2008.

- [MEL 00] C.A.B. Mello, R.D. Lins, "Image segmentation of historical documents", *Visual 2000*, Mexico City, Mexico, August 2000.
- [MEL 08] C.A.B. Mello, A.L. Oliveira, A. Sanchez, "Historical Document Image Binarization", *Proceedings of the International Conference on Computer Vision Theory and Applications*, vol. 1, pp. 108-113, Funchal-Portugal, 2008.
- [MEN 00] M. Mengucci, I. Granado, "Morphological Segmentation of text and figures in Renaissance books (XVI Century)", *Mathematical Morphology and its applications to image processing*, Springer US, pp. 397-404, 2000.
- [MEN 14] M. Liu, C. Li, W. Zhu, A. Lim, "A Morphology-Based Border Noise Removal Method for Camera-Captured Label Images", *Camera-Based Document Analysis and Recognition*, Springer International Publishing, pp. 126-138, 2014.
- [MOG 09] R.F. Moghaddam, M. Cheriet, "Word-spotting in historical document images", *Proceedings of the International Conference on Document Analysis and Recognition (ICDAR)*, pp. 511-515, 2009.
- [NAG 84] G. Nagy, S. Seth, "Hierarchical Representation of optically scanned documents", *Proceedings of the 7th International Conference on Pattern Recognition (ICPR)*, pp. 347-349, 1984.
- [NAG 88] G. Nagy, J. Kanai, M. Krishnamoorthy, M. Thomas, M. Viswanathan, "Two complementary techniques for digitized document analysis", *Proceedings of the ACM conference on Document processing systems*, pp. 169-176, New York-NY-USA, 1988.
- [NAG 95] G. Nagy, "Document Image Analysis: Automated Performance Evaluation", *Document Analysis Systems*, World Scientific, vol. 14, pp. 137-156, 1995.
- [NAG 00] G. Nagy, "Twenty years of document image analysis in PAMI", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, No. 1, pp. 38-62, 2000.
- [NAV 10] NAVigation into DOcument MASSes (NaviDoMass), <http://navidomass.univ-lr.fr/>. Visité le 17/03/2015.
- [NEW 99] W. Newman, C. Dance, A. Taylor, S. Taylor, M. Taylor, T. Aldhous, "CamWorks: A Video-based Tool for Efficient Capture from Paper Source Documents", *Proceedings of the IEEE International Conference on Multimedia Computing and Systems*, vol. 2, pp. 647-653, 1999.
- [NIB 86] W. Niblack, "An Introduction to Digital Image Processing", *Prentice Hall*, pp. 115-116, 1986.
- [NIC 04] S. Nicolas, T. Paquet, L. Heutte, "Text line segmentation in handwritten document using a production system", *Proceedings of the International Workshop on Frontiers in Handwriting Recognition (IWFHR)*, pp. 245-250, 2004.
- [NIC 09] A. Nicolaou, B. Gatos, "Handwritten text line segmentation by shredding text into its lines", *Proceedings of the International Conference on Document Analysis and Recognition (ICDAR)*, pp. 626-630, 2009.
- [NIR 11] N. Kaur, R. Kaur, "A review on various methods of image thresholding", *International journal on Computer Science and Engineering (IJCSE)*, vol. 3, No. 10, pp. 3441-3443, October 2011.
- [NOM 05] S. Nomura, K. Yamanaka, O. Katai, H. Kawakami, T. Shiose, "A novel adaptive morphological approach for degraded character image segmentation", *Pattern Recognition*, vol. 38, No. 11, pp. 1961-1975, 2005.
- [NTI 08] K. Ntirogiannis, B. Gatos, I. Pratikakis, "An Objective Evaluation Methodology for Handwritten Image Document Binarization Techniques", *Proceedings of the 11th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pp. 586-591, Montreal-Canada, August 2008.

- [NTI 08b] K. Ntirogiannis, B. Gatos, I. Pratikakis, “An Objective Evaluation Methodology for document Image Binarization Techniques”, *Proceedings of the 8th IAPR International Workshop on Document Analysis Systems (DAS)*, pp. 217-224, Nara-Japan, September 2008.
- [NTI 13] K. Ntirogiannis, B. Gatos, I. Pratikakis, “Performance Evaluation Methodology for Historical Document Image Binarization”, *IEEE Transactions on Image Processing*, vol. 22, No.2, pp. 595-609, 2013.
- [NTI 14] K. Ntirogiannis, B. Gatos, I. Pratikakis, “ICFHR2014 Competition on Handwritten Document Image Binarization (H-DIBCO 2014)”, *Proceedings of the 14th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pp. 809-813, 2014.
- [OCA] Open Content Alliance (OCA), <http://www.opencontentalliance.org/>. Visité le 17/03/2015.
- [OFF 14] Office québécois de la langue française, « *Le grand dictionnaire terminologique* » (en ligne), <http://www.granddictionnaire.com/>. Visité le 17/03/2015.
- [OGI 08] J.M. Ogier, “Ancient Document Analysis: A Set of New Research Problems”, *Colloque International Francophone sur l'Écrit et le Document (CIFED)*, Conférence invitée, Rouen-France, 2008.
- [OGO 93] L. O’Gorman, “The document spectrum for page layout analysis”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, No. 11, pp. 1162-1173, 1993.
- [OGO 94] L. O’Gorman, “Experimental Comparisons of Binarization and Multi-Thresholding Methods on Document Images”, *Proceedings of the 12th IAPR International Conference on Pattern Recognition (ICPR)*, vol. 2, pp. 395-398, Oct 1994.
- [OGO 94b] L. O’Gorman, “Binarization and multi-thresholding of document images using connectivity”, *CVGIP: Graphical Models and Image Processing*, vol. 56, No. 6, pp. 494-506, 1994.
- [OKU 99] O. Okun, M. Pietikainen, J. Sauvola, “Robust Skew Estimation on Low-Resolution Document Images”, *Proceedings of the 5th International Conference on Document Analysis and Recognition (ICDAR)*, pp. 621-624, 1999.
- [OTS 79] N. Otsu, “A threshold selection method from gray-level histograms”, *IEEE Transactions on Systems, Man and Cybernetics*, vol. 9, No. 1, pp. 62–66, 1979.
- [OUW 10] N. Ouwayed, “Segmentation en lignes de documents anciens : application aux documents arabes”, *Thèse de doctorat*, Université de Nancy 2, 11/06/2010.
- [OZT 99] E. Öztop, A.Y. Mulayim, V. Atalay, F. Yarman-Vural, “Repulsive attractive network for baseline extraction on document images”, *Signal Processing*, vol. 75, No. 1, pp. 1-10, 1999.
- [PAL 86] P.W. Palumbo, P. Swaminathan, S.N. Srihari, “Document image binarization: Evaluation of algorithms”, *Proceedings of Applications of Digital Image Processing IX, International Society for Optics and Photonics*, pp. 278-285, 1986.
- [PAP 00] N. Papamarkos, C. Strouthopoulos, I. Andreadis, “Multithresholding of color and gray-level images through a neural network technique”, *Image and Vision Computing*, vol. 18, No. 3, pp. 213-222, 2000.
- [PAP 03] N. Papamarkos, “A neuro-fuzzy technique for document binarisation”, *Neural Computing & Applications*, vol. 12, No. 3-4, pp. 190-199, 2003.
- [PAP 11] A. Papandreou, B. Gatos, “A Novel Skew Detection Technique Based on Vertical Projections”, *Proceedings of the International Conference on Document Analysis and Recognition (ICDAR)*, pp. 384-388, 2011.
- [PAR 10] R. Paredes, E. Kavallieratou, “ICFHR 2010 Contest: Quantitative Evaluation of Binarization Algorithms”, *Proceedings of the International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pp. 733-736, 2010.

- [PAS 96] Recherche d'ornement Passe-Partout, <http://www2.unil.ch/BCUTodai/app/Todai.do>. Visité le 17/03/2015.
- [PAT 10] T. Patnaik, S. Gupta, D. Arya, "Comparison of Binarization Algorithm in Indian Language OCR", *Proceedings of ASCNT*, pp. 61-69, Noida-India, 2010.
- [PEE 04] W. Peerawit, A. Kawtrakul, "Marginal Noise Removal from Document Images Using Edge Density", *Proceedings of 4th Information and Computer Engineering Postgraduate Workshop*, Phuket-Thailand, 2004.
- [PEN 12] X. Peng, H. Cao, K. Subramanian, R. Prasad, P. Natarajan, "Information Extraction from Historical Semi-Structured Handwritten Documents", *Proceedings of the 12th Annual workshop on Family History Technology*, pp. 20-24, 2012.
- [PEV 06] P.A. Pevzner, "Bio-informatique moléculaire : Une approche algorithmique", *Collection IRIS*, dirigé par Nicolas Puech, Traduction : Delphine Hachez, *Springer Editions*, Octobre 2006.
- [PHI 06] S. Philipp-Foliguet, L. Guiges, "Evaluation de la segmentation d'images: état de l'art, nouveaux indices et comparaison", *Traitement du signal*, vol. 23, No. 2, pp. 109 – 124, 2006.
- [POS 79] N. Postman, "Teaching as a conserving activity", *Instructor*, vol. 89, No 4, 1979.
- [POS 86] W. Postl, "Detection of linear oblique structures and skew scan in digitized documents", *Proceedings of the 8th International Conference on Pattern Recognition (ICPR)*, pp. 687–689, 1986.
- [PRA 78] W.K.Pratt, "Digital Image Processing John Wiley & Sons", *New York*, vol. 1191, pp. 491 – 556, 1978.
- [PRA 10] I. Pratikakis, B. Gatos, K. Ntirogiannis "H-DIBCO 2010 - Handwritten Document Image Binarization Competition", *Proceedings of the International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pp. 727-732, Kolkata - India, November 2010.
- [PRA 11] I. Pratikakis, B. Gatos, K. Ntirogiannis, "ICDAR 2011 Document Image Binarization Contest (DIBCO 2011)", *Proceedings of the International Conference on Document Analysis and Recognition (ICDAR)*, pp. 1506-1510, Beijing-China, September 2011.
- [PRA 12] I. Pratikakis, B. Gatos, K. Ntirogiannis, "ICFHR 2012 Competition on Handwritten Document Image Binarization (H-DIBCO 2012)", *Proceedings of the International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pp. 813-818, 2012.
- [PRA 13] I. Pratikakis, B. Gatos, K. Ntirogiannis, "ICDAR 2013 document image binarization contest (DIBCO 2013)", *Proceedings of the 12th International Conference on Document Analysis and Recognition (ICDAR)*, pp. 1471-1476, 2013.
- [PRE 66] J.M.S. Prewitt, M.L. Mendelsohn, "The analysis of cell images", *Annals of the New York Academy of Sciences*, vol. 128, No. 3, pp. 1035-1053, 1966.
- [PRE 10] N. Premchaiswadi, S. Yimgnagm, W. Premchaiswadi, "A scheme for salt and pepper noise reduction and its application for OCR systems", *WSEAS Transactions on Computers*, vol. 9, pp. 351–360, April 2010.
- [PUN 80] T. Pun, "A new method for gray-level picture thresholding using the entropy of the histogram", *Signal Processing*, vol. 2, No. 3, pp. 223-237, 1980.
- [RAB 13] V. Rabeux, "Évaluation de la qualité des documents anciens numérisés", *Thèse de doctorat*, Université de Bordeaux I, 06/03/2013.
- [RAM 10] M.A. Ramirez-Ortegeon, R. Rojas, "Unsupervised Evaluation Methods Based on Local Gray-Intensity Variances for Binarization of Historical Documents", *Proceedings of the 20th International Conference on Pattern Recognition (ICPR)*, pp. 2029-2032, Istanbul-Turkey, 2010.
- [RAM 10b] M.A. Ramirez-Ortegeon, E. Tapia, L.L. Ramirez-Ramirez, R. Rojas, E. Cuevas, "Transition pixel: A concept for binarization based on edge detection and gray-intensity histograms", *Pattern Recognition*, vol. 43, No. 4, pp. 1233-1243, 2010.

- [RID 78] T.W. Ridler, S. Calvard, "Picture thresholding using an iterative selection method", *IEEE transactions on Systems, Man and Cybernetics*, vol. 8, No. 8, pp. 630-632, 1978.
- [ROB 04] J.P. Robert, "Dictionnaire du Petit Robert", *Robert Edition*, Paris-France, pages 715, 1990.
- [ROS 03] P.L. Rosin, E. Ioannidis, "Evaluation of global image thresholding for change detection", *Pattern Recognition Letters*, vol. 24, No. 14, pp. 2345-2356, 2003.
- [SAH 88] P.K. Sahoo, S.A.K.C. Soltani, A.K.C. Wong, "A Survey of Thresholding Techniques", *Computer Vision, Graphics, and Image Processing*, vol. 41, No. 2, pp. 233-260, 1988.
- [SAI 07] Z. Saidane, C. Garcia, "Robust binarization for video text recognition", *Proceedings of the 9th International Conference on Document Analysis and Recognition (ICDAR)*, vol. 2, pp. 874-879, September 2007.
- [SAR 02] T. Sari, L. Souici, M. Sellami, "Off-line handwritten Arabic character segmentation algorithm: ACSA", *Proceedings of the 8th International Workshop on Frontiers in Handwriting Recognition (IWFHR)*, pp. 452-457, 2002.
- [SAR 11] T. Sari, A. Kefali, "Recognition-free Retrieval of Old Arabic Document Images", *Computación y Sistemas*, vol. 15, No. 2, pp. 165-208, 2011.
- [SAR 14] T. Sari, A. Kefali, H. Bahi, "Text Extraction from Historical Document Images by the Combination of Several Thresholding Techniques", *Advances in Multimedia*, vol. 2014, Article ID 934656, 10 pages, 2014.
- [SAR 15] T. Sari, A. Kefali, H. Bahi, "Structural feature-based evaluation method of binarization techniques for word retrieval in the degraded Arabic document images", *International Journal on Document Analysis and Recognition (IJDAR)*, 2015, doi: 10.1007/s10032-015-0254-y.
- [SAU 00] J. Sauvola, M. Pietikainen, "Adaptive document image binarization", *Pattern Recognition*, vol. 33, No. 2, pp. 225-236, 2000.
- [SAV 98] A.E. Savakis, "Adaptive document image thresholding using foreground and background clustering", *Proceedings of the International Conference on Image Processing (ICIP)*, pp.785-789, 1998.
- [SAY 73] K.M. Sayre, "Machine recognition of handwritten words: A project report", *Pattern Recognition*, vol. 5, No. 3, pp. 213-228, 1973.
- [SEH 05] A. Sehad, L. Mezai, L. Sekkai, M. Cheriet, "Méthode de détection de l'inclinaison des documents arabes imprimés basée sur la réduction des points et a transformée de Hough", *Proceedings of the 3rd International Conference: Sciences of Electronic, Technologies of Information and Telecommunications (SETIT)*, 2005.
- [SER 83] J. Serra, "Image Analysis and Mathematical Morphology," *Academic Press* (3rd edition), London, 1983.
- [SEZ 01] M. Sezgin, B. Sankur, "Selection of thresholding methods for non-destructive testing applications", *Proceedings of the International Conference on Image Processing (ICIP)*, vol. 3, pp. 764 - 767, 2001.
- [SEZ 01b] M. Sezgin, B. Sankur, "Comparison of thresholding methods for non-destructive testing applications", *accepted for publication in the International Conference on Image Processing (ICIP)*, 2001.
- [SEZ 04] M. Sezgin, B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation", *Journal of Electronic Imaging*, vol. 13, No. 1, pp. 146-165, 2004.
- [SHA 94] A.G. Shanbaga, "Utilization of Information Measure as a Means of Image Thresholding", *Graphical Models and Image Processing (CVGIP)*, vol. 56, No. 5, pp. 414-419, 1994.
- [SHA 00] G. Sharma, "Cancellation of show-through in duplex scanning", *Proceedings of the International Conference on Image processing (ICIP)*, vol. 2, pp. 609-612, 2000.

- [SHA 07] F. Shafait, J. Van Beusekom, D. Keysers, T.M. Breuel, "Page frame detection for marginal noise removal from scanned documents", *Proceedings of the 15th Scandinavian conference on Image analysis*, pp. 651 - 660, 2007.
- [SHA 09] F. Shafait, T.M. Breuel, "A simple and effective approach for border noise removal from document images", *Proceedings of the 13th International Multitopic Conference*, pp. 1-5, 2009.
- [SHA 12] D.V. Sharma, S. Singh, "An Analysis of Image Binarization Techniques for Natural Scene Images", *Proceedings of the International Conference on Recent Advances and Future Trends in Information Technology (iRAFIT)*, pp. 29-32, 2012.
- [SHI 04] Z. Shi, V. Govindaraju, "Line separation for complex document images using fuzzy run length", *Proceedings of the Document Image Analysis for Libraries*, pp. 306-306, 2004.
- [SHI 05] P. Shivakumara, G.H. Kumar, D.S. Guru, P. Nagabhushan, "A novel technique for estimation of skew in binary text document images based on linear regression analysis", *Sadhana*, vol. 30, No. 1, pp. 69-85, 2005.
- [SIL 06] J.M.M. da Silva, R.D. Lins, V.C. da Rocha, "Binarizing and filtering historical documents with back-to-front interference", *Proceedings of the ACM symposium on Applied computing (SAC)*, pp. 853-858, 2006.
- [SIL 09] J.M. Silva, R.D. Lins, G.P. e Silva, "Enhancing the Quality of Color Documents with Back-to-Front Interference", *Proceedings of the 6th International Conference on Image Analysis and Recognition (ICIAR)*, pp. 875-885, 2009.
- [SIN 11] T.R. Singh, S. Roy, O.I. Singh, T. Sinam, K. Singh, "A New Local Adaptive Thresholding Technique in Binarization", *International Journal of Computer Science Issues (IJCSI)*, vol. 8, No.2, pp.271-277, November 2011.
- [SMI 04] E. Smigiel, A. Belaid, H. Hamza, "Self-organizing Maps and Ancient Documents", *Proceedings of the 6th international workshop on Document Analysis Systems VI (DAS)*, pp.125-134, Florence - Italy, 2004.
- [SMI 05] H. Hamza, E. Smigiel, A. Belaid, "Neural based binarization techniques", *Proceedings of the 8th International Conference on Document Analysis and Recognition (ICDAR)*, pp. 317-321, 2005.
- [SOB 00] K. Sobottka, H. Kronenberg, T. Perroud, H. Bunke, "Text extraction from colored book and journal covers", *International Journal on Document Analysis and Recognition (IJ DAR)*, vol. 2, No. 4, pp. 163-176, June 2000.
- [SOL 99] Y. Solihin, C.G. Leedham, "Integral Ratio: A New Class of Global Thresholding Techniques for Handwriting Images", *IEEE Trans on Pattern Analysis and Machine Intelligence*, vol. 21, No. 8, pp.761-768, 1999.
- [SOM 11] H.M. Som, J.M. Zain, A.J. Ghazali, "Application of threshold techniques for readability improvement of jawi historical manuscript images", *Advanced Computing: An International Journal (ACIJ)*, vol.2, No.2, March 2011.
- [SOU 02] S. Souafi, "Contribution à la reconnaissance des structures des documents écrits : Approche probabiliste", *Thèse de doctorat*, Institut National des Sciences Appliquées de Lyon, 21 Septembre 2002.
- [STA 07] N. Stamatopoulos, B. Gatos, A. Kesidis, "Automatic Borders Detection of Camera Document Images", *Proceedings of the 2nd International Workshop on Camera-Based Document Analysis and Recognition*, pp. 71 - 78, Curitiba - Brazil, 2007.
- [STA 08] P. Stathis, E. Kavallieratou, N. Papamarkos, "An Evaluation Technique for Binarization Algorithms", *Journal of Universal Computer Science*, vol. 14, No. 18, pp. 3011-3030, 2008.
- [STA 08b] P. Stathis, E. Kavallieratou, N. Papamarkos, "An Evaluation Survey of Binarization Algorithms on Historical Documents", *Proceedings of the 19th International Conference on Pattern Recognition (ICPR)*, pp.1 - 4, 2008.

- [STE 95] S. Foster, R. Russel, J. Lyall, D. Marshall, “Mémoire du monde: principes directeurs pour la sauvegarde du patrimoine documentaire”, *document élaboré pour l’UNESCO au nom de l’IFLA*, Paris: UNESCO, 1995.
- [STO 92] G. Story, L. O’Gorman, D. Fox, L. Schaper, H. Jagadish, “The rightpages image-based electronic library for alerting and browsing”, *Computer*, vol. 25, No. 9, pp. 17–26, September 1992.
- [STR 91] K.C. Strasters, J.J.Gerbrands, “Three-dimensional image segmentation using a split, merge and group approach”, *Pattern Recognition Letters*, vol. 12, No. 5, pp. 307-325, 1991.
- [STR 02] C. Strouthopoulos, N. Papamarkos, A.E. Atsalakis, “Text extraction in complex color documents”, *Pattern Recognition*, vol. 35, No. 8, pp. 1743–1758, 2002.
- [SU 10] B. Su, S. Lu, C.L. Tan, “Binarization of Historical Document Images Using the Local Maximum and Minimum”, *Proceedings of the 9th IAPR International Workshop on Document Analysis Systems*, pp. 159-166, 2010.
- [SU 11] B. Su, S. Lu, C.L. Tan, “Combination of document image binarization techniques”, *Proceedings of the International Conference on Document Analysis and Recognition (ICDAR)*, pp. 22-26, 2011.
- [SUM 04] B. Sumengen, “Variational image segmentation and curve evolution on natural images”, *Thèse de doctorat*, University of California, Santa Barbara, 2004.
- [TAB 06] S. Tabbone, T.O. Nguyen, G. Masini, “Une méthode de binarisation hiérarchique floue”, *Proceedings of the Colloque International Francophone sur l’Ecrit et le Document (CIFED)*, pp. 43 - 48, 2006.
- [TAB 10] S.A. Tabatabaei, M. Bohlool, “Novel method for binarization of badly illuminated document images”, *Proceedings of the 17th International Conference on Image Processing (ICIP)*, pp. 3573-3576, Hong Kong, September 2010.
- [TAK 01] H.Takahashi, H. Hiroya, N. Koichi, “Image processing device for providing correction of a curvature distortion of a bound portion of a spread book”, *Brevet U.S. No. 6330050*, Washington, 2001.
- [TEN 05] B.T. Avila, R.D. Lins, “A fast orientation and skew detection algorithm for monochromatic document images”, *Proceedings of the ACM Symposium on Document engineering*, pp. 118–126, 2005.
- [TON 04] A. Tonazzini, L. Bedini, E. Salerno, “Independent Component Analysis for Document Restoration”, *International Journal on Document Analysis and Recognition (IJ DAR)*, vol. 7, No. 1, pp. 17-27, 2004.
- [TRE 04] Trésor de la langue française, France, 2004.
- [TRI 95] Ø.D. Trier, T. Taxt, “Evaluation of binarization methods for document images”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, No. 3, pp. 312-315, 1995.
- [TRI 95b] Ø.D. Trier, A.K. Jain, “Goal-Directed Evaluation of Binarization Methods”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, No. 12, pp. 1191-1201, 1995.
- [TRI 95c] Ø.D. Trier, T. Taxt, “Improvement of “Intergrated Function Algorithm” for binarization of document images”, *Pattern Recognition Letters*, vol. 16, No. 3, pp. 277-283, March 1995.
- [TRI 03] E. Trinh, “De la numérisation à la consultation des documents anciens”, *Thèse de doctorat*, Institut National des Sciences Appliquées de Lyon, 2003.
- [TRU 05] É. Trupin. “La reconnaissance d’images de documents: Un panorama”, *Traitement du signal*, vol. 22, No. 3, pp. 159-189, 2005.
- [TSA 85] W.H. Tsai, “Moment-preserving thresholding: A new approach”, *Computer Vision, Graphics, and Image Processing*, vol. 29, No. 3, pp. 377-393, 1985.

- [TSA 88] C.Tsallis, "Possible Generalization of Boltzmann-Gibbs statistics", *Journal of Statistical Physics*, 1988, vol. 52, No. 1-2, pp. 479-487.
- [TSA 02] C.M. Tsai, H.J. Lee, "Binarization of color document images via luminance and saturation color features", *IEEE Transaction on Image Processing*, vol. 11, No. 4, pp. 434 - 451, 2002.
- [VEL 80] F.R.D. Velasco, "Thresholding using the isodata clustering algorithm", *IEEE Transaction on System, Man and Cybernetics*, vol. 10, pp. 771-774, 1980.
- [WAG 74] R.A. Wagner, M.J. Fisher, "The string to string correction problem", *Journal of the ACM (JACM)*, vol. 21, No. 1, pp. 168 - 173, 1974.
- [WAG 14] M. Wagdy, F. Ibrahima, R. Dayang, "Border Noise Removal and Clean Up Based on Retinex Theory", *Proceedings of the 1st International Conference on Advanced Data and Information Engineering (DaEng-2013)*, pp. 345-352, 2014.
- [WAK 98] B. Waked, S. Bergler, C.Y. Suen, S. Khoury, "Skew Detection, Page Segmentation, and Script Classification of Printed Document Images", *Proceedings of the IEEE International conference on Systems, Man, and Cybernetics*, vol. 5, pp. 4470 - 4475, 1998.
- [WAK 11] T. Wakahara, K. Kita, "Binarization of Color Character Strings in Scene Images Using K-means Clustering and Support Vector Machines", *Proceedings of the International Conference on Document Analysis and Recognition (ICDAR)*, pp. 274 - 278, 2011.
- [WAN 94] J. Wang, J. Jean, "Segmentation of merged characters by neural networks and shortest path," *Pattern Recognition*, vol. 27, No. 5, pp. 649, 1994.
- [WAN 03] Q. Wang, T. Xia, C.L. Tan, L. Li, "Directional Wavelet Approach to Remove Document Image Interference", *Proceedings of the International Conference on Document Analysis and Recognition (ICDAR)*, pp. 736-740, Edinburgh, 2003.
- [WHI 83] J.M. White, G.D. Rohrer, "Image Thresholding for Optical Character Recognition and Other Applications Requiring Character Image Extraction", *IBM Journal of Research and Development*, vol. 27, No. 4, pp. 400 - 411, July 1983.
- [WHI 96] A. Whichello, H. Yan, "Linking Broken Character Borders with Variable Sized Masks to Improve Recognition", *Pattern Recognition*, vol. 29, No. 8, pp. 1429-1435, 1996.
- [WOL 02] C. Wolf, J. Jolion, F. Chassaing, "Extraction de texte dans des vidéos : le cas de la binarisation", *Proceedings of the 13^{ème} congrès francophone de Reconnaissance des Formes et Intelligence Artificielle (RFIA)*, pp. 145-152, 2002.
- [WOL 02b] C. Wolf, J.M. Jolion, "Détection et extraction de texte de la vidéo", *Proceedings of the Colloque international francophone sur l'écrit et le document (CIFED)*, pp. 215-224, 2002.
- [WOL 06] C. Wolf, "Document Ink bleed-through removal with two hidden Markov random fields and a single observation field", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, No. 3, pp. 431-447, 2010.
- [WOO 04] Y.W. Woo, "Performance Evaluation of Binarizations of scanned Insect Footprints", *Proceedings of the 10th International Conference on Combinatorial Image Analysis*, pp. 669-678, 2004.
- [WSH 09] S. Wshah, Z. Shi, V. Govindaraju, "Segmentation of Arabic handwriting based on both contour and skeleton segmentation", *Proceedings of the 10th International Conference on Document Analysis and Recognition (ICDAR)*, pp.793-797, Barcelona-Spain, 2009.
- [WU 02] C. Wu, G. Agam, "Document Image De-warping for Text/Graphics Recognition", *Proceedings of the Joint IAPR International Workshop on Structural, Syntactic, and Statistical Pattern Recognition*, pp. 348-357, 2002.
- [WU 03] S. Wu, A. Amin, "Automatic thresholding of gray-level using multistage approach", *Proceedings of the 7th International Conference on the Document Analysis and Recognition (ICDAR)*, pp. 493-497, 2003.

- [XIA 08] Y. Xiao, Z.G. Cao, T.X. Zhang, “Entropic thresholding based on gray-level spatial correlation histogram”, *Proceeding 19th International Conference on Pattern Recognition (ICPR)*, pp. 1-4, December 2008.
- [XIA 10] D. Xiang, H. Yan, X. Chen, Y. Cheng, “Offline Arabic handwriting recognition system based on HMM”, *Proceedings of the 3rd International Conference on Computer Science and Information Technology (ICCSIT)*, pp. 526-529, Chengdu-China, 2010.
- [XUE 12] J.H. Xue, Y.J. Zhang, “Ridler and Calvard’s, Kittler and Illingworth’s and Otsu’s methods for image thresholding”, *Pattern Recognition Letters*, vol. 33, No. 6, pp. 793–797, 2012.
- [YAM 04] A. Yamashita, A. Kawarago, T. Kanako, K.T. Miura, “Shape Reconstruction and Image restoration for Non-Flat Surfaces of Documents with a Stereo Vision System”, *Proceedings of the 17th International Conference on Pattern recognition (ICPR)*, pp. 482-485, Cambridge, Royaume-Uni, 2004.
- [YAN 86] S.D. Yanomitz, A.M. Bruckstein, “A new method for image segmentation”, *Computer Vision, Graphics and Image Processing*, vol.16, No.1, pp.82-95. 1986.
- [YAN 93] H. Yan, “Skew Correction of Document Images Using Interline Cross-Correlation”, *CVGIP: Graphical Models and Image Processing*, vol. 55, No. 6, pp. 538-543, 1993.
- [YAS 78] W.A. Yasnoff, W. Galbraith, J.W. Bacus, “Error measures for objective assessment of scene segmentation algorithms”, *Analytical and quantitative cytology*, vol. 1, No. 2, pp. 107-121, 1978.
- [YAS 80] Y. Yasuda, M. Dubois, T.S. Huang, “Data compression for check processing machines”, *Proceedings of the IEEE*, vol. 68, No. 7, pp. 874–885, 1980.
- [YEN 95] J.C. Yen, F.J. Chang, S. Chang, “A New Criterion for Automatic Multilevel Thresholding”, *IEEE Transactions on Image Processing*, vol. 4, No. 3, pp. 370-378, 1995.
- [YIN 08] F. Yin, C.L. Liu, “Handwritten text line segmentation by clustering with distance metric learning”, *Proceedings of the 11th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pp. 229–234, 2008.
- [YU 92] B. Yu, X. Lin, Y. Wu, B. Yuan, “Isothetic polygon representations for contours”, *CVGIP: Image Understanding*, vol. 56, No. 2, pp. 264–268, 1992.
- [YU 96] B. Yu, A.K. Jain, “A robust and fast skew detection algorithm for generic documents”, *Pattern Recognition*, vol. 29, No. 10, pp. 1599-1629, 1996.
- [YU 01] D.Yu, H.Yan, “Reconstruction of Broken Handwritten digits based on structural morphological features”, *Pattern Recognition*, No. 34, No. 2, pp. 235-254, 2001.
- [ZAC 77] G.W. Zack, W.E. Rogers, S.A. Latt, “Automatic measurement of sister chromatid exchange frequency”, *Journal of Histochemistry and Cytochemistry*, vol. 25, No. 7, pp. 741–753, 1977.
- [ZHA 92] Y.J. Zhang, J.J. Gerbrands, “Segmentation evaluation using ultimate measurement accuracy”, *Electronic Imaging 92, International Society for Optics and Photonics*, pp. 449-460, 1992.
- [ZHA 96] Y.J. Zhang, “A survey on evaluation methods for image segmentation”, *Pattern recognition*, vol. 29, No. 8, pp. 1335-1346, 1996.
- [ZHA 97] Y.J. Zhang, “Evaluation and comparison of different segmentation algorithms”, *Pattern recognition letters*, vol. 18, No. 10, pp. 963-974, 1997.
- [ZHA 01] Z. Zhang, C.L. Tan, “Restoration of images scanned from thick bound documents”, *Proceedings of the International Conference on Image Processing (ICPR)*, pp. 1074-1077, 2001.
- [ZHA 01b] Z. Zhang, C.L. Tan, “Recovery of Distorted Document Images from Bound Volumes”, *Proceedings of the 6th International Conference on Document Analysis and Recognition (ICDAR)*, pp. 429-433, 2001.

[ZHA 03] H. Zhang, J.E. Fritts, S.A. Goldman, “An entropy-based objective evaluation method for image segmentation”, *Electronic Imaging 2004, International Society for Optics and Photonics*, pp. 38-49, 2003.

[ZHA 03b] Z. Zhang, C.L. Tan, “Correcting Document Image Warping Based on Regression of Curved Text Lines”, *Proceedings of the International Conference on Document Analysis and Recognition (ICDAR)*, pp. 589-593, Edinburgh, 2003.

[ZHA 08] H. Zhang, J.E. Fritts, S.A. Goldman, “Image segmentation evaluation: A survey of unsupervised methods”, *Computer Vision and Image Understanding*, vol. 110, No. 2, pp. 260–280, 2008.

Liste des publications

Publications dans des revues internationales

- Toufik Sari, Abderrahmane Kefali, Halima Bahi, “Structural feature-based evaluation method of binarization techniques for word retrieval in the degraded Arabic document images”, *International Journal on Document Analysis and Recognition (IJ DAR)*, 2015, doi: 10.1007/s10032-015-0254-y.
- Abderrahmane Kefali, Toufik Sari, Halima Bahi, “Foreground-Background Separation by Feed-forward Neural Networks in Old Manuscripts”, *Informatica*, Vol. 38, No. 4, pp. 329–338, 2014.
- Toufik Sari, Abderrahmane Kefali, and Halima Bahi, “Text Extraction from Historical Document Images by the Combination of Several Thresholding Techniques”, *Advances in Multimedia*, Vol. 2014, Article ID 934656, 10 pages, 2014. doi:10.1155/2014/934656.
- Abderrahmane Kefali, Toufik Sari, Halima Bahi, “Text/ Background separation in the degraded document images by combining several thresholding techniques”, *WSEAS Transactions on Signal Processing*, Vol. 10, pp. 436–443, 2014.
- Toufik Sari, Abderrahmane Kefali, “Recognition-free Retrieval of Old Arabic Document Images”, *Computación y Sistemas*, Vol. 15 No. 2, pp 195–208, 2011.

Publications dans des conférences internationales

- Abderrahmane Kefali, Toufik Sari and Halima Bahi, “An MLP for binarizing images of old manuscripts”, *13th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pp. 247–251, Bari-Italy, 18–20 Septembre, 2012.
- Abderrahmane Kefali, chaouki Chemmam, “A semi-automatic approach of old Arabic documents indexing”, *3^{ème} Conférence Internationale sur l’Informatique et ses Applications (CIIA)*, Saida, 14–15 Décembre 2011.
- Abderrahmane Kefali, Toufik Sari, “Recherche d’images de documents arabes anciens sans reconnaissance du contenu”, *2^{ème} Conférence Internationale Image and Signal Processing and their Applications (ISPA)*, Biskra, 6–8 Décembre 2010.