

Université De BLIDA I
Faculté de Technologie
Département d'Electronique

MEMOIRE DE MAGISTER

en Electronique
Option : Signaux et Systèmes

SEGMENTATION DE BLOCS D'IMAGES DE DOCUMENTS IMPRIMES

Par

Mohamed Amine LOURCI

Devant le jury composé de :

Mr MELIANI Hamza	Pr Département d'électronique, UB1	Président
Mr GESSOUM Abderrezak	Pr Département d'électronique, UB1	Examineur
Mr HALIMI Mohamed	Dr CSC, Alger	Examineur
Mr NAMANE Abderrahmane	Pr Département d'électronique, UB1	Promoteur

La soutenance aura lieu le 11 juin 2015 à 9h 30 à la salle de conférences
(Pavillon 16)

ملخص

الهدف من هذا المشروع هو انشاء نظام يتعرف بطريقة آلية على البنية الفيزيائية للوثائق تحت عنوان تجزئة الكتل المكونة لصور الوثائق، و ذلك بالاعتماد على تطوير و تحسين خطوات المعالجات الأولية المطبقة بهدف تقليص التشويش عن الصورة و كذا اختزال كمية المعلومات المحتواة في الصورة المراد معالجتها قبل الشروع في مرحلة تصنيف الكتل حسب صنفين (كتل نصية و كتل غير نصية). التعرف على البنية الفيزيائية للوثائق يعتمد على طريقة هجينة و التي تجمع بين خوارزميتين مختلفتين الأولى تهدف الى الكشف عن الصور و المخططات و الرسومات...بالاعتماد على العمليات الرياضية لتغيير الأشكال (علم التشكل الرياضي). الخوارزمية الثانية تدعى خوارزمية الصقل أو التلميس و دورها هو استخراج السطور النصية باستعمال هذه الأخيرة.

الكلمات الجوهرية :

التجزئة ، بنية فيزيائية ، طريقة هجينة ، علم التشكل الرياضي ، التصنيف ، خوارزمية الصقل.

RESUME

Ce mémoire de magister a pour but la réalisation d'un système de reconnaissance automatique de structure physique d'un document, en se basant sur l'amélioration des étapes de prétraitement afin de minimiser le bruit et réduire la quantité d'information contenue dans l'image à traiter avant de commencer la classification des blocs selon deux classes (texte / non-texte).

La reconnaissance de la structure physique est basée sur une méthode hybride de segmentation qui combine deux algorithmes différents, le premier est un algorithme descendant appelé « Algorithme de Bloomberg » qui a pour but de détecter les images, les graphes, les dessins, ...en se basant sur les opérations de morphologie mathématique. Par contre, le second algorithme est classé parmi les méthodes de segmentation de type ascendant appelé RLSA son rôle est l'extraction des lignes textuelles en se basant sur les propriétés des lignes de texte.

Mots clés : Segmentation, descendant, ascendant, morphologie mathématique, Classification, Algorithme de Bloomberg, RLSA.

ABSTRACT

The main topic of this project is to construct a system of automatic detection of physical structure of document entitled Segmentation of Blocks Images of Documents. It is based on improvement preprocessing steps to minimize noise and reduce the amount of information in the image before starting the classification of blocks into two classes (text / non-text).

The recognition of the physical structure is based on a hybrid segmentation method that combines two different algorithms, the first is a top-down algorithm called "Bloomberg's algorithm" it is intended to detect images, graphs, drawings,... it is based on mathematical morphology operations. The second algorithm is classified as Bottom-up segmentation method and called RLSA, its role is in extracting textual lines based on the properties of lines of text.

Keywords: Segmentation, top-down, Bottom-up Classification, Bloomberg's algorithm, RLSA

REMERCIEMENT

Je tiens à remercier tout d'abord ALLAH qui m'a donné le courage et la patience jusqu'au bout de mes études, et qui m'a permis de réaliser mon travail.

C'est avec un immense plaisir d'exprimer mes remerciements les plus sincères à MR. NAMANE Abderrahmane qui m'a donné l'opportunité de découvrir le domaine de l'analyse et la reconnaissance de document à travers un projet dans son intégralité, qui, en tant que promoteur de mémoire, ainsi que pour l'inspiration, l'aide et le temps qu'il a bien voulu me consacrer et sans qui ce mémoire n'aurait jamais vu le jour.

Un grand merci aux membres du jury, qui ont accepté d'évaluer mon travail de magister.

Je remercie également mes chers parents et grands parents, ma source de courage, de soutien et de bonheur, merci pour tous ce que vous avez fait pour moi. A ma sœur, mes frères ainsi que toute ma famille, pour leur soutien moral compréhension et aide continu.

TABLE DES MATIERES

RESUME

REMERCIEMENTS

TABLE DES MATIERES

Liste des illustrations, graphiques et tableaux

GLOSSAIRE

INTRODUCTION	11
1. ETAT DE L'ART	13
1.1 Définition de la segmentation	15
1.2 Pourquoi on fait l'analyse de document.....	15
1.3 Les différentes approches de segmentation	16
1.3.1 Méthodes descendantes	16
1.3.2 Méthodes ascendantes	16
1.4 Algorithmes de segmentation de documents	18
1.4.1 Découpage X-Y	18
1.4.2 Division et fusion (split and merge)	21
1.4.3 Algorithme de Bloomberg	22
1.4.4 RLSA	24
1.4.5 Docstrum	27
1.4.6 Composantes connexes	27
1.4.7 Diagramme de Voronoi	27
2. TECHNIQUES DE TRAITEMENT	31
2.1 Filtrage	32
2.1.1 Filtrage linéaire	32
2.1.1.1 Filtrage moyen	34
2.1.1.2 Filtrage gaussien	35
2.1.2 Filtrage non linéaire	37
Filtre médian	37
2.2 Binarisation	40
2.2.1 Seuillage global	42
Méthode de Otsu	42
2.2.2 Seuillage local	45
2.2.2.1 Méthode de Bernsen	46
2.2.2.2 Méthode de Niblack	46
2.2.2.3 Méthode de Sauvola.....	46
2.2.2.4 Binarisation hiérarchique flou	47
2.3 Etiquetage	51
Algorithme	51
2.4 Opérations Morphologiques sur les images binaires	56
2.4.1 Dilatation	57
2.4.2 Erosion	58
2.4.3 Ouverture	59
2.4.4 Fermeture	66
2.4.5 Réduction avec seuillage.....	61

3. PRETRAITEMENT ET SEGMENTATION	
3.1 Aperçu de la méthode générale	64
3.2 Prétraitement	64
3.2.1 Filtrage	65
3.2.2 Binarisation	67
3.2.2.1 Principe de la méthode	68
3.2.2.2 Etape de seuillage global	68
3.2.2.3 Etape de raffinage	69
3.3 Segmentation	73
3.3.1 Pré-segmentation (Application de l'algorithme de Bloommberg amélioré	76
3.3.2 Segmentation finale	85
3.3.2.1 Constitution de lignes	87
3.3.2.2 Constitution de blocs	90
3.3.2.3 Etiquetage des blocs	91
3.3.2.4 Recherche des lignes	95
3.3.2.5 Classification	98
5. RESULTATS ET DISCUSSIONS	
4.1 Présentation de l'interface et Implémentation	101
4.1.1 Environnement logiciel	101
4.1.2 Description logicielle	101
4.2 Images utilisées	106
4.3 Résultats expérimentaux	109
4.3.1 Résultats du prétraitement	110
4.3.1.1 Résultats du filtrage Gaussien	110
4.3.1.2 Résultats de Binarisation	112
4.3.2 Résultats de pré-segmentation (algorithme de Bloomberg)	114
4.3.3 Résultats de segmentation finale	117
4.4 Evaluation quantitative des résultats	120
CONCLUSIONS ET PERSPECTIVES	123
BIBLIOGRAPHIE	
ANNEXE A	

LISTE DES ILLUSTRATIONS, GRAPHIQUES ET TABLEAUX

Figure 1.1	Exemples de structure physique (b) d'un document (a)	14
Figure 1.2	Structures logiques	15
Figure 1.3	Approche descendante et ascendante	17
Figure 1.4	Structure de page de type Manhattan	18
Figure 1.5	Exemple de découpage d'une image avec arbre de décomposition	19
Figure 1.6	exemple de document à problème	20
Figure 1.7	division	21
Figure 1.8	Représentation sous forme de Quadtree	21
Figure 1.9	carte des régions et le graphe d'adjacence	22
Figure 1.10	Agrégation itérative des blocs similaires au bloc 1.....	22
Figure 1.11	Exemple de document à problème avec Bloomberg	23
Figure 1.12	Segmentation de l'image de (Figure 2.6-a) après amélioration	24
Figure 1.13	Segmentation RLSA	25
Figure 1.14	Structure non-Manhattan	26
Figure 1.15	Composantes connexes et segmentation	27
Figure 1.16	Illustration de la définition d'un diagramme de Voronoï	28
Figure 1.17	Diagramme de Voronoï et diagramme de voisinage	29
Figure 2.1	Convolution locale	32
Figure 2.2	Masques 3*3 et 5*5 de filtre moyen	34
Figure 2.3	Filtrage moyen	35
Figure 2.4	Effet de lissage	35
Figure 2.5	Filtrage gaussien	37
Figure 2.6	Image originale et image bruitée Bruit Poivre et Sel	38
Figure 2.7	Filtrage médian	39
Figure 2.8	Seuil de recouvrement	41
Figure 2.9	Histogramme des niveaux de gris	41
Figure 2.10	Algorithme de binarisation	42
Figure 2.11	Choix du seuil sur l'histogramme de l'image à niveaux de gris	43
Figure 2.12	Binarisation par la méthode Otsu	44
Figure 2.13	Problème de seuillage global.....	45
Figure 2.14	Fonction d'appartenance linéaire	48
Figure 2.15	Fonction S de Zadeh	48
Figure 2.16	Résultat de la méthode hiérarchique flou	50
Figure 2.17	types de connexité	51
Figure 2.18	image binaire et sa matrice initiale correspondante	52
Figure 2.19	premier passage de l'étiquetage	54
Figure 2.20	deuxième passage de l'étiquetage	55
Figure 2.21	Carte finale des étiquètes	56
Figure 2.22	Exemples d'étiquetage	56
Figure 2.23	Formes de l'élément structurant	57
Figure 2.24	illustration de processus de dilatation	57
Figure 2.25	exemple de dilatation d'image	58
Figure 2.26	Illustration de processus de l'érosion	58
Figure 2.27	Exemple d'érosion d'image	59
Figure 2.28	Ouverture morphologique	60
Figure 2.29	Fermeture morphologique	60

Figure 2.30	Définition de réduction avec seuil	61
Figure 3.1	Méthode adoptée	64
Figure 3.2	Aperçu de la méthode de prétraitement.....	65
Figure 3.3	Filtrage d'image	66
Figure 3.4	Binarisation adoptée	67
Figure 3.5	première étape de binarisation	68
Figure 3.6	Structure tridimensionnelle de l'arbre quaternaire	70
Figure 3.7	Quadrillage de l'image intermédiaire	71
Figure 3.8	Résultat final de binarisation de l'image de la figure 3.5.a	73
Figure 3.9	Délimitation par des rectangles	73
Figure 3.10	Délimitation par des contours	74
Figure 3.11	Méthode de segmentation utilisée	75
Figure 3.12	Algorithme de Bloomberg	76
Figure 3.13	réduction de la taille de l'image	77
Figure 3.14	image à segmenter.....	78
Figure 3.15	problème de destruction des parties non textuelles	78
Figure 3.16	Résultats de réduction après bouchage des trous	79
Figure 3.17	étapes de remplissage.....	79
Figure 3.18	Exemple de remplissage d'une image	80
Figure 3.19	Algorithme de Bloomberg amélioré	80
Figure 3.20	Ouverture de l'image par un $ES = 5 \times 5$	81
Figure 3.21	Résidus d'une image	82
Figure 3.22	Expansion d'image.....	82
Figure 3.23	Localisation de blocs non-textuels	83
Figure 3.24	résultat final de pré-segmentation	83
Figure 3.25	problème de pré-segmentation	84
Figure 3.26	Algorithme RLSA	85
Figure 3.27	Segmentation RLSA	86
Figure 3.28	Variante de l'algorithme RLSA	86
Figure 3.29	Distances horizontales	87
Figure 3.30	Histogramme de distances horizontales	87
Figure 3.31	Distribution horizontale des espaces	88
Figure 3.32	Lissage horizontal.....	89
Figure 3.33	Exemple de distribution verticale des espaces	90
Figure 3.34	image en blocs.....	90
Figure 3.35	Premier passage de l'étiquetage	91
Figure 3.36	second passage de l'étiquetage	92
Figure 3.37	Premier passage de l'étiquetage de blocs	93
Figure 3.38	Détection de blocs d'image	94
Figure 3.39	Correction de sur-segmentation	95
Figure 3.40	Etiquetage des lignes	97
Figure 3.41	Segmentation finale	98
Figure 4.1	L'interface principale du logiciel	101
Figure 4.2	Sélection d'une image à traiter.....	103
Figure 4.3	Utilisation du Filtrage	104
Figure 4.4	Utilisation de Binarisation	104
Figure 4.5	Chois du sous menu Bloomberg's	105
Figure 4.6	Chois du sous menu RLSA.....	106
Figure 4.7	Images utilisées	108
Figure 4.8	Résultats du filtrage Gaussien	111

Figure 4.9	Résultats de Binarisation	113
Figure 4.10	Résultats de pré-segmentation	116
Figure 4.11	Résultats de segmentation finale.....	119
Figure 4.12	Image référentielle (G) et image à tester (T).....	120

GLOSSAIRE

RF	Reconnaissance de Formes
RS	Rectangle Structurant
RSH	Rectangle Structurant Horizontal
RSV	Rectangle Structurant Vertical
ES	Element Structurant
OCR	Optical Character Recognition
ARD	Analyse et Reconnaissance de Document
CCs	Composantes Connexes
PSF	Point Spread Function
P&S	Poivre et Sel
XY-C	XY Cut

INTRODUCTION

Depuis l'avènement de l'écriture, aux environs du III^{ème} millénaire en Mésopotamie, l'être humain n'a cessé d'améliorer ce moyen de communication [1].

En effet, plusieurs civilisations ont apporté leur savoir faire dans le domaine de l'écriture. Il y a eu d'abord l'écriture des hiéroglyphes des pharaons, l'écriture chinoise, l'écriture grecque, puis l'écriture arabe et romaine. Certaines de ces écritures ont disparu avec la destruction de leur civilisation alors que d'autres sont encore d'actualité. Malheureusement parmi les 3000 langues dénombrées dans le monde, seulement une centaine sont dotées d'un système d'écriture. Mais la percée majeure de l'écriture a atteint son apogée au 15^{ème} siècle avec l'invention de l'imprimerie par Gutenberg. Et c'est à partir de cette époque que l'écriture est entrée dans une nouvelle ère, à savoir l'ère du document imprimé.

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

L'être humain arrive avec aisance à reconnaître n'importe quel document. Cette faculté est assez basique pour un être humain ; en revanche elle pose jusqu'à ce jour encore des problèmes pour l'ordinateur. L'analyse et la reconnaissance d'images de documents englobent un ensemble de techniques informatiques avec comme but la reconstitution du contenu du document sous la forme de documents structurés, selon une forme définie par l'application en question. Les documents structurés couvrent deux catégories de documents : les documents imprimés et les documents manuscrits. Parmi les documents imprimés nous distinguons les documents à structures simples et les documents à structures complexes. Dans ce projet, nous nous intéressons aux documents à structures complexes.

Le développement des ordinateurs et des réseaux informatiques, et les avantages qu'ils présentent en matière de manipulation et de stockage d'informations écrites, font du support numérique un outil de plus en plus sollicité. Ainsi, le désir ou la nécessité de récupérer toujours plus d'informations provenant de documents papiers, même anciens, suppose des mécanismes de reconnaissance du document qui ne se limitent pas à l'identification des caractères mais qui vont jusqu'à l'analyse de sa structure (Segmentation) même. Cette demande a donné naissance, à un domaine de recherche particulier : l'analyse et la reconnaissance du document (ARD).

La segmentation est une étape importante pour l'extraction des informations qualitatives de l'image. C'est un processus qui consiste à découper une image en régions connexes présentant une homogénéité selon un certain critère, comme par exemple la couleur. L'union de ces régions doit redonner l'image initiale.

Problématique

De nos jours les organisations utilisent encore un grand nombre de documents papiers imprimés ou manuscrits qui nécessitent d'être représentés sous forme numérique et exploiter de la manière la plus efficace, sans recourir à une saisie manuelle. Il est souvent nécessaire d'accéder aux contenus des documents numérisés et de les modifier éventuellement. C'est l'objet de L'Analyse et la Reconnaissance des Documents (ARD).

Le problème de la segmentation peut se formuler comme suit :

Etant donnée l'image digitalisée d'une page, déterminer une partition géométrique (sous forme de blocs) de cette dernière de sorte à isoler tous les blocs qui la composent.

Objectif

L'objectif principal de ce projet est d'arriver à réaliser un système de segmentation de blocs d'images de documents en blocs textuels ou non-textuels.

Organisation du manuscrit

Notre travail est réparti en quatre chapitres :

- ❖ Dans le premier chapitre, nous présentons un état de l'art sur les méthodes de segmentation, afin d'en choisir celles qui sont adaptées au cas de notre étude.
- ❖ Le deuxième chapitre est dédié à la présentation des différentes techniques que nous avons utilisées dans notre travail.
- ❖ Une nouvelle approche de segmentation hybride est exposée au troisième chapitre, cette méthode basée sur une coopération de deux méthodes différentes dont la première cherche à extraire les composantes non textuelles et la deuxième cherche à extraire les lignes d'un texte.
- ❖ Le quatrième chapitre présentera les résultats d'application du système conçu, ainsi que les interprétations qui en découlent avec une brève présentation de l'interface utilisateur.

Enfin, une conclusion générale et des perspectives concernant le travail présenté achèvent ce document.

CHAPITRE 1

ETAT DE L'ART

La segmentation est une étape fondamentale dans le traitement d'image, en particulier le traitement d'image de document, parce qu'en fait, elle précède toute autre opération d'identification ou de classification. Cette étape dépend du type d'image qui diffère à la fois du système d'acquisition et du processus de formation d'image.

Dans le cas d'image de document, la segmentation est composée de la localisation et éventuellement la classification des éléments constituant le document, à différents niveaux de granularité. Ainsi, une première tâche de segmentation peut consister à localiser et à identifier les zones de texte et les zones de natures différentes.

La détermination de la structure des documents composites en blocs homogènes (photographies, graphiques, textes, sous blocs écrits avec des polices différentes ...) est un problème très complexe auquel on a souvent tenté de donner des solutions dans des cas bien particuliers (adresses postales, partitions musicales, plans cadastraux ...). L'organisation des constituants en blocs séparables pour la description du document constitue la structure physique. La reconnaissance de cette organisation et l'étiquetage des blocs en diverses catégories constituent la phase d'identification de la structure. Elle est connue sous le nom de structuration logique du document [2].

La reconnaissance des documents est scindée en deux parties :

- Reconnaissance de la structure physique :

La structure physique d'un document comporte des entités qui diffèrent d'un document à un autre. En revanche une entité qui reste commune à tous les documents est le texte. Les autres entités que nous retrouvons sont les images et les graphiques. En effet, ces dernières sont généralement insérées au sein des documents soit pour étoffer le document, soit pour expliquer une partie du texte.

La reconnaissance de structures physiques comprend deux étapes ; la détection et la classification des différentes zones de l'image. La détection comprend la segmentation qui consiste en la découpe, en zones (blocs) de l'image du document par des formes géométrique en 2D (rectangles ou polygones) (**Figure 1.1.b**). Dans l'étape de classification les formes géométriques, obtenues dans l'étape précédente, sont étiquetées en tant que texte, images et graphiques [3].

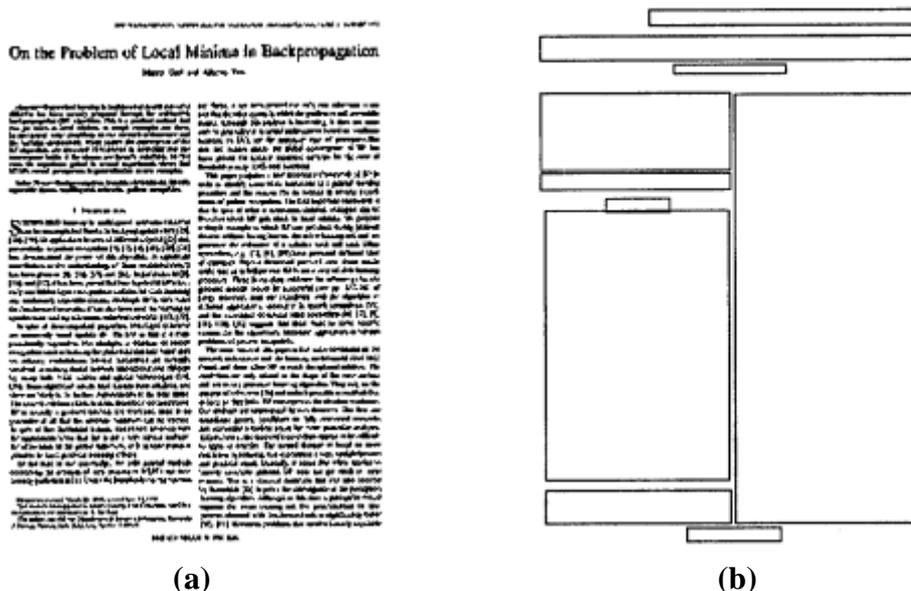


Figure 1.1 : Exemples de structure physique (b) d'un document (a)

- Reconnaissance de la structure logique :

La reconnaissance de structure de document consiste à retrouver la structure logique. C'est l'organisation de l'image du document en entités logiques : titres, chapitres ou paragraphes (**Figure 1.2**) telles qu'elles ont été perçues par l'auteur indépendamment de leur mise en page. La majorité des travaux se basent sur la structure physique pour aboutir à la structure logique [4].

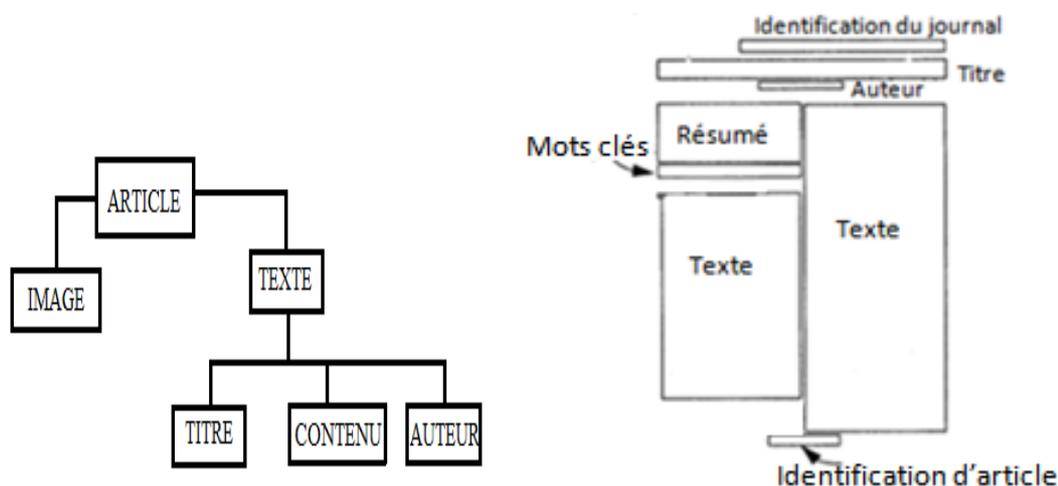


Figure 1.2 : Structures logiques [3]

Dans ce qui suit nous allons passer en revue l'état de l'art pour la reconnaissance de structures physiques (ou l'analyse d'image de document).

1.1 Définition de la segmentation

La segmentation est usuellement définie comme l'opération consistant à partitionner une image en régions homogènes [5]. Une région est alors définie comme un ensemble de pixels qui partagent une caractéristique commune comme l'intensité, la couleur, la texture, le mouvement, la distance, ... etc. [6]. Ces régions et contours sont supposés être pertinents, c'est-à-dire que les régions doivent correspondre aux parties significatives des objets du monde réel, et les contours à leurs contours apparents [7]

La segmentation sera validée si les conditions suivantes sont vérifiées :

- Tous les composants de l'image doivent être classés, autrement dit chaque pixel doit appartenir à une classe ;
- Les régions doivent être homogènes ;
- Les différentes régions doivent être séparées (disjointes).

Le contexte de notre étude repose plus particulièrement sur l'analyse des documents. L'objectif visé est de séparer les zones textuelles des zones non-textuelles

1.2 Pourquoi on fait l'analyse de document

L'utilisation des traitements de textes a permis d'uniformiser la présentation des documents et de faire ressortir les entités logiques essentielles du texte dans un format réutilisable ou de faire séparer les zones textuelles avec celles graphiques ou images. L'apport des standards a été considérable pour normaliser les modes de description des documents et proposer un cadre d'unification des interfaces avec les produits logiciels facilitant leur

archivage et de faire face à l'augmentation croissante du volume des documents des systèmes informatiques de plus en plus performants sont requis.

Il s'agit par exemple :

- * dans le domaine bancaire, de la lecture automatique des bulletins de versement
- * dans le domaine postal du tri automatique des courriers
- * dans le domaine bibliothécaire, de la classification automatique de documents
- * dans le domaine éditorial de la retro conversion automatique de documents papiers à une forme électronique c'est la récupération de l'existant.....etc.

Ces traitements procurent des avantages considérables aussi bien sur les plans économique organisationnel qu'écologique partage de l'information :

Accès plus rapide a l'information

Réduction des couts de saisie de nouvelles informations

Réduction des espaces de stockage

Réduction des frais lies à l'infrastructure nécessaire pour l'archivage des documents papiers

1.3 Les différentes approches de segmentation

Les méthodes de la reconnaissance de structures physiques, décrites dans la littérature peuvent être classées en deux grandes classes : les méthodes descendantes (Top - down), et les méthodes ascendantes (Bottom-up) [2][3]. Il existe aussi une troisième classe qui représente les méthodes mixte ou encore hybride [8]. L'approche mixte consiste à combiner le principe des deux approches descendantes et ascendantes [4].

1.3.1 Méthodes descendantes

Les méthodes descendantes permettent de segmenter l'image par des découpages successifs de grandes composantes (blocs de grandes tailles) en composantes plus petites. elles commencent par le niveau le plus élevé à savoir la page et descend d'un niveau à un autre jusqu'à arriver au niveau des caractères ou au niveau pixel. Ces méthodes sont performantes dans les situations où l'on connaît la structure a priori du document à analyser. C'est la raison pour laquelle, elles s'appliquent essentiellement à des documents très spécifiques et très hiérarchisés, tels les documents administratifs et scientifiques [2].

1.3.2 Méthodes ascendantes

Par contre les méthodes ascendantes permettent d'extraire des blocs de textes par la fusion de petites composantes (généralement des composantes connexes) jusqu'à obtenir des blocs plus larges. En effet, la dernière fusion permet la reconstitution de la page une fois les fusions du bas niveau ont été effectuées [3].

La **figure 1.3** illustre l'approche descendante et ascendante.

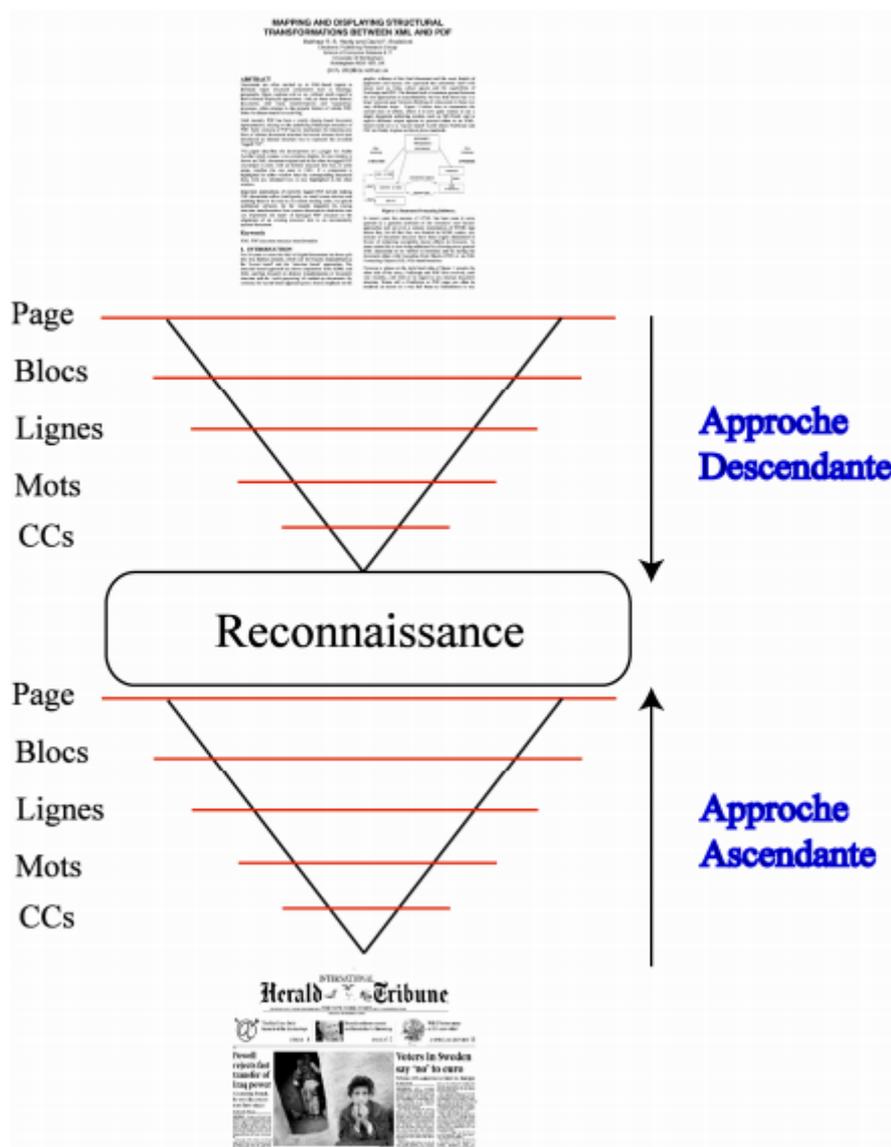


Figure 1.3 : Approche descendante et ascendante

1.4 Algorithmes de segmentation de documents

Nous proposons dans cette partie de faire une synthèse des différents travaux relevant de l'analyse de documents, ce que nous rappelons plus simplement la reconnaissance de la structure physique. Nous présenterons les méthodes liés à l'analyse d'images de document suivantes :

- Découpage X-Y
- Division et fusion (split and merge)
- Algorithme de Bloomberg
- RLSA
- Docstrum
- Composantes connexes
- Diagramme de Voronoi

1.4.1 Découpage X-Y

Cet algorithme a été introduit par **Nagy** en 1984 [8]. Il est plus approprié aux structures de type Manhattan « une structure de page de type Manhattan est celle dont les régions de la page sont toutes des rectangles et que les rectangles possèdent la même orientation ». En réalité ce nom fait référence à la disposition des bâtiments de quartier de Manhattan à New York.

La **figure 1.4** illustre une structure de page de type Manhattan segmentée par découpage X-Y [9].

<p>J. H. Greiner C. J. Kircher S. P. Klepner S. K. Lahiri A. J. Warnecke S. Basavalah E. T. Yen John M. Baker P. R. Broolous H.-C. W. Huang M. Murakami L. Ames</p>	
<p>Fabrication Process for Josephson Integrated Circuits</p>	
<p>A process for fabricating experimental Josephson integrated circuits is described that is based primarily on the use of vacuum-deposited Pb-ally and SiO₂ films patterned by photoresist stencil lift-off. The process has evolved from one previously reported, with changes having occurred in junction electrodes, tunnel barrier formation, layer patterning, device geometry, and minimum linewidths. Films of Pb-In(12 wt% Au) alloy (200-300 nm thick) are used for forming junction base electrodes, interferometer contacts, and interconnection lines. Tunnel barriers are formed on the base electrode films by thermal oxidation and subsequent sputter-etching in an oxygen plasma. Junction counter electrodes are formed from 400-nm-thick Pb-Bi(29 wt% Au) alloy films. Ground planes are formed from 300-nm-thick Ni films patterned by subtractive etching and insulated in part by a Nb₂O₅ layer formed by liquid condensation. Films of the inorganic compound AuAu₂Sb(42 nm thick) are used for forming terminating, load, and coupling resistors. The SiO₂ films are used for interlayer insulation, for defining junction areas in interferometers, and as protective coatings. Laser patterning is achieved mainly by means of photoresist lift-off masks. By utilizing this process, experimental logic and memory circuits containing ~100 interferometers with lines as small as 2.5 μm in width have been successfully fabricated.</p>	
<p>INTRODUCTION Josephson tunneling devices exhibit fast switching and low power dissipation [1], characteristics that make them attractive for future computer applications [2]. Work is in progress to explore the potential of circuits containing such devices. A process was developed that allowed several initial types of logic and memory circuits to be successfully fabricated [3]. Since that time, improvements have been made in the process, and other investigators have successfully made devices and circuits using adaptations of the process [4].</p>	<p>This paper describes the fabrication process used to prepare recent Josephson devices and circuits. The multi-layer, integrated circuits were formed on oxidized Si substrates primarily through use of vacuum-deposited thin films. Superconducting layers were generally Pb alloy and insulation layers, SiO₂. These layers were patterned using photoresist stencil lift-off methods and optical lithography with minimum linewidths of 2.5 μm. An overview of the fabrication process is presented, and preparation and properties of the various layers are discussed.</p>
<p>Copyright 1983 by International Business Machines Corporation. Copying is permitted without payment of royalty provided that the reproduction is done without alteration and (2) the Journal reference and IBM copyright notice are included on the first page. The title and abstract may be used without further permission in computer-based and other information-service systems. Permission to reproduce other contents should be obtained from the Publisher.</p>	

Figure 1.4 : exemples Structure de page de type Manhattan [9].

L'hypothèse de base repose sur le fait que les éléments structurés de la page sont généralement présentés dans des blocs rectangulaires. Mais aussi sur le fait que les blocs peuvent être divisés en groupes de telle sorte que les blocs qui sont adjacents l'un à l'autre, dans un groupe, ont une dimension en commun [10]. Le RXYC découpe l'image horizontalement puis verticalement, récursivement, le découpage se fait dans les creux (espaces blancs dans l'histogramme) des projections [11]. L'utilisation des projections horizontale et verticale se fait dans le but de trouver les espaces interligne. Une projection verticale (resp. horizontale) est la somme des valeurs des niveaux de gris des pixels sur une ligne (resp. colonne). Ces projections représentent donc, pour une ligne donnée, l'intensité totale des pixels. Si la valeur de la projection est faible, c'est qu'il y a beaucoup de pixels foncés sur la ligne.

Le résultat d'une telle segmentation peut être représenté dans un arbre X-Y (figure 1.5.e), dans lequel la racine correspond à la page toute entière et les feuilles représentent les blocs de la page et chaque niveau de l'arbre représente alternativement les résultats de la segmentation horizontale ou verticale [12]. la figure 1.5 représente un exemple de découpage avec son arbre correspondant. Dans la partie (a) nous montrons l'image originale et les parties (b), (c), (d) décrivent la position des coupes à différents niveaux de segmentation.

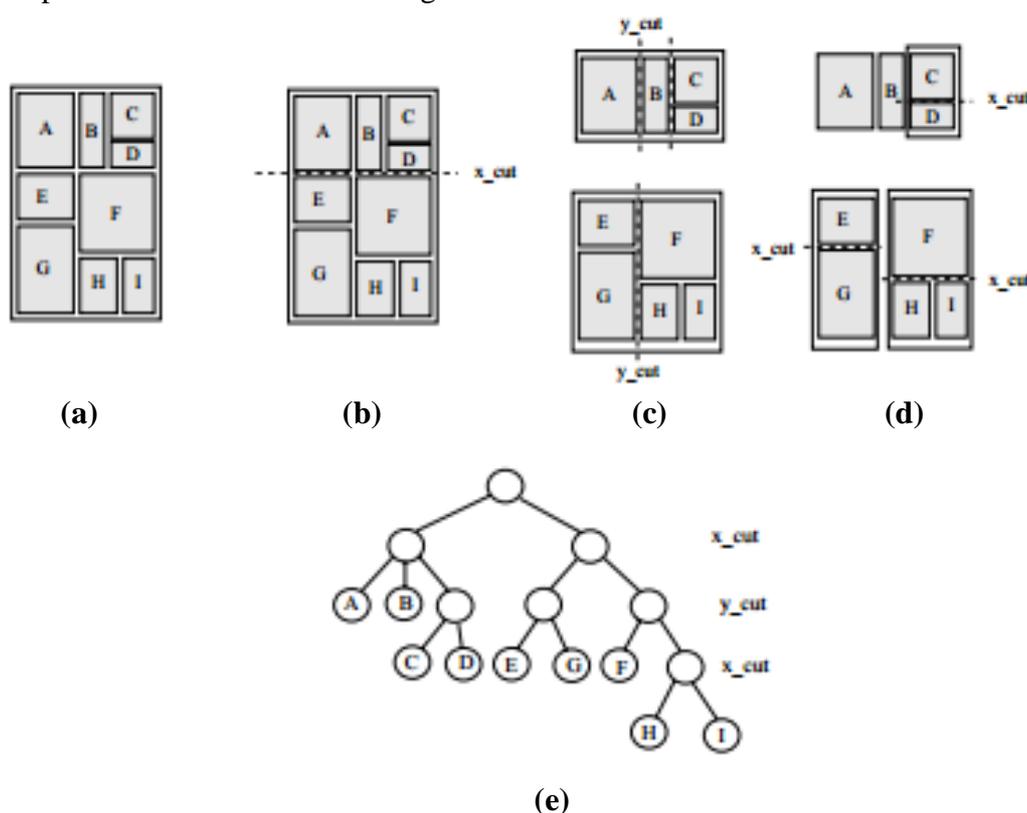


Figure 1.5 : Exemple de découpage d'une image avec arbre de décomposition

Cependant, cette technique a l'inconvénient majeur de poser des problèmes avec des documents dont la mise en page est particulière (figure 1.6). De même, dans le cas où les lignes ne seraient pas bien alignées, ou que le document serait légèrement tourné, la méthode ne peut plus trouver de creux dans les projections. Ces limitations ont conduit

quelques recherches pour améliorer la technique, dont une correction de l'inclinaison des lignes, mais les résultats ne sont pas très convaincants. Quelques recherches utilisent l'algorithme de découpage X-Y mixé avec d'autres méthodes [13].



Figure 1.6 : exemple de document à problème (structure non-Manhattan)

Cesarini et al, [14] proposent une modification sur l'arbre de découpage X-Y (the X-Y tree), le nouveau algorithme est nommé M-X-Y (the Modified X-Y tree). Les principales caractéristiques de M-X-Y sont le stockage des lignes de coupe en tant que feuilles de l'arbre et la description des relations d'adjacence entre les feuilles de cet arbre (inter-leaves).

Kritawan et al [13] présentent une méthode de segmentation de document qui utilise le découpage X-Y précédé par une détection de bordures dans l'image. Si deux blocs sont segmentés par un même contour, leur séparation est assurée par une projection verticale dans les espaces inter caractères.

Sukhvir et Palvinder [15] présentent une amélioration par rapport à l'algorithme de découpage X-Y pour bruit de frontière de page. Quand un bruit apparaît dans le document après la numérisation, l'algorithme segmente ce bruit en plusieurs segments, dans ce cas une mauvaise segmentation est obtenue. Le nouvel algorithme est composé de deux étapes, la première consiste à la suppression du bruit de frontière, et la deuxième c'est le processus de segmentation de page de document.

1.4.2 Division et fusion (split and merge)

L'algorithme split and merge a été présenté la première fois en 1974 par Pavlidis et Horowitz [16] [17]. Cette technique enchaîne les 2 phases suivantes:

1. Découper itérativement l'image jusqu'à avoir des blocs contenant exclusivement des pixels similaires.
2. Regrouper les blocs voisins s'ils sont similaires

Les deux phases sont nécessaires afin de garantir que les régions obtenues sont à la fois homogènes et également les plus grandes possibles. Chaque phase étant indépendante, on peut les étudier séparément [16].

❖ Division de l'image (split)

La méthode de découpage de l'image utilisée dans cet algorithme est basée sur l'utilisation de quadtree. Cette structure de données est un arbre quaternaire (**figure 1.8**) qui permet de stocker l'image à plusieurs niveaux de résolution [18]. Initialement, l'image est considérée comme une seule région. Par la suite, chaque région R_i est traitée, puis prédicat d'homogénéité est appliquée (niveau de gris, couleur, texture,...) [19]. Donc chaque noeud (bloc) du tétra-arbre est analysé de sorte à savoir s'il doit être divisé en quatre sous blocs et ce jusqu'à ce que chaque sous-bloc respecte le prédicat d'homogénéité [20].

Si la région R_i ne vérifie pas le prédicat, elle est alors divisée en plusieurs sous régions (généralement 4) (**figure 1.7**). L'opération est appliquée sur chaque sous régions, et l'algorithme de reconstruction s'arrête lorsque toutes les régions respectent le prédicat d'homogénéité (division de toutes les régions non homogènes) [21].

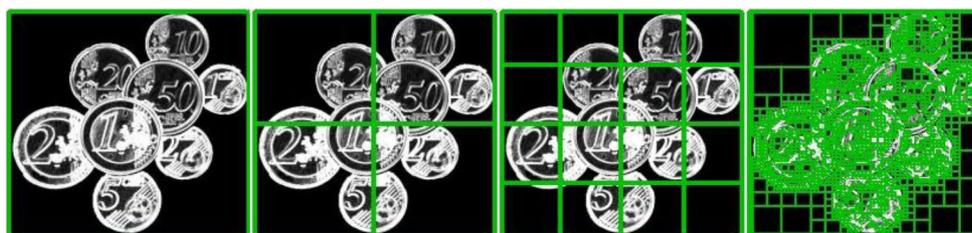


Figure 1.7 : exemple de division

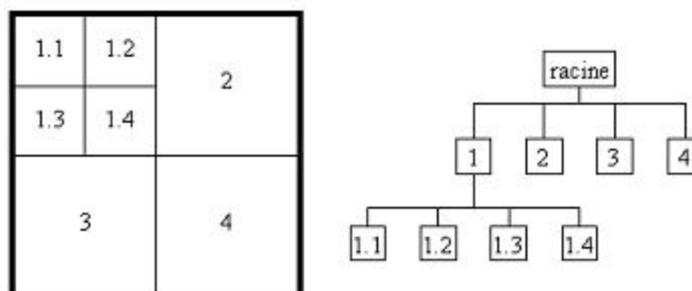


Figure 1.8 : Représentation sous forme de Quadtree

❖ Fusion des régions (merge)

La procédure de découpage décrite précédemment aboutit à un nombre de régions trop élevé. La cause fondamentale de cette sur-segmentation est que l'algorithme découpe les régions de manière arbitraire. Il se peut qu'il coupe de cette façon une zone homogène en deux ou quatre parties [16]. La solution, qui correspond à la phase merge de l'algorithme, est de procéder à une fusion de régions après le découpage, le prédicat d'homogénéité est vérifié pour les couples de régions voisines. Si deux régions vérifient ce prédicat, elles seront fusionnées. Pour réaliser cette fusion, il faut d'abord tenir à jour une liste des contacts entre régions on construit un graphe d'adjacence de régions (**figure 1.9**) (ou Region Adjacency Graph) en même temps que l'arbre de découpage.

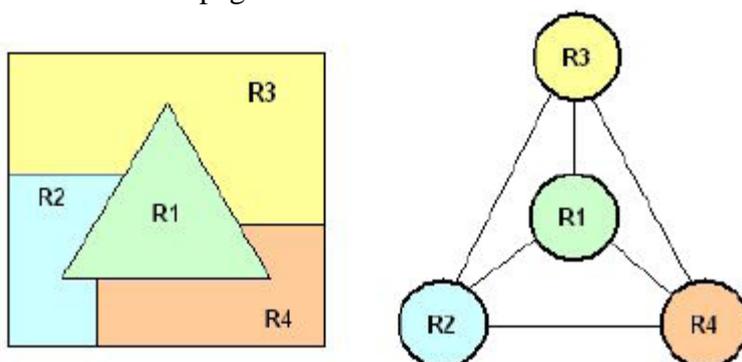


Figure 1.9 : La carte des régions et le graphe d'adjacence [16]

La figure suivante illustre un exemple de fusion de régions d'une image divisée :

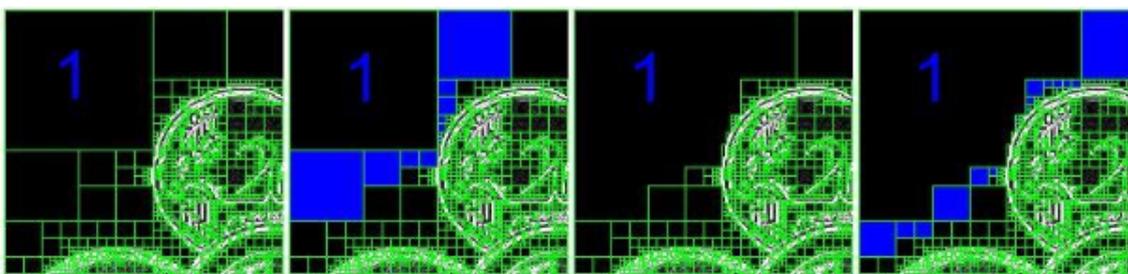


Figure 1.10 : Agrégation itérative des blocs similaires au bloc 1 [16]

1.4.3 Algorithme de Bloomberg

Bloomberg dans [22] présente une méthode de segmentation de pages d'images de documents en texte / non texte basées essentiellement sur la détection des parties non texte. Celle-ci comporte un ensemble d'opérations morphologiques (dilatation, érosion, ouverture, fermeture) et après réduction de la taille de l'image binaire ($\frac{1}{16}$ de la taille originale[22]).l'algorithme de segmentation élimine toute composante textuelle de l'image et ne laisse que des résidus (des parties non textuelles de l'image). Par la suite une opération d'expansion sur l'image résultante permettra l'extraction des blocs non textuels (image, graphique,...).

Bukhari et al [23] trouvent que l'algorithme de Bloomberg est incapable de séparer entre un texte et un dessin (ou bien une carte, un graphe) car les détails fins et les lignes des composantes non textuelles peuvent être détruites et considérés comme texte (**figure 1.11**) donc l'algorithme donne une mauvaise segmentation.

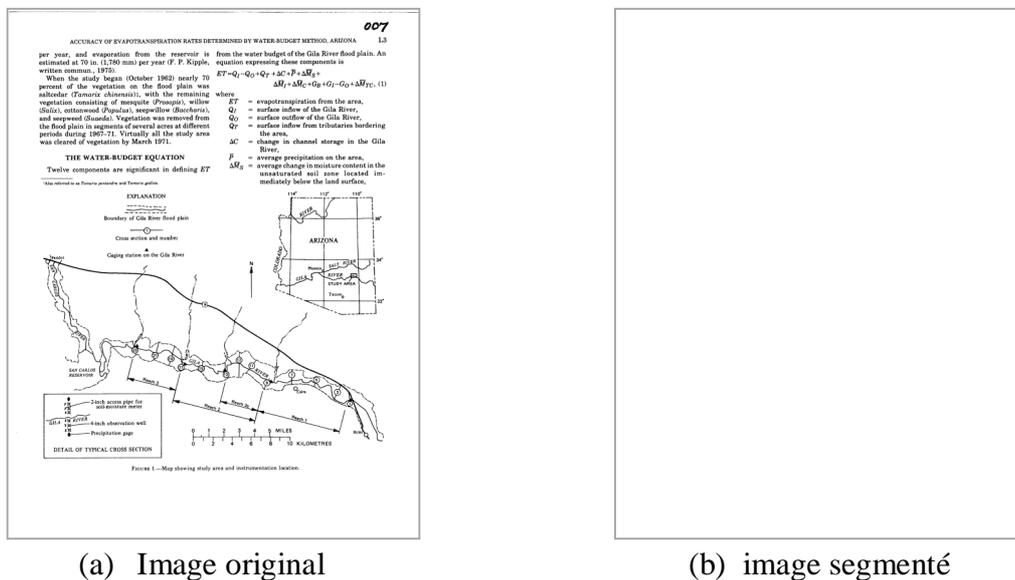
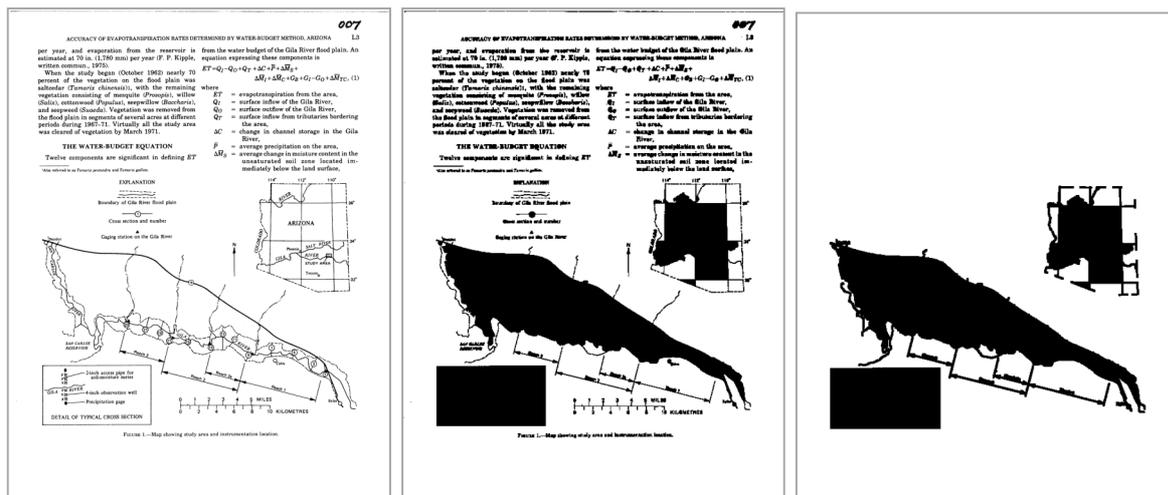


Figure 1.11 : exemple de document à problème avec Bloomberg [23]

Syed Saqib Bukhari, et al proposent une amélioration en deux versions :

- la première est de commencer par une opération morphologique de bouchage de trous dans l'image.
- la deuxième est d'ajouter une étape de reconstitution de lignes détruites. Cette opération doit précéder le bouchage cité précédemment.



(a) Image original

Reconstruction
et bouchage

(b) image segmenté
(partie non texte)

Figure 1.12 : segmentation de l'image après amélioration [23]

1.4.4 Algorithme RLSA

Le Run Length Smoothing or Smearing Algorithm (RLSA) est un filtre qui agglomère (smoothing / smearing) les séquences (Run-Length) de pixels noirs en fonction de la longueur des espaces. Il est aussi connu sous le nom Constrained Run-Length Algorithm (CRLA) [24]. La longueur des séquences dépend d'un paramètre appelé Contrainte C. Avec des valeurs croissantes de C, les caractères sont regroupés en mots, les mots en lignes [25]. C'est un algorithme assez connu permettant de segmenter un document-image en régions homogènes.

Le principe du RLSA (basique) est appliqué sur une séquence binaire ou les pixels blancs sont représentés par des 0 et les pixels noirs par des 1. L'algorithme transforme une séquence S en une séquence C selon les règles suivantes :

- les 0 dans S sont transformés en 1 dans C si le nombre de 0 adjacents est inférieur ou égale à une limite prédéfinie L.
- les 1 dans S restent inchangés dans C.

Par exemple avec L=4, la séquence S est transformée en la séquence C :

S : 00010000010100001000000011000
 C : 11110000011111111000000011111

Une amélioration du RLSA (basique) est la suivante :

- les 0 dans S sont transformés en X dans C si le nombre de 0 adjacents est inférieur ou égale à une limite prédéfinie L et que les deux valeurs bornant la séquence de 0 soient égales à X.

Par exemple avec L=5 et X=1 on transforme la séquence S en C :

S : 2211311000021332221110000012213

C : 2211311000021332221111111112213

Le principe de segmentation de document-image par RLSA est de noircir toute séquence de pixels blancs comprise entre deux pixels noirs, de longueur inférieure à un seuil donné. En pratique l'algorithme est appliqué horizontalement (**figure 1.13 .b**) et verticalement (**figure 1.13 .c**) sur l'image binaire originale avec des seuils éventuellement différents pour l'horizontale et la verticale, puis une opération « ET logique » est appliquée entre les deux images lissées obtenues (**figure 1.13 .d**). L'extraction des composantes connexes de l'image résultante permet d'obtenir les entités de la structure physique sur un niveau hiérarchique donné. On peut ainsi, en répétant la procédure avec des seuils de lissage horizontal et vertical différents, extraire itérativement les blocs de l'image, puis les lignes de texte et les mots. Ces seuils de lissage sont les seuls paramètres de l'algorithme RLSA. Ils contrôlent la manière dont les composantes sont fusionnées sur un niveau de segmentation prédéterminé. Par exemple pour segmenter des lignes de texte horizontales, droites et bien espacées, on pourra utiliser un seuil de lissage vertical nul et un seuil de lissage horizontal suffisamment grand pour combler les espaces inter-lettres et inter-mots [26].

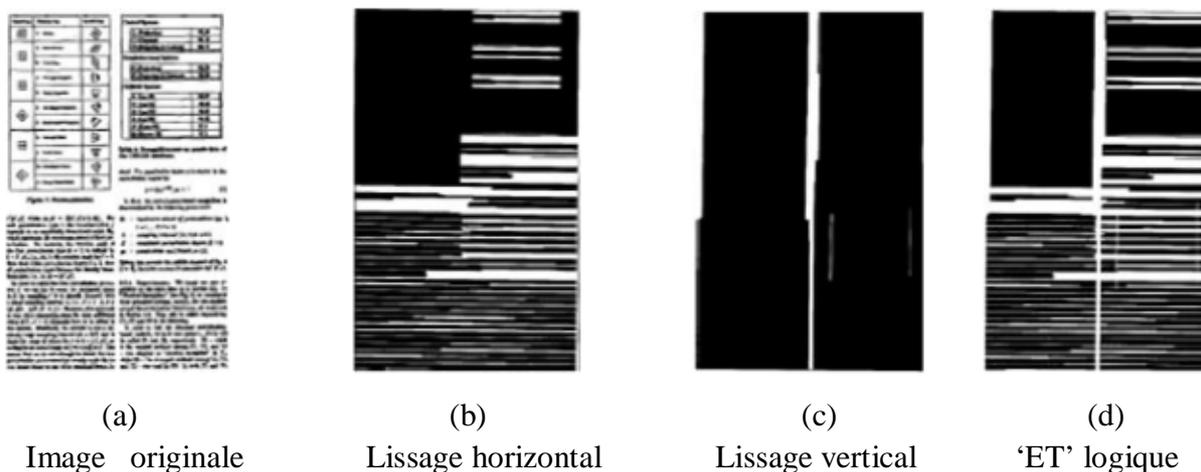


Figure 1.13 : segmentation RLSA [26]

Cet algorithme a l'avantage, en plus d'être très simple à mettre en œuvre, de ne requérir qu'un nombre limité d'itérations pour atteindre la segmentation complète du document. La principale difficulté dans l'utilisation de cet algorithme est le réglage des seuils de lissage qui est délicat. Il est nécessaire de déterminer les seuils adéquats pour chaque niveau de segmentation.

Le RLSA est particulièrement efficace l'ors qu'il s'agit de partitionner des documents ayant une mise en page de type Manhattan (textes / images alignés est séparés par des espaces blancs), mais s'avère moins fiable lorsqu'on est en présence de mises en page complexes (image imbriquée dans un texte par exemple). Son principale défaut est qu'il utilise uniquement des informations locales lors du lissage se qui peut entraîner l'apparition de liens entre le texte est l'image [27].

Hung-Ming Sun dans [24] présente une solution à ce problème en ajoutant des informations globales dans le processus de segmentation RLSA. Cet algorithme amélioré peut être appliqué à des documents ayant une mise en page non-Manhattan (**figure 1.14**). Il peut également extraire un texte entouré par une boîte (**figure 1.14**).

La nouvelle version de CRLA est effectuée sur une image en étiquette qui est dérivée à partir de l'image originale du document et a la même taille que l'image de document. Dans l'image étiquetée, chaque pixel peut avoir différentes valeurs; la valeur zéro signifie le fond et le non-zéro signifie non seulement l'avant-plan (les objets dans une image) mais aussi la taille de la composante [24]. Pour générer l'image d'étiquettes, un algorithme d'étiquetage en composantes connexes est utilisé pour localiser les composants dans l'image du document et de calculer leurs tailles.

Le travail de [28] propose une stratégie bottom-up générale pour aborder l'analyse de documents qui ont une mise en page non-Manhattan (peut-être rencontrée). C'est une variante de l'algorithme de lissage RLSA, appelé RLSO qui exploite l'opérateur logique « OR » à la place de « AND » [29].

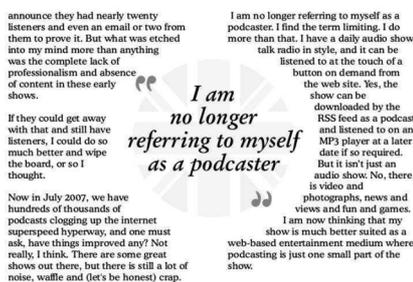


Figure 1.14 : structure non-Manhattan

1.4.5 Docstrume

O’Gorman a proposé d’utiliser uniquement le voisinage entre composantes pour localiser les zones de texte dans une image [30]. En (1992) il traite de ce qu’il appelle la technique docstrum pour déterminer la structure géométrique d’une page de document. Cette technique consiste à calculer les k plus proches voisins de chacune des composantes connectées noires de la page [31]. Chaque paire de voisins les plus proches possède un angle et une distance associée. En regroupant les composantes à travers les caractéristiques citées précédemment, les régions géométriques de structures physiques de la page peuvent être déterminées. La méthode proposée est indépendante du changement de l’orientation de la page mais aussi de l’espacement intertexte. Cependant, la valeur du k est dépendante de la structure de la page [3].

1.4.6 Composantes connexes

Ces méthodes considèrent chaque page comme un ensemble de composantes connexes. Sur la **figure 1.15**, on peut constater que la taille, la proximité et la position relative des composantes connexes peuvent être utilisées pour extraire la structure physique d’une page. Les rectangles circonscrits aux composantes connexes se chevauchent fréquemment dans les graphiques et rarement dans les textes. Ainsi, les zones graphiques correspondent aux composantes connexes de cet ensemble de rectangles circonscrits dont les dimensions (largeur ou hauteur) dépassent un seuil choisi souvent empiriquement par les experts [30].



Figure 1.15 : Composantes connexes et segmentation

Un bloc de texte est un ensemble de petites composantes connexes « proches ». Deux caractères sont dits voisins si la distance les séparant est inférieure à un espace maximal.

2.4.7 Diagramme de Voronoi

Le diagramme de Voronoï en tant que structure de représentation de données a été largement étudié dans le domaine de la géométrie et appliqué dans différentes disciplines. En ce qui concerne l’analyse d’images, le diagramme de Voronoï suscite un intérêt particulier. En effet, il représente un moyen efficace pour décrire, manipuler et interpréter des entités géométriques. Ce diagramme consiste à générer un partitionnement du plan en polygones (**figure 2.16**) et ceci dans un contexte de structure de données de type graphe [32].

Une des propriétés les plus connues des polygones de Voronoï en métrique euclidienne concerne leur convexité.

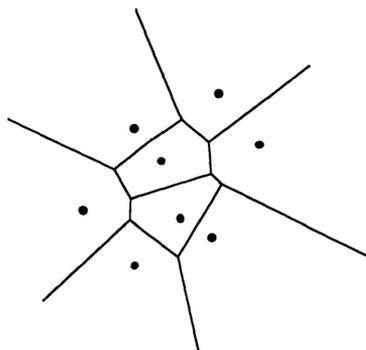


Figure 1.16 : Illustration de la définition d'un diagramme de Voronoï [32]

Ils existent deux approches principales pour la construction de diagramme de Voronoï en surface [33] :

- La première approche construit le diagramme zone de Voronoï en générant d'abord un schéma point de Voronoï pour un ensemble de points échantillonnés se trouvant sur les limites de composantes d'image et ensuite la suppression des bords générés à partir de points des mêmes composantes.
- La seconde approche construit le diagramme zone de Voronoï en étiquetant les composantes de l'image puis en appliquant des opérations morphologiques dont le but d'étendre leurs bordures jusqu'à ce que deux étiquettes différentes se rencontrent.

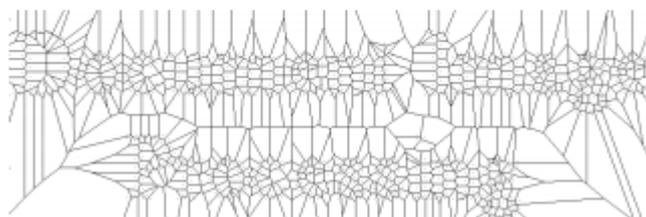
Kiochi Kise [34] propose d'utiliser un pavage de Voronoï plutôt que les composantes connexes. Le pavage permet de construire un graphe de voisinage simplifiant la localisation des lignes de texte par des regroupements des formes proches.

Kise [35] présente une méthode de segmentation de pages basée sur la surface approximée des diagrammes de Voronoï pour extraire les blocs de texte d'un document. La méthode repose sur les étapes suivantes : au début, un point du diagramme de Voronoï est construit à partir de l'ensemble des pixels noirs sur les contours des composantes connexes (**figure 1.17-b**). Ensuite, une surface est obtenue en éliminant du point du diagramme Voronoï tous les arêtes générées à partir d'une paire de points sur la même composante connexe (**figure 1.17-c**). Une caractéristique distinctive de cette méthode est qu'elle s'applique sur des images de documents possédant une structure de type Manhattan et ayant subi une rotation. Il est à noter que cette méthode est efficace pour l'extraction des zones de texte et possède un taux de reconnaissance comparable à celui obtenu par les méthodes basées sur l'analyse des composantes connexes.

La figure suivante illustre les différentes étapes de construction de diagramme de Voronoï et diagramme de voisinage.

Document Image Processing

(a) Image originale



(b) diagramme en points



(c) diagramme de Voronoi



(d) graphe de voisinage

Figure 1.17 : diagramme de Voronoi et diagramme de voisinage [33]

Discussion

Dans ce chapitre, nous avons passé en revue quelques méthodes de segmentation de blocs d'images de documents en explicitant les techniques les plus employées. Cependant chacune d'entre elles a ses qualités et ses défauts, et c'est sur cette base que doit se faire le choix de l'emploi de l'une ou l'autre de ces techniques.... ?

Malgré les travaux réalisés et les efforts effectués dans ce domaine de recherche ; il n'y a toujours pas de procédé général de traitement d'image de différentes classes de document à la fois avec précision et automatiquement.

Le but de ce mémoire est de proposer une méthode de segmentation de blocs d'images de documents composites, nous avons opté pour une méthode de classification hybride composée d'une phase de détection de zones non textuelles par utilisation de quelques opérations morphologiques et une deuxième phase de segmentation Bottom-up c'est le RLSA.

Dans la section suivante, nous présentons les différentes techniques de traitement effectuées sur l'image de document.

CHAPITRE 2

TECHNIQUES DE TRAITEMENT D'IMAGE UTILISEES

Le traitement d'images est un ensemble des traitements automatisés qui permet à partir d'image numérisée, de produire d'autres images numériques. Il s'agit donc d'un sous-ensemble du traitement de signal dédié aux images et aux données dérivées. Il constitue un outil puissant dans la recherche scientifique, son but ultime consiste à extraire le contenu informationnel (ou information pertinente : lignes de textes, image,...) des images en vue d'une prise de décision ou d'une action.

Le **traitement d'image** est la sous-branche du traitement du signal qui se consacre aux images. C'est aussi toute la partie de calcul d'une chaîne de vision artificielle. C'est une science au carrefour des mathématiques, de la physique, et de l'informatique. Ponctuellement on arrive à y glisser un peu de biologie [27].

En effet, on parle alors de photographie ou de télévision. Par contre dans la partie numérique on emploie le terme de traitement d'image, sous entendu numérique, et on fait souvent clairement référence au traitement du signal.

De ce fait, le traitement d'images est l'ensemble des méthodes et techniques opérant sur celles-ci, dans le but de rendre cette opération possible, plus simple, plus efficace et plus agréable, d'améliorer l'aspect visuel de l'image et d'en extraire des informations jugées pertinentes.

Dans ce chapitre nous allons détailler les techniques de traitement d'image utilisées dans notre projet qui sont :

- Filtrage (gaussien, moyen, médian)
- Binarisation (ou seuillage).
- Etiquetage logique.
- Opérations morphologiques (érosion, dilatation, ouverture,...)

2.1 Filtrage

Les applications actuelles de reconnaissance de l'écriture utilisent souvent des supports dont le fond n'a pas une couleur unie (un chèque par exemple). Il faut donc isoler le mot du motif d'arrière-plan. Une atténuation du bruit est ensuite effectuée. Ce bruit peut être déjà présent sur le support avant la numérisation ou être introduit par la phase d'acquisition ou les premiers prétraitements. Ce sont la encore des techniques de filtrage qui sont utilisées pour éliminer le bruit de type "poivre et sel" ou pour arrondir les contours. Un test sur la taille des motifs est aussi effectué pour éliminer les tâches trop petites pour faire partie d'un caractère [36]. Dans ce qui suit nous présentons le filtrage linéaire et le filtrage non linéaire.

2.1.1 Filtrage linéaire

Le filtrage consiste à réaliser une transformation locale, basée sur le voisinage d'un point (x, y) [37].

Un filtre est dit linéaire si la valeur du nouveau pixel est une combinaison linéaire des valeurs des pixels du voisinage.

Le filtrage d'une image est une opération de convolution à deux dimensions (**équation 2.1**).

Si f est la fonction à filtrer (ou à rehausser) et « g » la fonction de filtre spatial (ou **PSF** – Point Spread Function ou masque) on a :

$$f(x, y) * g(x, y) = F^{-1}\{ F(f(x, y)) \times F(g(x, y)) \} \quad (2.1)$$

Avec $F(g(x, y))$ est $G(u, v)$ fonction de transfert du filtre.

Une image numérique étant essentiellement discrète (pixels et niveaux de gris) nous allons présenter les filtres dans le cas discret. x et y sont des entiers (coordonnées des pixels) et f est à valeur entière (dans $\{0, \dots, 255\}$). Dans le cas discret l'opération de filtrage est donnée par les deux formules (**équation 2.2 et équation 2.3**).

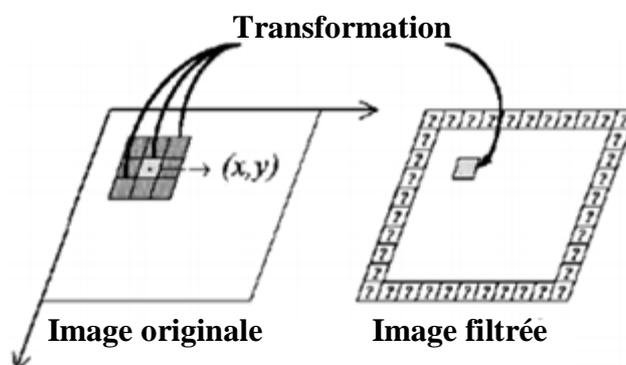


Figure 2.1 : Convolution locale

Le noyau de convolution (masque, PSF) du filtre k est à support compact inclus dans $[x_1, x_2] \times [y_1, y_2]$.

$$g(x, y) = (f * k)(x, y) = \sum_{i=x_1}^{x_2} \sum_{j=y_1}^{y_2} f(x-i, y-j)k(i, j) \quad (2.2)$$

Généralement le filtre est de dimension d_i impaire et est symétrique. Dans ce cas

$$[x_1, x_2] = \left[-\frac{d_1-1}{2}, \frac{d_1-1}{2} \right] \quad \text{et} \quad [y_1, y_2] = \left[-\frac{d_2-1}{2}, \frac{d_2-1}{2} \right]$$

w_1	w_2	w_3	← $y - 1$
w_4	w_5	w_6	← y
w_7	w_8	w_9	← $y + 1$
↑	↑	↑	
$x - 1$	x	$x + 1$	

Tableau 2.1 : Filtre (i,j) $d_1 = d_2 = 3$

Ici $d_1 = d_2 = d = 3$; les w_i (avec $i = 1, \dots, 9$) sont les poids du filtre par exemple $k(0,0) = w_5$.

Sur cet exemple on a :

$$\begin{aligned} g(x, y) = & w_1 f(x-1, y-1) + w_2 f(x, y-1) + w_3 f(x+1, y-1) \\ & + w_4 f(x-1, y) + w_5 f(x, y) + w_6 f(x+1, y) \\ & + w_7 f(x-1, y+1) + w_8 f(x, y+1) + w_9 f(x+1, y+1) \end{aligned} \quad (2.3)$$

Afin de conserver la moyenne de l'image f , la somme des éléments du filtre est normalisée à 1 (**équation 2.4**) :

$$\sum_i w_i = 1. \quad (2.4)$$

2.1.1.1 Filtre moyen

Le filtre **moyen** consiste à conformer un pixel avec son voisinage en le remplaçant par la **moyenne des niveaux de gris de son voisinage [38]**.

Ce type de filtre utilise la moyenne non pondérée des voisins et peut être mis sous la forme d'un masque 3*3 ou 5*5 comme suit :

$$\frac{1}{9} \cdot \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array} \quad \frac{1}{25} \cdot \begin{array}{|c|c|c|c|c|} \hline 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 1 \\ \hline \end{array}$$

Figure 2.2 : Masques 3*3 et 5*5 de filtre moyen

Il faut diviser tous les coefficients du masque **par leur somme**, pour que l'image de résultat soit bien normalisée entre 0 et 255 et dans ce cas la somme est de 9 d'où le $\frac{1}{9}$.

Lors de l'opération de débruitage, le masque ne changera pas de taille : ce sera toujours une matrice 3x3 (ou 5x5, 7x7...), qui se déplacera sur toute l'image.

Pour chaque pixel P de l'image, on effectuera alors la convolution entre le masque et la sous-image de même taille que le masque centrée sur P (pixel), et on remplacera la valeur du pixel P par le résultat obtenu.

Le filtre moyen :

- Permet de lisser l'image (smoothing)
- Remplace chaque pixel par la valeur moyenne de ses voisins
- Réduit le bruit
- Réduit les détails non-importants
- Brouille ou rend floue l'image

La (**figure 2.3**) illustre un exemple de résultats obtenus par filtrage moyen en utilisant différentes tailles du masque :

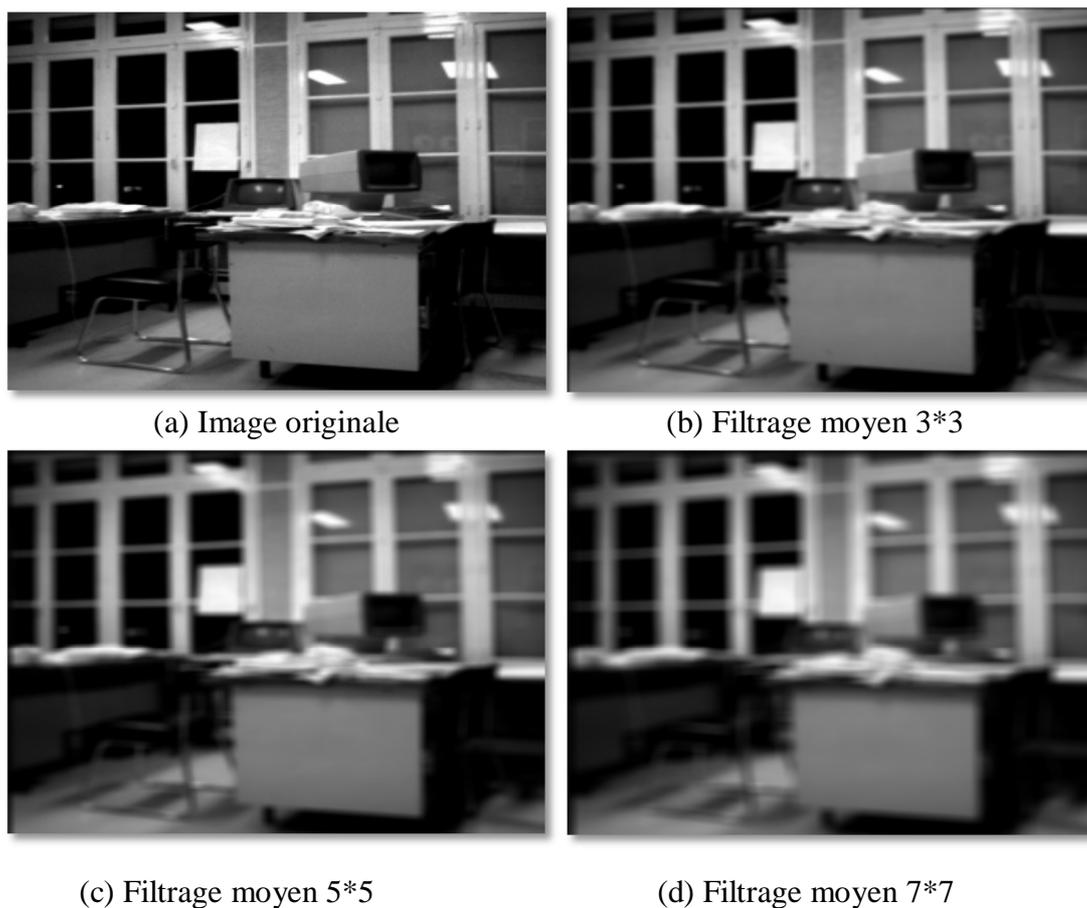


Figure 2.3: Filtrage moyen [39]

On remarque que plus la taille du masque augmente plus l'image devient floue.



Figure 2.4: Effet de lissage

2.1.1.2 Filtre gaussien

Il s'agit de filtre que l'on utilise pour lisser une image, mais en tenant compte du fait que plus un pixel du voisinage est proche du pixel actuellement traité, plus il a des chances de faire partie du même objet, et donc d'avoir la même couleur [38].

Le filtre gaussien est un filtre isotrope spécial avec des propriétés mathématiques bien précises puisqu'il se base sur la fonction gaussienne (**équation 2.5**). Dans le traitement d'image on s'intéresse à des données à deux dimensions x et y d'où on introduit une fonction gaussienne à deux dimensions $G(x, y)$ [40] :

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2.5)$$

Si par exemple $\sigma = 0,8$ on a le filtre 3×3 suivant

$$\begin{array}{|c|c|c|} \hline G(-1, -1) & G(0, -1) & G(1, -1) \\ \hline G(-1, 0) & G(0, 0) & G(1, 0) \\ \hline G(-1, 1) & G(0, 1) & G(1, 1) \\ \hline \end{array} \simeq \frac{1}{16} \cdot \begin{array}{|c|c|c|} \hline 1 & 2 & 1 \\ \hline 2 & 4 & 2 \\ \hline 1 & 2 & 1 \\ \hline \end{array}$$

Tableau 2.2 : Filtre gaussien 3×3

et $\sigma = 1$ pour un filtre 5×5 donne environ

$$\frac{1}{300} \cdot \begin{array}{|c|c|c|c|c|} \hline 1 & 4 & 6 & 4 & 1 \\ \hline 4 & 18 & 30 & 18 & 4 \\ \hline 6 & 30 & 48 & 30 & 6 \\ \hline 4 & 18 & 30 & 18 & 4 \\ \hline 1 & 4 & 6 & 4 & 1 \\ \hline \end{array}$$

Tableau 2.3 : Filtre gaussien 5×5

Idéalement on devrait prévoir un filtre de taille

$$(6\sigma - 1) \times (6\sigma - 1)$$

En générale pour réduire le bruit a l'aide de filtre gaussien on doit utiliser un écart-type < 1 .

Il faut diviser tous les coefficients du masque **par leur somme**, pour que l'image de résultat soit bien normalisée entre 0 et 255.

L'effet du filtre gaussien est assez similaire à celui du filtre moyen, excepté que la moyenne est pondérée et donc les pixels situés au centre ont un poids plus important que ceux situés sur les extrémités.

Un inconvénient évident du filtrage gaussien est qu'il ne lisse pas uniquement le bruit, mais il gomme aussi les contours, les rendant difficilement identifiables. Le filtre gaussien donnera un meilleur lissage et une meilleure réduction du bruit que le filtre moyen.

La figure suivante illustre les résultats obtenus par filtrage Gaussien en utilisant différentes tailles du masque :

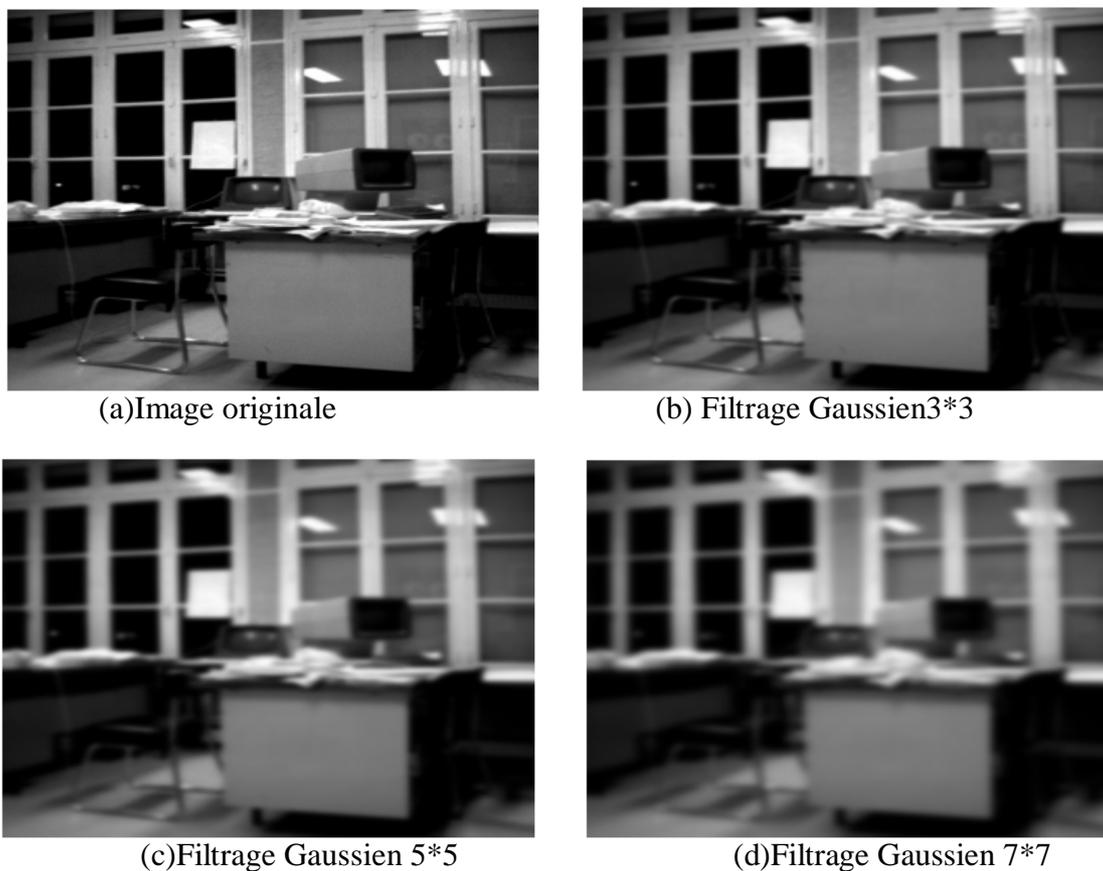


Figure 2.5 : Filtrage gaussien

On remarque que lorsque la taille du filtre augmente l'image devient floue car le nombre de pixels voisins augmente donc on introduit de nouvelles valeurs au pixel considéré.

2.1.2 Filtrage non linéaire

Les filtres non linéaires ne sont pas des filtres de convolution [37].

Si le filtre ne peut pas être exprimé par une combinaison linéaire, il est appelé " non-linéaire ". Les filtres non-linéaires sont plus complexes à mettre en œuvre que les filtres linéaires. Cependant les résultats obtenus avec les filtres non-linéaires sont très souvent de meilleure qualité que ceux obtenus par les filtres linéaires.

Filtre médian

Dans ce filtre on remplace le pixel par la médiane (la valeur du milieu) de son voisinage. On considère une suite de pixels X_i et une fenêtre glissante H . on définit H_i la fenêtre H centrée sur le pixel X_i et contenant $2 \times N + 1$ pixels avec $H_i = \{ X_{i+d} ; X_{i+d} \in H_i \}$ [40] :

Si on utilise une fenêtre de taille 3×3 (9 pixels) alors $N = 1$.

Soit S la suite ordonnée des pixels composants H_i on définit l'élément médian par le pixel de position « med » donnée par la (**équation 2.6**)

$$\text{med} = N + 1 \quad (2.6)$$

Prenons par exemple cette image :

30	10	20
10	250	25
20	25	30

Tableau 2.4 : image à niveau de gris

On va prendre toutes les valeurs de ce voisinage et les classer (par ordre croissant ou décroissant)

									bruit
									↓
10	10	20	20	25	25	30	30	250	
				↑					
				médiane					

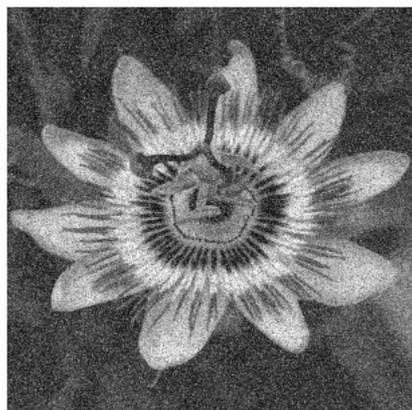
Tableau 2.5 : Classement par ordre croissant

Dans notre exemple la valeur 250 est remplacée par 25.

Ce filtre est utilisé pour contrer l'effet « Poivre et Sel » (P&S) dans l'image (**figure 2.6**).



(a)



(b)

Figure 2.6 : Image originale (a) et image bruitée (b) par Bruit Poivre et Sel [40]

Caractéristiques du filtre médian :

- Robuste, du moment que les valeurs bruitées ou aberrantes n'affectent pas le résultat final.
- Efficace lorsque l'image est dégradée par un bruit impulsionnel consistant dans des variations brusques des pixels isolés.
- Efficace sur le bruit gaussien.
- Inefficace sur des images binaires bruitées

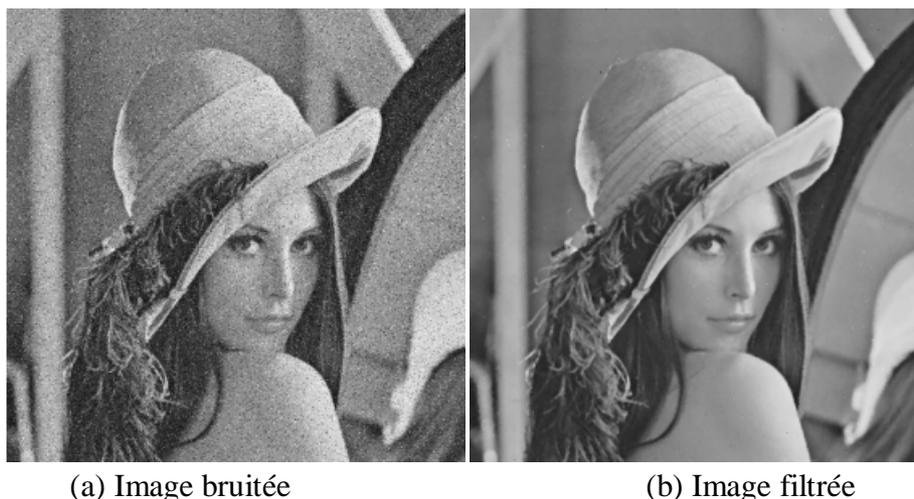
Lorsqu'on prend le voisinage d'un pixel du bord de l'image, il y a un risque de dépasser ces bords ce problème est appelé « effet de bord ». Il n'existe pas de méthode standard pour y remédier. On peut choisir de ne pas s'occuper des pixels du bord en les laissant de côté (en les laisse tels quels).

Soient désormais b_{inf} et b_{sup} avec $1 \leq b_{inf} \leq med \leq b_{sup} \leq (2 \times N + 1)$, qui définissent une sous-suite de S telle que la sortie du filtre soit X_i si ce dernier appartient à la sous-suite. L'élément médian si non. Ces deux paramètres permettent alors de définir un compromis entre conservation des détails et suppression du bruit. Un tel filtre est qualifié de Filtre médian relâché.

Les deux images sur la figure suivante montrent le résultat de l'opération de filtrage médian (**figure2.7.b**) sur l'image originale (**figure2.7.a**). La comparaison des deux images est réalisée par le calcul du rapport suivant :

$$\frac{\text{nombre de pixels bruités}}{\text{nombre de pixels de l'image}}$$

Pour l'image (a) bruitée c'est égale à 0.26 et pour l'image (b) filtrée ce rapport est réduit à 0.07.



(a) Image bruitée

(b) Image filtrée

Figure 2.7 : Filtrage médian [40]

2.2 Binarisation

La binarisation est la technique de classification la plus simple où les pixels de l'image sont partagés par un seuil S en deux classes : ceux qui appartiennent au fond et ceux qui appartiennent à la scène (l'objet). L'image est alors séparée en deux classes de façon à ce que l'information comprise entre 0 et S est retenue et l'autre non, ou vice-versa. Le fond de l'image contient généralement du texte mais également du graphique qui sont quelques fois assez proches, le texte risque donc de fusionner des différentes parties du graphique à cause du changement faible d'intensité des pixels du fond et ceux de l'objet. Une binarisation efficace permet de séparer clairement le fond et l'objet. Cette étape prépare l'image entrée pour l'étape de reconnaissance [41].

Travailler avec des images binaires est plus simple, dans une certaine mesure, qu'avec des images en niveaux de gris, pour les raisons suivantes :

- On peut très facilement identifier des objets à condition qu'ils soient bien séparés (on appelle ça "l'analyse par composantes connexes") ;
- on a beaucoup moins d'informations à traiter, donc les calculs sont globalement plus rapides ;
- si notre image est bien binarisée, les algorithmes sont dans l'ensemble très robustes.

Évidemment, il y a un certain nombre d'inconvénients à binariser une image comme :

- il faut choisir un seuil correct, ce qui n'est pas toujours facile.
- on passe d'un espace de couleurs de dimension 255 à un nouvel espace de dimension 2 : on perd énormément d'informations qui pourraient s'avérer utiles !

La plupart des algorithmes de binarisation sont très sensibles au bruit, on est souvent obligé de lisser l'image en prétraitement.

Définition de l'histogramme

On définit l'histogramme des niveaux de gris d'une image comme étant la fonction $h : [0 \dots L-1]$ qui associe à chaque niveau de gris entre 0 et $L-1$ la quantité de pixels de l'image qui possèdent cette intensité lumineuse [42].

L'histogramme d'une image peut être représenté par un vecteur dont chaque composante est un nombre de pixels de niveau de gris correspondant à son indice. Il permet de fournir effectivement une estimation de la densité de probabilité des valeurs des pixels sur l'image observée [43].

$h(i) = n_i$, $i = 0, 1, \dots, L - 1$ où n_i le nombre de pixels de niveau de gris i dans l'image.

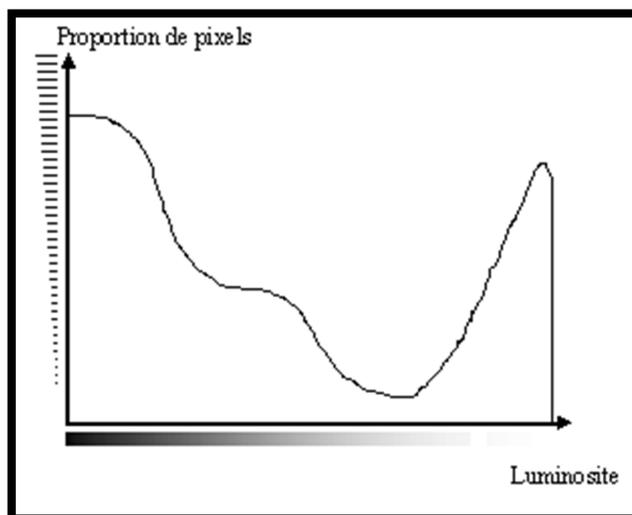


Figure 2.8: Histogramme des niveaux de gris

Seuillage

Il consiste à déterminer la valeur du seuil à partir duquel tous les pixels ayant un niveau de gris inférieur à cette valeur sont représentés par un zéro « 0 » (le noir), et tous les pixels de niveau de gris supérieur auront la valeur un « 1 » (le blanc).

Le seuil doit être calculé d'une manière adéquate en cherchant la valeur du niveau de gris qui se trouve dans la vallée entre les deux pics de l'histogramme du niveau de gris (**figure 2.9**), car les composantes du texte liées des traits fins peuvent se déconnecter, ce qui modifie la forme originale du texte [44].

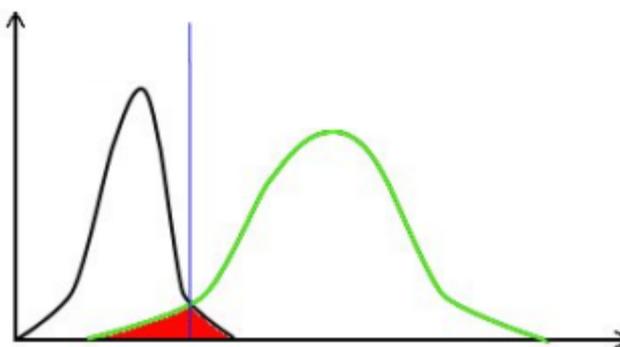


Figure 2.9 : Seuil de recouvrement

Soit l'image I (M x N), supposons que $f(x, y)$ représente le niveau de gris du pixel aux coordonnées (x, y) , $0 \leq x \leq M$, $0 \leq y \leq N$ et S est le seuil choisi, les pixels de l'objet sont ceux ayant le niveau de gris inférieur à S et les autres ayant le niveau de gris supérieur à S sont des pixels du fond. Alors, l'image binarisée g (**équation 2.7**) est déterminée par les pixels (x, y) dont la valeur est :

$$g(x, y) = \begin{cases} 1 & \text{si } f(x, y) > S \\ 0 & \text{si } f(x, y) \leq S \end{cases} \quad (2.7)$$

Il existe deux grandes techniques de binarisation (sélection du seuil s):

- Binarisation (seuillage) global
- Binarisation local

2.2.1 Seuillage global :

Le seuillage global consiste à partitionner l'image en deux classes grâce à un seuil optimal qui est calculé à partir d'une mesure globale sur toute l'image. L'histogramme est une mesure utilisée le plus souvent dans les méthodes de seuillage. Dans ce cas, le seuil attendu est celui qui correspond à la vallée de l'histogramme (**figure2.11**), celui qui distingue le plus possible les deux classes : fond et objet [43].

Dans la méthode de binarisation globale un seuil unique est calculé à partir d'une mesure globale sur toute l'image (**équation 2.8**). Il nous permet de décider l'appartenance d'un pixel à l'objet ou au fond (**figure2.10**).

$$S = \frac{1}{M \times N} \times \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} f(i, j) \quad (2.8)$$

Entrée : Image avec niveaux de gris I en format BMP ;

Sortie : Image noire et blanche (1 et 0) I' ;

Début

Copier le contenu de I dans I' ;

H = Histogramme (I') ;

\bar{X}_{histo} = seuil

Pour chaque pixel P de I' **faire**

Si $P < \bar{X}_{\text{histo}}$ alors P = 1 ; % rendre le pixel noir

Sinon p = 0 ; % rendre le pixel blanc

Fin

Figure 2.10 : Algorithme de binarisation

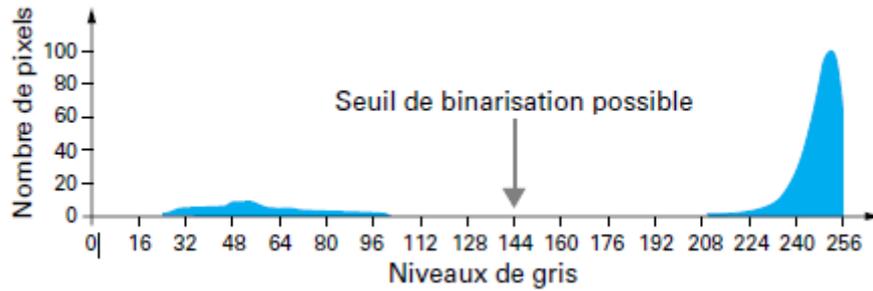


Figure 2.11: Choix du seuil sur l’histogramme de l’image à niveaux de gris [45]

La méthode décrite dans [46] peut être tenue comme représentative de cette approche (binarisation globale) ; cette méthode cherche à maximiser la variance entre deux classes.

Méthode de Otsu

Le but de cet algorithme est la binarisation d’images à niveaux de gris (**figure 2.12**). Ceci revient à séparer les pixels de l’image en deux classes, la première ayant un niveau maximal (typiquement 255) et la seconde un niveau minimal (0). Cette méthode de binarisation nécessite au préalable le calcul de l’histogramme. Puis, la séparation en deux classes est effectuée [46]. Le principe est de trouver un seuil optimal qui maximise la différence entre deux classes. Il est effectué en se basant sur la variance. Le seuil optimal (**équation 2.10**) S est celui qui maximise la fonction suivante :

$$\mu(t) = \frac{\delta_B^2(t)}{\delta_T^2}$$

$$S_{\text{optimal}} = \arg_{t \in [\min, \max]} \max \mu(t) \quad (2.9)$$

δ_T^2 (**équation 2.10**) et $\delta_B^2(t)$ (**équation 2.12**) sont respectivement la variance totale de l’image, la variance inter-classes (between-class variance)

$$\delta_T^2 = \sum_{i=0}^{\max} (i - m_T)^2 \quad (2.10)$$

$$m_T = \sum_{i=0}^{\max} i \times P_i \quad (2.11)$$

m_T : la moyenne totale des intensités de tous les points dans l’image (**équation 2.11**).

$$\delta_B^2(t) = P_{\text{fond}}(t) \times \delta_{\text{fond}}^2(t) - P_{\text{objet}}(t) \times \delta_{\text{objet}}^2(t) \quad (2.12)$$

Ou P_i : la probabilité d’occurrence du niveau de gris i dans l’image.

$$P_i = \frac{\text{nombre de pixels dont le niveau de gris } i}{\text{nombre de pixels dans l'image}} = \frac{h(i)}{M \times N} \quad (2.13)$$

$P_{fond}(t), P_{objet}(t)$ données par les deux **équations 2.14 et 2.15** représentent : la somme des probabilités d'occurrence des niveaux de gris des pixels du fond et celle de l'objet en prenant le seuil t .

$$P_{objet}(t) = \sum_{i=t}^{max} P_i \quad (2.14)$$

$$P_{fond}(t) = \sum_{i=0}^{t-1} P_i = 1 - P_{objet}(t) \quad (2.15)$$

m_{fond}, m_{objet} données par les deux **équations 2.16 et 2.17** représentent : la moyenne des pixels appartenant au fond et celle des pixels de l'objet.

$$m_{objet}(t) = \frac{\sum_{i=t}^{max} i \times P_i}{P_{objet}} \quad (2.16)$$

$$m_{fond}(t) = \frac{\sum_{i=0}^{t-1} i \times P_i}{P_{fond}} \quad (2.17)$$

$\delta_{fond}^2(t), \delta_{objet}^2(t)$ données par les deux **équations 2.18 et 2.19** représentent : la variance de la classe fond et la variance de la class objet.

$$\delta_{objet}^2(t) = \frac{\sum_{i=t}^{max} (i - m_{objet})^2 \times P_i}{P_{objet}} \quad (2.18)$$

$$\delta_{fond}^2(t) = \frac{\sum_{i=0}^{t-1} (i - m_{fond})^2 \times P_i}{P_{fond}} \quad (2.19)$$

[min, max] est l'intervalle dynamique de l'image.



(a)Image originale

(b) image binarisée

Figure 2.12: Binarisation par la méthode Otsu [45]

Cette méthode est simple à implanter et donne de bons résultats en général. Cependant, dans les cas des images de documents, les résultats ne sont pas nets, deux différents objets peuvent être confondus. Dans le cas d'une mauvaise illumination du document elle ne permet pas de trouver un seuil global (**figure 2.13**).

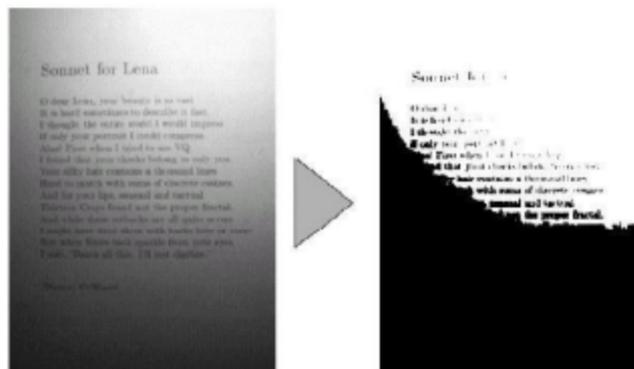


Figure 2.13: Problème de seuillage global

2.2.2 Seuillage local :

Dans ce type de seuillage, le seuil de binarisation dépend essentiellement du voisinage du pixel considéré. Donc chaque pixel doit être comparé au seuil calculé par sa région.

Le seuillage adaptatif consiste à faire varier le seuil localement, en fonction des niveaux de gris de l'information (texte) et du fond. Plusieurs solutions de seuillage adaptatif ont été proposées pour restituer correctement le texte dans tous les cas. On distingue des méthodes basées sur un découpage de l'image en régions [45]. Donc La classification d'un pixel dépend non seulement du pixel soi-même mais aussi de ses informations locales, on peut prendre comme exemple la moyenne des pixels de son voisinage qui est prise en compte lorsqu'on construit l'histogramme de deux dimensions [43].

La table suivante représente le voisinage de pixel de coordonnées (x,y) :

$f(x-1,y-1)$	$f(x-1,y)$	$f(x-1,y+1)$
$f(x,y-1)$	$f(x,y)$	$f(x,y+1)$
$f(x+1,y-1)$	$f(x+1,y)$	$f(x+1,y+1)$

Table 2.6 : Voisinage d'un pixel $f(x,y)$

Parmi les technique de binarisation locale existantes on trouve :

2.2.2.1 Methode de Bernsen

le seuil est calculé mathématiquement par l'équation **2.20** :

$$S(i, j) = \frac{\max(i, j) + \min(i, j)}{2} \quad (2.20)$$

Avec :

- $S(i, j)$: seuil à appliquer pour le point i, j
- $\text{Max}(i, j)$: valeur du niveau de gris maximal dans une fenêtre centré - en (i, j) de taille $N \times M$
- $\text{Min}(i, j)$: valeur du niveau de gris minimal dans une fenêtre centré en (i, j) de taille $N \times M$
- N et M appartenant à \mathbb{N} .

Cependant, cette technique est très sensible au bruit du fond. À cause de la prise en compte du maximum et du minimum uniquement, dans le cas où la fenêtre est uniquement sur du fond, le bruit sera interprété comme objet, car le seuil sera bas [26].

2.2.2.2 Methode de Niblack

Niblack proposa une méthode similaire sur le principe, mais prenant en compte d'autres paramètres. Dans sa méthode, le seuil est calculé ainsi :

$$S(i, j) = \mu(i, j) + k \times \sigma(i, j)$$

Avec :

- $S(i, j)$: seuil à appliquer pour le point i, j
- $\sigma(i, j)$: valeur de l'écart type dans une fenêtre centré en i, j de taille $N \times M$
- $\mu(i, j)$: valeur moyenne des niveaux de gris dans la même fenêtre
- k : constante fixée le plus généralement à 0, 2 [26]
- N et M appartenant à \mathbb{N} .

Cette méthode est victime du problème cité précédemment mais elle est plus efficace.

2.2.2.3 Methode de Sauvola

Sauvola proposa une amélioration de la méthode de Niblack afin de réduire sa sensibilité au bruit sur le fond. Le seuil est alors calculé ainsi :

$$S(i, j) = \mu(i, j) + k \times ((\sigma(i, j)R) - 1)$$

avec les mêmes notations que pour Niblack et R étant une constante permettant d'ajuster la dynamique de l'écart type. Le gain apporté par cette pondération permet à la méthode d'être très satisfaisante dans le cas de documents mal éclairés. [26]

2.2.2.4 Binarisation hiérarchique flou :

Soit Ω un ensemble de n éléments, $\Omega = \{x_1, x_2, \dots, x_n\}$. Supposons qu'on a besoin de chercher des éléments satisfaisant une propriété quelconque α . L'ensemble Ω se divise en deux sous-ensembles A et B . A contient des éléments possédant α , tant dis que les autres appartiennent au sous-ensemble B , le complément de A dans Ω . En vue de la logique classique, un élément n'appartient à qu'un sous-ensemble, A ou B . Ça veut dire qu'un élément n'a que deux possibilités, soit il a cette propriété, soit il ne la possède pas absolument. Cependant, il est possible qu'il existe dans Ω des éléments qu'on ne sait pas toujours s'ils satisfont α ou qu'ils ne la possèdent qu'avec un certain degré. Dans ce cas, il vaut mieux prendre le sous-ensemble flou pour représenter ces informations [43].

Sous-ensemble flou :

Un sous-ensemble flou A (équation 2.21) de l'espace observée Ω est caractérisé par une fonction d'appartenance $\mu_A(x)$ qui associe un élément x de Ω avec un nombre réel, $\mu_A(x)$ dans l'intervalle $[0, 1]$ et qui quantifie le degré d'appartenance de l'élément x au sous ensemble A . Généralement, un sous-ensemble flou est défini comme une collection des paires en ordre $(\mu_A(x), x)$.

$$A = \{ \mu_A(x_i) / x_i, i = 1, \dots, n \} \quad (2.21)$$

Fonction d'appartenance :

$$\mu_A : x \in \Omega \longrightarrow \mu_A(x) \in [0,1]$$

Chaque élément dans un sous-ensemble A possède un degré qui estime dans quelle mesure l'appartenance de l'élément dans A . Ce degré est déterminé par la fonction d'appartenance. Il existe des différentes fonctions. La plus simple est la fonction linéaire (figure 2.14) qui correspond à l'équation 2.22 :

$$\mu_A(x) = \begin{cases} 0 & x \leq a \\ \frac{x-a}{c-a} & a < x < c \\ 1 & x \geq c \end{cases} \quad (2.22)$$

et celle qui est la plus connue et la plus utilisée est la fonction S de Zadeh (figure 2.15) qui est donnée par l'équation 2.23 :

$$\mu_A(x) = S_Z(x) = \begin{cases} 0 & x \leq a \\ 2 \left(\frac{x-a}{c-a} \right)^2 & a < x \leq b \\ 1 - 2 \left(\frac{c-x}{c-a} \right)^2 & b < x \leq c \\ 1 & x > c \end{cases} \quad (2.23)$$

où (a, c) la région floue, b un point au milieu de a et c , $b = (a+c) / 2$.

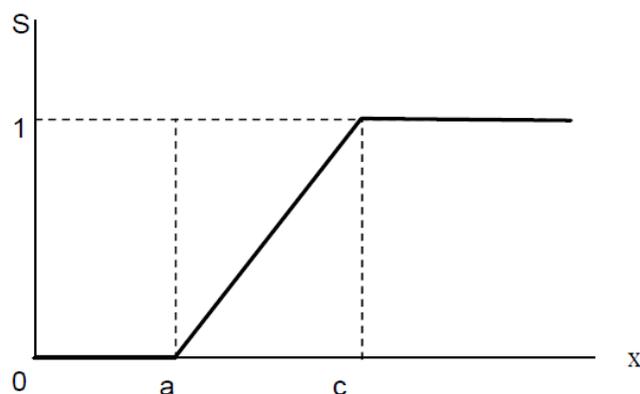


Figure 2.14: Fonction d'appartenance linéaire

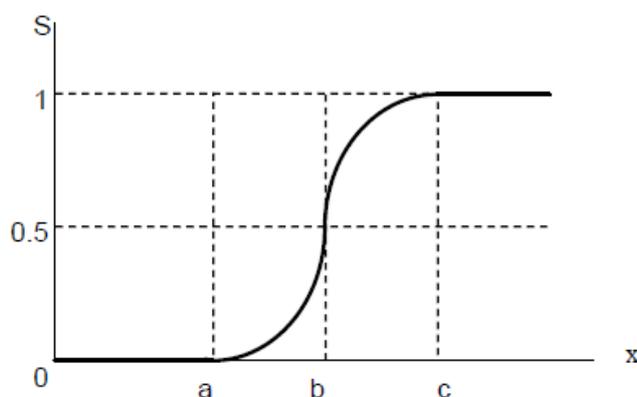


Figure 2.15: Fonction S de Zadeh

Application de la logique floue dans la binarisation d'images :

Pour le seuillage des images, le but principal est d'obtenir deux classes « blanc » et « noir » à partir de l'image originale à niveaux de gris. Cependant, il n'y a aucune assurance pour une classification grâce à un seuil quelconque. La question se pose toujours si un point est vraiment « noir » ou « blanc » ? La théorie de sous-ensembles flous est devenue une solution. En définissant un intervalle flou, la fonction d'appartenance nous permet d'obtenir deux sous ensembles flous représentant le « noir » et le « blanc » de l'image.

Décrite dans [47], cette méthode est comme une représentante de l'approche locale adaptative. Elle est basée sur un principe hiérarchique pour résoudre le problème d'éclairage non uniforme. Sous l'hypothèse que l'image ne contient que deux classes : l'objet et le fond, le principe de cette méthode est de récupérer le plus possible des pixels à la classe objet.

La méthode se compose de 2 étapes :

◆ Fuzzification :

- Construction de l'arbre quaternaire : l'image originale est divisée consécutivement en quatre sous images de taille de plus en plus petite en évaluant le critère d'homogénéité. Chaque sous-image est associée à un nœud de l'arbre quaternaire. Si une sous-image satisfait le critère d'homogénéité, la division n'est plus nécessaire, elle devient un nœud terminal dans l'arbre. Dans le 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.

La condition pour que le critère d'homogénéité soit satisfait sur une région, c'est qu'il n'y a plus de « différence significative » entre cette région et ses quatre filles. Cette condition est vérifiée par le test statistique de Fisher.

$f \leq F_{3;(4k-1)}^\alpha$: Sous image est homogène.

$f > F_{3;(4k-1)}^\alpha$: Sous image est non homogène.

f : l'estimation du critère d'homogénéité sur la sous-image évaluée.

$F_{3;(4k-1)}^\alpha$: La valeur prédéfinie de la distribution F avec 3 et $4(k-1)$ degrés de liberté. Les valeurs de la distribution F sont indiquées dans un tableau de Fisher (**Tableau 2.6**) (voir annexe A).

- Calcul des degrés d'appartenance : Les degrés d'appartenance de tous les pixels sont calculés à chaque niveau de l'arbre. L'équation 2.24 donne le degré d'appartenance μ d'un pixel (x, y) au niveau de division k de l'arbre.

$$\mu_{(x,y)}^k = S\left((x, y) ; \text{moyenne} - \text{ecart type}; \text{moyenne}; \text{moyenne} + \text{ecart type}\right) \quad (2.24)$$

où la moyenne et l'écart type sont déterminés dans la région contenant pixel (x, y) au niveau k .

◆ Défuzzification :

- Décision : Après avoir fait des différentes évaluations de l'appartenance de chaque pixel à une des deux classes, la fonction d'agrégation t-conorme de Zadeh est affectée à la détermination de la mesure d'appartenance finale à la classe objet (équation 2.25).

$$\begin{aligned} \mu_F^f(x, y) &= h\left(\mu_O^0(x, y), \mu_O^1(x, y), \dots, \mu_O^{l-1}(x, y)\right) \\ &= \max\left(\mu_O^0(x, y), \mu_O^1(x, y), \dots, \mu_O^{l-1}(x, y)\right) \end{aligned} \quad (2.25)$$

et le degré d'appartenance final au fond est :

$$\mu_F^f(x, y) = 1 - \mu_O^f(x, y)$$

- Défuzzification : il s'agit de mettre au point des pixels à deux classes.

$$\mu_O^f(x, y) > \mu_F^f(x, y) \Rightarrow (x, y) \in \text{classe objet}$$

$$\mu_O^f(x, y) \leq \mu_F^f(x, y) \Rightarrow (x, y) \in \text{classe fond}$$

L'affectation de cette méthode à l'image de la (**figure 2.6.a** ; page 36) donne le résultat représenté par la (**figure 2.16**).

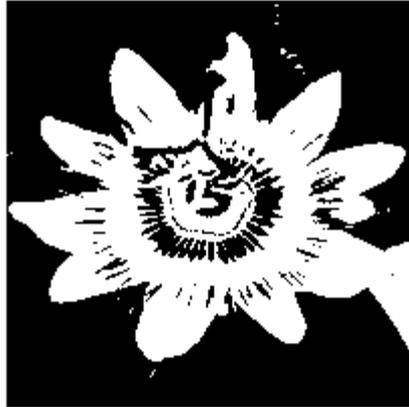


Figure 2.16: Résultat de la méthode hiérarchique flou

La binarisation réalisée dans ce travail se divise en deux étapes :

- La première utilise un seuil global calculé de la manière suivante :

$$S = \frac{1}{M \times N} \times \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} f(i, j)$$

Si $f(x, y) > S$ alors $g(x, y) = 255$.

Si $f(x, y) \leq S$ alors ; garder le niveau de gris.

- La deuxième étape est une application de la méthode hiérarchique flou sur le résultat obtenu par la première.

2.3 Etiquetage :

L'étiquetage des pixels (L'extraction des composantes connexes, procédure également appelée capture des connexités), est largement utilisé en Reconnaissance des Formes (RdF) pour segmenter les images binaires. La technique consiste à regrouper les pixels voisins dans un ensemble appelé composante connexe. Chaque ensemble est disjoint des autres et peut ensuite être aisément isolé (**figure 2.22**). La 4-connexité est distinguée de la 8-connexité suivant que le critère de voisinage comprend les 4 ou les 8 voisins d'un pixel.

Il existe deux principales méthodes pour accomplir cette tâche :

- la première est basé sur une procédure de suivi de contour : en parcourant le contour d'un objet et en revenant au point de départ, une composante connexe est délimitée, à l'exclusion cependant des contours intérieurs correspondant aux éventuels trous.

- la seconde méthode procède par une propagation d'un étiquetage des pixels lorsque l'on effectue un balayage des lignes et des colonnes de l'image [44].

Nous avons utilisé un algorithme de ce type fonctionnant en plusieurs passes (2 ou 4 passes), suivant le critère de 4-connexité.

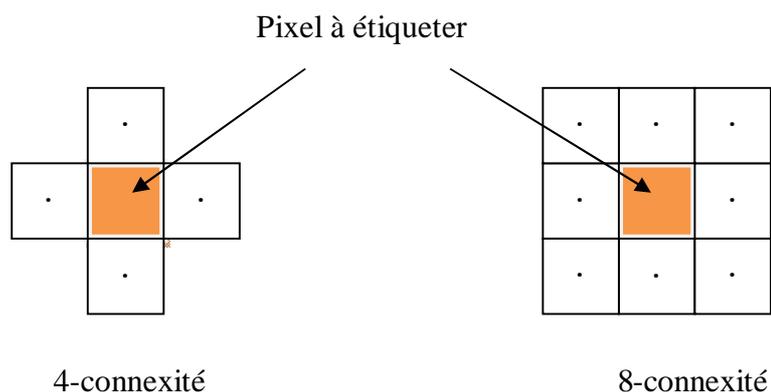


Figure 2.17: types de connexité

Algorithme : Approche par double parcours (two-pass)

Pour l'explication de cette méthode on note **B** l'image binaire à étiqueter et par **L** le résultat obtenu, c'est la matrice (ou carte) des étiquettes [48].

Premier parcours de l'image, dans le sens classique (raster-scan order) :

À chaque pixel de valeur « 1 » dans B, on affecte

- la plus petite étiquette parmi celles de ses voisins haut et gauche
ou
- une nouvelle étiquette si aucun de ces 2 voisins n'est encore étiqueté

Second parcours de l'image, dans le sens inverse :

- à chaque pixel précédemment étiqueté, on affecte la plus petite étiquette parmi la sienne et celle de ses voisins bas et droite.

Prenons l'exemple de l'image **B** de la **figure 2.20** pour détailler cet algorithme :

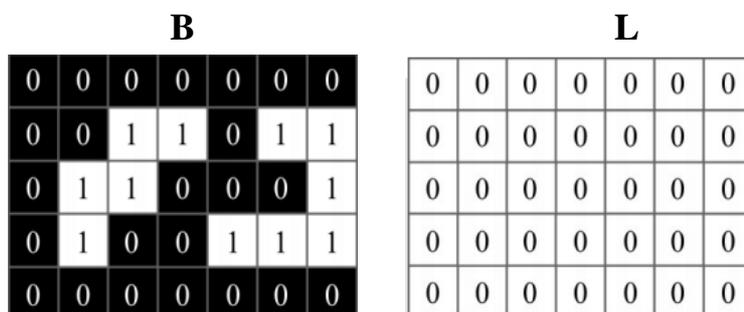


Figure 2.18 : image binaire et sa matrice initiale correspondante

1^{er} parcours :

- Initialisation nbLabels=0

<ul style="list-style-type: none"> • Étape 1 les voisins haut et gauche du premier pixel à 1 dans B ne sont pas encore étiquetés \Rightarrow nouvelle étiquette (nbLabels=1) 		
<ul style="list-style-type: none"> • Étape 2 le voisin gauche du pixel suivant à 1 dans B est déjà étiqueté à 1 \Rightarrow affecter cette même étiquette au pixel 		
<ul style="list-style-type: none"> • Étape 3 nouvelle étiquette (nbLabels=2) 		

<ul style="list-style-type: none"> Étape 4 affecter au pixel l'étiquette de son voisin gauche 	<table border="1"> <tbody> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>1</td><td>1</td></tr> <tr><td>0</td><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td></tr> <tr><td>0</td><td>1</td><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> </tbody> </table>	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	1	0	1	1	0	0	0	0	1	0	1	0	0	1	1	1	1	0	0	0	0	0	0	0	0	<table border="1"> <tbody> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>2</td><td>2</td><td>2</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> </tbody> </table>	0	0	0	0	0	0	0	0	0	0	1	1	0	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0																																																																																			
0	0	1	1	0	1	1	1																																																																																			
0	1	1	0	0	0	0	1																																																																																			
0	1	0	0	1	1	1	1																																																																																			
0	0	0	0	0	0	0	0																																																																																			
0	0	0	0	0	0	0	0																																																																																			
0	0	1	1	0	2	2	2																																																																																			
0	0	0	0	0	0	0	0																																																																																			
0	0	0	0	0	0	0	0																																																																																			
0	0	0	0	0	0	0	0																																																																																			
0	0	0	0	0	0	0	0																																																																																			
<ul style="list-style-type: none"> Étape 5 nouvelle étiquette (nbLabels=3) 	<table border="1"> <tbody> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>1</td><td>1</td></tr> <tr><td>0</td><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td></tr> <tr><td>0</td><td>1</td><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> </tbody> </table>	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	1	0	1	1	0	0	0	0	1	0	1	0	0	1	1	1	1	0	0	0	0	0	0	0	0	<table border="1"> <tbody> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>2</td><td>2</td><td>2</td></tr> <tr><td>0</td><td>3</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> </tbody> </table>	0	0	0	0	0	0	0	0	0	0	1	1	0	2	2	2	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0																																																																																			
0	0	1	1	0	1	1	1																																																																																			
0	1	1	0	0	0	0	1																																																																																			
0	1	0	0	1	1	1	1																																																																																			
0	0	0	0	0	0	0	0																																																																																			
0	0	0	0	0	0	0	0																																																																																			
0	0	1	1	0	2	2	2																																																																																			
0	3	0	0	0	0	0	0																																																																																			
0	0	0	0	0	0	0	0																																																																																			
0	0	0	0	0	0	0	0																																																																																			
0	0	0	0	0	0	0	0																																																																																			
<ul style="list-style-type: none"> Étape 6 les voisins haut et gauche du pixel suivant à 1 dans B sont étiquetés différemment ⇒ affecter au pixel l'étiquette minimale 	<table border="1"> <tbody> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>1</td><td>1</td></tr> <tr><td>0</td><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td></tr> <tr><td>0</td><td>1</td><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> </tbody> </table>	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	1	0	1	1	0	0	0	0	1	0	1	0	0	1	1	1	1	0	0	0	0	0	0	0	0	<table border="1"> <tbody> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>2</td><td>2</td><td>2</td></tr> <tr><td>0</td><td>3</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> </tbody> </table>	0	0	0	0	0	0	0	0	0	0	1	1	0	2	2	2	0	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0																																																																																			
0	0	1	1	0	1	1	1																																																																																			
0	1	1	0	0	0	0	1																																																																																			
0	1	0	0	1	1	1	1																																																																																			
0	0	0	0	0	0	0	0																																																																																			
0	0	0	0	0	0	0	0																																																																																			
0	0	1	1	0	2	2	2																																																																																			
0	3	1	0	0	0	0	0																																																																																			
0	0	0	0	0	0	0	0																																																																																			
0	0	0	0	0	0	0	0																																																																																			
0	0	0	0	0	0	0	0																																																																																			
<ul style="list-style-type: none"> Étape 7 affecter au pixel l'étiquette de son voisin haut 	<table border="1"> <tbody> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>1</td><td>1</td></tr> <tr><td>0</td><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td></tr> <tr><td>0</td><td>1</td><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> </tbody> </table>	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	1	0	1	1	0	0	0	0	1	0	1	0	0	1	1	1	1	0	0	0	0	0	0	0	0	<table border="1"> <tbody> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>2</td><td>2</td><td>2</td></tr> <tr><td>0</td><td>3</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>2</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> </tbody> </table>	0	0	0	0	0	0	0	0	0	0	1	1	0	2	2	2	0	3	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0																																																																																			
0	0	1	1	0	1	1	1																																																																																			
0	1	1	0	0	0	0	1																																																																																			
0	1	0	0	1	1	1	1																																																																																			
0	0	0	0	0	0	0	0																																																																																			
0	0	0	0	0	0	0	0																																																																																			
0	0	1	1	0	2	2	2																																																																																			
0	3	1	0	0	0	0	2																																																																																			
0	0	0	0	0	0	0	0																																																																																			
0	0	0	0	0	0	0	0																																																																																			
0	0	0	0	0	0	0	0																																																																																			
<ul style="list-style-type: none"> Étape 8 affecter au pixel l'étiquette de son voisin haut 	<table border="1"> <tbody> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>1</td><td>1</td></tr> <tr><td>0</td><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td></tr> <tr><td>0</td><td>1</td><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> </tbody> </table>	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	1	0	1	1	0	0	0	0	1	0	1	0	0	1	1	1	1	0	0	0	0	0	0	0	0	<table border="1"> <tbody> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>2</td><td>2</td><td>2</td></tr> <tr><td>0</td><td>3</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>2</td></tr> <tr><td>0</td><td>3</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> </tbody> </table>	0	0	0	0	0	0	0	0	0	0	1	1	0	2	2	2	0	3	1	0	0	0	0	2	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0																																																																																			
0	0	1	1	0	1	1	1																																																																																			
0	1	1	0	0	0	0	1																																																																																			
0	1	0	0	1	1	1	1																																																																																			
0	0	0	0	0	0	0	0																																																																																			
0	0	0	0	0	0	0	0																																																																																			
0	0	1	1	0	2	2	2																																																																																			
0	3	1	0	0	0	0	2																																																																																			
0	3	0	0	0	0	0	0																																																																																			
0	0	0	0	0	0	0	0																																																																																			
0	0	0	0	0	0	0	0																																																																																			

<ul style="list-style-type: none"> Étape 9 nouvelle étiquette (nbLabels=4) 		
<ul style="list-style-type: none"> Étape 10 affecter au pixel l'étiquette de son voisin gauche 		
<ul style="list-style-type: none"> Étape 11 les voisins haut et gauche du pixel suivant à 1 dans B sont étiquetés différemment ⇒ affecter au pixel l'étiquette minimale 		

Figure 2.19 : premier passage de l'étiquetage

2 eme parcours :

<ul style="list-style-type: none"> Image et carte initiale 		
<ul style="list-style-type: none"> 1ère étiquette modifiée Affecter au pixel l'étiquette de son voisin droit car elle est inférieure à la sienne 		

<ul style="list-style-type: none"> • 2ème étiquette modifiée Affecter au pixel l'étiquette de son voisin droit car elle est inférieure à la sienne 	<pre> 0 0 0 0 0 0 0 0 0 1 1 0 1 1 0 1 1 0 0 0 1 0 1 0 0 1 1 1 0 0 0 0 0 0 0 </pre>	<pre> 0 0 0 0 0 0 0 0 0 1 1 0 2 2 0 3 1 0 0 0 2 0 3 0 0 2 2 2 0 0 0 0 0 0 0 </pre>
<ul style="list-style-type: none"> • (étiquette non modifiée) 	<pre> 0 0 0 0 0 0 0 0 0 1 1 0 1 1 0 1 1 0 0 0 1 0 1 0 0 1 1 1 0 0 0 0 0 0 0 </pre>	<pre> 0 0 0 0 0 0 0 0 0 1 1 0 2 2 0 3 1 0 0 0 2 0 3 0 0 2 2 2 0 0 0 0 0 0 0 </pre>
<ul style="list-style-type: none"> • (étiquette non modifiée) 	<pre> 0 0 0 0 0 0 0 0 0 1 1 0 1 1 0 1 1 0 0 0 1 0 1 0 0 1 1 1 0 0 0 0 0 0 0 </pre>	<pre> 0 0 0 0 0 0 0 0 0 1 1 0 2 2 0 3 1 0 0 0 2 0 3 0 0 2 2 2 0 0 0 0 0 0 0 </pre>
<ul style="list-style-type: none"> • 3ème étiquette modifiée Affecter au pixel l'étiquette de son voisin droit car elle est inférieure à la sienne et à celle de son voisin bas 	<pre> 0 0 0 0 0 0 0 0 0 1 1 0 1 1 0 1 1 0 0 0 1 0 1 0 0 1 1 1 0 0 0 0 0 0 0 </pre>	<pre> 0 0 0 0 0 0 0 0 0 1 1 0 2 2 0 1 1 0 0 0 2 0 3 0 0 2 2 2 0 0 0 0 0 0 0 </pre>
<ul style="list-style-type: none"> • Carte à l'issue du 2^{ème} parcours Aucune autre étiquette n'est plus modifiée 	<pre> 0 0 0 0 0 0 0 0 0 1 1 0 1 1 0 1 1 0 0 0 1 0 1 0 0 1 1 1 0 0 0 0 0 0 0 </pre>	<pre> 0 0 0 0 0 0 0 0 0 1 1 0 2 2 0 1 1 0 0 0 2 0 3 0 0 2 2 2 0 0 0 0 0 0 0 </pre>

Figure 2.20: deuxièm passage de l'étiquetage

L'opération d'étiquetage (en deux passages) se répète jusqu'à ce qu'il n'y ait plus de changement sur l'image.

0	0	0	0	0	0	0
0	0	1	1	0	2	2
0	1	1	0	0	0	2
0	1	0	0	2	2	2
0	0	0	0	0	0	0

Figure 2.21 : Carte finale des étiquètes



Figure 2.22 : Exemples d'étiquetage

2.4 Operations Morphologiques sur les images binaires :

Le terme « Morphologie » est utilisé en Biologie, qui veut dire l'étude de la structure des animaux et des plantes. Le même terme peut être appliqué sur les objets d'une image dont le but est d'étudier leurs structures. Ils existent deux opérations basiques et essentielles : la dilatation et l'érosion [49].

La définition de la dilatation est de rendre un objet plus large, vaste ou encore plus ouvert.

L'érosion vise à détruire graduellement la surface d'un objet. Pour réaliser l'érosion d'une image on applique des filtres appelés SE (Structuring Element) [50].

2.4.1 Dilatation

L'opération de dilatation est effectuée par balayage de l'élément structurant SE (Structuring Element) sur l'image A manière similaire à la convolution. Elle est complètement décrite par les étapes suivantes [51] :

- 1) Si l'origine de l'élément structurant (centre de la fenêtre SE) coïncide avec un pixel « blanc », pas de changement et déplacer le SE vers le prochain pixel.
- 2) Si l'origine de l'élément structurant (SE) coïncide avec un pixel « noir » dans l'image, alors tous les pixels voisins appartenant à la fenêtre SE prennent la valeur « noir ».

La dilatation d'une image A par un élément structurant B est symbolisée par : $A \oplus B$.

Un élément structurant SE peut avoir plusieurs formes (**figure 2.23**). Les formes les plus utilisées sont (carré, disque, ligne horizontale ou verticale,...), la figure suivante :

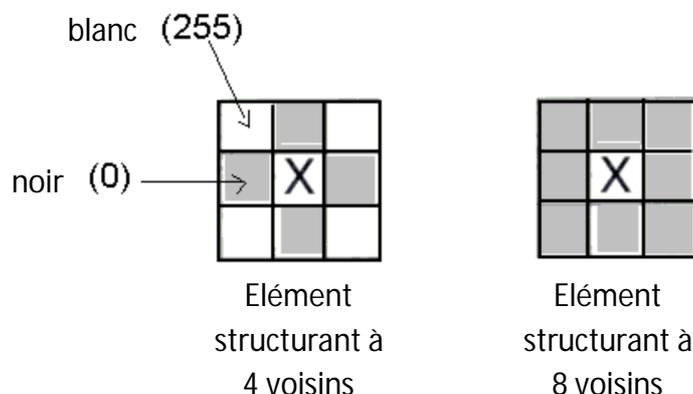


Figure 2.23 : Formes de l'élément structurant

Un exemple de dilatation d'une forme quelconque est présenté sur la (**figure 2.24**). On note qu'après l'opération de dilatation tous les pixels « noirs » de l'image originale sont conservés. Elle provoque l'expansion de tous les contours et aussi le remplissage des petits trous (**figure 2.25**) de l'image originale [51].

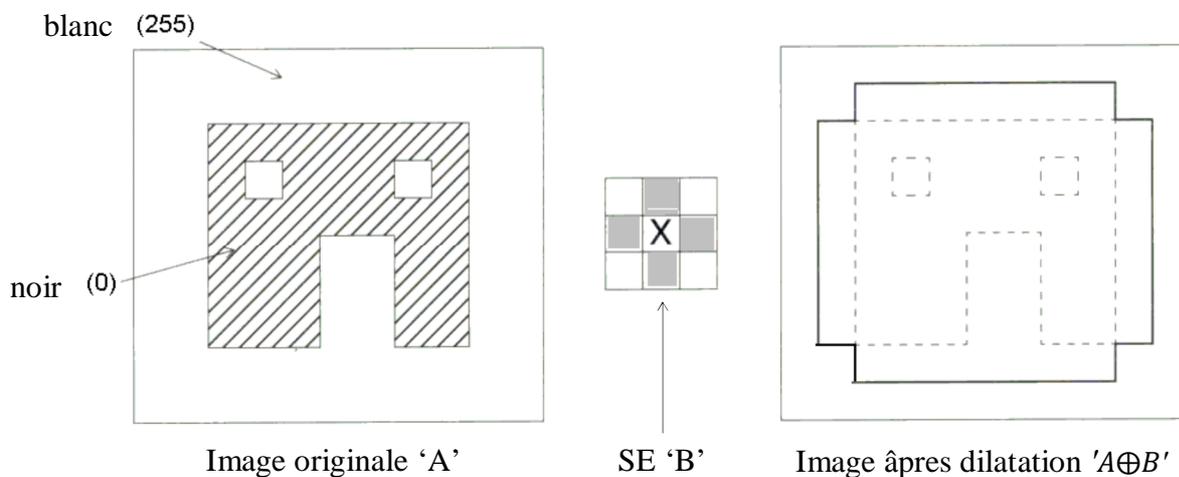


Figure 2.24 : Illustration de processus de dilatation



Image originale

image après dilatation

Figure 2.25 : Exemple de dilatation d'image

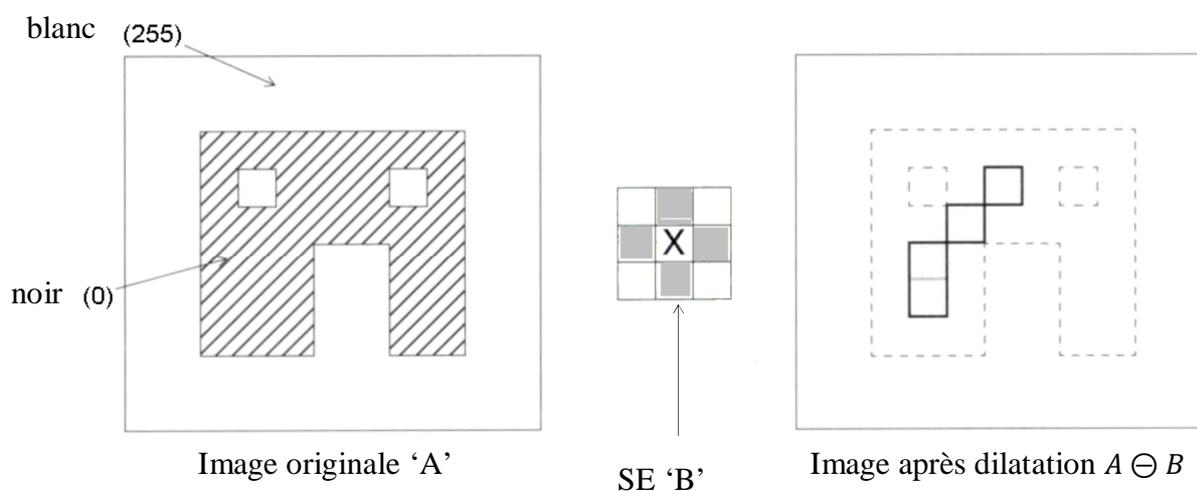
2.4.2 Erosion

Cette opération est similaire à la dilatation. Les pixels sont transformés en « blanc » au lieu de « noir ». Donc balayer l'image par le SE et suivre les étapes [38] :

- 1) Si l'origine de l'élément structurant (centre de la fenêtre SE) coïncide avec un pixel « blanc », pas de changement et déplacer le SE vers le prochain pixel.
- 2) Si l'origine de l'élément structurant (SE) coïncide avec un pixel « noir » dans l'image, et il existe au moins un seul pixel blanc dans la fenêtre SE alors le pixel prend la valeur « blanc ».

L'érosion d'une image A par un élément structurant B est symbolisée par : $A \ominus B$.

Un exemple d'érosion d'une forme quelconque est présenté par la **figure 2.26**.

**Figure 2.26 : Illustration de processus de l'érosion**

Le résultat de ce traitement revient à transformer en pixels du fond les pixels objets en contact direct avec des pixels du fond. Cela revient à "raboter" d'une couche de pixels les masques des objets [52]. Un exemple d'érosion d'une image est illustré sur la (figure 2.27).

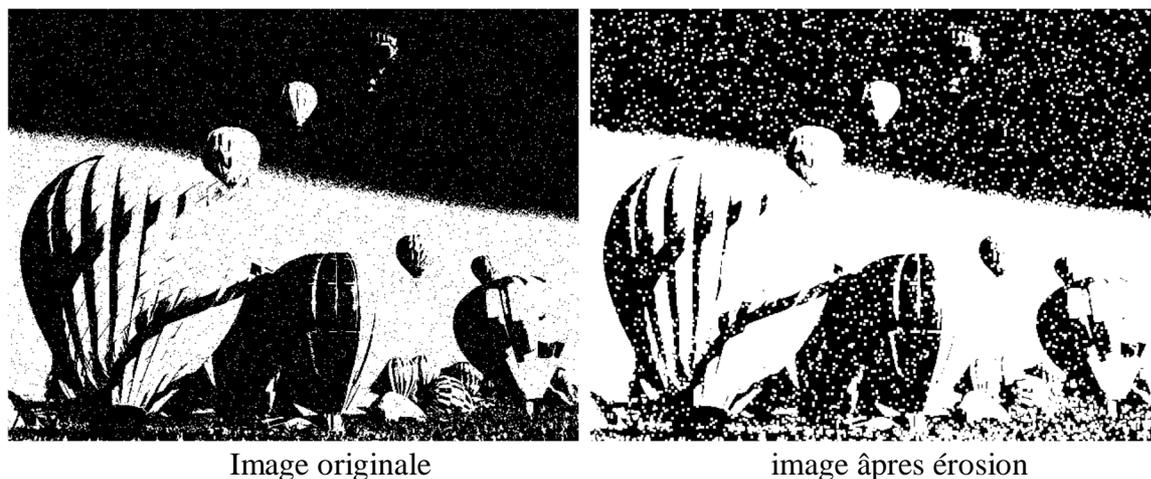


Figure 2.27 : Exemple d'érosion d'image

La combinaison des deux opérations basiques dilatation et érosion donne d'autres opérations plus complexes. Les combinaisons les plus utilisées en filtrage morphologique sont l'ouverture et la fermeture de l'image.

2.4.3 Ouverture

L'ouverture consiste à appliquer à l'image binaire une érosion d'une profondeur donnée suivie d'une dilatation de profondeur identique. L'érosion va permettre d'éliminer les petites structures isolées. Par exemple, sur la **figure 2.28**, l'érosion permet de gommer les petits objets et ne laisser subsister que le masque du noyau. La dilatation permet de retrouver les dimensions originales du noyau [37].

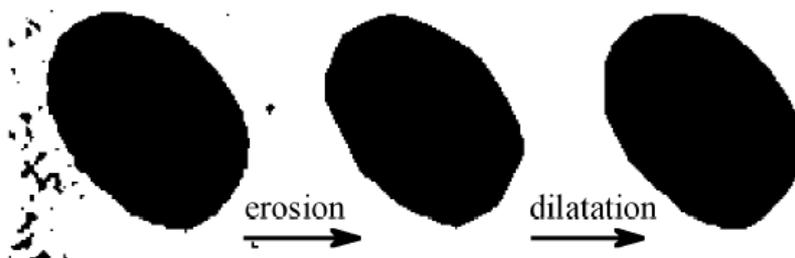


Figure 2.28 : Ouverture morphologique

2.4.4 Fermeture

La fermeture consiste à appliquer à l'image binaire une dilatation d'une profondeur donnée suivie d'une érosion de profondeur identique. La dilatation va permettre de boucher les trous. Par exemple, sur la **figure 2.29**, la dilatation comble tous les trous dans le masque du noyau. Cependant on constate que le noyau a maintenant une forme empâtée et surdimensionnée. L'érosion qui suit permet de compenser cet épaissement. On constate que le masque du noyau a retrouvé ses dimensions originales alors que son contour apparaît moins échancré. La fermeture a aussi effectué un lissage du contour [52].



Figure 2.29 : Fermeture morphologique

Discussion

Dans ce chapitre, nous avons passé en revue quelques techniques de traitement d'images de documents en présentant des techniques de prétraitement (filtrage, binarisation), et d'autres techniques de traitement de haut niveau comme l'étiquetage et les opérations morphologiques (dilatation, érosion, fermeture, ouverture, réduction).

En raison de la multitude de sources de bruits, et surtout de la multitude d'effets de ces bruits sur une image, il n'existe pas de technique de restauration générale d'image adaptée à toutes les situations. Il existe de nombreuses recherches permettant d'approcher ce traitement automatisé du bruit mais, en général, ces techniques cherchent à adapter les traitements, en fonction d'estimations calculées à partir de l'image. La méthode de segmentation proposée est basée sur les différentes techniques de traitement

Dans la section suivante nous présentons notre méthode proposée pour la segmentation d'image de document qui utilise les différentes techniques de traitement citées dans ce chapitre, cette méthode comprend la détection des zones d'intérêt (ou blocs) et leur classification.

CHAPITRE 3

PRETRAITEMENT ET SEGMENTATION

Dans l'analyse et la reconnaissance de documents (ARD), le prétraitement est toujours une des premières étapes utilisées. Elle a donc une grande influence sur la performance des étapes suivantes et sur le résultat final. L'image brute est une structure simple dans son organisation mais complexe dans son contenu. Ceci est dû à plusieurs facteurs, d'une part la grande quantité d'information qui y réside et d'autre part le nombre important de processus indépendants de l'utilisateur qui tendent à y introduire des distorsions indésirables, donc au dégrader.

L'image acquise par un capteur est toujours accompagnée de parasites [53], Parmi les éléments de perturbation on trouve :

- L'éclairage qui peut engendrer des ombres, des reflets, peut être très puissant et saturer la caméra ou bien trop faible et rendre l'image trop sombre.
- La qualité optique de l'objectif : aberration, astigmatisme, etc...., et celle de prise de vue : mise au point.
- Le bruit électronique inhérent au capteur qui engendre une imprécision dans la restitution des niveaux de gris.

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

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

Dans cette section nous détaillons notre méthode, nous présentons dans l'ordre les stratégies de prétraitement et préparation, puis les techniques appliquées en matière de la segmentation.

3.1 Aperçu de la méthode générale

La méthode de segmentation que nous avons développée est composée de trois phases : le prétraitement, l'initialisation de la segmentation, et la finalisation de segmentation, comme le montre la **figure 3.1**. La première phase consiste à filtrer le bruit dans l'image et de passer de l'image à niveau de gris à une image à deux niveaux seulement (fond ou objet). Pour effectuer la deuxième phase, on implémente une méthode de classification descendante, c'est une analyse d'image par des opérations morphologiques. Cette deuxième étape permet d'effectuer une segmentation initiale. Dans la troisième étape on procède l'image par une segmentation plus précise, en utilisant un lissage horizontal et vertical des composantes suivi par un étiquetage, cette dernière phase corrige aussi la segmentation initiale, ce qui renforce la robustesse de la méthode.

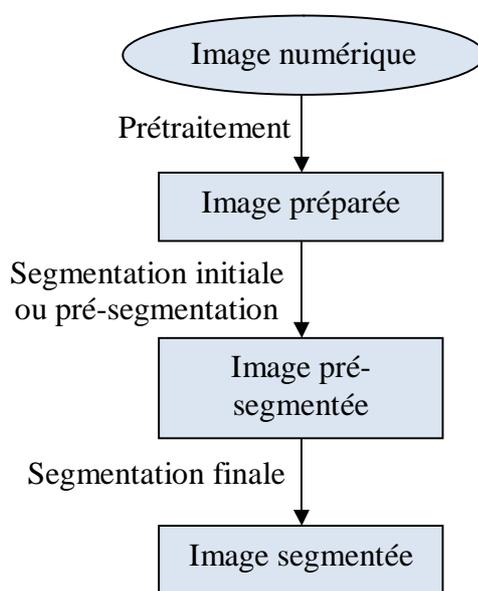


Figure 3.1 : Méthode adoptée

3.2 Prétraitement :

Le rôle du prétraitement dans sa définition la plus générale est de remédier aux dégradations ayant affecté l'image et de rendre cette image mieux adaptée à une application particulière. Donc le prétraitement consiste à améliorer la qualité de l'image en éliminant les défauts dus à l'éclairage et au processus d'acquisition [54].

L'étape suivante est de réaliser un traitement de l'image capturée à préparer pour une analyse ultérieure. Ce traitement comprend deux parties [55] :

- Filtrage : la réduction de bruit
- Binarisation : réduction des niveaux de gris ou transformation d'une image à niveaux de gris en une image binaire.

Comme illustré par la figure suivante.

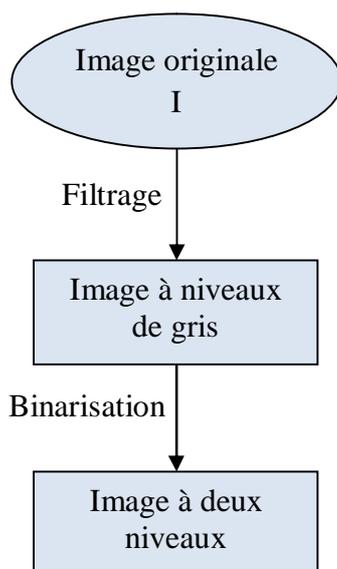


Figure 3.2 : Aperçu de la méthode de prétraitement

3.2.1 Filtrage :

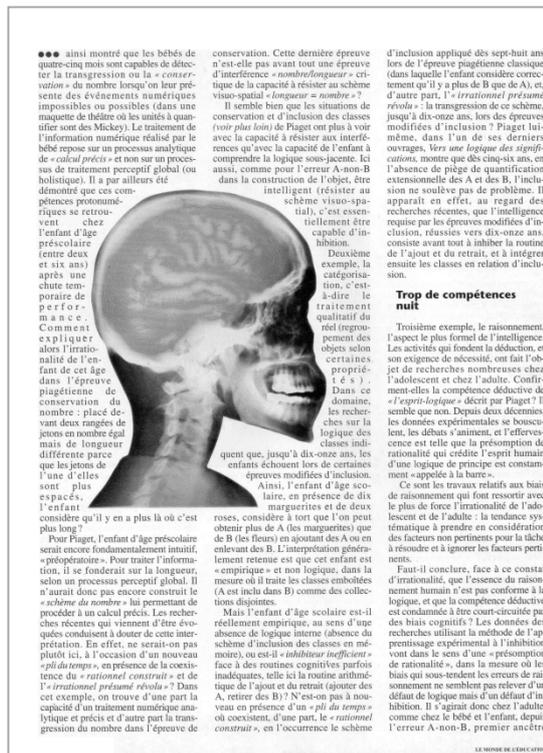
Le principe du filtrage est de modifier la valeur des pixels d'une image, généralement dans le but d'améliorer son aspect. En pratique, il s'agit de créer une nouvelle image en se servant des valeurs des pixels de l'image d'origine.

Le type de filtrage utilisé est Gaussien, il permet de réduire le bruit et offre un bon lissage, avec la conservation des contours, la **figure3.3** illustre l'application de cette procédure sur une image de document et la comparaison de son résultat avec les résultats obtenus par filtrage médian et filtrage moyen. Le filtrage moyen est très simple à implémenter et rapide mais peut être robuste, en comparant la **figure3.3.b** et la **figure3.3.d** nous remarquons que le filtrage gaussien élimine bien le bruit présent sur la **figure3.3.a** 'fond en bleu en haut de l'image' par contre le filtrage moyen garde la couleur du fond.

Le filtrage médian (**figure3.3.c**) élimine la couleur du fond mais il rabote le texte.



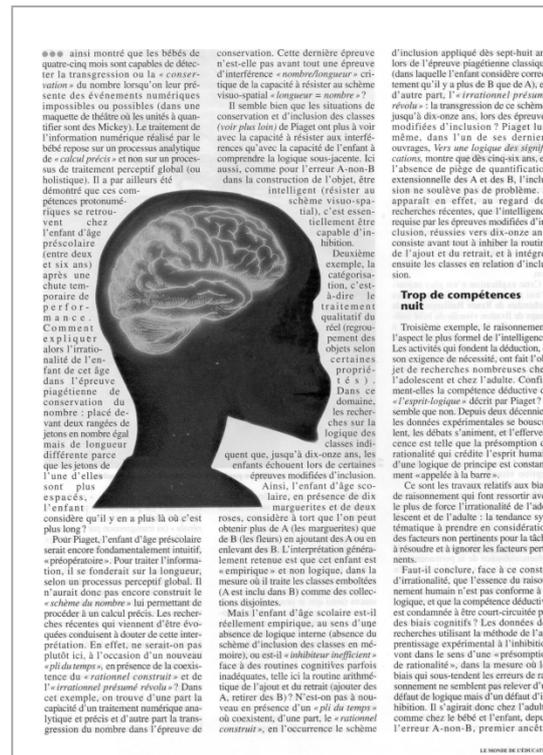
(a) Image originale



(b) filtrage gaussien



(c) filtrage médian



(d) filtrage moyen

Figure 3.3 : Filtrage d'image

3.2.2 Binarisation :

L'histogramme de l'image de document contient deux modes : un fort correspondant au fond et un faible correspondant à l'objet. Cependant ce qui est important est celui de l'objet. La frontière entre l'objet et le fond n'est pas toujours claire surtout dans les zones où les caractères et les objets non textuels sont proches. Donc, obtenir l'objet dont les composants sont clairs et nets est le but final de notre méthode.

La méthode proposée peut être considérée comme la combinaison de l'approche globale et l'approche locale. Elle se compose de deux étapes. Utilisant la technique de seuillage global, la première étape a pour but d'éliminer la plupart du fond qui domine l'image observée et de garder la partie importante contenant l'objet. La deuxième étape consiste à raffiner le résultat de l'étape précédente pour rendre l'objet plus net.

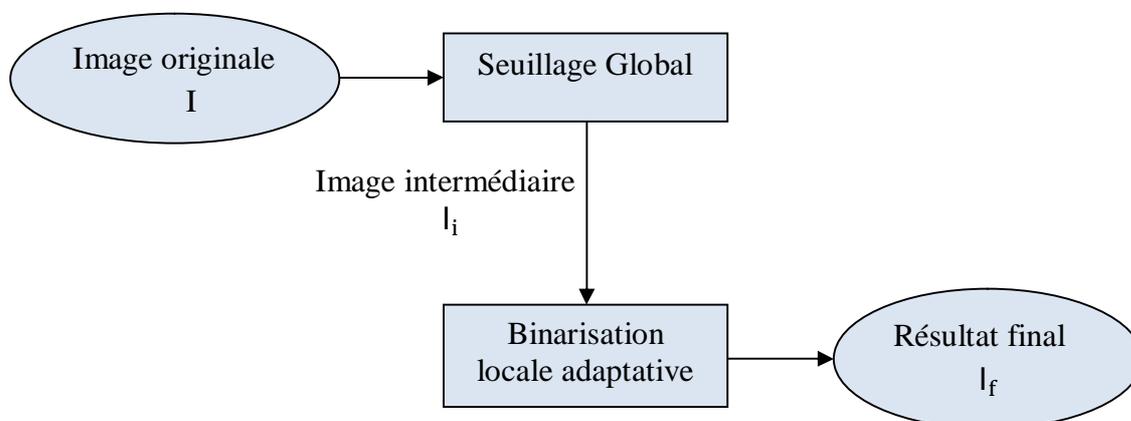


Figure 3.4 : Binarisation adoptée

Dans les parties ci-dessous, on va prendre ces notations suivantes :

$g(x, y)$: le niveau de gris du pixel (x, y) de l'image originale I .

$g_i(x, y)$: le niveau de gris du pixel (x, y) de l'image intermédiaire I_i .

$g_f(x, y)$: le niveau de gris du pixel (x, y) du résultat final I_f .

3.2.2.1 Etape de seuillage global

Une méthode de seuillage global nous aide à chercher un seuil pour toute l'image. Dans ce cas, nous avons choisi comme solution possible, d'utiliser un seuil global représenté par la moyenne des intensités des pixels de l'image originale I (**équation 3.1**).

$$S = \frac{1}{M \times N} \times \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} g(x, y) \quad (3.1)$$

Au lieu de mettre le résultat de cette étape comme une image noire et blanche, on va garder la valeur originale des pixels appartenant à l'objet pour obtenir une image intermédiaire (**équation 3.2**) (**figure 3.5.b**). Si $g_i(x, y)$ est l'intensité lumineuse du pixel (x, y) de cette image, alors :

$$g_i(x, y) = \begin{cases} 255 & \text{si } g(x, y) > S \\ g(x, y) & \text{si } g(x, y) \leq S \end{cases} \quad (3.2)$$

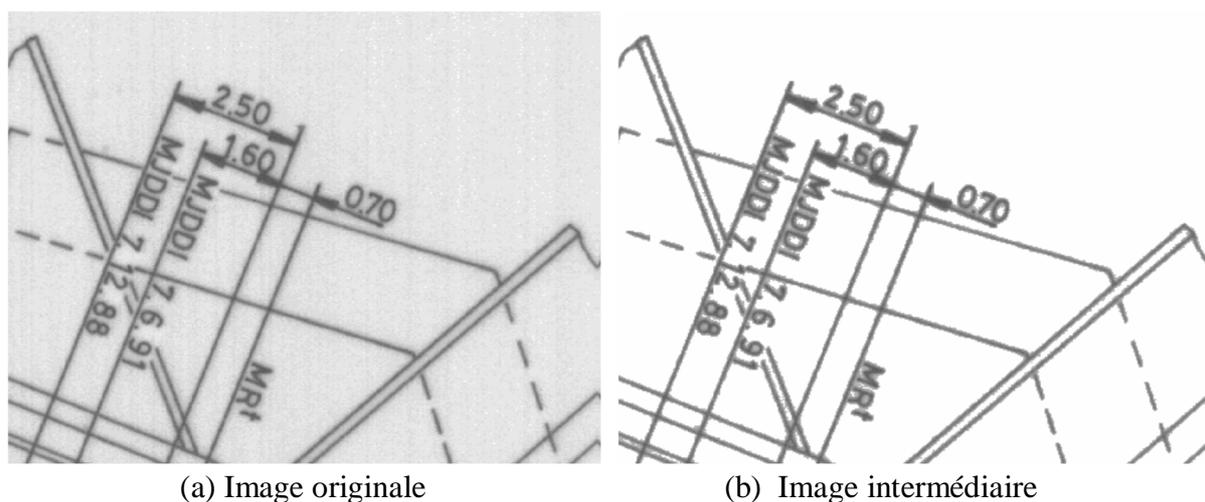


Figure 3.5: première étape de binarisation

3.2.2.2 Etape de raffinement

L'image obtenue après la première étape a bien gardé la partie qui nous intéresse. Cependant, l'objet n'est pas vraiment net, les parties différentes de l'objet ne sont pas clairement distinguées car une minorité des pixels qui auraient dû appartenir au fond sont attribués à l'objet. En général, ce sont des pixels aux frontières objet-fond. C'est pourquoi on a besoin d'un autre traitement pour éliminer ces pixels. Dans cette deuxième étape, l'opération n'est effectuée que sur l'objet obtenu à partir de la première étape, c.à.d, on manipule sur l'image intermédiaire I_i mais sans compter les pixels du fond (ceux dont le niveau de gris est égal à 255).

Si l'on essaie de chercher un seuil à effectuer globalement sur I_i , il risque de perdre des parties de l'objet dont les intensités sont moins fortes que ceux des autres. Cela vient du fait que l'illumination n'est pas forcément constante sur l'image. C'est la raison pour laquelle on doit chercher une méthode de seuillage adaptatif qui permet de tenir compte des informations locales pour diminuer l'effet ci-dessus. Cette méthode est basée sur le principe de l'arbre quaternaire et la théorie de sous-ensembles flous. L'image à traiter sera décomposée de plus en plus en sous-images de taille petite en évaluant le critère d'homogénéité. Une image dont ce critère n'est pas satisfait sera divisée en 4 sous images.

Le processus appliqué, afin de réaffecter un pixel qui est déjà classé comme l'objet dans I_i à la classe fond ou à la classe objet, se compose de 3 sous-étapes :

- Construction de l'arbre quaternaire.
- Calcul de degrés d'appartenance de chaque pixel à chaque niveau de l'arbre.
- Décision de degré final d'un pixel pour le classer au fond ou à l'objet.

Soit O_{R_i} l'ensemble des pixels portant la valeur originale d'une région rectangle (**équation 3.3**) (une sous-image) quelconque R_i de l'image intermédiaire I_i . On peut considérer R_i comme un nœud de l'arbre et R_0 comme la racine I_i .

$$O_{R_i} = \{(x, y) \in R_i | g_i(x, y) \neq 255\}, R_i \subset I_i \quad (3.3)$$

Parce qu'on ne s'intéresse que sur l'ensemble O_{R_i} , toutes les notions et les formules concernant la région R_i ne sont appliquées que sur les pixels dans O_{R_i} .

Construction de l'arbre quaternaire

La hiérarchie associée à l'image I_i de taille $M \times N$ est construite en divisant successivement cette image en sous-images de taille de plus en plus petite.

- L'image I_i est prise comme la racine de l'arbre qu'on va construire. Elle correspond à un nœud au niveau 0 (**figure 3.6**).
- Les nœuds au niveau k sont créés par des nœuds décomposables au niveau $k-1$. Les nœuds décomposables sont ceux qui ne satisfont pas le critère d'homogénéité. Un nœud décomposable au niveau k est divisé en 4 nœuds au niveau $k+1$. Ceux qui ne sont pas décomposables représentent des nœuds terminaux (des feuilles) de l'arbre. Ce processus est répété jusqu'à ce qu'il n'y a plus de nœuds décomposables.

Quand le processus de subdivision s'arrête, l'image originale est représentée par des nœuds terminaux.

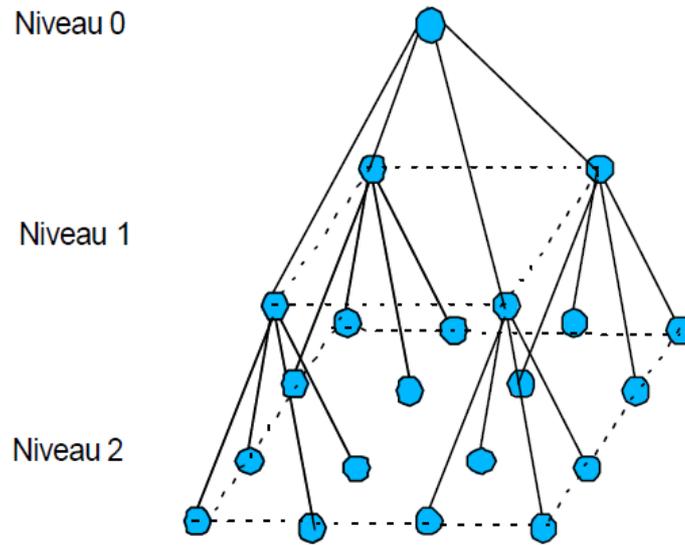


Figure 3.6 : Structure tridimensionnelle de l'arbre quaternaire

Le critère d'arrêt est le suivant :

On ne décompose jamais une région R_i dont tous les pixels sont déjà classés comme un fond, c'est à dire que l'ensemble O_{R_i} est vide. La région R_i sera représentée par un nœud terminal. Pour d'autres régions, la décision dépend de la relation entre la région et ses quatre filles correspondantes (**équation 3.4**). En principe, on ne découpe plus une région s'il n'y a pas de différence significative entre la moyenne de la région mère et celles de ses quatre filles, ainsi que entre leurs variances. Alors, on doit prédéfinir un seuil afin de définir la « différence significative ». Pour éviter le problème de choisir le seuil, un test statistique de Fisher est utilisé pour vérifier le critère d'arrêt [56]. Ce test nous permet de comparer les moyennes et les écarts-types entre la mère R_i et les quatre filles $R_{i1}, R_{i2}, R_{i3}, R_{i4}$.

$$\begin{cases} \text{Hypothèse null } \sigma_1 = \sigma_2 = \sigma_3 = \sigma_4 = \sigma \text{ et } m_1 = m_2 = m_3 = m_4 = m \\ \text{Hypothèse alternative } \exists j \in \{1, 2, 3, 4\} \mid \sigma_j \neq \sigma \text{ ou } m_j \neq m \end{cases} \quad (3.4)$$

Où

$\sigma_j, j \in \{1, 2, 3, 4\}$ et σ sont successivement les écarts-types calculés sur les données de 4 filles et de la mère.

$m_j, j \in \{1, 2, 3, 4\}$ et m sont les moyennes correspondantes.

$\sigma_j, m_j, j \in \{1, 2, 3, 4\}$ et σ et m sont calculées sur $O_{R_{ij}}, j \in \{1, 2, 3, 4\}$ et O_{R_i}

Supposons que les quatre sous-images filles de l'image mère sont indépendantes et présentent des distributions des niveaux de gris normales et identiques, le test d'homogénéité f de Fisher (**équation 3.5**) a une distribution $F_{p; n-p-1}^\alpha$.

$$f = \frac{k \sum_{j=1}^4 \frac{(m_j - m)^2}{3}}{\sum_{j=1}^4 \sum_{k=1}^K \frac{(X_{jk} - m_j)^2}{4 * (K - 1)}} \quad (3.5)$$

K : le nombre de pixels dans chaque sous image

X_{jk} : le niveau de gris du kème pixel de la sous-image j .

p : le degré de liberté, dans ce cas, $p = 3 = \text{le nombre de sous ensemble} - 1$

n : le nombre total de pixels de l'image mère = $4K$

a : le niveau de confiance (*confidence level*)

La décision d'homogénéité d'une région dépend de la comparaison f et $F_{p; n-p-1}^{\alpha}$

$f \leq F_{p; n-p-1}^{\alpha}$: L'hypothèse H_0 est « vrai ». La région est homogène

$f > F_{p; n-p-1}^{\alpha}$: L'hypothèse H_1 est « vrai ». La région est hétérogène.

Il est bien évident qu'un test statistique n'apporte de signification que si la taille de l'échantillon est suffisamment grande. Il faut donc déterminer la taille minimale de l'échantillon pour appliquer le test. Il nous aide à éviter le problème de sur-découpage.

En bref, la décomposition d'une région R_i s'arrête si une des deux conditions suivantes est satisfaite. R_i deviendra un nœud terminal.

1) $\text{Card}(O_{R_i}) \leq \text{taille}_{\min}$ ou bien

2) $\text{Card}(O_{R_i}) > \text{taille}_{\min}$ et $f \leq F_{3; n-2}^{\alpha}$

$\text{Card}(O_{R_i})$: la cardinalité de l'ensemble O_{R_i}

taille_{\min} : La taille minimale de l'échantillon pour l'application du test statistique.

A l'implémentation, nous avons choisi 40 comme une valeur expérimentale de taille_{\min} . Cette valeur représente la moyenne entre le nombre de pixels des espaces inter-caractère et le nombre de pixels des trous se trouvant à l'intérieur du texte.



Figure 3.7 : Quadrillage de l'image intermédiaire

Calcul des degrés d'appartenance de chaque pixel

À chaque nœud de l'arbre, si la région correspondante n'est pas homogène, la théorie de l'ensemble flou sera appliquée pour la classification de ses données en deux sous-ensembles flous F (fond) et O (objet) en évaluant leurs degrés d'appartenance. Cela signifie que ces

degrés d'appartenance à la classe objet $\mu_O^k(x, y)$ et à la classe fond $\mu_F^k(x, y)$ de chaque pixel sont calculés pour chaque niveau k de l'arbre.

Etant une fonction la plus souvent utilisée, la fonction S de Zadeh est prise pour calculer le degré d'appartenance à la classe fond d'un pixel. Supposons que $\mu_F^k(x, y)$ et $\mu_O^k(x, y)$ sont successivement le degré d'appartenance à la classe fond et celui à la classe objet du pixel (x, y) ayant le niveau de gris $g_i(x, y)$, ils sont déterminés par les deux **équations** (3.6) et (3.7) :

$$\mu_F(x, y) = S_Z(g_i(x, y), a, b, c) = \begin{cases} 0 & g_i(x, y) \leq a \\ 2 \left(\frac{g_i(x, y) - a}{c - a} \right)^2 & a < g_i(x, y) \leq b \\ 1 - 2 \left(\frac{c - g_i(x, y)}{c - a} \right)^2 & b < g_i(x, y) \leq c \\ 1 & g_i(x, y) > c \end{cases} \quad (3.6)$$

$$\mu_O(x, y) = 1 - S_Z(g_i(x, y), a, b, c) \quad (3.7)$$

Avec : $b = (a + c)/2$

Pour les estimations des paramètres a, b, c, on prend des propriétés locales des régions (des nœuds). Sur l'intervalle dynamique de la région, l'intervalle (moyenne – écart-type, moyenne + écart-type) est considéré comme la bande d'incertitude, b est la moyenne de cette bande.

Les degrés d'appartenance d'un pixel (x, y) au niveau k sont déterminés par :

$$\mu_F^k(x, y) = S_Z(g_R(x, y), m - \sigma, m, m + \sigma) ; \mu_O^k(x, y) = 1 - \mu_F^k(x, y) \quad (3.8)$$

m et σ sont la moyenne et l'écart-type calculés sur une sous-image correspondante à un nœud au niveau k .

Décision de degré d'appartenance final

Après avoir calculé les degrés d'appartenance d'un pixel à tous les niveaux, il faut prendre une décision : parmi eux quelle est la valeur qui va décider la classification du pixel ? c'est donné par les **équations 3.9** et **3.10**.

$$\begin{aligned} \mu_O^f(x, y) &= h(\mu_O^0(x, y), \mu_O^1(x, y), \dots, \dots, \mu_O^l(x, y)) \\ &= \max(\mu_O^0(x, y), \mu_O^1(x, y), \dots, \dots, \mu_O^l(x, y)) \end{aligned} \quad (3.9)$$

$$\mu_F^f(x, y) = 1 - \mu_O^f(x, y) \quad (3.10)$$

Alors, un pixel (x, y) va appartenir à la classe fond F si $\mu_F^f(x, y) > \mu_O^f(x, y)$, si non il est mis comme un pixel de l'objet.

Le résultat final I_f est:

$$I_f = \{(x, y), x = 0, \dots, M - 1; y = 0, \dots, N - 1; g_f(x, y) \in \{0, 255\}\}$$

$$g_f(x, y) = \begin{cases} 0 & \mu_F^f(x, y) \leq \mu_O^f(x, y) \\ 255 & \mu_F^f(x, y) > \mu_O^f(x, y) \end{cases}$$

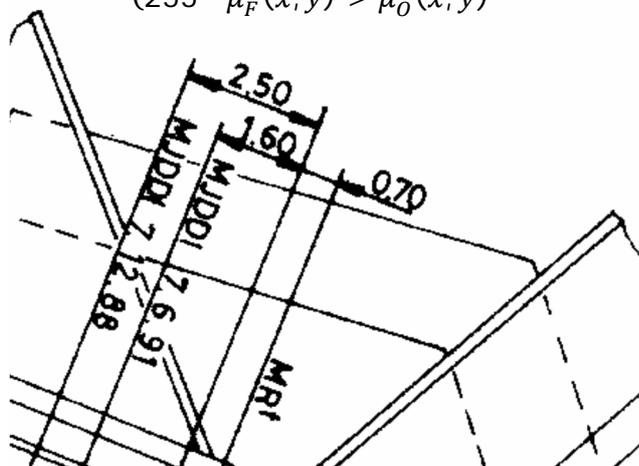


Figure 3.8 : Résultat final de binarisation de l'image de la figure 3.5.a

3.3 Segmentation :

Comme il a été mentionné précédemment, la reconnaissance de la structure physique d'un document est de segmenter l'image de ce document en blocs, la segmentation consiste à localiser les régions du texte et les régions non textuelles, et les délimiter par des rectangles circonscrits [45] (figure 3.9) ou bien par leurs contours (figure 3.10).

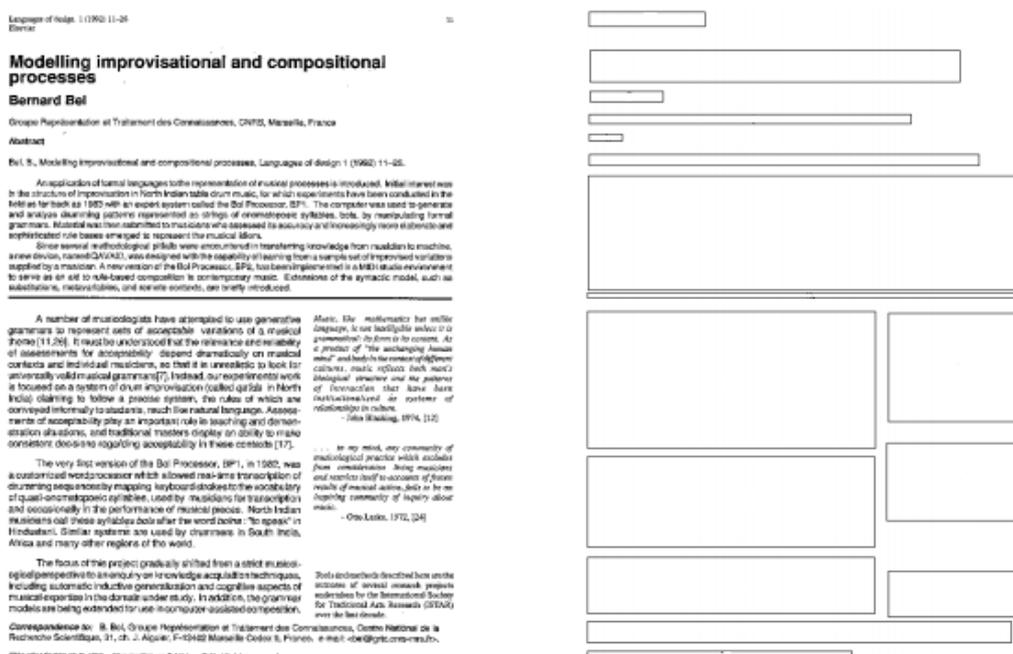


Figure 3.9 : délimitation par des rectangles

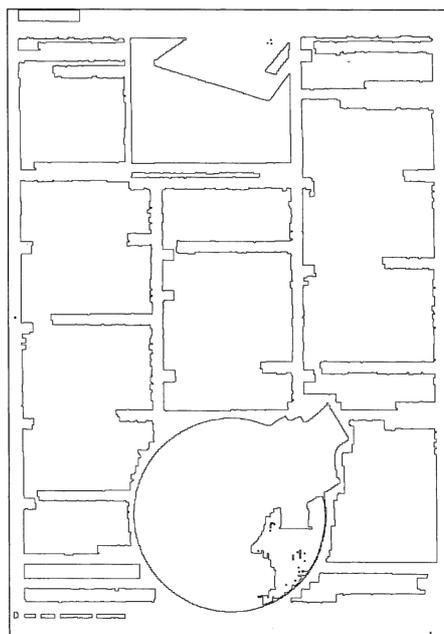


Figure 3.10 : Délimitation par des contours

La reconnaissance de la structure physique des documents est la partie la plus compliquée, car elle présente une richesse et une variété au niveau des composantes. Notre étude est consacrée à résoudre ce problème dans le cas des documents à structure complexe [53].

La (figure 3.11) illustre la démarche proposé pour la segmentation de documents:

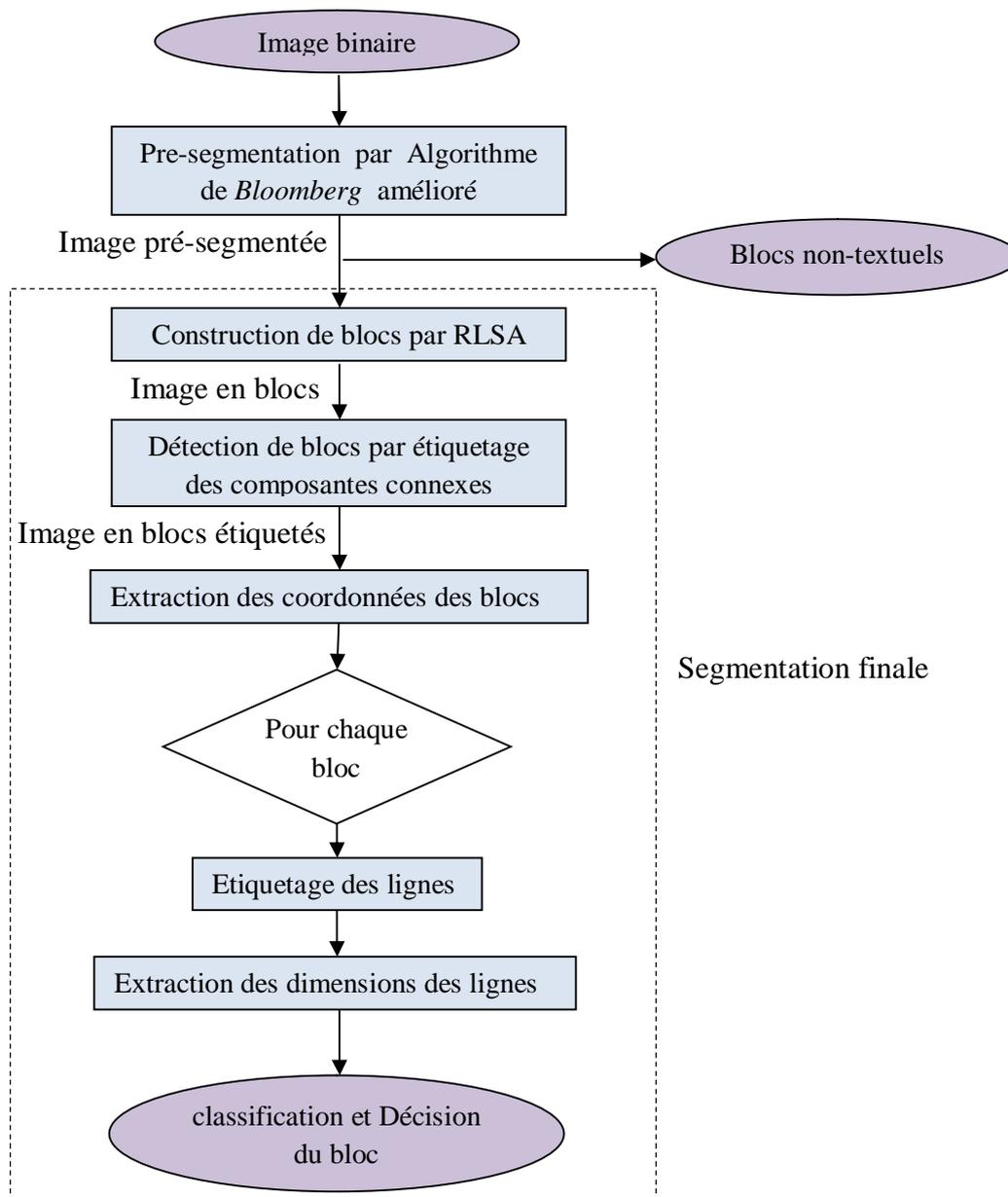


Figure 3.11 : Méthode de segmentation utilisée

3.3.1 Pré-segmentation (Application de l'algorithme de Bloomberg amélioré)

L'algorithme de Bloomberg est un algorithme de segmentation d'images de documents en blocs texte/image utilisant les opérations morphologique de base. L'image peut être fermée par un élément structurant (ES) dont le but de solidifier les composante non textuelles comme elle peut être ouverte pour expulser les parties du texte et conserver les composantes des parties non-textuelles. Les portions résiduelles ou les grains des parties non-texte peuvent utiliser pour régénérer ces composantes [23]. Les étapes à suivre pour réaliser cet algorithme sont organisées dans le diagramme de la **figure 3.12**.

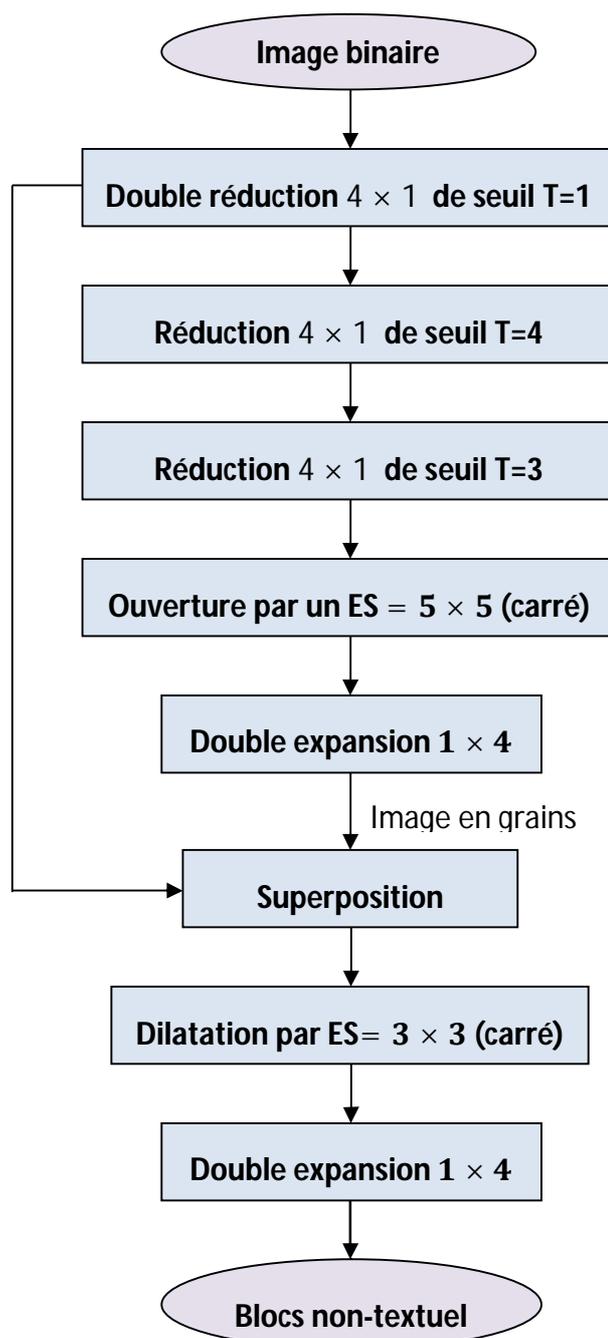


Figure 3.12 : Algorithme de Bloomberg[23]

En partant d'une image de document composite binaire, nous appliquons le "Bloomberg's Algorithm" qui consiste à faire une double réduction 4×1 avec seuillage de seuils $T_1 = T_2 = 1$ Comme le montre la **figure 3.13.b**, ensuite de la même façon une réduction 4×1 avec seuil $T_3 = 4$ suivie d'une quatrième réduction de dimension 4×1 toujours mais avec seuil $T_4 = 3$.

Sur la figure suivante on donne un exemple de double réduction de limage (a) avec seuils $T_1 = T_2 = 1$ (on peut dire aussi que c'est une seule réduction (16×1). L'image résultante (b) n'est pas en dimensions réelles.



(a) Image binaire

(b) image réduite 16×1

Figure 3.13 : réduction de la taille de l'image

Dans ce qui suit, on détaillera les différentes opérations nécessaires pour la segmentation d'image en représentant les résultats des opérations effectuées, on prend l'image représentée sur la **figure 3.14** comme image originale à segmenter.

Under the option EDIT select PASTE SPECIAL. A dialog box will open, select paste picture, then click OK. Your figure should now be in the Word Document.

If you are preparing images in TIFF, EPS, or PS format, note the following. High-contrast line figures and tables should be prepared with 600 dpi resolution and saved with no compression, 1 bit per pixel (monochrome), with file names in the form of "fig3.tif" or "table1.tif."

Photographs and grayscale figures should be prepared with 300 dpi resolution and saved with no compression, 8 bits per pixel (grayscale).

Sizing of Graphics

Most charts graphs and tables are one column wide (3 1/2 inches or 21 picas) or two-column width (7 1/16 inches, 43 picas wide). We recommend that you avoid sizing figures less than one column wide, as extreme enlargements may distort your images and result in poor reproduction. Therefore, it is better if the image is slightly larger, as a minor reduction in size should not have an adverse affect the quality of the image.

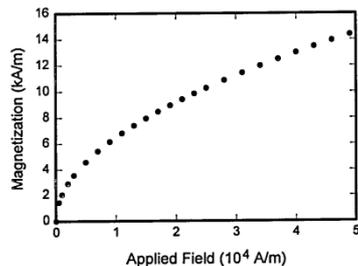


Fig. 1. Magnetization as a function of applied field. Note that "Fig." is abbreviated. There is a period after the figure number, followed by two spaces. It is good practice to explain the significance of the figure in the caption.

fonts are as follows: Open Type Fonts: Times Roman, Helvetica, Helvetica Narrow, Courier, Symbol, Palatino, Avant Garde, Bookman, Zapf Chancery, Zapf Dingbats, and New Century Schoolbook.

Print Color Graphics Requirements

IAENG accepts color graphics in the following formats: EPS, PS, TIFF, Word, PowerPoint, Excel, and PDF. The resolution of a RGB color TIFF file should be 400 dpi.

Web Color Graphics

IAENG accepts color graphics in the following formats: EPS, PS, TIFF, Word, PowerPoint, Excel, and PDF. The resolution of a RGB color TIFF file should be at least 400 dpi.

Your color graphic will be converted to grayscale if no separate grayscale file is provided. If a graphic is to appear in print as black and white, it should be saved and submitted as a black and white file. If a graphic is to appear in print or on IAENG web site in color, it should be submitted as RGB color.

Graphics Checker Tool

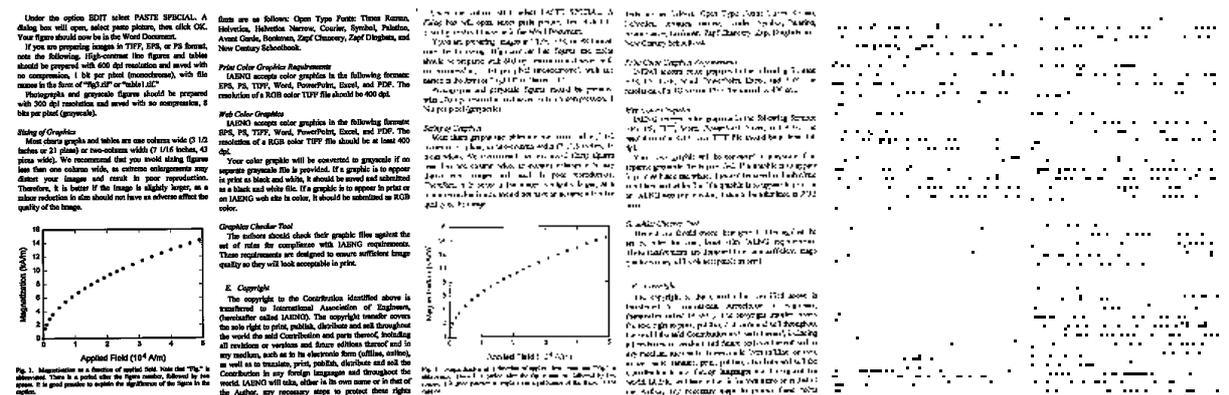
The authors should check their graphic files against the set of rules for compliance with IAENG requirements. These requirements are designed to ensure sufficient image quality so they will look acceptable in print.

E. Copyright

The copyright to the Contribution identified above is transferred to International Association of Engineers, (hereinafter called IAENG). The copyright transfer covers the sole right to print, publish, distribute and sell throughout the world the said Contribution and parts thereof, including all revisions or versions and future editions thereof and in any medium, such as in its electronic form (offline, online), as well as to translate, print, publish, distribute and sell the Contribution in any foreign languages and throughout the world. IAENG will take, either in its own name or in that of the Author, any necessary steps to protect these rights

Figure 3.14 : image à segmenter

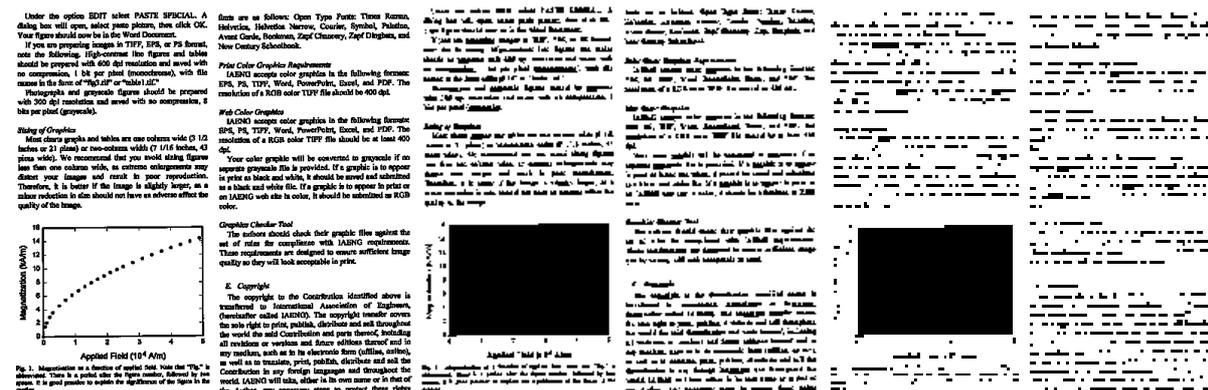
Le but de l'algorithme utilisé est de détruire les parties (composantes) textuelles de l'image de document et ne laisser que les parties non-textuelles (dessins ; tableaux ; graphes ; etc.). Un problème est rencontré l'ors de la réduction de la taille de l'image est qu'une composante non-textuelles peut être détruite si elle présente une faible épaisseur (un contour isolé par exemple ou bien une ligne d'un tableau) dans ce cas l'algorithme échoue pour la discrimination entre les composantes textuelles et non-textuelles (voir la figure 3.15).



(a) Image réduite 16 × 1 T=1 (b) Image réduite 4 × 1 T=4 (c) Image réduite 4 × 1 T=3

Figure 3.15 : problème de destruction des parties non textuelles

La solution proposée dans [23], consiste à introduire à cette algorithmme une étape qui suit immédiatement la double réduction 4×1 . La modification de l’algorithme est donnée par le diagramme de la **figure 3.19**. Cette étape, appelée ‘bouchage des trous’ de l’image permet de remplir les vides situés à l’intérieur des objets et vise à augmenter l’intensité (en terme de pixels noirs) des composantes non-textuelles. si on compare le résultat présenté par la **figure 3.16.b** avec celui de la **figure 3.16.c** on remarque que la partie non-textuelle est conservée pour les étapes suivantes de ce traitement contrairement aux composantes textuelles.



(a) Image réduite 16×1 T=1 (b) Image réduite 4×1 T=4 (c) Image réduite 4×1 T=3
Figure 3.16 : Résultats de réduction après bouchage des trous

❖ Méthode utilisé pour le bouchage :

L’objectif est de remplir les zones vides à l’intérieur des objets par le noir, le principe utilisé dans cette opération est comme suit :

On partant d’une image binaire avec objet en ‘noir’ et fond en ‘blanc’.

- 1) Balayage de l’image, si le pixel est noir et son voisin inferieur est blanc, le pixel inferieur est mis bleu.
- 2) Balayage de l’image, s il ya une séquence de pixels bleus (un ou plus) bornés par deux pixels noirs alors la séquence est mise noire.
- 3) les deux étapes sont répétées jusqu’à ce qu’aucune modification n’apparisse voire **figure 3.17**.

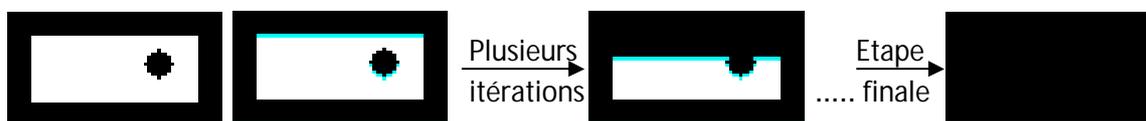


Figure 3.17 : étapes de remplissage

Un exemple de bouchage de trous d’un bloc image est donné par la figure ci-dessous

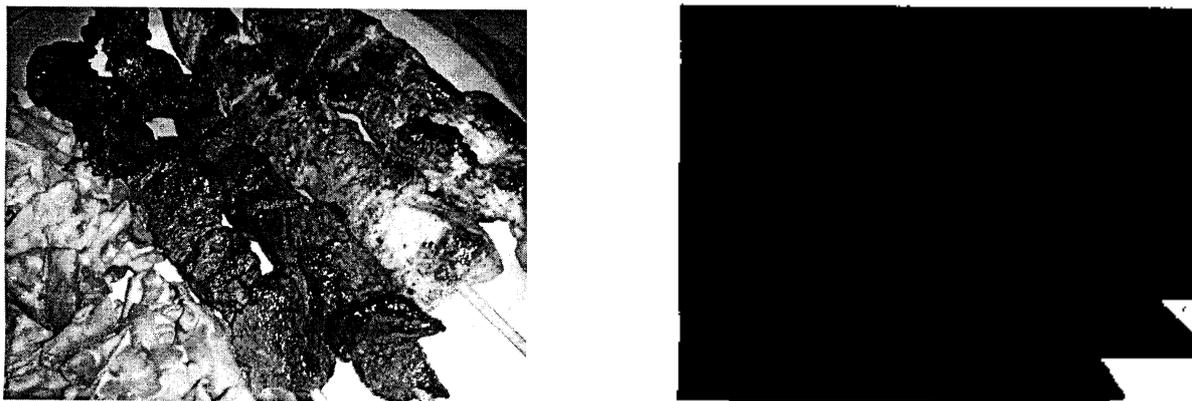


Figure 3.18 : Exemple de remplissage d'une image

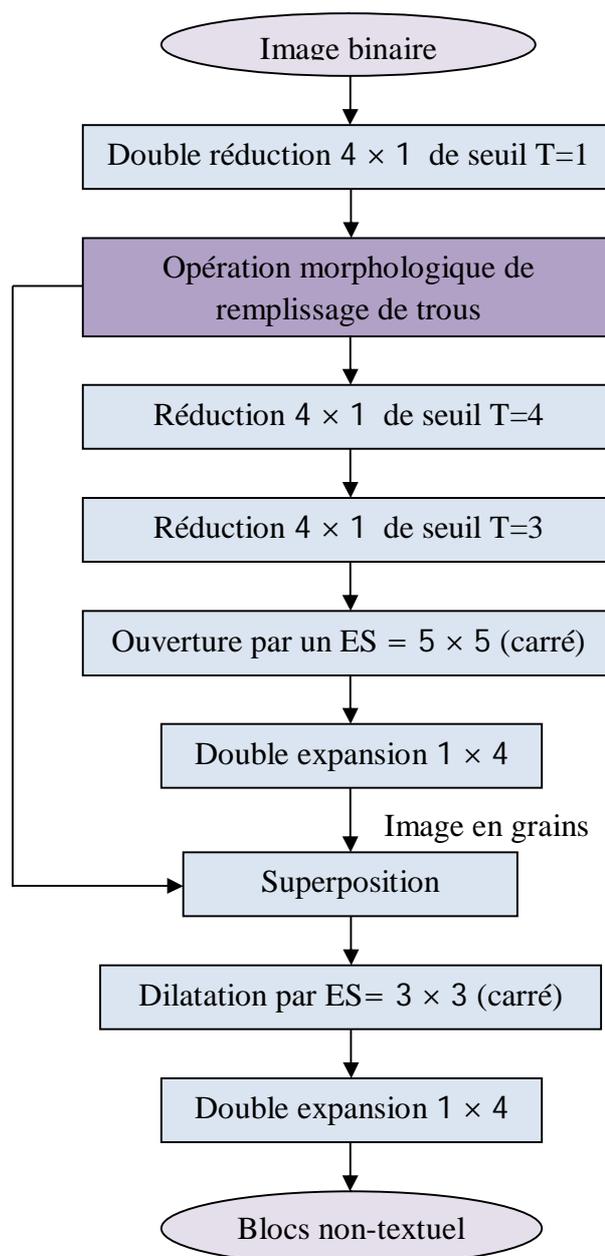


Figure 3.19: Algorithme de Bloomberg amélioré [23]

Après les étapes de réduction de l'image de documents selon les seuils indiqués précédemment avec introduction du remplissage ; on procède par une ouverture de l'image par un élément structurant carré de taille 5×5 [23].

L'opération morphologique d'ouverture est divisée en deux étapes chacune réalise une opération morphologique de base et s'organise comme suit :

L'image d'entrée considérée est l'image réduite.

- 1) une seule itération d'érosion par un ES (carré) de taille 5×5 .
- 2) une seule itération de dilatation de l'image par un ES de taille 5×5 .

Les résultats de l'érosion et l'ouverture de l'image visualisé sur la **figure 3.16.c** sont visualisées sur les deux images de la **figure 3.20.b** et la **figure 3.20.c** respectivement.

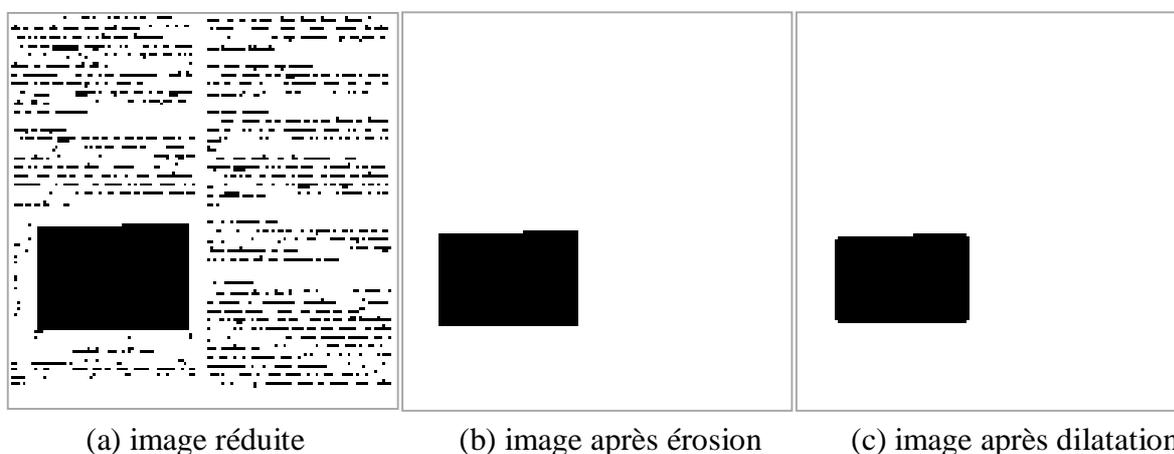
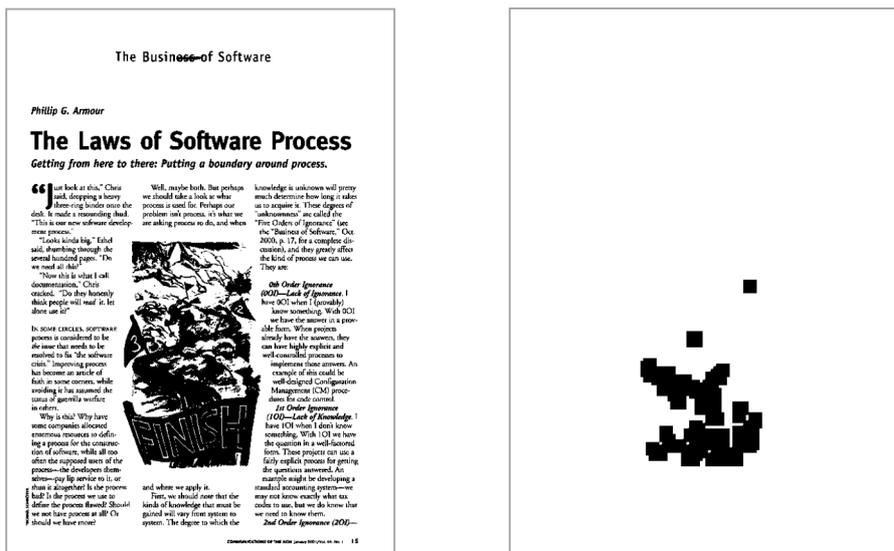


Figure 3.20 : Ouverture de l'image par un ES = 5×5

Il est clair sur le diagramme de la **figure 3.19** que les étapes précédentes réalisent une compression de la taille de l'image, le but est de rapprocher les composantes non textuelles et les maintenir. Les étapes suivantes réalisent une décompression du résultat, le but est de revenir à la taille origine de l'image et détecter les blocs non-textuels.

Comme indique le diagramme, la phase présente implique une double dilatation sur l'image chacune de dimension 1×4 . A ce stade de traitement le résultat obtenu (des quatre réductions successives + remplissage ; ouverture par ES 5×5 et double expansion 1×4) donne une image en grains (**figure 3.21.c**) qui représentent les résidus des parties non textuelles de l'image de document.



(b) Image originale (c) image en grains
Figure 3.21 : Résidus d'une image

L'expansion de l'image représentée sur la **figure 3.20.c** donne les résultats représentés par la figure suivante :

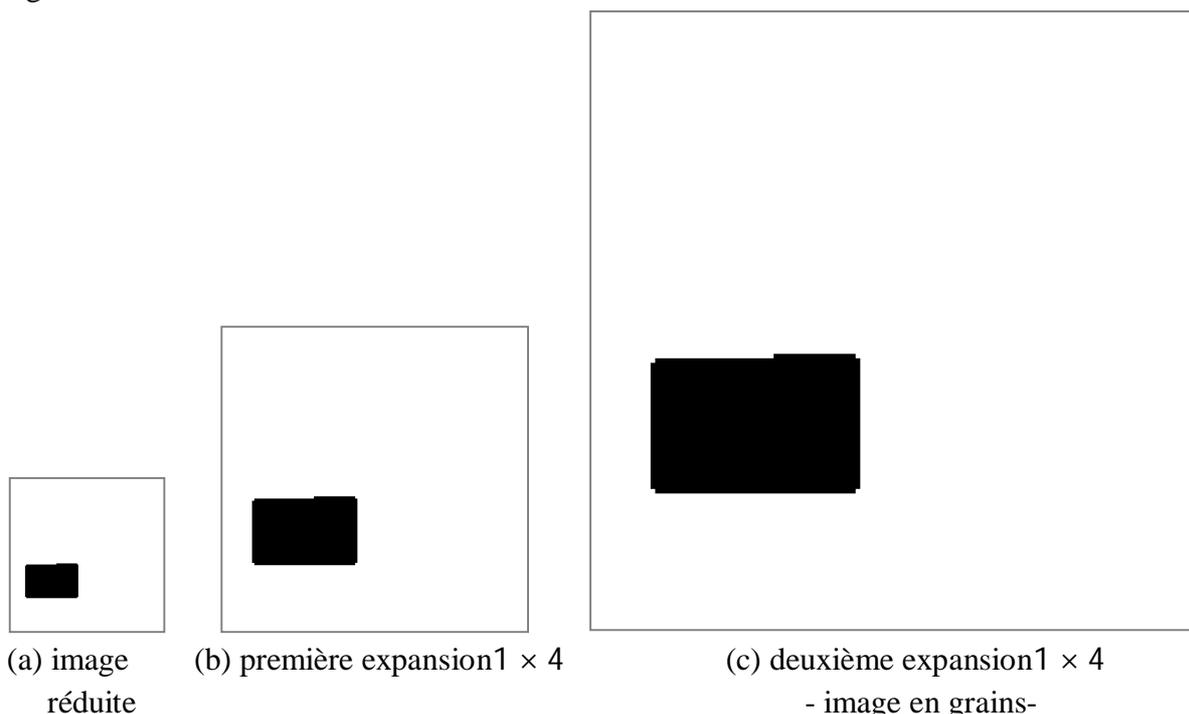


Figure 3.22 : Expansion d'image

La **figure 3.23** visualise la localisation de bloc non-textuel de l'exemple traité (**figure 3.14**).

La récupération de la taille originale des blocs localisés nécessite une opération de dilatation par un $ES = 3 \times 3$ suivie d'une autre opération de double expansion 1×4 chacune. Le résultat final « image pre-segmentée » est visualisé par la **figure 3.24**.

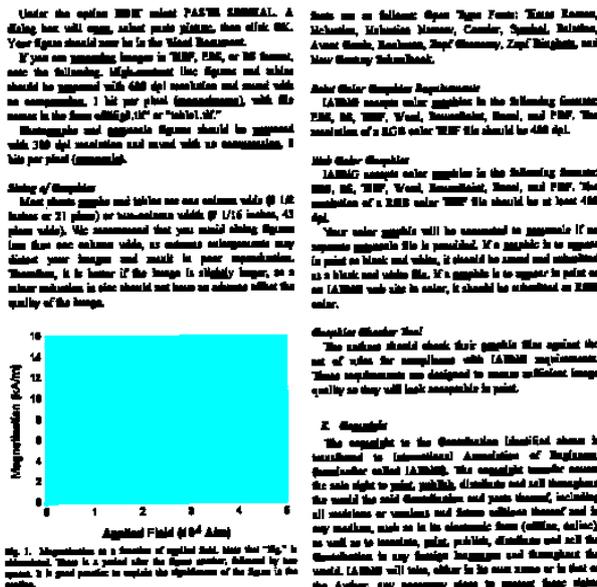
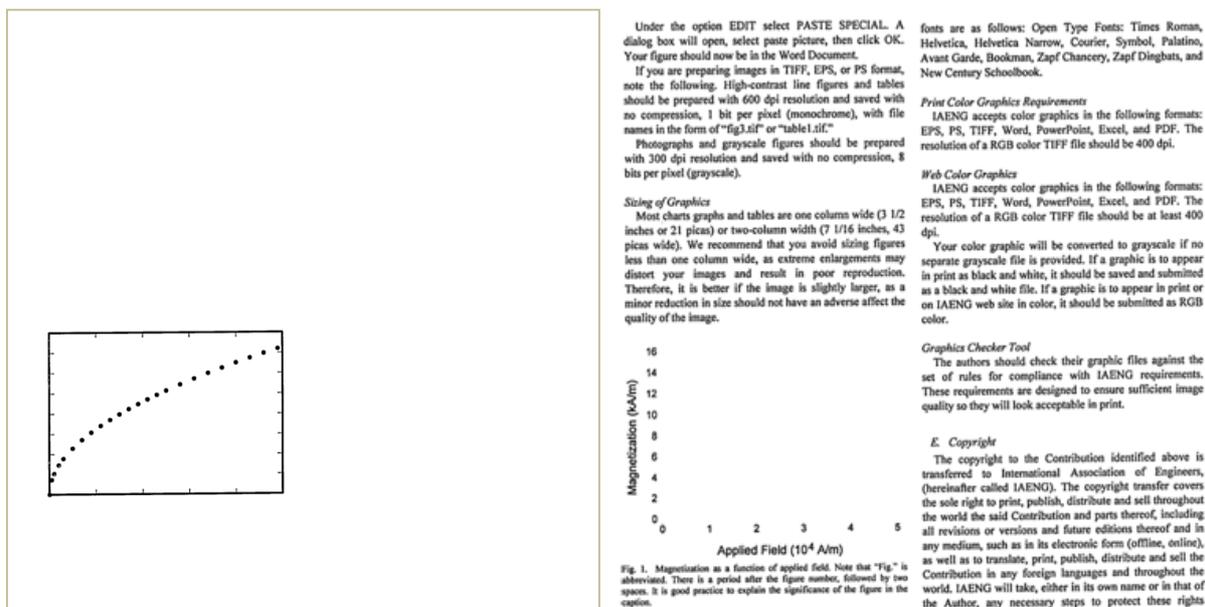


Figure 3.23 : Localisation de blocs non-textuels



(a) partie non texte

(b) image préparée à la segmentation finale

Figure 3.24 : résultat final de pré-segmentation

Un mauvais fonctionnement de cet algorithme est rencontré expérimentalement, on rappelle que le principe est d'écraser les composantes textuelles ensuite rehausser les composantes non-textuelles. Dans le cas où l'image de document contient une partie non-textuelle avec une taille faible ou bien des petites composantes isolées les unes des autres, lors du rehaussement (ouverture de l'image) l'algorithme les détruit. Dans la **figure 3.25.a** on voit que la composante encadrée en rouge est effacée après ouverture de l'image (**figure 3.25.c**) et tout le contenu est considéré comme texte.

Pour remédier à l'insuffisance rencontrée avec l'algorithme de Bloomberg on fait appel à l'algorithme de segmentation RLSA.

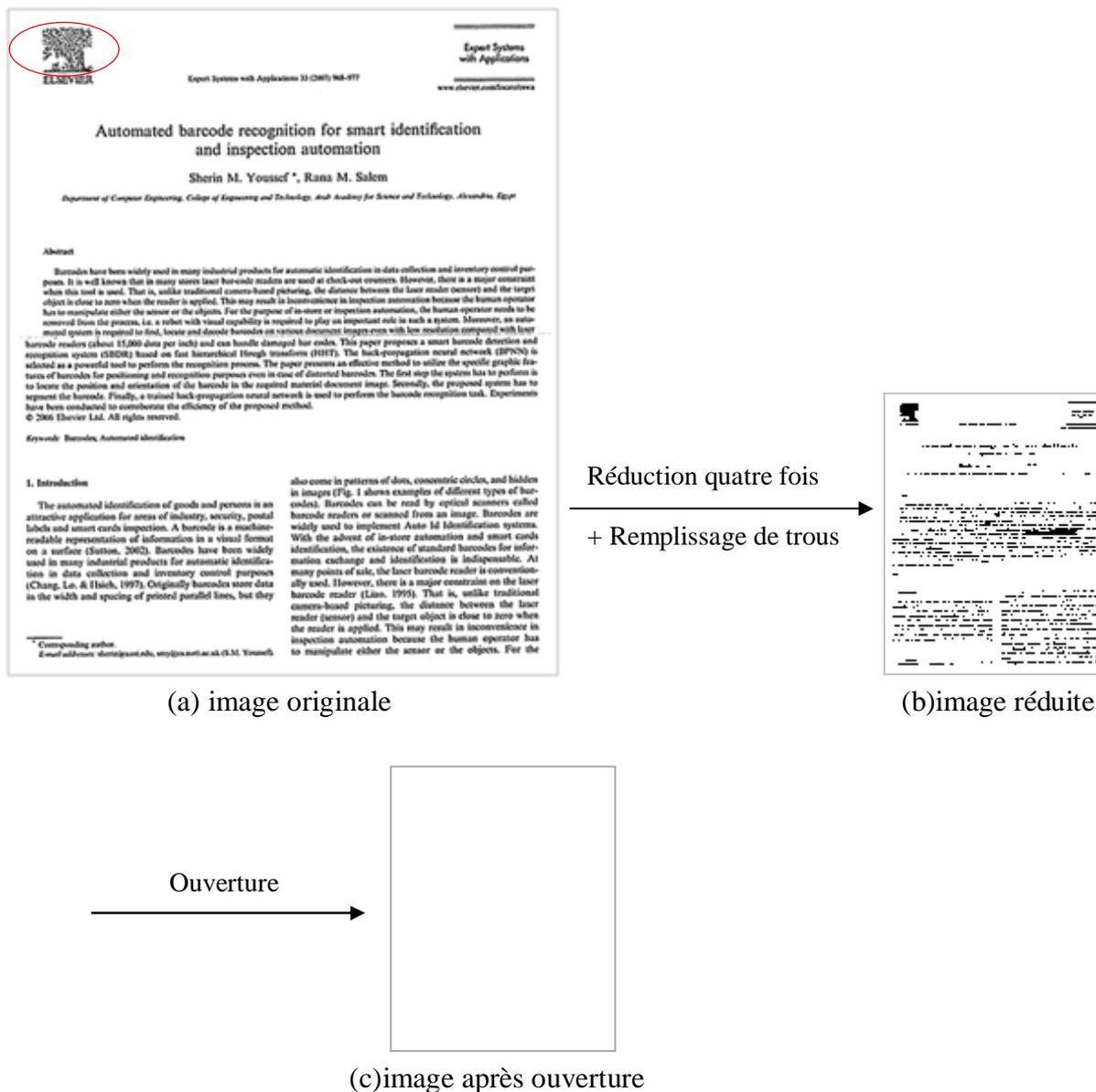


Figure 3.25 : problème de pré-segmentation

Dans la section suivante, la segmentation finale sera détaillée étape par étape avec visualisation de quelques résultats obtenus.

3.3.2 Segmentation finale

En partant d'une image de document composite, nous appliquons la technique Run Length Smoothing Algorithm (RLSA) qui consiste à faire un double lissage unidirectionnel de l'image de documents à segmenter selon des seuils à calculer. Les blocs sont obtenus en appliquant l'opérateur logique "ET" sur les deux images résultant respectivement d'un lissage horizontal et d'un lissage vertical [57].

La technique implémentée pour la création de blocs est une variante de l'algorithme RLSA qui utilise elle-même un lissage horizontal avec un seuil calculé mais la différence réside au lissage vertical, ce dernier est appliqué sur l'image résultant du premier lissage (au lieu de l'appliqué sur l'image originale) voir **figure 3.28**. Et le résultat donne directement l'image en blocs.

Dans ce travail la segmentation par RLSA est appliquée sur l'image résultant de la première segmentation déjà réalisé (la sortie de la pré-segmentation par Algorithme de Bloomberg est l'entrée de RLSA).

La **figure 3.26** et la **figure 3.28** illustrent respectivement le diagramme de l'algorithme RLSA et sa variante qui sont utilisées dans notre programme:

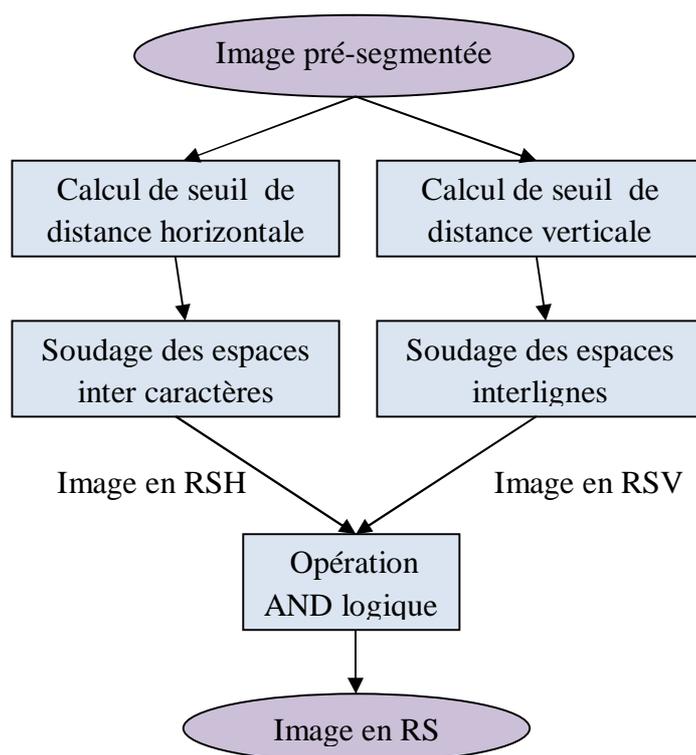
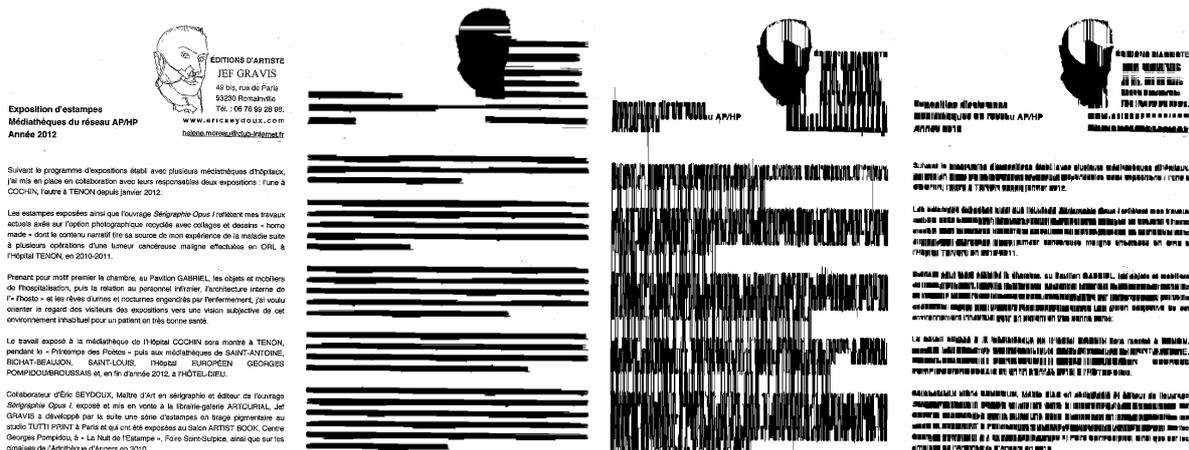


Figure 3.26: Algorithme RLSA

La figure si dessous illustre un exemple de segmentation d'un document.



(a) image originale (b) lissage horizontal (c) lissage vertical (d) image segmentée

Figure 3.27 : Segmentation RLSA

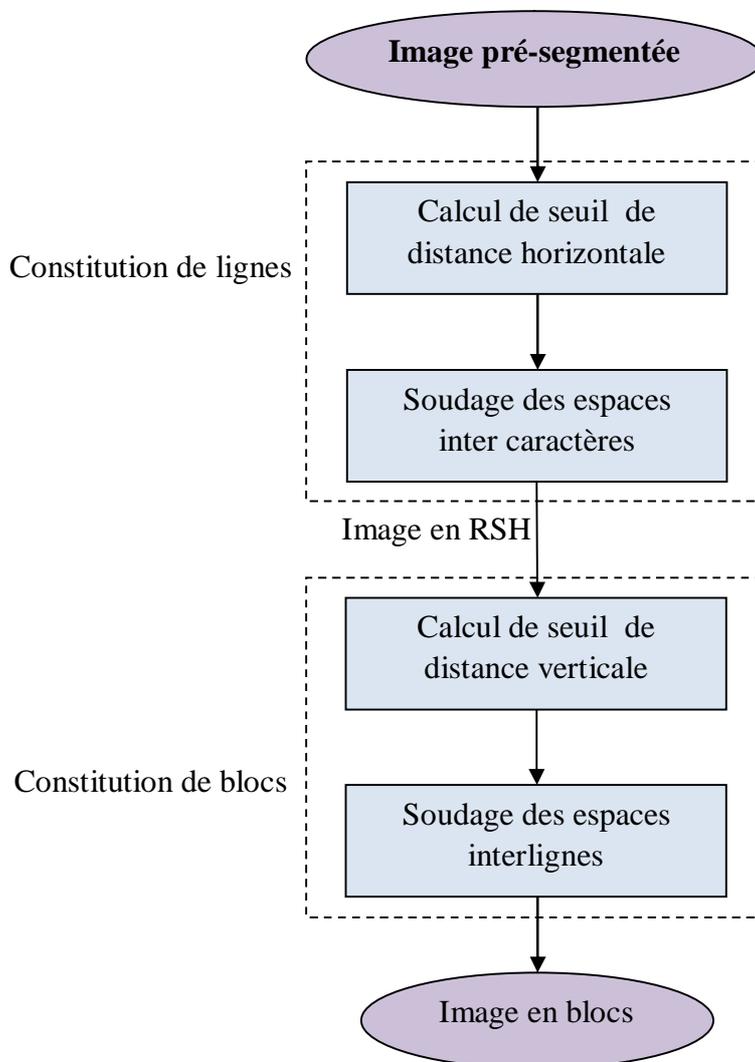


Figure 3.28: Variante de l'algorithme RLSA

Par la suite on détaillera la technique RLSA et on visualisera les résultats obtenus par l'application de cette technique.

3.3.2.1 Constitution de lignes

Cette méthode est basée sur le calcul d'une distance seuil, en utilisant les espaces inter caractères, inter-mots c'est-à-dire faire un soudage en utilisant les distances afin de supprimer, l'espace blanc entre les mots d'une ligne.

Principe

1ere étape : Constitution de l'histogramme de distances horizontales. Pour toutes les lignes de l'image ; calculer la distance (le nombre de pixels blancs) entre deux pixels noirs successifs (**figure 3.29**).

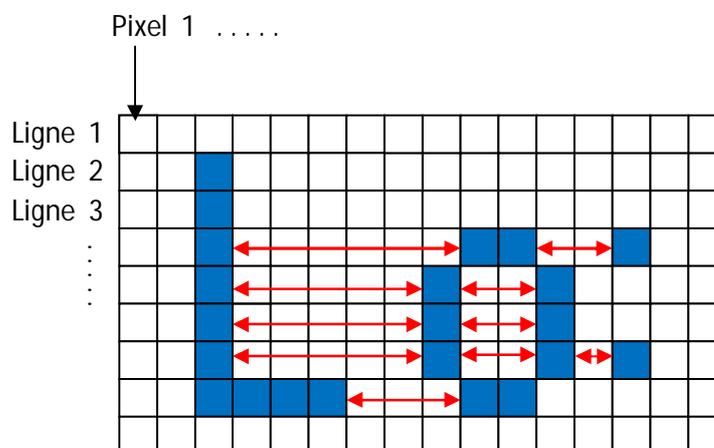


Figure 3.29 : Distances horizontales

Pour la figure si dessus on obtient l'histogramme suivant :

0	1	4	1	1	3	1	...
---	---	---	---	---	---	---	-----

Figure 3.30 : Histogramme de distances horizontales

Pour une image de document les valeurs des distances résultantes sont concentrées autour de deux valeurs. la plus fréquente c'est la distance inter-caractère (α) et l'autre représente la distance inter-mot (β); voir **figure 3.31**

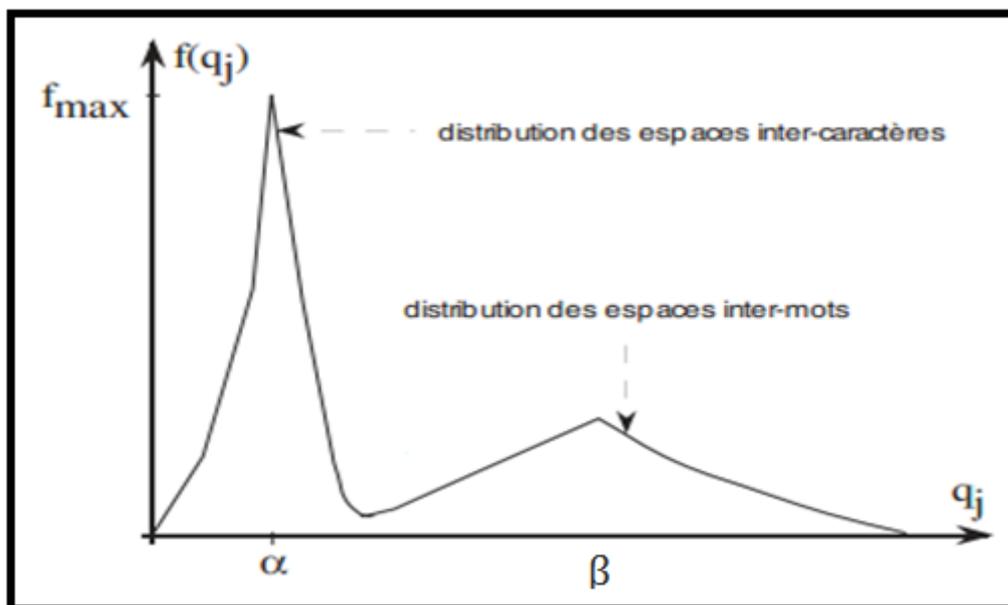


Figure 3.31 : Distribution horizontale des espaces

2eme étape : Lissage horizontale de l'image de document (ou constitution de lignes). Il s'agit ensuite de balayer l'image ligne par ligne de gauche à droite et dès qu'on trouve une distance d le teste suivant est effectué :

Si la distance $d \leq \beta$, alors soudage (ajout des pixels noirs dans d).

Sinon, ne pas souder.

Le résultat de cette étape est une image structurée en lignes ou rectangles noirs disposés horizontalement appelés les rectangles structurant horizontaux RSH.

La **figure 3.32** représente le résultat de lissage horizontal sur l'image pré-segmentée (**figure 3.24.b**).

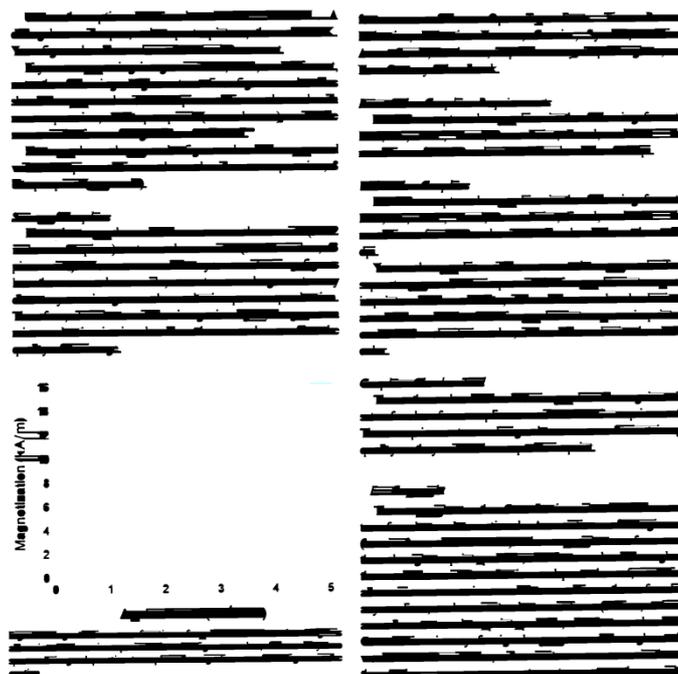


Figure 3.32: Lissage horizontal

3.3.2.2 Constitution de blocs

La constitution de blocs d'un document basée dans notre méthode sur le calcul d'une distance seuil, en utilisant les espaces interlignes, c'est-à-dire faire un remplissage (noire) dans les espaces blancs qui respect une condition sur le seuil choisi.

Principe

1ère étape : Calculer la distance interligne D de l'histogramme de distances verticales, dans ce cas il faut calculer les distances (nombre de pixels blanc) entre deux pixels noirs mais dans le sens vertical.

L'interligne

Interligne désigne l'espace entre deux lignes consécutives d'un texte. Cette définition de l'interligne présente une propriété intéressante : celle d'être constante dans un document mono fonte.

Pour une image de document les valeurs des distances résultantes sur l'histogramme de distances verticales sont concentrées autour d'une valeur D (la plus fréquente) c'est la distance interligne ; voir **figure 3.33**.

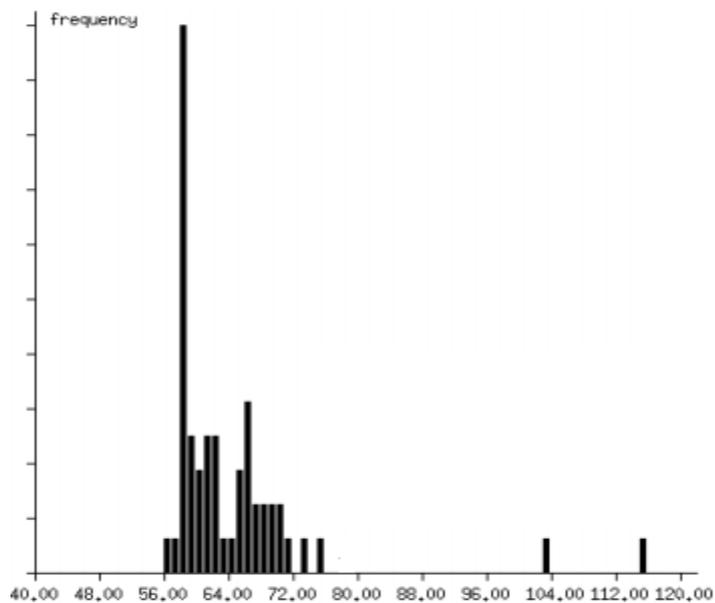


Figure 3.33: Exemple de distribution verticale des espaces

2eme étape : Lissage vertical de l'image de document. Balayer de nouveau l'image et dès qu'on trouve une distance d le teste suivant est effectué :

Si la distance $d \leq D$, alors soudage (ajout des pixels noirs dans d).

Sinon, ne pas souder.

Le lissage horizontal de l'image de la **figure 3.32** donne le résultat suivant :



Figure 3.34 : image en blocs

La **figure 3.34** Montre la fiabilité de la technique RLSA au niveau du regroupement d'objets connexes proches les uns par rapport aux autres dans un même bloc.

3.3.2.3 Étiquetage des blocs

Cette opération permet de donner la même étiquette aux pixels d'un même bloc par l'algorithme des composantes connexes à 4-connexions (Approche par double parcours ou (two-pass) citée dans le chapitre 2). Donc de faire la séparation des blocs **figure 3.38** on donnant à chaque bloc une couleur (étiquette) différentes des autres blocs constituant l'image.

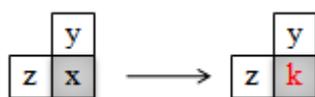
Principe

1ère étape : Le premier passage consiste à balayer l'image ligne par ligne (de haut en bas) et pour chaque ligne le balayage se fait de gauche à droite (en commençant par le premier pixel de la première ligne) en cherchant les pixels objets. On considère que l'étiquette d'un pixel noir est $eti_q=1$ et pour un pixel blanc $eti_q=0$. Les étiquettes x , y , z et k peuvent prendre la valeur 0 ou 1 ou autre.

- Si le pixel est noir $p=1$, alors tester ses deux pixels voisins gauche (pg) et haut (ph) :

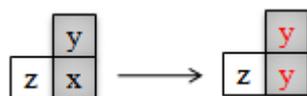
- si $pg == ph == 0$:

$p =$ nouvelle étiquette (k)



- si $ph == 1$ et $pg == 0$:

$eti_q p = eti_q ph$



- si $ph == 0$ et $pg == 1$:

$eti_q p = eti_q pg$



- si $ph == pg == 1$:

$eti_q p = \min\{eti_q ph, eti_q pg\}$

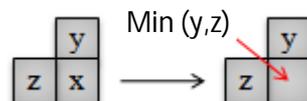


Figure 3.35 : Premier passage de l'étiquetage

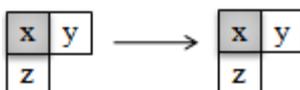
La **figure 3.37** illustre le résultat de l'étiquetage au premier passage des blocs de l'image visualisée sur la **figure 3.34**.

2eme étape : second Balayage de l'image ligne par ligne (de bas en haut) de droite à gauche en commençant par le dernier pixel de la dernière ligne en cherchant les pixels objets qui sont déjà étiquetés **figure 3.36**.

- Si le pixel est noir $p=1$, alors tester ses deux pixels voisins droit (pd) et bas (pb) :

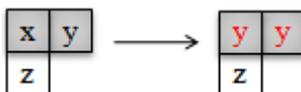
- si $pd == pb == 0$:

Garder l'étiquette



- si $pd == 1$ et $pb == 0$:

etiq p = etiq pd



- si $pd == 0$ et $pb == 1$:

etiq p = etiq pb



- si $pd == pb == 1$:

etiq p = $\min\{\text{etiq pd}, \text{etiq pb}\}$

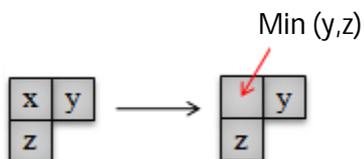


Figure 3.36 : second passage de l'étiquetage

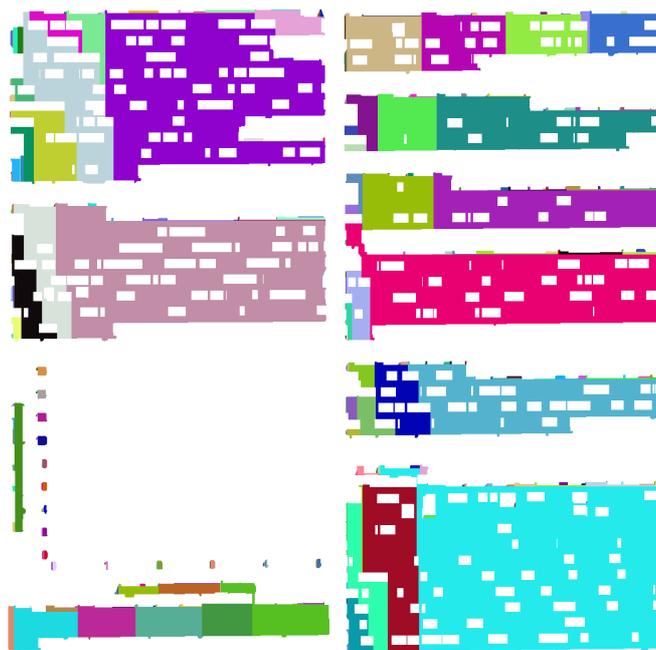


Figure 3.37 : Premier passage de l'étiquetage de blocs

Après cette étape on remarque qu'il existe toujours des étiquètes différentes (différence de couleur) dans un même bloc de l'image (voir **figure 3.37**), par contre un bloc doit être représenté par la même couleur, il est nécessaire de rectifier l'étiquetage par des retours au premier et second passage. Pour chaque ligne (ou colonne) dans l'image, s'il existe deux pixels adjacents pix1 et pix2 d'un même bloc qui ont des étiquètes différentes ($\text{pix1} = \text{etq1}$ et $\text{pix2} = \text{etq2}$ avec $\text{etq1} \neq \text{etq2}$), le changement est effectué par la relation (3.13) suivant l'algorithme :

$$\text{pix1} = \text{pix2} = \min \{ \text{etq1} ; \text{etq2} \} \quad (3.13)$$

*/******retour au premier passage******/*

Pour chaque ligne (suivant le premier passage)

Si $\text{pix1} \neq \text{blanc}$ et $\text{pix2} \neq \text{blanc}$

{

Si $\text{etiq1} \leq \text{etiq2}$

$\text{Pix2} = \text{etiq1}$;

Si non

$\text{Pix1} = \text{etiq2}$;

```
}
```

Pour chaque colonne (suivant le premier passage)

```
Si pix1 != blanc et pix2 != blanc
```

```
{
```

```
Si etiq1 <= etiq2
```

```
  Pix2 = etiq1 ;
```

```
Si non
```

```
  Pix1 = etiq2 ;
```

```
}
```

```
/****** fin du retour au premier passage******/
```

```
/******retour au second passage******/
```

Pour effectuer le retour au second passage le même algorithme est appliqué seulement que le balayage des lignes se fait de droite à gauche ainsi que le balayage des colonnes commence de bas de l'image vers le haut.

```
/****** fin du retour au second passage******/
```

L'algorithme se répète jusqu'à ce qu'aucun changement ne reste, ce qui donne à chaque bloc une couleur (étiquète) uniforme (**figure 3.38**).

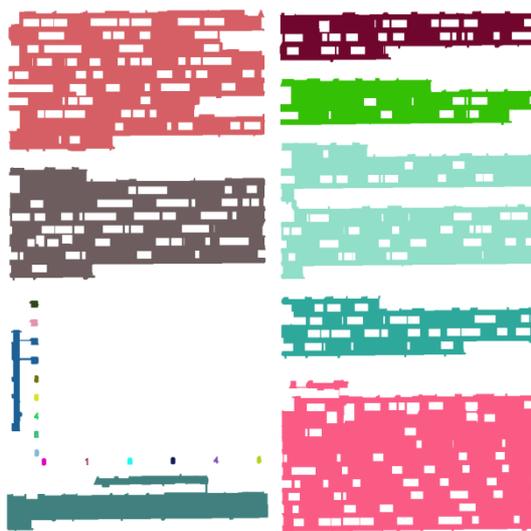
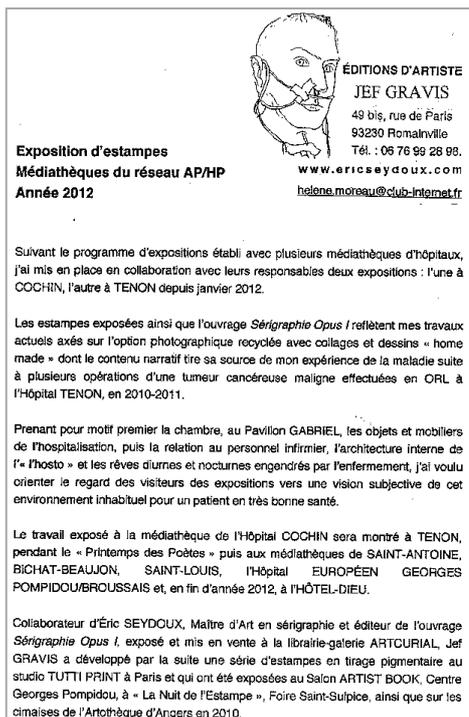


Figure 3.38 : Détection de blocs d'image

3.3.2.4 Recherche des lignes

Une fois la détection des blocs réalisée, il faut détecter maintenant les lignes dans un même bloc. Nous utilisons pour cela un deuxième étiquetage de composantes connexes dans chaque bloc.

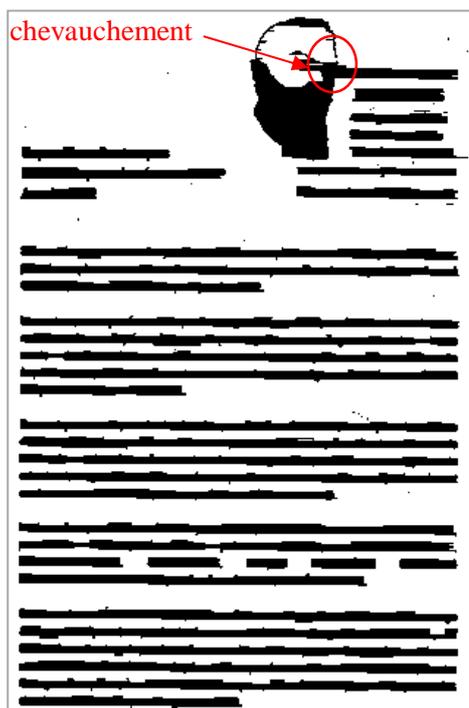
L'algorithme RLSA comme indiqué précédemment est basé sur le calcul des distances horizontalement et verticalement pour fusionner respectivement les caractères d'une même ligne d'un bloc et les lignes d'un même bloc de texte. Le calcul de distances influe fortement la segmentation, dans le cas où les distances sont faibles que nécessaire on obtient une sous-segmentation (découpage d'une ligne en plusieurs parties), et dans le cas contraire on obtient une sur-segmentation, ce dernier cas présente un grand problème pour la segmentation car deux composantes de classes différentes peuvent être chevauchées (**figure 3.39.a**). La solution à ce problème est donnée par l'algorithme RLSA représenté sur la **figure 3.26**, l'opérateur « AND logique » réalise la séparation entre deux composantes de blocs différents. La **figure 3.39** illustre un exemple de soudage d'une ligne de texte avec un bloc (non-texte) voisin et la correction de cette sur-segmentation.



(a) image originale



(b) image en blocs avec chevauchement



(c) image en lignes soudées



(d) image en RS

Figure 3.39 : Correction de sur-segmentation

L'étiquetage des lignes d'un bloc est atteint en suivant les étapes :

- 1) Construire les rectangles structurants (RS) de l'image.
- 2) Calcul des coordonnées de chaque bloc détecté.
- 3) Utiliser ces coordonnées pour localisées les blocs sur l'image en RS.
- 4) Etiquetage des lignes correspondantes à chaque bloc.

Nous avons besoin de connaître avec la plus grande précision possible les dimensions des RS pour arriver à une bonne classification de ces RS (ligne de texte/composante non textuelle). contrairement au sur-segmentation dans le cas d'une sous-segmentation, le document peut contenir des lignes incomplètes (**figure 3.40.b**) qui peuvent générer des espaces blancs importants entre lignes voisines dans certaines régions, ce n'est pas un problème pour la classification car elle est basée seulement sur l'épaisseur des RS.

Les résultats obtenus par ce dernier étiquetage sont visualisés par la **figure 3.40** pour les deux cas :

- 1) Etiquetage des RSH (**figure 3.40.a**)
- 2) Etiquetage des RS (**figure 3.40.b**)

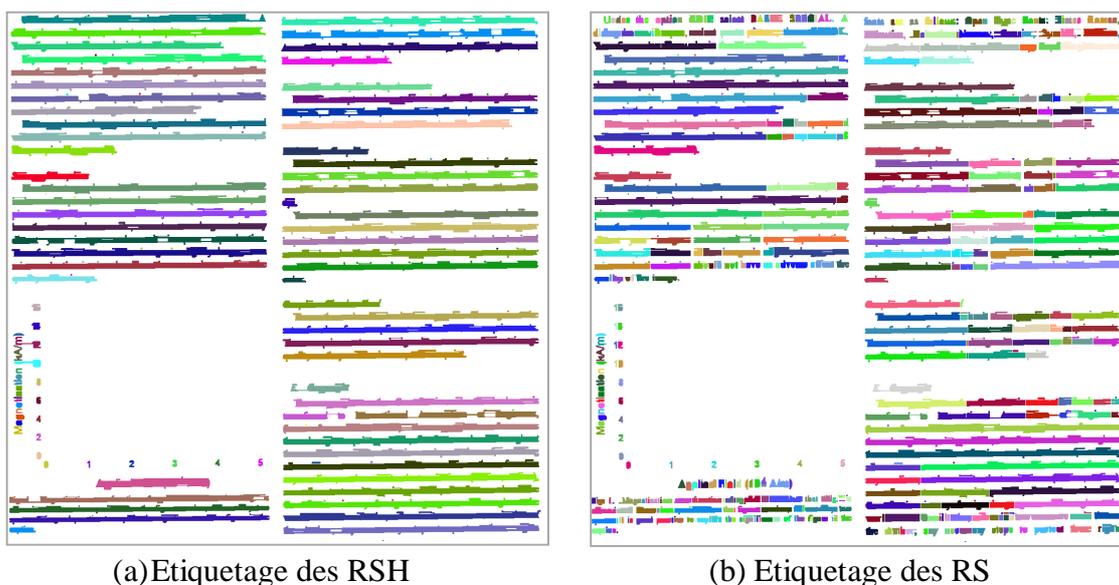


Figure 3.40 : Etiquetage des lignes

On remarque que l'étiquetage des RSH donne un nombre de ligne relativement minimal mais avec une sur-segmentation; c'est une mauvaise segmentation. Dans le deuxième cas le nombre de lignes augmente mais les épaisseurs des lignes sont conservées, donc une segmentation meilleure avec un temps d'exécution relativement élevé, ce temps dépend du nombre de divisions (ou de rectangles structurant) sur une même ligne (rectangle structurant horizontal).

3.3.2.5 Classification

A pour but de distinguer entre les composantes textuelles et les composantes non textuelles (une composante est un rectangle structurant RS), cette classification comprend les étapes suivantes :

- Calcul des épaisseurs des rectangles structurants.
- Calcul de la moyenne des épaisseurs des RS dans un bloc.
- Décision

Les épaisseurs des RS sont calculées en utilisant leurs coordonnées qui sont fournies grâce à l'opération de l'étiquetage.

La moyenne des épaisseurs est calculée selon la formule :

$$\text{moy} = \frac{\sum \text{épaisseurs de lignes}}{\text{nombre de lignes}}$$

Cette étape de segmentation finale vise à extraire les composantes textuelles et aussi à chercher les composantes non-textuelles mal reconnus par la première segmentation.

La classification d'une composante se fait par comparaison de sa propre épaisseur avec un seuil relatif à l'épaisseur moyenne du bloc.

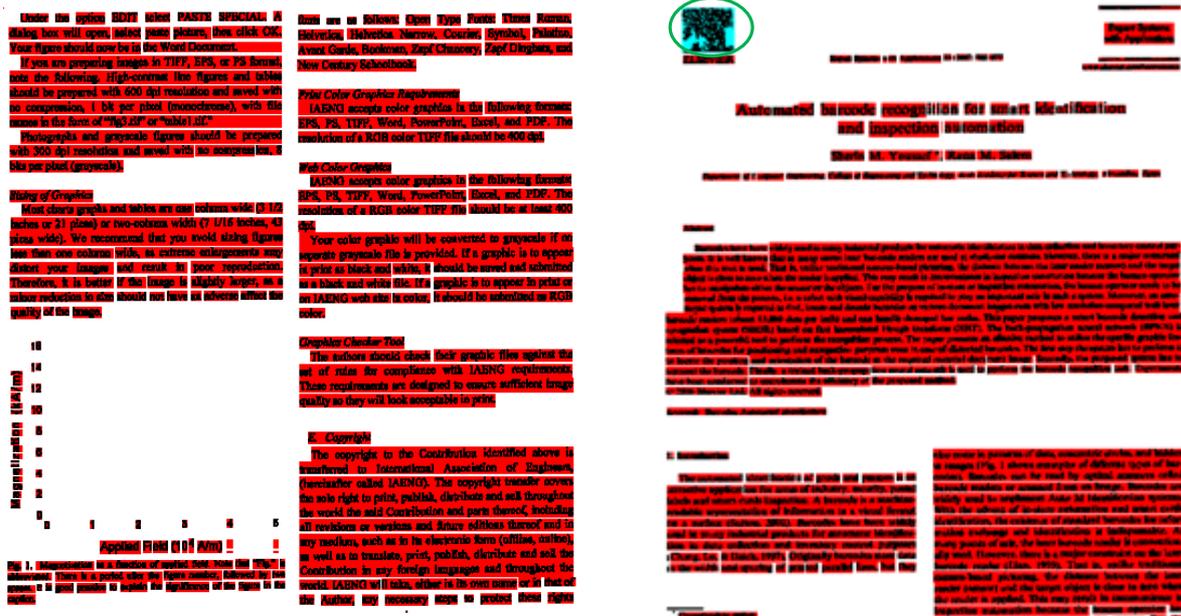
Nous avons fixé le seuil expérimentalement à $1,6 \times \text{moy}$ (c'est le rapport entre l'épaisseur maximale et l'épaisseur moyenne calculé pour tous les ES de l'image). Avec cette valeur on obtient une bonne séparation entre composantes des deux classes différentes, donc :

Si un RS possède une épaisseur $> 1,6 \times \text{moy} \Rightarrow$ composante non-textuelle

Si non c'est une partie du texte.

Pour une meilleure visualisation des résultats on utilise une coloration de blocs.

La (figure 3.41) illustre un exemple d'application sur les deux images représentées sur les (figure 3.14, page 75) et (figure 3.25.a, page 81) respectivement. Les blocs de texte portent une couleur rouge et les espaces des blocs non-texte sont représentés en bleu.



(a) Extraction de texte

(b) Correction et Extraction de texte

Figure 3.41 : Segmentation finale

Discussion

Nous avons présenté dans cette partie l'approche qui vise à extraire les blocs d'image d'un document à structure complexe et leur classification. D'abord, les zones de nature non-texte sont détectées en utilisant une destruction des composantes textuelle par opérations morphologique dont l'ensemble de ces opérations constitue l'algorithme de Bloomberg. Ensuite, les composants textuels sont détectés en utilisant le Run Length Smoothing Algorithm avec un étiquetage des composantes (rectangles structurants).

Finalement notre document numérique est segmenté en deux classes, chaque bloc est jugé soit un texte ou non-texte. A partir de cette classification on peut commencer un traitement sur le texte (OCR).

Dans le chapitre suivant, nous présentons les résultats obtenus par notre propre méthode à chaque étape de segmentation avec discussions.

CHAPITRE 4

RESULTATS ET DISCUSSIONS

Dans ce chapitre, nous allons faire la présentation des résultats détaillés de la segmentation, appliquée sur différentes images (structure simple, structure complexes, mise en page de type Manhattan et non Manhattan). Une interprétation et une discussion des différents résultats suivront.

Nous considérons un document qui ne comporte que du texte comme étant un document simple par opposition à un document composite qui pourra comporter en plus du graphique. Les documents fictifs sont des documents issus d'un logiciel de traitement d'images par contre les documents réels sont des documents administratifs numérisés.

Pour pouvoir expérimenter notre démarche et vérifier le bon fonctionnement de l'algorithme implémenté, nous avons constitué une base d'images de documents composites. Cette base contient deux types de documents : 10 documents fictifs (5 documents simples et 5 autres composites) et 20 documents réels (10 documents simples et 10 autres composites).

4.1 Présentation de l'interface et Implémentation

Il est nécessaire de fournir à l'utilisateur un outil, permettant de visualiser les images obtenues par les segmentations décrites aux chapitres précédents. Pour cela, une interface utilisateur a été développée, et intitulée « Segmentation de blocs d'images de documents ». Nous allons, tout d'abord, présenter brièvement le langage de programmation utilisé pour la réalisation du logiciel, puis détailler les étapes de sa réalisation, et les différentes structures qui le constituent, ainsi que les résultats obtenus par les différents algorithmes utilisés.

4.1.1 Environnement logiciel

Ce travail a été réalisé au sein du laboratoire, avec un ordinateur doté, du processeur pentium R Dual Corps CPU T4200 @ 2.00GHz, de mémoire vive RAM de 2 GB.

4.1.2 Description logiciel

Le logiciel « Segmentation de blocs d'images de documents » est une application simple pour n'importe quel utilisateur, il peut l'aider à classer les textes et les images des documents en quelque secondes.

Les images utilisées sont de taille 1078x768 et le temps d'exécution de ce logiciel est satisfaisant (inférieurs à 3 secondes).

L'interface principale du logiciel (**Figure 4.1**) comprend un menu regroupant les différentes procédures du traitement disponible.

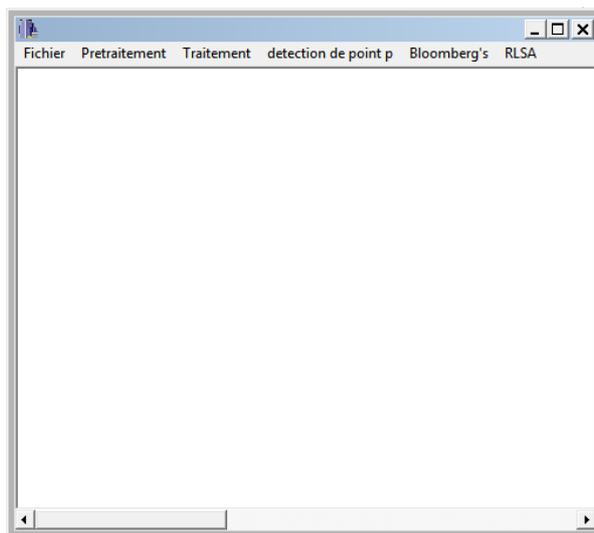
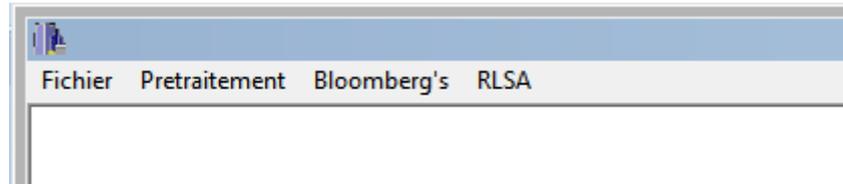


Figure 4.1 : L'interface principale du logiciel

❖ La barre menuue

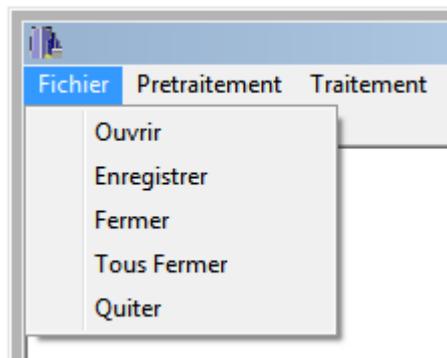
C'est une barre contenant des articles que l'on peut sélectionner à partir de la souris. L'activation d'un article fait appel, soit à une fonction, soit à une boîte de dialogue. Dans la barre Menu de l'interface principale, on trouve quatre sous menu : Fichier, Prétraitement, Bloomberg's et RLSA.



➤ Fichier

Le sous menu Fichier de l'interface principale contient les commandes suivantes :

- Ouvrir
- Enregistrer
- Fermer
- Tous Fermer
- Quitter



Ce menu a pour but de diriger et organiser les images dans l'interface de logiciel, ouvrir un nouveau projet, et de charger l'image à traitée (**Figure 4.2**), comme le montre la figure suivante :

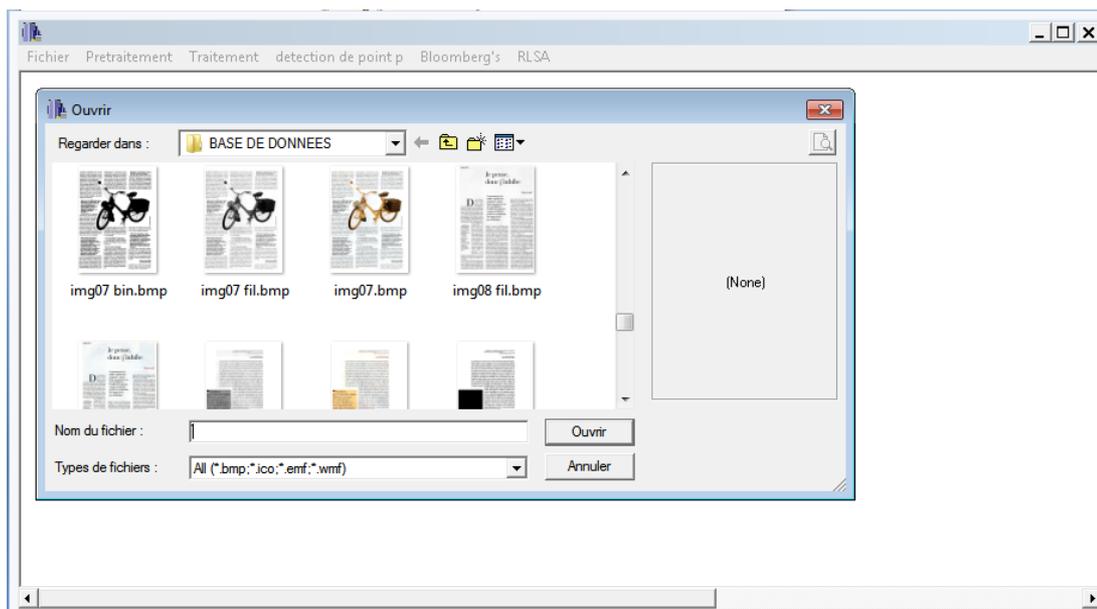


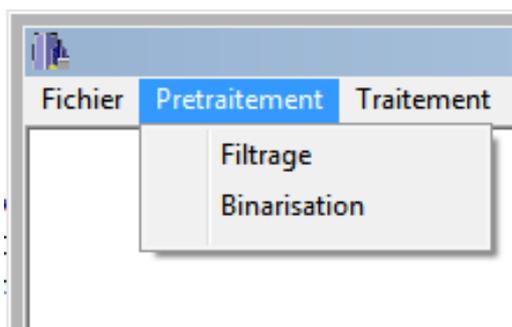
Figure 4.2 : Sélection d'une image à traiter

Ça permet aussi d'enregistrer les résultats obtenus, d'effacer l'image d'entrée ou l'image de résultat. La commande Quitter : permet de fermer le logiciel (Ctrl+f4).

➤ Prétraitement

Dans cette commande on trouve les méthodes de prétraitement utilisées pour la préparation de l'image à la segmentation. Pour fixer le Prétraitement on doit choisir entre deux programmes :

- Filtrage
- Binarisation



Pour pouvoir effectuer le prétraitement sur une image déjà sélectionnée, on sélectionne le sous menu Prétraitement, et on clique sur la commande Filtrage, une fenêtre doit apparaître sur l'interface (résultat de filtrage) (**figure 4.3**). On clique de nouveau sur Prétraitement et on choisit la commande Binarisation(**figure 4.4**).

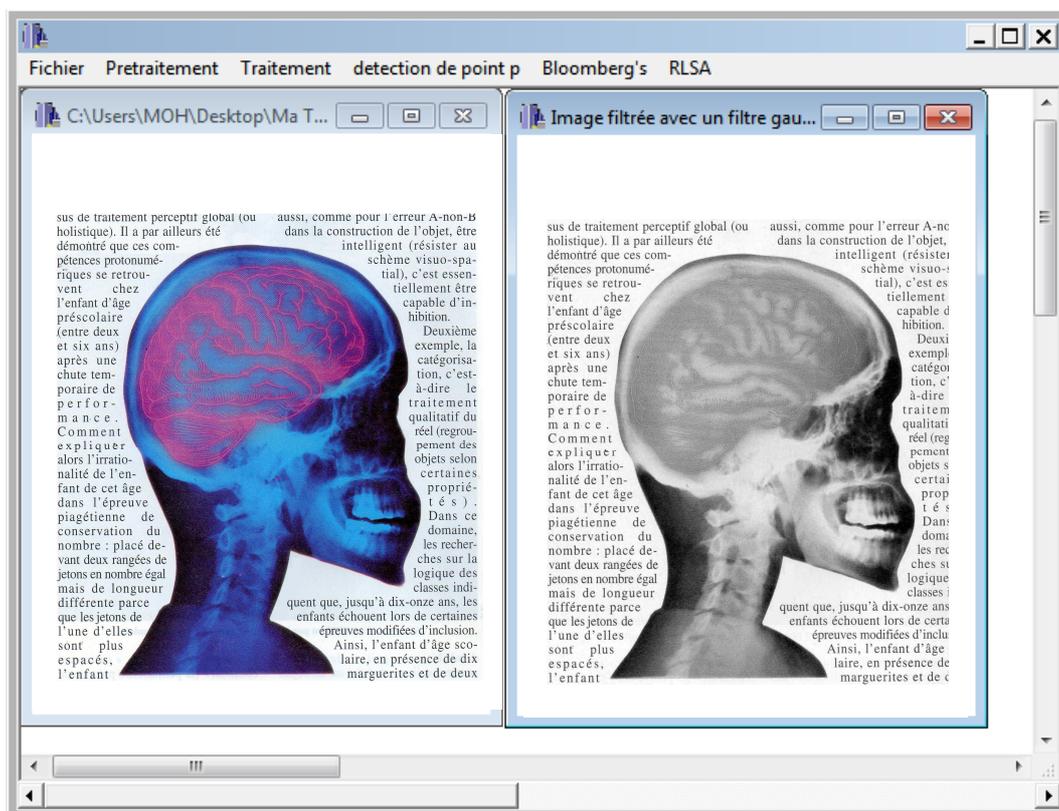


Figure 4.3 : Utilisation du Filtrage

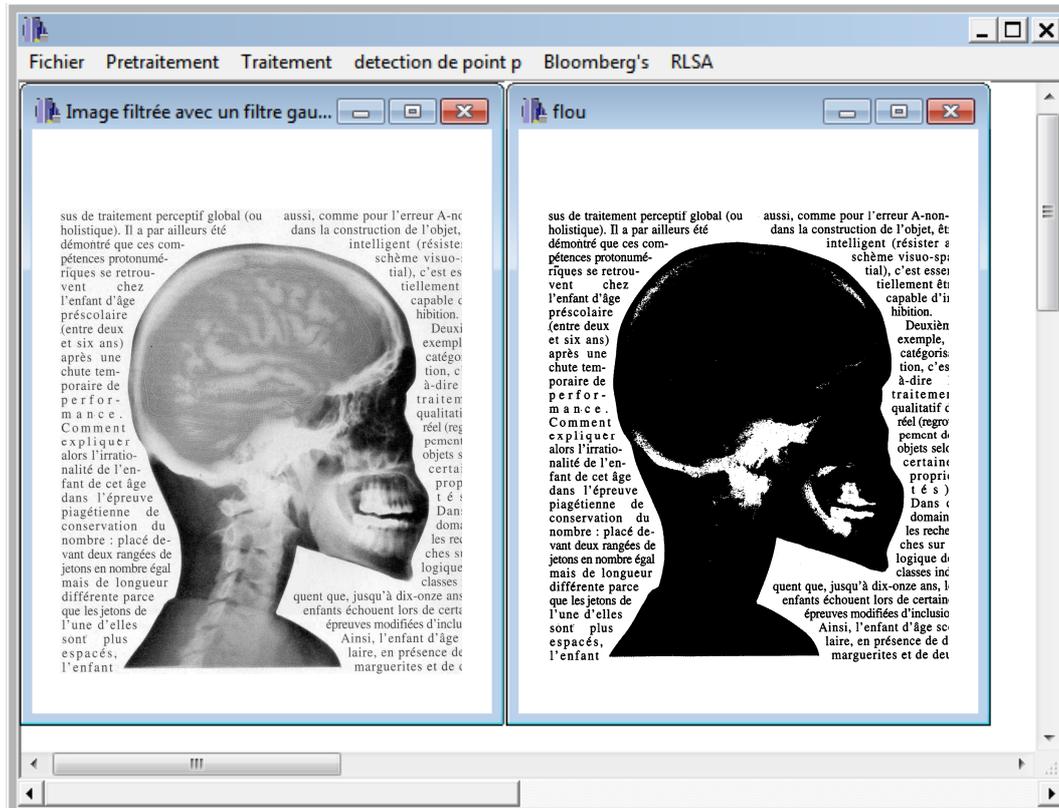


Figure 4.4 : Utilisation de Binarisation

➤ Bloomberg's

Le choix de sous menu Bloomberg's donne directement le résultat de pré-segmentation

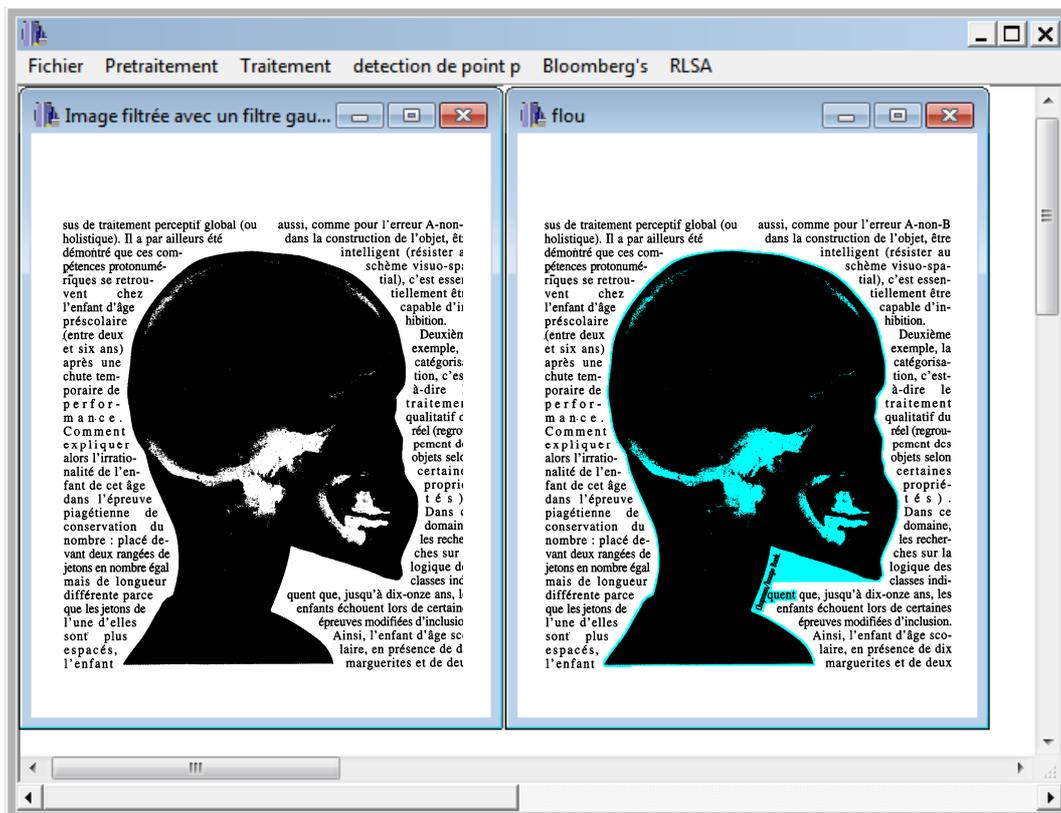


Figure 4.5 : Choix du sous menu Bloomberg's

➤ RLSA

En sélectionnant le sous menu RLSA, il va être affiché le résultat de segmentation finale c'est une image de document dont les la distinction entre les blocs textuels et non-textuels par couleurs différentes.

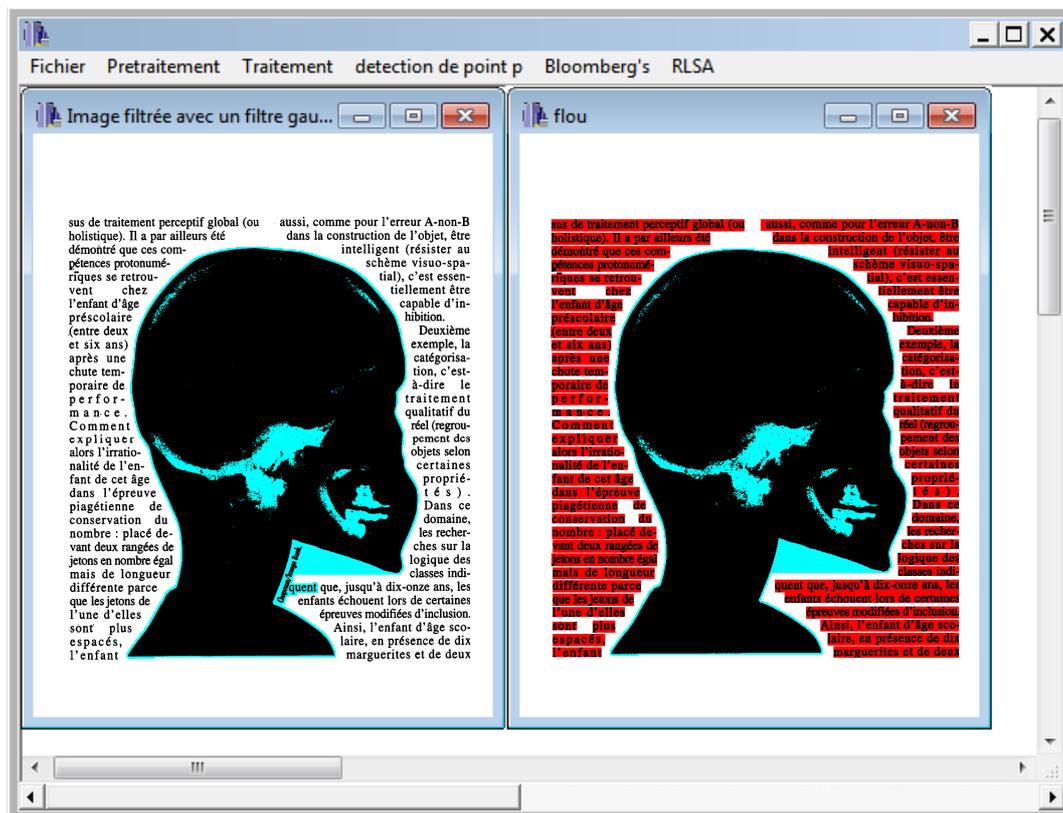


Figure 4.6 : Choix du sous menu RLSA

4.2 Images utilisées

Dans la suite de ce mémoire, nous allons utiliser des images de documents simples (textuels) et aussi de documents composites avec des structures différentes (de type Manhattan et non-Manhattan), nous présentons ci-dessous ces images avec les propriétés de chacune.

Le socio-constructivisme

La construction d'un savoir, bien que personnelle, dépend du contexte social. Le socio-constructivisme remet en cause les modèles psychologiques de développement cognitif et insiste sur les dimensions sociales contribuant à la formation des compétences.

Les recherches portant sur la notion de conceptions prélabées ou représentations initiales ne datent pas d'aujourd'hui et permettent de dégager trois grands modèles de réduction.

LE PREMIER est le modèle transmissif appelé aussi «magistral» ou «frontal» où le cerveau de l'élève est considéré telle une «craie morte» au niveau de laquelle viendraient s'imprimer les informations. C'est ce qu'on appelle aussi le modèle des «vessies communicants».

LE SECOND est appelé comportementaliste (behavoriste), modèle où l'apprentissage résulte d'une suite de conditionnements «stimulus-réponse». Les connaissances sont définies en termes de comportements observables attendus en fin d'apprentissage.

LE TROISIÈME est le modèle constructiviste et socio-constructiviste au niveau duquel les conceptions initiales ne sont pas simplement le point de départ du processus d'apprentissage mais

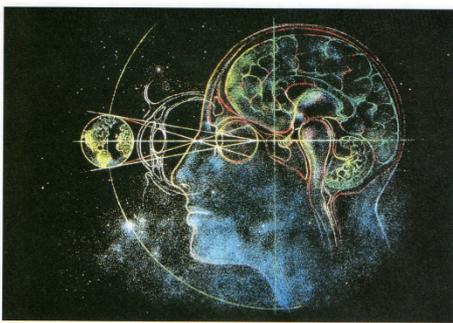
sont au cœur de celui-ci. Ainsi, l'élève construit lui-même ses connaissances car le développement intellectuel est un processus interne qui se réalise par étapes successives. Il ne peut, donc, «assimiler» de nouvelles connaissances que si ses structures mentales le permettent. Il serait donc insensé de vouloir «gaver» les élèves de savoirs qu'ils ne sont pas en mesure d'intégrer.

Dans cette perspective, l'enseignant doit impérativement respecter ce processus de développement personnel de l'élève et s'adapter à ses besoins. Son rôle consiste à observer, à diagnostiquer et à pratiquer ainsi une évaluation formative et une pédagogie différenciée puisqu'un nouveau savoir n'est effectif que s'il se construit en s'intégrant au réseau conceptuel de l'apprenant. Ausubel parle de «ponts cognitifs» car c'est la manière dont l'élève assimile les nouvelles connaissances qui est primordiale. Il peut y avoir alors des apprentissages significatifs en interaction avec les acquis préalables ou mécaniques sans liens avec les prérequis. Concrètement, afin que les enfants soient acteurs de leurs apprentissages, il est souhaitable que l'enseignant propose, stimule, observe, écoute et accepte que les élèves découvrent, explorent et entreprennent. En outre, il aide les élèves à verbaliser, à formuler, à sélectionner des indices, émettre des hypothèses et prendre conscience du problème qu'ils seront en mesure de résoudre. Par ailleurs, il gère la communication afin que les apprenants explicitent leurs

démarches, justifient leurs choix, les argumentent et tolèrent ceux d'autrui. Enfin, il propose des exercices adaptés pour qu'ils consolident, maîtrisent, approfondissent et réinvestissent les nouvelles connaissances.

En définitive, dans cette optique socio-constructiviste, apprendre c'est élaborer soi-même ses connaissances tout en passant nécessairement par une phase d'interaction, voire de conflit socio-cognitif avec autrui qui est d'autant plus efficace si l'enseignant est capable de gérer cette situation et de permettre à chacun de passer d'un niveau interpersonnel à un niveau intrapersonnel ; ce qui n'est incontestablement la phase d'appropriation et d'élaboration d'un langage intérieurisé. Il apparaît ainsi clairement qu'un système éducatif voulant offrir à tous les mêmes chances doit prendre en compte le contexte social, ce qui permettra un développement individuel de l'intelligence considérée comme une forme particulière d'adaptation. Adaptation nécessaire à l'ère des valeurs universelles, loin des schémas et modèles éducatifs rectilignes car pour progresser, il est impératif d'opérer constamment des «feedback», de se remettre en question et surtout avoir constamment à l'esprit que l'élève n'est pas une «page vierge» qu'il faut absolument remplir mais, surtout et avant tout, une personne «unique» qui a sa propre réalité car comme l'écrivit Watzlawick : «*Toute réalité est toujours inventée et c'est l'on construit soi-même sa propre réalité.*»

Image 1



Si passionnante que soit le géographe du cerveau, il reste à appréhender les mécanismes neurologiques des tâches cognitives.

dans une reconnaissance particulière (formes, couleurs, visages, mouvements, etc.). Très régionalisée, cette organisation cérébrale est assez similaire chez l'homme et chez le singe. À une exception près, mais elle est notable : il existe sans doute, chez l'homme, une aire spécifiquement consacrée à la reconnaissance des mots écrits. Une aire si sélective qu'elle ne s'active pas lorsque sont présentées, à la place de mots, des lettres assemblées au hasard... Plus généralement, que sait-on aujourd'hui des bases cérébrales du langage (parlé ou écrit), dont l'apprentissage et le traitement restent l'un des domaines de prédilection des cognitivistes ? Grâce aux travaux de Paul Broca sur l'aphasie, on sait depuis plus d'un siècle que cette tâche est principalement dévolue à l'hémisphère gauche. Mais du langage à l'écriture, où est la passerelle ? Comment le cerveau gère-t-il ces deux fonctions, distinctes et pourtant étroitement liées ? Pour en savoir plus, l'équipe d'imagerie neuro-fonctionnelle du service hospitalier Frédéric-Joliot (Orsay) eut récemment l'idée d'examiner, chez des sujets japonais, la localisation du langage écrit en fonction du système d'écriture employé : soit les caractères «kanjis», morphogrammes ayant chacun une identité graphique et une signification ; soit les «kanas», symboles syllabiques plus proches de notre système alphabétique. «Les ré-

ultats ont été plus subtils que prévu. Quelle que soit l'écriture, l'hémisphère gauche prédomine», résume Denis le Bihan, directeur de ces recherches à Orsay. L'IRM a toutefois révélé deux différences. D'une part, l'utilisation des kanjis implique une activité cérébrale globale deux à trois fois plus grande que celle des kanas. D'autre part, les kanjis sollicitent, en plus des zones habituelles de l'hémisphère gauche, certaines zones pariétales et temporales de l'hémisphère droit. Au-delà de cette passionnante géographie du cerveau, il reste à appréhender, dans leur détail, les mécanismes neurologiques impliqués dans les tâches cognitives. Le cerveau humain se compose d'un extraordinaire réseau de cellules nerveuses, les neurones, dont le nombre est estimé à environ mille milliards. Chacun échange des informations avec ses voisins sous forme de courant électrique, par le biais de diverses connexions (environ dix mille par neurone), axones, dendrites et synapses. Les chercheurs en ont maintenant convaincus : c'est dans la plasticité de ces synapses, dans leur faculté à s'organiser durablement en réseaux qui résident les véritables secrets des phénomènes d'apprentissage et de mémoire. Seule certitude : dans ce fouillis de neurones repose un trésor, un concentré d'esprit qu'il est convenu d'appeler

«intelligence». Tout à la fois «faculté de connaître, de comprendre», «ensemble des fonctions mentales ayant pour objet la connaissance conceptuelle et rationnelle» et «aptitude d'un être vivant à adapter à des situations nouvelles», tantôt considérée comme spécifique à l'homme, tantôt reconnue comme partagée par l'ensemble des animaux supérieurs, la notion n'en continue pas moins de se dérober à l'entendement. Ici s'arrête la science, la commencent les suppositions. Y compris les plus harmonieuses. Il y a quelques années, une très sérieuse étude américaine, publiée dans la revue *Nature*, suggérait ainsi que l'écoute assidue de Mozart rendait transitoirement plus intelligent. En soumettant un groupe de trente-six personnes aux tests de QI, les chercheurs de l'université de Californie (Irvine) se sont aperçus que leurs résultats étaient en moyenne de 8 à 9 points supérieurs lorsque les tests étaient précédés de l'audition de la sonate pour deux pianos en D majeur (K488). «On peut imaginer qu'en écoutant une telle musique on stimule des réseaux neuronaux qui jouent un rôle important dans la connaissance», avaient supposé les chercheurs. Personne, depuis lors, n'a confirmé l'hypothèse... ni ne l'a contredite. ■

À lire : Le Cerveau en action, sous la direction de Stanislas Dehaene, PUF (1997), 290 p., 198 F.

Image 3



Expert Systems with Applications 33 (2007) 968–977

Expert Systems with Applications
www.elsevier.com/locate/eswa

Automated barcode recognition for smart identification and inspection automation

Sherin M. Youssef*, Rana M. Salem

Department of Computer Engineering, College of Engineering and Technology, Arab Academy for Science and Technology, Alexandria, Egypt

Abstract

Barcodes have been widely used in many industrial products for automatic identification in data collection and inventory control purposes. It is well known that in many stores laser bar-code readers are used at check-out counters. However, there is a major constraint when this tool is used. That is, unlike traditional camera-based picturing, the distance between the laser reader (sensor) and the target object is close to zero when the reader is applied. This may result in inconvenience in inspection automation because the human operator has to manipulate either the sensor or the objects. For the purpose of in-store or inspection automation, the human operator needs to be removed from the process, i.e. a robot with visual capability is required to play an important role in such a system. Moreover, an automated system is required to find, locate and decode barcodes on various document images even with low resolution compared with laser barcode readers (about 15,000 dots per inch) and can handle damaged bar codes. This paper proposes a smart barcode detection and recognition system (SBDRC) based on fast hierarchical Hough transform (HHT). The back-propagation neural network (BPNN) is selected as a powerful tool to perform the recognition process. The paper presents an effective method to utilize the specific graphic features of barcodes for positioning and recognition purposes even in case of distorted barcodes. The first step the system has to perform is to locate the position and orientation of the barcode in the required material document. Secondly, the proposed system has to segment the barcode. Finally, a trained back-propagation neural network is used to perform the barcode recognition task. Experiments have been conducted to corroborate the efficiency of the proposed method.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Barcodes; Automated identification

1. Introduction

The automated identification of goods and persons is an attractive application for areas of industry, security, postal letters and smart cards inspection. A barcode is a machine-readable representation of information in a visual format on a surface (Sutton, 2002). Barcodes have been widely used in many industrial products for automatic identification in data collection and inventory control purposes (Chang, Lo, & Hsieh, 1997). Originally barcodes store data in the width and spacing of printed parallel lines, but they

* Corresponding author.

E-mail addresses: sherin@aat.edu.eg, smy@cs.net.eg (S.M. Youssef).

also come in patterns of dots, concentric circles, and hidden in images (Fig. 1 shows examples of different types of barcodes). Barcodes can be read by optical scanners called barcode readers or scanned from an image. Barcodes are widely used to implement Auto Id Identification systems. With the advent of in-store automation and smart cards identification, the existence of standard barcodes for information exchange and identification is indispensable. At many points of sale, the laser barcode reader is conventionally used. However, there is a major constraint on the laser barcode reader (Liao, 1995). That is, unlike traditional camera-based picturing, the distance between the laser reader (sensor) and the target object is close to zero when the reader is applied. This may result in inconvenience in inspection automation because the human operator has to manipulate either the sensor or the objects. For the

image 2

Under the option EDIT select PASTE SPECIAL. A dialog box will open, select paste picture, then click OK. Your figure should now be in the Word Document.

If you are preparing images in TIFF, EPS, or PS format, note the following. High-contrast line figures and tables should be prepared with 600 dpi resolution and saved with no compression, 1 bit per pixel (monochrome), with file names in the form of "fig3.tif" or "table.tif".

Photographs and grayscale figures should be prepared with 300 dpi resolution and saved with no compression, 8 bits per pixel (grayscale).

Sizing of Graphics

Most charts graphs and tables are one column wide (3 1/2 inches or 21 picas) or two-column width (7 1/16 inches, 43 picas wide). We recommend that you avoid sizing figures less than one column wide, as extreme enlargements may distort your images and result in poor reproduction. Therefore, it is better if the image is slightly larger, as a minor reduction in size should not have an adverse affect the quality of the image.

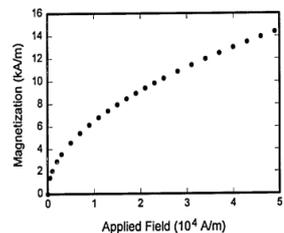


Fig. 1. Magnetization as a function of applied field. Note that "Fig." is abbreviated. There is a period after the figure number, followed by two spaces. It is good practice to explain the significance of the figure in the caption.

fonts are as follows: Open Type Fonts: Times Roman, Helvetica, Helvetica Narrow, Courier, Symbol, Palatino, Avant Garde, Bookman, Zapf Chancery, Zapf Dingbats, and New Century Schoolbook.

Print Color Graphics Requirements

IAENG accepts color graphics in the following formats: EPS, PS, TIFF, Word, PowerPoint, Excel, and PDF. The resolution of a RGB color TIFF file should be 400 dpi.

Web Color Graphics

IAENG accepts color graphics in the following formats: EPS, PS, TIFF, Word, PowerPoint, Excel, and PDF. The resolution of a RGB color TIFF file should be at least 400 dpi.

Your color graphic will be converted to grayscale if no separate grayscale file is provided. If a graphic is to appear in print as black and white, it should be saved and submitted as a black and white file. If a graphic is to appear in print or on IAENG web site in color, it should be submitted as RGB color.

Graphics Checker Tool

The authors should check their graphic files against the set of rules for compliance with IAENG requirements. These requirements are designed to ensure sufficient image quality so they will look acceptable in print.

E. Copyright

The copyright to the Contribution identified above is transferred to International Association of Engineers, (hereinafter called IAENG). The copyright transfer covers the sole right to print, publish, distribute and sell throughout the world the said Contribution and parts thereof, including all revisions or versions or future editions thereof and in any medium, such as in its electronic form (offline, online), as well as to translate, print, publish, distribute and sell the Contribution in any foreign languages and throughout the world. IAENG will take, either in its own name or in that of the Author, any necessary steps to protect these rights

image 4

L'art contemporain

nels et l'attachement à la terre natale, suggère bel et bien, comme le démontre Philippe Dagen dans son ouvrage *La Haïne de l'art, « une généalogie de la pensée réactionnaire française »*. Vingt ans plus tard, *Le Monde* « une avant-garde officielle et indifférente à tout ce qui n'est pas son dogme ». Jean Clair en vient à formuler un cri d'alarme : « Cette sollicitude de l'État a conduit à faire des artistes des assistés sociaux, avec peu préparés à la lutte sur un marché libre et mondialisé. » Et ce par-tisan d'un libéralisme économique qui ne devrait pas égarer la culture agreste que, pour réussir à s'imposer sur les scènes internationales, rien ne vaut l'« exaltation » des « valeurs culturelles propres », comme l'on fait, selon lui, les écoles allemande, italienne et londonienne, et non pas l'« internationalisme », source de tous les maux. Marc Fumaroli écrivait aussi qu'il s'en pré-« nous ont artises qu'à un certain discours international sur l'art ».

Cette polémique qui s'est cristallisée autour de déclarations prétextées « L'art contemporain est nul » ou « Il n'y a plus d'artistes en France » — exclut et condamne à priori les artistes et les œuvres. Que l'art n'ait plus droit de cité est révélateur d'une crise qui n'est pas celle de l'art mais d'un décalage entre le pouvoir croissant des experts en charge de gérer les normes du bon goût et des œuvres reléguées en cinquième roue du carrosse culturel.

Le Difficile Partage du gâteau est un projet d'artiste français dont l'œuvre, depuis une dizaine d'années, oppose aux rouages bureaucratiques des spéculations financières et intellectuelles du système culturel l'humilité de ses tableaux aux proportions réduites. Situés à une hauteur de huit et 40 de large.

Difficile Partage du gâteau est caracté-ristique de ce qui s'appelle autrefois les genres mineurs ou nature morte vul-gaire peinte d'ambiance, ou des dis-crètes réflexions sur le savoir, connues au XVIII^e siècle sous le nom de vanitas. Sur la toile, la masse compacte d'un baba s'élève sobrement d'une assiette blanche posée sur une assiette marron, dont le seul ornement est une vague liseret doré. La tension dramati-que de la scène est entretenue par sept fourchettes acérées qui s'apprennent à le déchiqueter. La facture désinvolte et naïve de la toile mêle inextricablement

Une palette d'actions pour les jeunes

Peut-on apprendre à aimer l'art contemporain aux jeunes? Sylvie Cohen, documentaliste, et Alain Nahmias, plasticien, en sont convaincus. C'est pourquoi ils passent aujourd'hui leur temps à démarcher des partenaires institutionnels pour pouvoir créer en Ile-de-France le premier centre d'art contemporain pour les enfants. La délégation aux arts plastiques et la Fondation Coprin ont d'ores et déjà donné leur accord de principe. Reste à trouver des terrains et des mentions importantes pour mener à bien un projet qui nécessite d'abord un lieu, « si possible dans un parc qui soit déjà un but de promenade familiale ».

Sylvie Cohen et Alain Nahmias envisagent une structure d'offres d'œuvres originales représentatives des tendances actuelles, des circuits de visite seraient conçus avec des conseillers pédagogiques. Afin d'éveiller le regard et encourager la création de l'enfant, un système d'ateliers et de conférences compléterait le dispositif. Pourquoi un lieu spécifique? « D'abord pour que les œuvres soient accrochées à hauteur des enfants, ensuite pour que cela s'approprient le centre. Selon Sylvie Cohen et Alain Nahmias, l'art contemporain est « un art ludique, interactif, plus facile d'accès que les autres ». Un avis qui partage Philippe Coubertegues, enseignant en arts plastiques détaché à la Galerie nationale du Jeu de Paume : « La littérature critique et analytique crée un écart de la perception. On peut avoir une approche plastique et sensible, sans connaissances historiques approfondies. Les enfants le comprennent plus que les adultes, qui se sentent davantage

De plus, le service culturel du Jeu de Paume s'est résolument attaché à développer les activités en direction des scolaires et, faisant, à apposer à preuve que le fossé entre l'art contemporain et le public n'est pas incommensurable, pour peu que des actions concertées favorisent le contact direct avec les œuvres. Un des principes retenus est celui de partenariats avec des collectivités qui offrent lieu à des sorties au Jeu de Paume et à des ateliers en milieu scolaire. Au dernier, plus de deux cents enseignants ont été accueillis pour des visites commentées. Une centaine de professeurs de toutes disciplines ont été invités à participer à des ateliers de formation pédagogique. Au printemps prochain, la Galerie nationale du Jeu de Paume devrait organiser un colloque international sur la médiation entre les institutions d'art contemporain et l'enseignement. M.S.

le bon et le mauvais goût, l'acidité reuse des couleurs artificielles et les brousses mineurs. Cet effet est accentué par l'usage alterné de la peinture à l'huile, médium noble et traditionnel avec des nuances profondes de peinture acrylique, aux tons artificiels. Ainsi, le tableau allégorique du goût contemporain entend jeter quelque confusion quant aux recettes bien établies de l'authentique tradition et de l'empire modernité. En présentant l'œuvre d'art comme un gâteau agressé par sept fourchettes, Mayaux fait allusion à ce péché capital qui consiste à réduire vulgairement l'expérience esthétique à une dégustation.

Wassily Kandinsky, l'inventeur de la peinture abstraite, se plaignait déjà de ces regards concupiscent qui goûtent la peinture comme un pâté. Le peintre Paul Klee n'avait guère non plus de considération pour ces esprits pressés qui en restent à « brouter » la surface de la toile. Selon Mayaux, l'art appartient au domaine du simulacre, et le but d'une œuvre est de jeter un soupçon salubre sur les apparences pour les engager à le dépasser. L'artiste Marcel Duchamp disait : « Ce sont les regards qui font l'art ». Dans le tableau de Philippe Mayaux, ces regards se sont transformés en fourchettes, et la contemplation s'est muée en consommation. Cette absence de regard favorise l'aveuglement — « Il n'y a plus d'artistes », *L'art contemporain est nul* —, un discours sans objet, une querelle d'experts dont le seul objectif est de prouver l'existence de leurs expertises dans ce qu'il est désormais convenu d'appeler *Le Difficile Partage du gâteau*.

(1) La Haïne de l'art, Philippe Dagen, Grasset, 2009, 198 p., 18 €. (2) L'art contemporain, Yves Michard, 1982, coll. « Universitaires », 198 p., 28 €. (3) *Le Monde* de l'art, « L'art contemporain est nul », par Jean-Claude Guillemin, par hasard et par exemple, Jacques Grégoire, Grasset, coll. « Figures », 215 p., 12 €. (4) *Le Monde* de l'art contemporain, Anne-Cécile Guille, Grasset, coll. « La couleur des idées », 177 p., 19 €.

periods during 1967-71. Virtually all the study area was cleared of vegetation by March 1971.

THE WATER-BUDGET EQUATION

Two water parameters are significant in defining ET

(Also referred to as Panovir, panovirus and Panovirus gelation)

EXPLANATION

Boundary of Gila River flood plain

Cross section and number

Gauging station on the Gila River

ARIZONA

STUDY AREA

San Carlos Reservation

San Carlos

the area, ΔC = change in channel storage in the Gila River, F = average precipitation on the area, ΔS_{un} = average change in moisture content in the unsaturated soil zone located immediately below the land surface.



Image 5

image 6

ainsi montré que les bébés de quatre-vingt mois sont capables de détecter la transgression ou la « conservation » du nombre lorsqu'on leur présente des événements numériques impossibles ou possibles (dans une maquette de théâtre où les unités à quantifier sont des Mickey). Le traitement de l'information numérique réalisé par le bébé repose sur un processus analytique de « calcul précis » sur un processus de traitement perceptif global (ou holistique). Il par ailleurs été démontré que ces compétences proumées, se retrouvent chez l'enfant d'âge préscolaire (entre deux et six ans) après une chute temporaire de performance. Comment expliquer alors l'irrationalité de l'enfant de cet âge dans l'épreuve piagetienne de conservation du nombre ? placé devant deux rangées de jetons en nombre égal mais de longueur différente parce que les jetons de l'une d'elles sont plus espacés, l'enfant considère qu'il y en a plus là où c'est plus long ? Pour Piaget, l'enfant d'âge préscolaire serait encore fondamentalement intuitif « préopérateur ». Pour traiter l'information, il se fonderait sur la longueur, selon un processus perceptif global. Il n'aurait donc pas encore construit le « schéma du nombre » lui permettant de procéder à un calcul précis. Les recherches récentes qui viennent d'être évoquées conduisent à douter de cette interprétation. En effet, ne serait-on pas plutôt ici, à l'occasion d'un nouveau « *plâta temps* », en présence de la coexistence de « rationnel construit » et de l'« *irrational présumé révolu* » ? Dans cet exemple, on trouve d'une part la capacité d'un traitement numérique analytique et précis et d'autre part la transgression du nombre dans l'épreuve de

conservation. Cette dernière épreuve n'est-elle pas avant tout une épreuve d'« *interférence* » « *nombre/longueur* » critique de la capacité à résister au schéma visuo-spatial « *longueur = nombre* » ? Il semble bien que les situations de conservation et d'inclusion des classes (*voilà plus loin* de Piaget ont plus à voir avec la capacité à résister aux interférences qu'avec la capacité de l'enfant à comprendre la logique sous-jacente. Ici aussi, comme pour l'erreur A-non-B dans la construction de l'objet, être intelligent n'est pas capable d'inhibition. Deuxième exemple, la catégorisation, c'est-à-dire le traitement de la qualité du réel (regroupement des objets selon certaines propriétés) (e s). Dans ce domaine, les recherches sur la logique des classes indifférentes, jusqu'à dix-neuf ans, les enfants échouent lors de certaines épreuves modifiées d'inclusion. Ce sont les travaux relatifs aux biais de raisonnement qui font ressortir avec le plus de force l'irrationalité de l'adolescent et de l'adulte : la tendance systématique à prendre en considération des facteurs non pertinents pour la tâche à résoudre et à ignorer les facteurs pertinents. Faut-il conclure, face à ce constat d'irrationalité, que l'essence du raisonnement humain n'est pas conforme à la logique, et que la compétence déductive est condamnée à être court-circuitée par des biais cognitifs ? Les données des recherches utilisant la mesure de l'apprentissage expérimental à l'inhibition vont dans le sens d'une présomption de rationalité, dans la mesure où les biais qui sous-tendent les erreurs de raisonnement ne semblent pas relever d'un défaut de logique mais d'un défaut d'inhibition. Il s'agit donc chez l'adulte, comme chez le bébé et l'enfant, depuis l'erreur A-non-B, premier ancêtre

d'inclusion appliqué des sept-huit ans lors de l'épreuve piagetienne classique (dans laquelle l'enfant compare correctement l'1 y a plus de B que de A, et d'autre part, l'1 « *irrational présumé révolu* » : la transgression de ce schéma, jusqu'à dix-neuf ans, lors des épreuves modifiées d'inclusion ? Piaget lui-même, dans l'un de ses derniers ouvrages, *vers une logique des significations*, montre que le raisonnement de l'absence de piège de quantification extensionnelle des A et des B, l'inclusion soulevée de problème, il apparaît en effet, au regard des recherches récentes, que l'intelligence requise par les épreuves modifiées d'inclusion, réussies vers dix-neuf ans, consiste avant tout à inhiber la routine de l'ajout et du retrait, et à intégrer ensuite les classes en relation d'inclusion. Trop de compétences nul

Le public en appelle alors au rôle critique des institutions qui doivent dédicier les valeurs culturelles et ne pas s'associer à ce qu'il considère comme une entreprise de démolition. Le conflit esthétique vite alors à la politique et la morale. C'est pourquoi les débats les plus violents sont liés à des commandes publiques. Les colonnes de Buren en sont un exemple : le rejet esthétique d'une proposition heurte les attentes traditionnelles se combinant à une réaction politique suscitée par une initiative ministérielle. Celle-ci a été en effet à contre-courant de ce que l'on exige d'une commande publique, à savoir un minimum de consensus. Selon vous, les différends entre défenseurs et détracteurs de l'art contemporain relèvent moins de divergences d'appréciation esthétique que de différences de catégories cognitives. Les quelles sont-elles ?

Au siècle dernier, il s'agissait de décider si c'était beau ou non, si les peintres savaient peindre. En somme, c'était une question de goût. Aujourd'hui, on est en face d'un désaccord fondamental sur les catégories de classement : est-ce ou non de l'art ? Le problème n'est pas d'évaluer la valeur esthétique d'une œuvre, mais de définir sa nature même. Les conflits ne peuvent donc pas se résoudre lorsqu'on ne parle pas de la même chose : pis, si se creusent. Autre rupture historique : la place fondamentale qu'occupent aujourd'hui les institutions artistiques (centres d'art, musées, FRAC), notamment depuis quinze ans. Ce sont elles qui, par leur politique d'achat, accèdent les avant-gardes en temps direct, sans attendre plusieurs générations. Quels sont les effets de ce face-à-face ?

Il met entre parenthèses les marchés privés. Les œuvres de jeunes artistes peuvent être achetées et exposées par des structures d'Etat sans être passées par le circuit des collectionneurs et des galeries. Depuis la politique de Jack Lang, la prise en charge institutionnelle de l'avant-garde s'est accélérée. Il en résulte des effets positifs mais aussi des effets pervers. L'avant-garde est traditionnellement associée à une position progressive, marginale par rapport à l'établissement. Aujourd'hui, les institutions sont de son côté. Or celui qui désire conserver une posture de fronteur doit aller à contre-courant de ce qui est admis par les institutions, autrement dit contre la reconnaissance. Les oppositions droite/gauche, conservatisme/progressisme, tradition/innovation sont inversées. Depuis Pourquoi la pédagogie est-elle si difficile auprès des adultes ?

En raison de la nature d'œuvres profondément transgressives. Pour évaluer l'intérêt et la pertinence des propositions, juger que celle-ci est redondante ou qu'elle est radicalement neuve, il faut avoir une culture de l'art contemporain. Or, faute d'enseignement artistique, cette culture n'existe pas. De surcroît, il faut souvent sortir d'une posture d'admiration et d'émotion immédiate pour entrer dans une démarche plus intellectuelle afin de comprendre ce que fait l'œuvre à nos catégories mentales. A force de surenchère, d'une extension toujours plus grande des frontières de l'art, l'art contemporain ne court-il pas à une impasse, à ce moment fatal où auront été expérimentés tous les processus de déconstruction des critères qui définissent traditionnellement l'œuvre ?

C'est un risque. Encourager systématiquement, qui permet de comprendre les renversements d'opinions et pourquoi les spécialistes et responsables de l'art contemporain sont dans une position si inconfortable sur le plan éthique et politique ; ils se sont engagés dans le soutien à l'avant-garde, à l'innovation difficile, et leurs détracteurs les présentent comme des apprentis-châssés réactionnaires défiant les œuvres les plus standards, à l'image des promoteurs de l'art pompier au siècle dernier. Propos recueillis par Macha Scellius



Image 7

image 8

Figure 4.7 : Images utilisées

Chaque image présentée dans la Figure 5.4, prendra le même nom dans tout le reste de ce document.

- **Image1** : Image de document textuel de structure Manhattan, de taille 2472×3425 ;
- **Image2** : Image de document composite de structure Manhattan, de taille 2509×09 contenant un petit bloc d'image;
- **Image3** : Image de document composite, de structure Manhattan, de taille 2190×3125 , contenant une image;
- **Image4** : image de document composite, de structure Manhattan, de taille 2472×3425 , contenant un graphe;
- **Image 5** : Image de document textuel de structure non Manhattan, de taille 2172×3114 , contenant un bloc encadré par les autres blocs;
- **Image6** : Image de document graphique de structure non Manhattan, de taille 621×560 ;
- **Image7** : Image de document composite de structure non Manhattan, de taille 2232×3125 , contenant une image;
- **Image8** : Image de document composite de structure non Manhattan, de taille 2256×3125 , contenant une image;

4.3 Résultats expérimentaux

Dans cette partie, les méthodes de segmentation proposées ont été mise en œuvre, dans le contexte de la détection des blocs constituant l'image de document, et de la discrimination entre blocs textuels et non-textuels.

Cette partie est composée des étapes suivantes :

- Présentation des résultats obtenus après prétraitement, à savoir le filtrage gaussien, la binarisation en deux étapes (seuillage global suivi par un seuillage local).
- Présentation des résultats obtenus après Pré-segmentation par extraction des blocs non textuels en utilisant l'algorithme de Bloomberg.
- Puis les résultats obtenus dans la segmentation finale, qui comprend l'approche par extraction de lignes de texte.

4.3.1 Résultats du prétraitement

4.3.1.1 Résultats du filtrage Gaussien

Pour une meilleure distinction entre les différentes zones adjacentes des images de documents, et pour réduire le bruit, il est nécessaire de filtrer ces images avant de les utiliser dans d'autres traitements tels que la segmentation, pour cela, nous présentons ci-dessous, l'effet du filtre Gaussien de taille 3×3 .

Le socio-constructivisme

La construction d'un savoir, bien que personnelle, dépend du contexte social. Le socio-constructivisme remet en cause les modèles psychologiques du développement cognitif et insiste sur les dimensions sociales contribuant à la formation des compétences.

Les recherches portant sur la notion de conceptions préalables ou représentations initiales ne datent pas d'aujourd'hui et permettent de dégager trois grands modèles de l'éducation.

LE PREMIER est le modèle transmissif appelé aussi «magistral» ou « frontal » où le cerveau de l'élève est considéré telle une «craie morte» au niveau de laquelle viendraient s'imprimer les informations. C'est ce qu'on appelle aussi le modèle des « vases communicants ».

LE SECOND est appelé comportementaliste (behaviouriste), modèle où l'apprentissage résulte d'une suite de conditionnements «stimulus-réponse ». Les connaissances sont définies en termes de comportements observables attendus en fin d'apprentissage.

LE TROISIÈME est le modèle constructiviste et socio-constructiviste au niveau duquel les conceptions initiales ne sont pas simplement le point de départ du processus d'apprentissage mais

sont au cœur de celui-ci. Ainsi, l'élève construit lui-même ses connaissances car le développement intellectuel est un processus interne qui se réalise par étapes successives. Il ne peut, donc, « assimiler » de nouvelles connaissances que si ses structures mentales le permettent. Il serait donc insensé de vouloir «gaver» les élèves de savoirs qu'ils ne sont pas en mesure d'intégrer.

Dans cette perspective, l'enseignant doit impérativement respecter ce processus de développement personnel de l'élève et s'adapter à ses besoins. Son rôle consiste à observer, à diagnostiquer et à pratiquer ainsi une évaluation formative et une pédagogie différenciée puis qu'un nouveau savoir n'est effectif que s'il se construit en s'intégrant au réseau conceptuel de l'apprenant. Ausubel parle de « ponts cognitifs » car c'est la manière dont l'élève assimile les nouvelles connaissances qui est primordiale. Il peut y avoir alors des apprentissages significatifs en interaction avec les acquis préalables ou mécaniques sans liens avec les prérequis. Concrètement, afin que les enfants soient acteurs de leurs apprentissages, il est souhaitable que l'enseignant propose, stimule, observe, écoute et accepte que les élèves découvrent, explorent et entreprennent. En outre, il aide les élèves à verbaliser, à formuler, à sélectionner des indices, émettre des hypothèses et prendre conscience du problème qu'ils seront en mesure de résoudre. Par ailleurs, il gère la communication afin que les apprenants énoncent leurs

démarches, justifient leurs choix, les argumentent et tolèrent ceux d'autrui. Enfin, il propose des exercices adaptés pour qu'ils consolident, maîtrisent, approfondissent et réinvestissent les nouvelles connaissances.

En définitive, dans cette optique socio-constructiviste, apprendre c'est élaborer soi-même ses connaissances tout en passant nécessairement par une phase d'interaction, voire de conflit socio-cognitif avec autrui qui est d'autant plus efficace si l'enseignant est capable de gérer cette situation et de permettre à chacun de passer d'un niveau interpersonnel à un niveau intrapersonnel ; ce qui facilite incontestablement la phase d'appropriation et l'élaboration d'un langage intériorisé. Il apparaît ainsi clairement qu'un système éducatif voulant offrir à tous les mêmes chances doit prendre en compte le contexte social, ce qui permettra un développement individuel de l'intelligence considérée comme une forme particulière d'adaptation. Adaptation nécessaire à l'ère des valeurs universelles, loin des schémas et modèles éducatifs rectilignes car pour progresser, il est impératif d'opérer constamment des «feed-back», de se remettre en question et surtout avoir constamment à l'esprit que l'élève n'est pas une « page vierge » qu'il faut absolument remplir mais, surtout et avant tout, une personne «unique » qui a sa propre réalité car comme l'écrit Watzlawick : « Toute réalité est toujours inventée » et « l'on construit soi-même sa propre réalité ».

Image 1 filtrée

Image 2 filtrée



Expert Systems with Applications 33 (2007) 968–977

Expert Systems with Applications

www.elsevier.com/locate/eswa

Automated barcode recognition for smart identification and inspection automation

Sherin M. Youssef^a, Rana M. Salem

^a Department of Computer Engineering, College of Engineering and Technology, Arab Academy for Science and Technology, Alexandria, Egypt

Abstract

Barcodes have been widely used in many industrial products for automatic identification in data collection and inventory control purposes. It is well known that in many stores laser barcode readers are used at check-out counters. However, there is a major constraint when this tool is used. That is, unlike traditional camera-based picturing, the distance between the laser reader (sensor) and the target object is close to zero when the reader is applied. This may result in inconvenience in inspection automation because the human operator has to manipulate either the sensor or the objects. For the purpose of in-store or inspection automation, the human operator needs to be removed from the process, i.e. a robot with visual capability is required to play an important role in such a system. Moreover, an automated system is required to find, locate and decode barcodes on various document images even with low resolution compared with laser barcode readers (about 15,000 dots per inch) and can handle damaged bar codes. This paper proposes a smart barcode detection and recognition system (SBDR) based on fast hierarchical Hough transform (HHT). The back-propagation neural network (BPNN) is selected as a powerful tool to perform the recognition process. The paper presents an effective method to utilize the specific graphic features of barcodes for positioning and recognition purposes even in case of distorted barcodes. The first step the system has to perform is to locate the position and orientation of the barcode in the required material document image. Secondly, the proposed system has to segment the barcode. Finally, a trained back-propagation neural network is used to perform the barcode recognition task. Experiments have been conducted to corroborate the efficiency of the proposed method.

© 2006 Elsevier Ltd. All rights reserved.

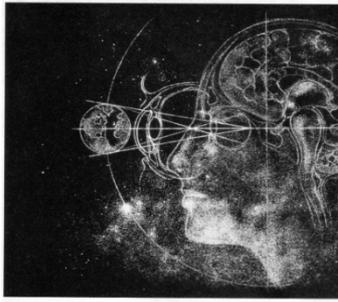
Keywords: Barcode; Automated identification

1. Introduction

The automated identification of goods and persons is an attractive application for areas of industry, security, postal labels and smart cards inspection. A barcode is a machine-readable representation of information in a visual format on a surface (Sutton, 2002). Barcodes have been widely used in many industrial products for automatic identification in data collection and inventory control purposes (Chang, Lo, & Hsieh, 1997). Originally barcodes store data in the width and spacing of printed parallel lines, but they

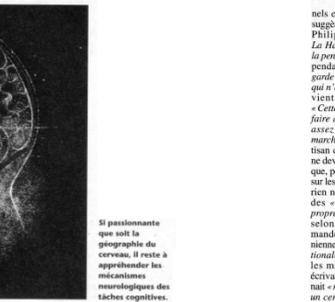
also come in patterns of dots, concentric circles, and hidden in images (Fig. 1 shows examples of different types of barcodes). Barcodes can be read by optical scanners called barcode readers or scanned from an image. Barcodes are widely used to implement Auto Id identification systems. With the advent of in-store automation and smart cards identification, the existence of standard barcodes for information exchange and identification is indispensable. At many points of sale, the laser barcode reader is conventionally used. However, there is a major constraint on the laser barcode reader (Liao, 1995). That is, unlike traditional camera-based picturing, the distance between the laser reader (sensor) and the target object is close to zero when the reader is applied. This may result in inconvenience in inspection automation because the human operator has to manipulate either the sensor or the objects. For the

^a Corresponding author.
E-mail addresses: sherin@qaat.edu.eg, smy@cc.nott.ac.uk (S.M. Youssef).



dans une reconnaissance particulière (formes, couleurs, visages, mouvements, etc.). Très régionaliste, cette organisation cérébrale est assez similaire chez l'homme et chez le singe. A une exception près, mais elle est notable : il existe sans doute, chez l'homme, une aire spécifiquement consacrée à la reconnaissance des mots écrits. Une aire si sélective qu'elle ne s'active pas lorsque sont présentées, à la place de mots, des lettres assemblées au hasard.

Plus généralement, que sait-on aujourd'hui des bases cérébrales du langage (parlé ou écrit), dont l'apprentissage et le traitement restent l'un des domaines de prédilection des cognitivistes ? Grâce aux travaux de Paul Broca sur l'aphasie, on sait depuis plus d'un siècle que cette tâche est principalement dévolue à l'hémisphère gauche. Mais du langage à l'écriture, où est la passerelle ? Comment le cerveau gère-t-il ces deux fonctions, distinctes et pourtant étroitement liées ? Pour en savoir plus, l'équipe d'imagerie neuro-fonctionnelle du service hospitalier Frédéric-Joliot (Orsay) est récemment l'idée d'examiner, chez des sujets japonais, la localisation du langage écrit en fonction du système d'écriture employé : soit les caractères « kanjis », morphogrammes ayant chacun une identité graphique et une signification ; soit les « kanas », symboles syllabiques plus proches de notre système alphabétique. « Les résul-



ats ont été plus subtils que prévu. Quelle que soit l'écriture, l'hémisphère gauche prédomine », résume Denis le Bihan, directeur de ces recherches à Orsay. L'IRM a toutefois révélé deux différences. D'une part, l'utilisation des kanjis implique une activité cérébrale globale deux à trois fois plus grande que celle des kanas. D'autre part, les kanjis sollicitent, en plus des zones habituelles de l'hémisphère gauche, certaines zones pariétales et temporales de l'hémisphère droit.

Au-delà de cette passionnante géographie du cerveau, il reste à appréhender, dans leur détail, les mécanismes neurologiques impliqués dans les tâches cognitives. La comment les difficultés du cerveau humain se composent d'un extraordinaire réseau de cellules nerveuses, les neurones, dont le nombre est estimé à environ mille milliards. Chacune échange des informations avec ses voisins sous forme de courant électrique par le biais de diverses connexions (environ dix par neurone), axones, dendrites et synapses. Les chercheurs en sont maintenant convaincus : c'est dans la plasticité de ces synapses, dans leur faculté à s'organiser durablement en réseaux que résident les véritables secrets des phénomènes d'apprentissage et de mémoire.

Seule certitude : dans ce fouillis de neurones repose un trésor, un concentré d'esprit qui n'est convenu d'appeler



« intelligence ». Tout à la fois « faculté de connaître, de comprendre », « ensemble des fonctions mentales conçues pour objecter la connaissance conceptuelle et rationnelle » et « aptitude d'un être vivant à adapter à des situations nouvelles », tantôt considérée comme spécifique à l'homme, tantôt reconnue comme partagée par l'ensemble des animaux supérieurs, la notion n'en continue pas moins de se dérobant à l'entendement. Ici s'arrête la science, la commentent les suppositions. Y compris les plus harmonieuses.

Il y a quelques années, une très sérieuse étude américaine, publiée dans la revue *Nature*, suggérait ainsi que l'écoïde assidue de Mozart rendait transitoirement plus intelligent. Et soumettant un groupe de trente-six personnes aux tests de QI, les chercheurs de l'université de Californie (Irving) se sont aperçus que leurs résultats étaient en moyenne de 8 à 9 points supérieurs lorsque les tests étaient précédés de l'audition de la sonate pour deux pianos en D mineur (K488). « On peut imaginer un stimule des réseaux neuronaux qui jouent un rôle important dans la connaissance », avaient supposé les chercheurs. Personne, depuis lors, n'a confirmé l'hypothèse... ni ne l'a contredit. ■

L'art contemporain

Wassily Kandinsky, l'inventeur de la peinture abstraite, se plaignait déjà de ces regards concupiscent qui folloient la peinture comme un plié. Le peintre Paul Klee n'avait guère non plus de considération pour ces esprits pressés qui se restent à « brouter » la surface de la toile. Selon Mayaux, l'art appartient au domaine du simulacre et du but d'une œuvre est de jouer un soupçon sautatoire sur les apparences pour les engager à se dépasser. L'artiste Marcel Duchamp disait : « Ce sont des regards qui font l'art. » Dans le tableau de Philippe Mayaux, ces regards se sont transformés en fourchettes, et la contemplation s'est muée en consommation. Cette absence de regard favorise l'aveuglement - « Il n'y a plus d'artistes », dit-il, « L'art contemporain est nul » - un discours sans objet, une querelle d'experts dont le seul objectif est de prouver l'excellence de leurs expertises dans ce qu'il est désormais convenu d'appeler *Le Difficile Partage du gâteau*. ■

Une palette d'actions pour les jeunes

Peut-on apprendre à aimer l'art contemporain aux jeunes ? Sylvie Cohen, documentaliste, et Alain Nahmias, plasticien, en sont convaincus. C'est pourquoi ils passent aujourd'hui leur temps à démarcher des partenariats institutionnels pour pouvoir créer en Ile-de-France le premier centre d'art contemporain pour les enfants. La dérogation aux arts plastiques et la Fondation Copernic d'ores et déjà donne leur accord de principe. Reste à trouver des partenariats importants pour mener à bien un projet qui nécessite d'abord un lieu, si possible dans un parc qui soit déjà un but de promenade familiale.

Cette polémique qui s'est cristallisée autour de collections pérennes - « L'art contemporain est nul » ou « il n'y a plus d'artistes », dit-il, « L'art contemporain est nul » - exclut et condamne a priori les artistes et les œuvres. Que l'art a plus droit de cité est révélateur d'une crise qui n'est pas celle de l'art mais d'un discours sur l'art. Ce phénomène traduit un décalage entre le pouvoir croissant des experts en charge de gérer les normes du bon goût et des œuvres reléguées en cinquième roue du carrosse culturel.

« Cette polémique qui s'est cristallisée autour de collections pérennes - « L'art contemporain est nul » ou « il n'y a plus d'artistes », dit-il, « L'art contemporain est nul » - exclut et condamne a priori les artistes et les œuvres. Que l'art a plus droit de cité est révélateur d'une crise qui n'est pas celle de l'art mais d'un discours sur l'art. Ce phénomène traduit un décalage entre le pouvoir croissant des experts en charge de gérer les normes du bon goût et des œuvres reléguées en cinquième roue du carrosse culturel.

Image 3 filtrée

ainsi montré que les bébés de quatre-vingt mois sont capables de détecter la transmission ou la « conservation » du nombre lorsqu'on leur présente des événements numériques impossibles ou possibles (dans une maquette de théâtre ou les unités à quantifier sont des Mickey). Le traitement de l'information numérique réalisé par le bébé repose sur un processus analytique de « calcul précis » et non sur un processus de traitement perceptif global (ou holistique). Il a par ailleurs été démontré que ces compétences prénumériques se retrouvent chez l'enfant d'âge préscolaire (entre deux et six ans) après une chute temporaire de performance. Comment expliquer alors l'irrationalité de l'enfant de cet âge dans l'épreuve piagetienne de conservation du nombre : place devant deux rangées de jouons en nombre égal mais de longueur différente parce que les jouons de l'une d'elles sont plus espacés, l'enfant considère qu'il y en a plus là qu'il est plus long ?

Pour Piaget, l'enfant d'âge préscolaire serait encore fondamentalement intuitif, « préopérateur ». Pour traiter l'information, il se fonderait sur la longueur, selon un processus perceptif global. Il aurait donc pas encore construit le « schéma du nombre » lui permettant de procéder à un calcul précis. Les recherches récentes qui viennent d'être évoquées conduisent à douter de cette interprétation. En effet, ne serait-on pas plutôt ici, à l'occasion d'un nouveau « pli du temps », en présence de la coexistence de « rationnel construit » et de l'« intuitif présumé révolu » ? Dans cet exemple, on trouve d'une part la capacité d'un traitement numérique analytique et précis et d'autre part la transmission du nombre dans l'épreuve de

Image 5 filtrée

conservation. Cette dernière épreuve n'est-elle pas avant tout une épreuve d'inférence « non/bonheur » critique de la capacité à résister au schéma sensé des événements numériques impossibles ou possibles (dans une maquette de théâtre ou les unités à quantifier sont des Mickey). Le traitement de l'information numérique réalisé par le bébé repose sur un processus analytique de « calcul précis » et non sur un processus de traitement perceptif global (ou holistique). Il a par ailleurs été démontré que ces compétences prénumériques se retrouvent chez l'enfant d'âge préscolaire (entre deux et six ans) après une chute temporaire de performance. Comment expliquer alors l'irrationalité de l'enfant de cet âge dans l'épreuve piagetienne de conservation du nombre : place devant deux rangées de jouons en nombre égal mais de longueur différente parce que les jouons de l'une d'elles sont plus espacés, l'enfant considère qu'il y en a plus là qu'il est plus long ?

Pour Piaget, l'enfant d'âge préscolaire serait encore fondamentalement intuitif, « préopérateur ». Pour traiter l'information, il se fonderait sur la longueur, selon un processus perceptif global. Il aurait donc pas encore construit le « schéma du nombre » lui permettant de procéder à un calcul précis. Les recherches récentes qui viennent d'être évoquées conduisent à douter de cette interprétation. En effet, ne serait-on pas plutôt ici, à l'occasion d'un nouveau « pli du temps », en présence de la coexistence de « rationnel construit » et de l'« intuitif présumé révolu » ? Dans cet exemple, on trouve d'une part la capacité d'un traitement numérique analytique et précis et d'autre part la transmission du nombre dans l'épreuve de

Image 8 filtrée

le public en appelle alors au rôle critique des institutions qui doivent défendre les valeurs culturelles et ne pas s'associer à ce qu'il considère comme une entreprise de démolition. Le conflit esthétique vise alors à la politique et la morale. C'est pourquoi les débats les plus vifs sont liés à des commandes publiques. Les commandes publiques, depuis un exemple : le rejet esthétique d'une proposition heurtant les attitudes traditionnelles se consomme dans une réaction politique suscitée par une initiative ministérielle. Celle-ci a été en effet à contre-courant de ce que l'on exige d'une commande publique, à savoir un minimum de consensus.

Selon vous, les différends entre défenseurs et détracteurs de l'art contemporain révèlent-ils moins de divergences d'appréciation esthétique que de différences de catégories cognitives. Les querelles sont donc insolubles ?

Au siècle dernier, il s'agissait de décréter si c'était beau ou non, si les peintres savaient peindre. En somme, c'était une question de goût. Aujourd'hui, on est en face d'un désaccord fondamental sur les catégories de classement : est-ce ou non de l'art ? Le problème n'est pas d'évaluer la valeur esthétique d'une œuvre, mais de définir sa nature même. Les conflits ne peuvent donc pas se résoudre lorsqu'on ne parle pas de la même chose ; pas, ils se résout.

Autre rupture historique : la place fondamentale qu'occupent aujourd'hui les institutions artistiques (centres d'art, musées, FRAC), notamment depuis quinze ans. Ce sont elles qui, par leur politique d'achat, accèdent les avant-gardes en temps direct, sans attendre plusieurs générations. Quels sont les effets de ce face-à-face ?

Il met entre parenthèses les marchés privés. Les œuvres de jeunes artistes peuvent être achetées et exposées par des structures d'Etat sans être passées par le circuit des collectionneurs et des galeries. Depuis la politique de Jack Lang, la prise en charge institutionnelle de l'avant-garde s'est accélérée. Il en résulte des effets positifs mais aussi des

Image 7 filtrée

effets pervers. L'avant-garde est traditionnellement associée à une position progressive, marginale par rapport à son époque. Aujourd'hui, les institutions désirent conserver une posture de franchise tout aller à contre-courant de ce qui est admis par les institutions, autrement dit contre la reconnaissance. Les positions innovatrices se sont positionnées contre l'art contemporain au sens où celui-ci est perçu comme un art officiel, académique.

quelques années, des personnes invoquant l'authenticité artistique et la marginalité novatrice se sont positionnées contre l'art contemporain au sens où celui-ci est perçu comme un art officiel, académique.

C'est la position de Jean Clair... C'est aussi celle de Jean Baudrillard, de Jean-Philippe Domecq, de Marc Fumaroli, qui s'érigent contre le rôle des institutions.

Est-ce le cœur du débat ? Non, mais c'est son sous-basement implicite, qui permet de comprendre les renversements d'opinions et pourquoi les spécialistes et observables de l'art contemporain sont dans une position si inconfortable sur le plan éthique et politique ; ils se sont engagés dans le soutien à l'avant-garde, à l'innovation difficile, et leurs détracteurs les présentent comme des appartements des réactionnaires défendant les œuvres plus standards, à l'image des promoteurs de l'art pompier au siècle dernier.

Pourquoi la pédagogie est-elle si difficile auprès des adultes ?

En raison de la nature d'événement transgressives. Pour évaluer l'intérêt et la pertinence des propositions, juger que celle-ci est redondante ou qu'elle est radicalement neuve, il faut avoir une culture de l'enseignement artistique, ce qui n'est pas le cas de la majorité des adultes soucieux d'une posture d'adulte formateur et d'émotion immédiate pour

Image 7 filtrée

Image 8 filtrée

Figure 4.8 : Résultats du filtrage Gaussien

Interprétation

La **figure 4.8** donne une illustration sur l'influence du filtrage. Toutes les images filtrées montrent une différence des régions lisses et homogènes, par rapport aux images originales. Nous avons fait varier les dimensions du masque, les meilleurs résultats ont été obtenus pour la taille 3×3 .

4.3.1.2 Résultats de Binarisation

La binarisation consiste à séparer le fond de l'objet. Une méthode de binarisation coopérative en deux étapes a été employée. Les résultats expérimentaux obtenus sur des images de documents en terme de séparation de l'objet et le fond sont satisfaisantes, on obtient une meilleure conservation des contours des objets et on garde les détails significatifs (on évite la fusion des formes) Nous avons appliqué cette opération sur un ensemble d'images, et nous présentons les résultats obtenus sur la **figure 4.9**.

Les valeurs des paramètres de la binarisation hiérarchique flou sont :

$$\alpha = 0.975, \quad p = 3$$

Le socio-constructivisme

La construction d'un savoir, bien que personnelle, dépend du contexte social. Le socio-constructivisme remet en cause les modèles psychologiques du développement cognitif et insiste sur les dimensions sociales contribuant à la formation des compétences.

Les recherches portant sur la notion de conceptions préalables ou représentations initiales ne datent pas d'aujourd'hui et permettent de dégager trois grands modèles de l'éducation.

LE PREMIER est le modèle transmissif appelé aussi «magistral» ou « frontal » où le cerveau de l'élève est considéré telle une «cire molle» au niveau de laquelle viendraient s'imprimer les informations. C'est ce qu'on appelle aussi le modèle des « vases communicants ».

LE SECOND est appelé comportementaliste (béhavioriste), modèle où l'apprentissage résulte d'une suite de conditionnements «stimulus-réponse ». Les connaissances sont définies en termes de comportements observables attendus en fin d'apprentissage.

LE TROISIÈME est le modèle constructiviste et socio-constructiviste au niveau duquel les conceptions initiales ne sont pas simplement le point de départ du processus d'apprentissage mais

sont au cœur de celui-ci. Ainsi, l'élève construit lui-même ses connaissances car le développement intellectuel est un processus interne qui se réalise par étapes successives. Il ne peut, donc, « assimiler » de nouvelles connaissances que si ses structures mentales le permettent. Il serait donc insensé de vouloir «gaver» les élèves de savoirs qu'ils ne sont pas en mesure d'intégrer.

Dans cette perspective, l'enseignant doit impérativement respecter ce processus de développement personnel de l'élève et s'adapter à ses besoins. Son rôle consiste à observer, à diagnostiquer et à pratiquer ainsi une évaluation formative et une pédagogie différenciée puisqu'un nouveau savoir n'est effectif que s'il se construit en s'intégrant au réseau conceptuel de l'apprenant. Ausubel parle de « ponts cognitifs » car c'est la manière dont l'élève assimile les nouvelles connaissances qui est primordiale. Il peut y avoir alors des apprentissages significatifs en interaction avec les acquis préalables ou mécaniques sans liens avec les prérequis. Concrètement, afin que les enfants soient acteurs de leurs apprentissages, il est souhaitable que l'enseignant propose, stimule, observe, écoute et accepte que les élèves découvrent, explorent et entreprennent. En outre, il aide les élèves à verbaliser, à formuler, à sélectionner des indices, émettre des hypothèses et prendre conscience du problème qu'ils seront en mesure de résoudre. Par ailleurs, il gère la communication afin que les apprenants explicitent leurs

démarches, justifient leurs choix, les argumentent et tolèrent ceux d'autrui. Enfin, il propose des exercices adaptés pour qu'ils consolident, maîtrisent, approfondissent et réinvestissent les nouvelles connaissances.

En définitive, dans cette optique socio-constructiviste, apprendre c'est élaborer soi-même ses connaissances tout en passant nécessairement par une phase d'interaction, voire de conflit socio-cognitif avec autrui qui est d'autant plus efficace si l'enseignant est capable de gérer cette situation et de permettre à chacun de passer d'un niveau interpersonnel à un niveau intrapersonnel ; ce qui facilite incontestablement la phase d'appropriation et l'élaboration d'un langage intériorisé. Il apparaît ainsi clairement qu'un système éducatif voulant offrir à tous les mêmes chances doit prendre en compte le contexte social, ce qui permettra un développement individuel de l'intelligence considérée comme une forme particulière d'adaptation. Adaptation nécessaire à l'ère des valeurs universelles, loin des schémas et modèles éducatifs rectilignes car pour progresser, il est impératif d'opérer constamment des « feedback », de se remettre en question et surtout avoir constamment à l'esprit que l'élève n'est pas une « page vierge » qu'il faut absolument remplir mais, surtout et avant tout, une personne «unique» qui a sa propre réalité car comme l'écrivit Watzlawick : « Toute réalité est toujours inventée » et « l'on construit soi-même sa propre réalité ».

Image 1 binarisée



Expert Systems with Applications 33 (2007) 948–977

Expert Systems with Applications

www.elsevier.com/locate/eswa

Automated barcode recognition for smart identification and inspection automation

Sherin M. Youssef*, Rana M. Salem

Department of Computer Engineering, College of Engineering and Technology, Assiut Academy for Science and Technology, Assiut, Egypt

Abstract

Barcodes have been widely used in many industrial products for automatic identification in data collection and inventory control purposes. It is well known that in many stores laser bar-code readers are used at check-out counters. However, there is a major constraint when this tool is used. That is, unlike traditional camera-based picturing, the distance between the laser reader (sensor) and the target object is close to zero when the reader is applied. This may result in inconvenience in inspection automation because the human operator has to manipulate either the sensor or the objects. For the purpose of in-store or inspection automation, the human operator needs to be removed from the process, i.e. a robot with visual capability is required to play an important role in such a system. Moreover, an automated system is required to find, locate and decode barcodes on various document images even with low resolution compared with laser barcode readers (about 15,000 dots per inch) and can handle damaged bar codes. This paper proposes a smart barcode detection and recognition system (SBDRC) based on fast hierarchical Hough transform (HHT). The back-propagation neural network (BPNN) is selected as a powerful tool to perform the recognition process. The paper presents an effective method to utilize the specific graphic features of barcodes for positioning and recognition purposes even in case of distorted barcodes. The first step the system has to perform is to locate the position and orientation of the barcode in the required material document image. Secondly, the proposed system has to segment the barcode. Finally, a trained back-propagation neural network is used to perform the barcode recognition task. Experiments have been conducted to corroborate the efficiency of the proposed method.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Barcodes; Automated identification

1. Introduction

The automated identification of goods and persons is an attractive application for areas of industry, security, postal labels and smart cards inspection. A barcode is a machine-readable representation of information in a visual format on a surface (Sutton, 2002). Barcodes have been widely used in many industrial products for automatic identification in data collection and inventory control purposes (Chang, Lo, & Hsieh, 1997). Originally barcodes store data in the width and spacing of printed parallel lines, but they

also come in patterns of dots, concentric circles, and hidden in images (Fig. 1 shows examples of different types of barcodes). Barcodes can be read by optical scanners called barcode readers or scanned from an image. Barcodes are widely used to implement Auto Id Identification systems. With the advent of in-store automation and smart cards identification, the existence of standard barcodes for information exchange and identification is indispensable. At many points of sale, the laser barcode reader is conventionally used. However, there is a major constraint on the laser barcode reader (Liso, 1995). That is, unlike traditional camera-based picturing, the distance between the laser reader (sensor) and the target object is close to zero when the reader is applied. This may result in inconvenience in inspection automation because the human operator has to manipulate either the sensor or the objects. For the

* Corresponding author.
E-mail addresses: sherin@assut.edu.eg, smy@assut.ac.eg (S.M. Youssef).

Image 2 binarisée

Interprétation

A travers ces tests, la méthode de binarisation adoptée a démontré son efficacité dans la binarisation des images de documents. C'est une bonne solution pour résoudre le problème de classification des pixels dans la zone floue à la frontière fond-objet.

4.3.2 Résultats de pré-segmentation (algorithme de Bloomberg)

Comme on a dit précédemment, l'algorithme de Bloomberg est basé sur des opérations morphologiques. Ces opérations consistent en : une réduction pour réduire la taille des objets, un remplissage des trous s'ils en existent, une ouverture pour éliminer les petits objets, puis une expansion pour revenir à la taille originale de l'objet.

L'opération de réduction est appliquée quatre fois en utilisant les seuils : $T_1 = 1, T_2 = 1, T_3 = 4, T_4 = 3$.

L'élément structurant utilisé est de type carré de taille 5×5 pour l'ouverture et 3×3 pour la dilatation

Pour vérifier le bon fonctionnement de l'algorithme de Bloomberg, nous l'avons appliqué sur différents types de documents images avec une variété de textes et non textuels composants avec des structures de type Manhattan et autre de type non Manhattan.

La figure 5.11 illustre les résultats de pré-segmentation obtenus :

Le socio-constructivisme

La construction d'un savoir, bien que personnelle, dépend du contexte social. Le socio-constructivisme remet en cause les modèles psychologiques du développement cognitif et insiste sur les dimensions sociales contribuant à la formation des compétences.

Les recherches portant sur la notion de conceptions préalables ou représentations initiales ne datent pas d'aujourd'hui et permettent de dégager trois grands modèles de l'éducation.

LE PREMIER est le modèle transmissif appelé aussi «magistral» ou « frontal » où le cerveau de l'élève est considéré telle une «craie molle» au niveau de laquelle viendraient s'imprimer les informations. C'est ce qu'on appelle aussi le modèle des « vases communicants ».

LE SECOND est appelé comportementaliste (behaviouriste), modèle où l'apprentissage résulte d'une suite de conditionnements «stimulus-réponse». Les connaissances sont définies en termes de comportements observables attendus en fin d'apprentissage.

LE TROISIÈME est le modèle constructiviste et socio-constructiviste au niveau duquel les conceptions initiales ne sont pas simplement le point de départ du processus d'apprentissage mais

sont au cœur de celui-ci. Ainsi, l'élève construit lui-même ses connaissances car le développement intellectuel est un processus interne qui se réalise par étapes successives. Il ne peut, donc, « assimiler » de nouvelles connaissances que si ses structures mentales le permettent. Il serait donc insensé de vouloir «igner» les élèves de savoirs qu'ils ne sont pas en mesure d'intégrer.

Dans cette perspective, l'enseignant doit impérativement respecter ce processus de développement personnel de l'élève et s'adapter à ses besoins. Son rôle consiste à observer, à diagnostiquer et à pratiquer ainsi une évaluation formative et une pédagogie différenciée puis qu'un nouveau savoir n'est effectif que s'il se construit en s'intégrant au réseau conceptuel de l'apprenant. Ausubel parle de « ponts cognitifs » car c'est la manière dont l'élève assimile les nouvelles connaissances qui est primordiale. Il peut y avoir alors des apprentissages significatifs en interaction avec les liens avec les prérequis. Concrètement, afin que les enfants soient acteurs de leurs apprentissages, il est souhaitable que l'enseignant propose, stimule, observe, écoute et accepte que les élèves découvrent, explorent et entreprennent. En outre, il aide les élèves à verbaliser, à formuler, à sélectionner des indices, émettre des hypothèses et prendre conscience du problème qu'ils seront en mesure de résoudre. Par ailleurs, il gère la communication afin que les apprenants explicitent leurs

démarches, justifient leurs choix, les argumentent et tolèrent ceux d'autrui. Enfin, il propose des exercices adaptés pour qu'ils consolident, maîtrisent, approfondissent et réinvestissent les nouvelles connaissances.

En définitive, dans cette optique socio-constructiviste, apprendre c'est élaborer soi-même ses connaissances tout en passant nécessairement par une phase d'interaction, voire de conflit socio-cognitif avec autrui qui est d'autant plus efficace si l'enseignant est capable de gérer cette situation et de permettre à chacun de passer d'un niveau interpersonnel à un niveau intrapersonnel ; ce qui facilite incontestablement la phase d'appropriation et l'élaboration d'un langage intériorisé. Il apparaît ainsi clairement qu'un système éducatif voulant offrir à tous les mêmes chances doit prendre en compte le contexte social, ce qui permettra un développement individuel de l'intelligence considérée comme une forme particulière d'adaptation. Adaptation nécessaire à l'ère des valeurs universelles, loin des schémas et modèles éducatifs rectilignes car pour progresser, il est impératif d'opérer constamment des «feed-back», de se remettre en question et surtout avoir constamment à l'esprit que l'élève n'est pas une «page vierge» qu'il faut absolument remplir mais, surtout et avant tout, une personne «unique» qui a sa propre réalité car comme l'écrivit Watzlawick : « Toute réalité est toujours inventée » et « l'on construit soi-même sa propre réalité ».



Expert Systems with Applications 33 (2007) 948–977

Expert Systems with Applications

www.elsevier.com/locate/eswa

Automated barcode recognition for smart identification and inspection automation

Sherin M. Yousef^a, Rana M. Salem

^aDepartment of Computer Engineering, College of Engineering and Technology, Arab Academy for Science and Technology, Alexandria, Egypt

Abstract

Barcodes have been widely used in many industrial products for automatic identification in data collection and inventory control purposes. It is well known that in many stores laser bar-code readers are used as check-out counters. However, there is a major constraint when this tool is used. That is, unlike traditional camera-based picturing, the distance between the laser reader (sensor) and the target object is close to zero when the reader is applied. This may result in inconvenience in inspection automation because the human operator has to manipulate either the sensor or the objects. For the purpose of in-store or inspection automation, the human operator needs to be removed from the process. I.e. a robot with visual capability is required to play an important role in such a system. Moreover, an automated system is required to find, locate and decode barcodes on various document images even with low resolution compared with laser barcode readers (about 15,000 dots per inch) and can handle damaged bar codes. This paper proposes a smart barcode detection and recognition system (SBDRC) based on fast hierarchical Hough transform (FHT). The back-propagation neural network (BPNN) is selected as a powerful tool to perform the recognition process. The paper presents an effective method to utilize the specific graphic features of barcodes for positioning and recognition purposes even in case of distorted barcodes. The first step the system has to perform is to locate the position and orientation of the barcodes in the required material document image. Secondly, the proposed system has to segment the barcodes. Finally, a trained back-propagation neural network is used to perform the barcode recognition task. Experiments have been conducted to corroborate the efficiency of the proposed method.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Barcode; Automated identification

1. Introduction

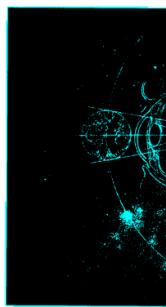
The automated identification of goods and persons is an attractive application for areas of industry, security, postal labels and smart cards inspection. A barcode is a machine-readable representation of information in a visual format on a surface (Barton, 2002). Barcodes have been widely used in many industrial products for automatic identification in data collection and inventory control purposes (Chang, Lo, & Fiaish, 1997). Originally barcodes store data in the width and spacing of printed parallel lines, but they

also come in patterns of dots, concentric circles, and hidden in images (Fig. 1 shows examples of different types of barcodes). Barcodes can be read by optical scanners called barcode readers or scanned from an image. Barcodes are widely used to implement Auto Id identification systems. With the advent of in-store automation and smart cards identification, the existence of standard barcodes for information exchange and identification is indispensable. At many points of sale, the laser barcode reader is conventionally used. However, there is a major constraint on the laser barcode reader (Liao, 1995). That is, unlike traditional camera-based picturing, the distance between the laser reader (sensor) and the target object is close to zero when the reader is applied. This may result in inconvenience in inspection automation because the human operator has to manipulate either the sensor or the objects. For the

^{*} Corresponding author.
E-mail address: sherin@ast.edu_egy@cc.sci.aust.edu.eg (S.M. Yousef).

(a) image 1 pré-segmentée

(b) image 2 pré-segmentée



dans une reconnaissance particulière (formes, couleurs, visages, mouvements, etc.). Très régionalisée, cette organisation cérébrale est assez similaire chez l'homme et chez le singe. A une exception près, mais elle est notable : il existe sans doute, chez l'homme, une aire spécifiquement consacrée à la reconnaissance des mots écrits. Une aire si sélective qu'elle ne s'active pas lorsque sont présentées, à la place de mots, des lettres assemblées au hasard... Plus généralement, que sait-on aujourd'hui des bases cérébrales du langage (parlé ou écrit), dont l'apprentissage et le traitement restent l'un des domaines de prédilection des cognitivistes ? Grâce aux travaux de Paul Broca sur l'aphasie, on sait depuis plus d'un siècle que cette tâche est principalement dévolue à l'hémisphère gauche. Mais du langage à l'écriture, ou la possibilité ? Comment le cerveau gère-t-il ces deux fonctions, distinctes et pourtant étroitement liées ? Pour en savoir plus, l'équipe d'imagerie neuro-fonctionnelle du service hospitalier Frédéric-Joliot (Orsay) est récemment l'idée d'examiner, chez des sujets japonais, la localisation du langage écrit en fonction du système d'écriture employé : soit les caractères « kanji », morphogrammes ayant chacun une identité graphique et une signification ; soit les « kana », symboles syllabiques plus proches de notre système alphabétique. « Les ré-

ultats ont été plus subtils que prévu. Quelle que soit l'écriture, l'hémisphère gauche prédomine », résume Denis Le Bilhan, directeur de ses recherches à Orsay. L'IRM a toutefois révélé deux différences. D'une part, l'utilisation des kanjis implique une activité cérébrale globale deux à trois fois plus grande que celle des kana. D'autre part, les kanjis sollicitent, en plus des zones habituelles de l'hémisphère gauche, certaines zones pariétales et temporales de l'hémisphère droit. Au-delà de cette passionnante géographie du cerveau, il reste à appréhender, dans leur détail, les mécanismes neurologiques impliqués dans les tâches cognitives. Là commencent les difficultés. Le cerveau humain se compose d'un extraordinaire réseau de cellules nerveuses, des neurones, dont le nombre est estimé à environ mille milliards. Chacun échange des informations avec ses voisins sous forme de courant électrique, par le biais de diverses connexions (dendrites et synapses). Les chercheurs en sont maintenant convaincus : c'est dans la plasticité de ces synapses, dans leur faculté à s'organiser durablement en réseaux que résident les véritables secrets des phénomènes d'apprentissage et de mémoire. Seule certitude : dans ce fouillis de neurones repose un trésor, un concentré d'expériences qu'il est convenu d'appeler

« intelligence ». Tout à la fois « faculté de connaître, de comprendre », « ensemble des fonctions mentales ayant pour objet la connaissance conceptuelle et rationnelle » et « à adapter à des situations nouvelles », tantôt considérée comme spécifique à l'homme, tantôt reconnue comme partagée par l'ensemble des animaux supérieurs, la notion n'en continue pas moins de se dérobant à l'entendement. Ici s'arrête la science, là commencent les suppositions. Y compris les plus harmonieuses. Il y a quelques années, une très sérieuse étude américaine, publiée dans la revue *Nature*, suggérait ainsi que l'écoute assidue de Mozart rendait transitoirement plus intelligent. En soumettant un groupe de treize personnes aux tests de QI, les chercheurs de l'Université de Californie (Irvine) se sont aperçus que leurs résultats étaient en moyenne de 8 à 9 points supérieurs lorsque les tests étaient précédés de l'audition de la sonate pour deux pianos de D major (K488). « On peut imaginer qu'en écoutant une telle musique on stimule des réseaux neuronaux qui jouent un rôle important dans la connaissance », avaient supposé les chercheurs. Personne, depuis lors, n'a confirmé l'hypothèse... ni ne l'a contredite.
 à voir : *Le Cerveau en action*, sous la direction de Stanislas Dehaene, PUF (1997), 290 p., 148 f.

« Si passionnante que soit la géographie du cerveau, il reste à appréhender les mécanismes neurologiques des tâches cognitives. »

Under the option EDIT select PASTE SPECIAL. A dialog box will open, select paste picture, then click OK. Your figure should now be in the Word Document. If you are preparing images in TIFF, EPS, or PS format, note the following. High-contrast line figures and tables should be prepared with 600 dpi resolution and saved with no compression, 1 bit per pixel (monochrome), with file names in the form of "fig3.tif" or "table1.tif." Photographs and grayscale figures should be prepared with 300 dpi resolution and saved with no compression, 8 bits per pixel (grayscale).

Sizing of Graphics
Most charts graphs and tables are one column wide (3 1/2 inches or 21 picas) or two-column width (7 1/16 inches, 43 picas wide). We recommend that you avoid sizing figures less than one column wide, as extreme enlargements may distort your images and result in poor reproduction. Therefore, it is better if the image is slightly larger, as a minor reduction in size should not have an adverse affect the quality of the image.

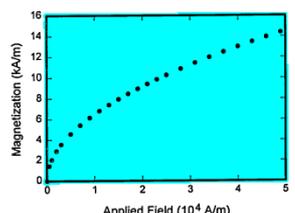


Fig. 1. Magnetization as a function of applied field. Note that "Fig." is abbreviated. There is a period after the figure number, followed by two spaces. It is good practice to explain the significance of the figure in the caption.

fonts are as follows: Open Type Fonts: Times Roman, Helvetica, Helvetica Narrow, Courier, Symbol, Palatino, Avant Garde, Bookman, Zapf Chancery, Zapf Dingbats, and New Century Schoolbook.

Print Color Graphics Requirements
IAENG accepts color graphics in the following formats: EPS, PS, TIFF, Word, PowerPoint, Excel, and PDF. The resolution of a RGB color TIFF file should be at least 400 dpi.

Web Color Graphics
IAENG accepts color graphics in the following formats: EPS, PS, TIFF, Word, PowerPoint, Excel, and PDF. The resolution of a RGB color TIFF file should be at least 400 dpi.

Your color graphic will be converted to grayscale if no separate grayscale file is provided. If a graphic is to appear in print as black and white, it should be saved and submitted as a black and white file. If a graphic is to appear in print or on IAENG web site in color, it should be submitted as RGB color.

Graphics Checker Tool
The authors should check their graphic files against the set of rules for compliance with IAENG requirements. These requirements are designed to ensure sufficient image quality so they will look acceptable in print.

E. Copyright
The copyright to the Contribution identified above is transferred to International Association of Engineers, (hereinafter called IAENG). The copyright transfer covers the sole right to print, publish, distribute and sell throughout the world the said Contribution and parts thereof, including all revisions or versions and future editions thereof and in any medium, such as in its electronic form (offline, online), as well as to translate, print, publish, distribute and sell the Contribution in any foreign languages and throughout the world. IAENG will take, either in its own name or in that of the Author, any necessary steps to protect these rights

(c) image 3 pré-segmentée

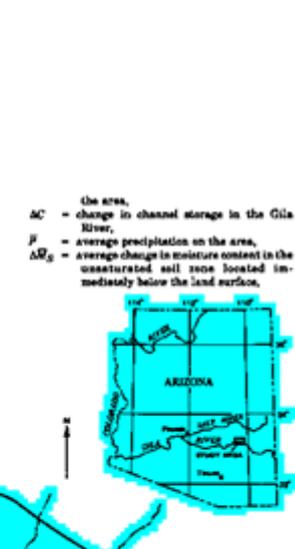
(d) image 4 pré-segmentée

nels et l'attachement à la terre natale, suggère bel et bien, comme le démontre Philippe Dagen dans son ouvrage *La Haine de l'art, une généalogie de la pensée réactionnaire française*. Villipendant dans *Le Monde* « une avant-garde officielle et indifférente à tout ce qui n'est pas son dogme ». Jean Clair en vient à formuler un grief majeur : « Cette sollicitude de l'Etat a conduit à faire des artistes des assistants sociaux, assez peu préparés à la lutte sur un marché libre et mondialisé. » Et ce parti-d'un libéralisme économique qui ne devrait pas gêner la culture ajoute que, pour réussir à s'imposer sur les scènes internationales, rien ne vaut l'« exaltation » des « valeurs culturelles propres », comme l'ont fait, selon lui, les écoles allemande, italienne et londonienne, et non pas l'« internationalisme », source de tous les maux. Marc Fumaroli écrivait aussi qu'il s'en prenait « moins aux artistes qu'à un certain discours international sur l'art ».

Une palette d'actions pour les jeunes
Peut-on apprendre à aimer l'art contemporain aux jeunes ? Sylvie Cohen et Alain Nahmias envisagent une structure d'exposition d'œuvres originales représentatives des tendances actuelles, dont les circuits de visite seraient conçus avec des conseillers pédagogiques. Afin d'éveiller le regard et encourager la création des enfants, un système d'ateliers et de conférences compléterait le dispositif. Pourquoi un lieu spécifique ? « D'abord pour que les œuvres soient accrochées à hauteur des enfants, ensuite pour que ce soit un lieu de rencontre, un lieu de partage. Un avis qui partage Philippe Coubetergues, enseignant en arts plastiques détaché à la Galerie nationale du Jeu de paume : « La littérature critique et sensible, sans connaissances historiques approfondies, les enfants le comprennent plus que les adultes, qui se sentent davantage investis d'un jugement esthétique. » Depuis six ans, le service culturel du Jeu de paume s'est résolument attaché à développer les activités en direction des scolaires et, ce faisant, à apporter la preuve que le fossé entre l'art contemporain et le public n'est pas incommensurable, pour peu que des actions concertées favorisent le contact direct avec les œuvres. Un des principes retenus est celui de partenariats avec des collectivités qui donnent lieu à des sorties au jeu de paume et à des ateliers en milieu scolaire. Un dernier, plus de deux cents enseignants ont ainsi suivi les séances de préparation le mercredi, à la suite desquelles 6 746 élèves, en grande majorité résidents dans la périphérie parisienne, ont été accueillis pour des visites commentées. Et une centaine de professeurs de toutes disciplines ont été reçus en stage de formation pédagogique. Au printemps prochain, la Galerie nationale du Jeu de paume devrait organiser un colloque international sur la médiation entre les institutions d'art contemporain et l'enseignement. ■ M. S.

L'art contemporain
Le bon et le mauvais goût, l'acidité rieuse des couleurs artificielles et les bruns solennels. Cet effet est accentué par l'usage allénié de la peinture à l'huile, médium noble et traditionnel aux nuances profondes, et de la peinture acrylique, aux tons artificiels. Ainsi, le tableau allégorique du goût contemporain entredit jeter quelque confusion, quant aux recettes bien établies de l'authentique tradition et de l'impure modernité. En présentant l'œuvre d'un homme un gîteau agrès par sept fourchettes, Mayaux fait allusion à ce péché capital qui consiste à réduire vulgairement l'expérience esthétique à une dégustation.

Wassily Kandinsky, l'inventeur de la peinture abstraite, se plaignait déjà de ces regards conceptuels qui gèrent la peinture comme un pôle. Le peintre Paul Klee n'avait guère non plus de considération pour ces esprits pressés qui en restent à « broder » la surface de la toile. Selon Mayaux, l'art appartient au domaine du simulacre, et le but d'une œuvre est de jeter un soupçon salutaire sur les apparences pour les engager à se dépasser. L'artiste Marcel Duchamp disait : « Ce sont les regards qui sont les regards ». Dans le tableau de Philippe Mayaux, ces regards se sont transformés en fourchettes, et la contemplation s'est mue en consommation. Cette absence de regard favorise l'aveuglement. « Il n'y a plus d'artistes ». « L'art contemporain est nul », un discours sans objet, une querelle d'experts dont le seul objectif est de prouver l'excellence de leurs expertises dans ce qu'il est désormais convenu d'appeler *Le Difficile Partage du gâteau*. ■



La Difficile Partage du gâteau
Le titre d'une peinture de Philippe Mayaux, jeune artiste français dont l'œuvre, depuis une dizaine d'années, oppose aux rouages tentaculaires des spéculations financières et intellectuelles du système culturel l'humilité de ses tableaux aux proportions réduites. 50x40 centimètres de haut et 40 de large.

(1) *La Haine de l'art*, Philippe Dagen, Grasset, 1988.
(2) *Le Ciel de l'art contemporain*, Yves Michaud, PUF, coll. « Interférences philosophiques », 283 p., 98 f.
(3) *Le monde de l'art*, sous la direction de Jean Clair, Cahiers de la Sorbonne, 1989, p. 51.
(4) *Le monde de l'art*, sous la direction de Jean Clair, Cahiers de la Sorbonne, 1989, p. 51.
(5) *Le monde de l'art*, sous la direction de Jean Clair, Cahiers de la Sorbonne, 1989, p. 51.
(6) *Le monde de l'art*, sous la direction de Jean Clair, Cahiers de la Sorbonne, 1989, p. 51.

Le monde de l'art, sous la direction de Jean Clair, Cahiers de la Sorbonne, 1989, p. 51.

(e) image 5 pré-segmentée

(f) image 6 pré-segmentée

●●● ainsi montré que les bébés de quatre-vingt mois sont capables de détecter la transgression ou la « conservation » du nombre lorsqu'on leur présente des événements numérotés impossibles ou possibles (dans une maquette de théâtre où les unités à quantifier sont des Mickey). Le traitement de l'information numérique réalisé par le bébé repose sur un processus analytique de « calcul précis » et non sur un processus de traitement perceptif global (ou holistique). Il a par ailleurs été démontré que ces compétences proto-numériques se retrouvent chez l'enfant d'âge préscolaire (entre deux et six ans) après une chute temporaire de performance.

Comment expliquer alors l'irrationalité de l'enfant de cet âge dans l'épreuve piagetienne de conservation du nombre : placé devant deux rangées de jetons en nombre égal mais de longueur différente parce que les jetons de l'une d'elles sont plus espacés, l'enfant considère qu'il y en a plus là où c'est plus long ?

Pour Piaget, l'enfant d'âge préscolaire serait encore fondamentalement intuitif, « préopérateur ». Pour traiter l'information, il se fonderait sur la longueur, selon un processus perceptif global. Il n'aurait donc pas encore construit le « schéma du nombre » lui permettant de procéder à un calcul précis. Les recherches récentes qui viennent d'être évoquées conduisent à douter de cette interprétation. En effet, ne serait-on pas plutôt ici, à l'occasion d'un nouveau « pli du temps », en présence de la coexistence d'un « rationnel construit » et de l'« irrationnel présumé révolu » ? Dans cet exemple, on trouve d'une part la capacité d'un traitement numérique analytique et précis et d'autre part la transgression du nombre dans l'épreuve de

conservation. Cette dernière épreuve n'est-elle pas avant tout une épreuve d'inférence « nombre/longueur » critique de la capacité à résister au schéma visuo-spacial « longueur = nombre » ? Il semble bien que les situations de conservation et d'inclusion des classes (voir plus loin) de Piaget ont en fait avec la capacité à résister aux interférences qu'avec la capacité de l'enfant à comprendre la logique sous-jacente. Ici aussi, comme pour l'erreur A-non-B dans la construction de l'objet, être intelligent (résister au schéma visuo-spacial), c'est essentiellement être capable d'inhibition.

Deuxième exemple, la catégorisation, c'est-à-dire le traitement qualitatif du réel (regroupement des objets selon certaines propriétés). Dans ce domaine, les recherches sur la logique des classes indiquent que, jusqu'à dix-onze ans, les enfants échouent lors de certaines épreuves modifiées d'inclusion.

Ainsi, l'enfant d'âge scolaire, en présence de dix marguerites et de deux roses, considère à tort que l'on peut obtenir plus de A (les marguerites) que de B (les fleurs) en ajoutant des A ou en enlevant des B. L'interprétation généralement retenue est que cet enfant est « empirique » et non logique, dans la mesure où il traite les classes emboîtées (A est inclu dans B) comme des collections disjointes.

Mais l'enfant d'âge scolaire est-il réellement empirique, au sens d'une absence de logique interne (absence du schéma d'inclusion des classes en mémoire), ou est-il « inhibiteur inefficace » face à des routines cognitives parfois inadéquates, telle ici la routine arithmétique de l'ajout et du retrait (ajouter des A, retirer des B) ? N'est-on pas à nouveau en présence d'un « pli du temps » où coexistent, d'une part, le « rationnel construit », en l'occurrence le schéma

d'inclusion appliqué dès sept-huit ans jusque dans l'adolescence prépubertaire (dans laquelle l'enfant considère correctement qu'il y a plus de B que de A), et, d'autre part, l'« irrationnel présumé révolu » : la transgression de ce schéma, jusqu'à dix-onze ans, lors des épreuves modifiées d'inclusion ? Piaget lui-même, dans l'un de ses derniers ouvrages, *Vers une logique des significations*, montre que dès cinq-six ans, en l'absence de piège de quantification extensionnelle des A et des B, l'inclusion ne soulève pas de problème. Il apparaît en effet, au regard des recherches récentes, que l'intelligence requise pour les épreuves modifiées d'inclusion, réussies vers dix-onze ans, consiste avant tout à inhiber la routine de l'ajout et du retrait, et à intégrer dans le raisonnement.

Trop de compétences nuit

Troisième exemple, le raisonnement, l'aspect le plus formel de l'intelligence. Les activités qui fondent la déduction, et son exigence de nécessité, ont fait l'objet de recherches nombreuses chez l'adolescent et chez l'adulte. Confrontées-elles à la compétence déductive de « l'épistémologie » décrit par Piaget ? Il semble que non. Depuis deux décennies, les données expérimentales se bousculent, les débats s'aiguisent, et l'effervescence est telle que la présomption de rationalité qui caractérise l'esprit humain d'une logique de principe est constamment « appelée à la barre ».

Ce sont les travaux relatifs aux biais de raisonnement qui font ressortir avec le plus de force l'irrationalité de l'adolescent et de l'adulte : la tendance systématique à prendre en considération des facteurs non pertinents pour la tâche à résoudre et à ignorer les facteurs pertinents.

Faut-il conclure, face à ce constat d'irrationalité, que l'essence du raisonnement humain n'est pas conforme à la logique, et que la compétence déductive est condamnée à être court-circuitée par des biais cognitifs ? Les données des recherches utilisant la méthode de l'apprentissage expérimental à l'inhibition vont dans le sens d'une « présomption de rationalité », dans la mesure où les biais qui sous-tendent les erreurs de raisonnement ne semblent pas relever d'un défaut de logique mais d'un défaut d'inhibition. Il s'agirait donc chez l'adulte, comme chez le bébé et l'enfant, depuis l'erreur A-non-B, premier ancêtre

Le public en appelle alors au rôle civique des institutions qui doivent défendre les valeurs culturelles et ne pas s'associer à ce qu'il considère comme une entreprise de démolition. Le conflit esthétique vise alors à la politique et à la morale. C'est pourquoi les débats les plus violents sont liés à des commandes publiques. Les colonnes de Buron en sont un exemple : le rejet esthétique d'une proposition heurtant les attitudes traditionnelles se combine à une réaction politique suscitée par une initiative ministérielle. Celle-ci a été en effet à contre-courant de ce que l'on exige d'une commande publique, à savoir un minimum de consensus.

Selon vous, les différends entre défenseurs et destructeurs de l'art contemporain relèvent-ils moins de divergences d'appréciation esthétique que de différences de catégories cognitives. Les querelles sont donc insolubles ?

Au siècle dernier, il s'agissait de décider si c'était beau ou non, si les peintres savaient peindre. En somme, c'était une question de goût. Aujourd'hui, on est en face d'un désaccord fondamental sur les catégories de classement : est-ce ou non de l'art ? Le problème n'est pas d'évaluer la valeur esthétique d'une œuvre, mais de définir sa nature même. Les conflits ne peuvent donc pas se résoudre lorsqu'on ne parle pas de la même chose : pas, ils se creusent.

Autre rupture historique : la place fondamentale qu'occupent aujourd'hui les institutions artistiques (centres d'art, musées, FRAC), notamment depuis quinze ans. Ce sont elles qui, par leur politique d'achat, accèdent aux avant-gardes en temps direct, sans attendre plusieurs générations. Quels sont les effets de ce face-à-face ?

Il met entre parenthèses les marchés privés. Les œuvres de jeunes artistes peuvent être achetées et exposées par des structures d'Etat sans être passées par le circuit des collectionneurs et des galeries. Depuis la politique de Jack Lang, la prise en charge institutionnelle de l'avant-garde s'est accélérée. Il en résulte des effets positifs mais aussi des

effets pervers. L'avant-garde est traditionnellement associée à une position progressiste, marginale par rapport à l'établissement. Aujourd'hui, les institutions sont de son côté. Or celui qui désire conserver une posture de franc-tireur doit aller à contre-courant de ce qui est admis par les institutions, autrement dit contre la reconnaissance, cette culture n'existe pas. De surcroît, il faut souvent sortir d'une posture d'admiration et d'émotion immédiate pour

Pourquoi la pédagogie est-elle si difficile auprès des adultes ?

En raison de la nature d'œuvres fondamentalement transgressives. Pour évaluer l'intérêt et la pertinence des propositions, juger que celle-ci est redoutable ou qu'elle est radicalement neuve, il faut avoir une culture de l'art contemporain. Or, faute d'enseignement artistique, cette culture n'existe pas. De surcroît, il faut souvent sortir d'une posture d'admiration et d'émotion immédiate pour

A force de surenchère, d'une extension toujours plus grande des frontières de l'art, l'art contemporain ne court-il pas à une impasse, à ce moment fatal où aurait été expérimentés tous les processus de déconstruction des critères qui définissent traditionnellement l'œuvre ?

C'est un risque. Encourager systématiquement ce qui est transgressif conduit à une limite où un dépassement des frontières civiquement et légalement acceptables. Or il y a toute une série d'œuvres qui ne transgressent plus de frontières et font un travail fort intéressant.

Propos recueillis par Macha Séry

(1) Les Représentations de l'art, le public et les valeurs, Séry, Jacqueline Chambon, coll. « Représentations », 1987. Le Japon de l'art contemporain : vingt-cinq ans plus tard, Editions de Minuit, coll. « Poésie », 1982, 148 p., 18 francs 10.



(g) image 7 pré-segmentée (h) image 8 pré-segmentée

Figure 4.10 : Résultats de pré-segmentation

Interprétation

Sur la figure 4.10.a, on utilise l'image de document « image1 » (Les caractéristiques de cette image sont indiquées précédemment). Cette image ne contient pas de composantes non textuelles donc l'algorithme ne détecte rien (l'image de sortie et l'image d'entrée sont identiques).

La même interprétation pour l'image (e) de la figure 4.10.

Sur la figure 4.10.b, on utilise l'image de document « image2 » ; pour cette image on a une mauvaise pré-segmentation tous les blocs sont classés textes par contre il existe un bloc non-texte (l'image d'un arbre). L'opération de dilatation élimine cette image « l'image de l'arbre » et l'algorithme la considère comme composante textuelle. Ce problème est expliqué en chapitre 3 (page 81) voir figure 3.25.

Sur la figure 4.10.c, on utilise l'image de document « Image3 », présentant un bloc non textuel (image) au dessus du texte, à droite du bloc image se trouvent aussi des lignes de texte. La grande masse de pixels noirs dans ce bloc facilite la création de masque d'image par l'algorithme de Bloomberg et la détection de l'image. La couleur bleu représente le masque d'image et le bloc est inclus dans ce masque.

Sur la **figure 4.10.d**, on utilise une image de document « **Image 4** », cette image contient un graphe entouré par des blocs textuels, le manque de bloc non textuel est plus claire.

La **figure 4.10.f**, illustre le résultat de pré-segmentation de l'image « **Image6** », c est une image graphique présentant des composantes textuelles. Dans ce type d'image la pré-segmentation est plus difficile car les traits minces peuvent être détruits par l'algorithme mais l'utilisation du remplissage (bouchage) donne un résultat optimal de pré-segmentation avec un temps d'exécution relativement élevé grâce au remplissage de trous.

Sur les deux figures **figure 4.10.g** et **figure 4.10.h**, on utilise respectivement les images de documents « **Image 7** » et « **Image 8** », la séparation texte/image est très précise aucune composante non textuelle n'est attachée aux images segmentées, l'algorithme donne des résultats satisfaisantes pour la segmentation des image inclus dans du texte.

Pour la majorité des images de documents utilisées, on remarque qu'on a obtenu une bonne détection de parties non textuelles, sauf que l'on observe quelques composantes mal segmentés, qui sont assignés à la classe texte, dans des cas particuliers lorsque le bloc non-textuelles est constitué de composantes séparées avec des dimensions comparables aux celles des caractères, comme le cas de la **figure 4.10.b**. Pour cela, on fait appel à un autre algorithme de segmentation qui utilise des opérations de soudage entre composantes, afin de remédier à ce problème, c'est l'algorithme RLSA utilisé dans la segmentation finale. La

4.3.3 Résultats de segmentation finale

Cette étape de segmentation finale réalise deux tâches : la détection des lignes de texte, et la correction de la segmentation initiale donc c'est une étape essentielle pour la détection des zones textuelle et complémentaire pour la détection des zones de nature non-textuelle.



(e) image 5 segmentée



(f) image 6 segmentée



(g) image 7 segmentée



(h) image 8 segmentée

Figure 4.11 : Résultats de segmentation finale

4.4 Evaluation quantitative des résultats

Dans cette section, notre objectif est d'évaluer les résultats de la méthode développée dans cette thèse. Nous définissons une fonction de précision appelée « P » (**équation 4.1**) qui donne le degré de ressemblance des résultats produits par l'algorithme développé. Cette fonction est calculée en comparant les blocs obtenus à des blocs référentiels de l'image de référence qui sont prédéfinis manuellement par un opérateur expert. On appelle « image G » l'image de référence et « image T » l'image obtenue par l'application de cette méthode.

$$P = \frac{G \cap T}{G \cup T} \quad (4.1)$$

Avec :

- $G \cap T$: nombre de pixels communs entre les deux images G et T.
- $G \cup T$: nombre de pixels des deux images G et T.



Image G **Image T**
Figure 4.12 : Image référentielle (G) et image à tester (T)

Pour une image contenant des blocs textes et des blocs non-texte on calcule la précision relative à chaque type de blocs indépendamment c'est-à-dire, on calcule « P_{text} » et « P_{image} ». Dans ce cas le degré de ressemblance de l'image entière « P » prend la moyenne des deux degrés (image ; texte).

Les résultats de l'évaluation sont organisés dans le (**tableau 4.1**) suivant :

Image	Image1	Image2	Image3	Image4	Image5	Image6	Image7	Image8
Deg _{text}	0.953	0.895	0.922	0.905	0.950	0.920	0.947	0.934
Deg _{image}	\	0.801	0.992	0.987	\	0.911	0.913	0.886
Deg	0.953	0.848	0.957	0.946	0.950	0.915	0.930	0.910

Tableau 4.1 : Résultats d'évaluation

La moyenne des degrés de similarité des images testées calculée à partir de ce tableau est égale à « Moy=0.926 » en utilisant la (équation 4.2) suivante:

$$\text{Moy} = \frac{\sum_{i=1}^n \text{Deg}_i}{n} \quad (4.2)$$

L'écart-type des degrés de similarité des images est égal à « $\sigma=0.0347$ » en utilisant la la (équation 4.3) suivante:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (\text{Deg}_i - \text{Moy})^2}{n}} \quad (4.3)$$

Nous avons atteint des résultats prometteurs d'après le tableau ci-dessus, ce qui montre bien l'efficacité de la méthode développée.

Interprétation

Le résultat de la segmentation finale en utilisant l'algorithme RLSA avec l'étiquetage des composantes connexes s'avère très bon. Dans les images (c), (d), (e), (f), (g), (h) de la **figure 4.11**, aucune ligne n'est chevauchée avec une image aussi pas de problème de sur-segmentation entre lignes (inter-blocs ou intra-blocs). On conclut alors que la coopération a abouti à de meilleurs résultats.

Pour l'image (a) de la **figure 4.11**, on remarque une mauvaise segmentation de quelques caractères attachés au « L » parce que ce dernier présente une grande épaisseur donc il est classer comme image.

Pour l'image (b) de la **figure 4.11**, on observe une correction de pré-segmentation le bloc image (de l'arbre) est initialement considéré comme texte mais la segmentation finale corrige son classification en bloc non-textuel.

Discussion

L'application des méthodes de segmentation d'images de documents demeure encore un domaine de recherche très vaste. Nous avons présenté dans ce mémoire, une méthode basées sur la coopération de plusieurs méthodes de différentes approches. La coopération est implémentée séquentiellement, elle commence par une stratégie de segmentation descendante en se basant sur les caractéristiques d'une image, et termine par une segmentation ascendante en exploitant les principales caractéristiques des lignes de texte qui sont les distances inter-caractères, inter-mots, interlignes et les épaisseurs des lignes.

Nous avons mis en œuvre une méthode de segmentation hybride qui utilise les propriétés d'une image et des lignes de texte à la fois. Nous avons utilisés une méthode de binarisation hiérarchique floue pour la classification des pixels aux classes objet ou fond cette binarisation donne une bonne classification des pixels de frontière fond-objet, et une méthode de lissage horizontalement et verticalement dont le but est de constituer les blocs constituant le document numérique et aussi les rectangles structurants de l'image, par la suite un étiquetage logique de composantes connexes est appliquée pour la détection des blocs ensuite, nous avons utilisé un deuxième étiquetage des rectangles structurant à l'intérieur des blocs.

CONCLUSION ET PERSPECTIVES

Le travail de recherche proposé dans ce mémoire rentre dans le cadre de l'analyse d'images, en particulier dans l'étude d'une méthode de segmentation, appliquée sur des images de documents imprimés, pour faire la discrimination entre les zones de texte et les zones de nature non-textuelle (image, graphe, dessin, carte, ...).

Dans ce travail de magister, nous nous sommes principalement intéressés à l'automatisation de la segmentation de blocs d'images de documents imprimés. Nous avons ainsi proposé un nouveau système de reconnaissance de structure physique de document qui se base sur une combinaison de deux méthodes différentes de segmentation l'une est descendante et l'autre est ascendante.

En premier lieu, nous avons présenté la problématique liée à ce travail et l'intérêt de faire l'analyse d'image de document . Nous avons décrit, les problèmes de segmentation liés à la complexité des documents. Et d'autre part, nous avons effectué une étude bibliographique de méthodologies existantes, pour résoudre notre problème.

Nous avons présenté ensuite en détail, quelques techniques de traitement d'images, avant de les mettre en œuvre. La troisième partie est le cœur de notre travail ; dans cette partie, on a décrit toutes les étapes de la mise en œuvre de notre approche. Nous avons présenté ensuite dans la dernière partie les résultats obtenus.

Il est apparu très clairement que le prétraitement, tel que le filtrage du bruit et la binarisation est un préalable incontournable. Les objectifs de ce mémoire ont donc été formulés comme suit : concevoir en premier lieu une méthode robuste de prétraitement puis, introduire les résultats obtenus dans la phase de segmentation.

La phase de prétraitement est débutée par un filtrage Gaussien, pour atténuer le bruit, issu de l'acquisition, cette opération transforme l'image en couleurs en une image à niveaux de gris. Les algorithmes étudiés dans ce travail sont applicables sur des images binaires, une deuxième opération de binarisation est effectuée pour passer de l'image à niveaux de gris en une image à deux niveaux seulement. Ces opérations de prétraitement ont pour but d'améliorer la qualité d'image et de restreindre la zone de travail avant la segmentation. Nous avons extrait par la suite les zones non-textuelles, les zones restantes peuvent être de nature textuelles ou non. Puis, nous avons effectué une détection des lignes de texte.

La méthode de segmentation proposée utilise une approche descendante et une approche ascendante. La méthode descendante segmente des cas difficiles ou les zones non-textuelles sont proches des lignes de texte, mais elle présente un inconvénient majeur, elle est incapable de détecter les composants non-textes de faible dimension. Pour remédier à ce problèmes, une méthode ascendante a été proposée, mais celles-ci souffre aussi du problème de sur-segmentation (chevauchement entre composantes).

La méthode proposée combine ces deux approches, en utilisant les avantages de l'une pour résoudre les défauts de l'autre. Cette coopération séquentielle a deux phases principales : pré-segmentation par une méthode descendante, et segmentation finale par une méthode ascendante.

Au niveau de la reconnaissance, l'architecture adoptée semble être performante puisqu'elle permet un traitement efficace.

Une limite de l'approche peut être observé est que l'algorithme RLSA peut échouer dans le cas d'une forte inclinaison de lignes de texte, toutefois, ce problème peut être corrigé on utilisant une correction d'inclinaison.

Les travaux futurs visent à segmenter les tables et les lignes.

ANNEXE A : Table de la loi de Fisher

La table donne les valeurs d'une loi de Fisher à $(n_1; n_2)$ degrés de liberté pour une probabilité 0,95 et pour une probabilité égale à 0,975.

$n_2 \backslash n_1$	1	2	3	4	5	6	7	8	9	10	12	14
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	243.90	245.36
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.42
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.71
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.87
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.64
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.96
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.53
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.24
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.03
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.86
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.74
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.64
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.55
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.48
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.42
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.37
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.33
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.29
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.26

Table 2.7 : Table de Fisher pour une probabilité 0,95.

$n_2 \backslash n_1$	1	2	3	4	5	6	7	8	9	10	12	14
1	647.79	799.48	864.15	899.60	921.83	937.11	948.20	956.64	963.28	968.63	976.72	982.55
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.41	39.43
3	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.34	14.28
4	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.75	8.68
5	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.46
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.30
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	4.60
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20	4.13
9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.87	3.80
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.55
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.43	3.36
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.21
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25	3.15	3.08
14	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21	3.15	3.05	2.98
15	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.96	2.89
16	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05	2.99	2.89	2.82
17	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98	2.92	2.82	2.75
18	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87	2.77	2.70
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88	2.82	2.72	2.65

Table 2.8 : Table de Fisher pour une probabilité 0,975.

BIBLIOGRAPHIE

- [1] S.D.Hassen « Classification de Blocs de documents», Université de Blida 1 (2012).
- [2] Véronique EGLIN, Stéphane BRES et Hubert EMPTOZ « Structuration de documents par repérage de zones d'intérêt » « Document segmentation by interest areas detection » Laboratoire de Reconnaissance de Formes et Vision RFV, INSA de Lyon (1999)
- [3] Karim Hadjar, « Une étude de l'évolutivité des modèles pour la reconnaissance de documents arabes dans un contexte interactif » Thèse de doctorat. Département d'Informatique de l'Université de Fribourg (Suisse). (2006)
- [4] Dalel Ketata et Maher Khemakhem « Un survol sur l'analyse et la reconnaissance de documents : imprimé, ancien et manuscrit » FSEGS, Université de Sfax BP 1088, 3018 Sfax, Tunisie. (2010)
- [5] Rémi Vieux, Jean-Philippe Domenger, « Segmentation non supervisée d'images de document en paragraphes », LaBRI - Université de Bordeaux - CNRS UMR 5800 - 341 Cours de la libération,33400 Talence (2012)
- [6] Hatt, M « Détermination automatique des volumes fonctionnels en imagerie d'émission pour les applications en oncologie » Thèse de doctorat. L'Université Occidentale et de Telecom Bretagne. (2008)
- [7] Lecoeur, J « Segmentation d'IRM cérébrales Multidimensionnelles par coupe de graphe » Thèse de doctorat. L'université de Rennes I. (2010)
- [8] G. Nagy ,S. S'eth. "Hierarchical representation of optically scanned documents," in *Proc. oJ ICPR*,pp.347-349,1984.
- [9] Mukkai Krishnamoorthy, George Nagy, Sharad seth, Mahish Viswanathan « Syntactic segmentation and labeling of digitized pages from technical journals » IEEE transactions on pattern analysis and machine intelligence Vol. 15,NO.7,July 1993
- [10] Francesca Cesarini, Marco Lastri, Simone Marinai, and Giovanni Soda« Page Classification for Meta-data Extraction from Digital Collections » Dipartimento di Sistemi e Informatica - Università di Firenze, Via S.Marta, 3 - 50139 Firenze - Italy

- [11] Yves Rangoni, Abdel Belaïd « Segmentation de documents composites par une technique de recouvrement des espaces blancs» LORIA Campus scientifique BP 239 F-54506 Vandœuvre-lès-Nancy Cedex (2013)
- [12] Nagy G., Seth S., Viswanathan M., « A prototype document image analysis system for technical journals », Computer, vol. 25, p. 10-22, 1992
- [13] Boontee Kruatrachue, Narongchai Moongfangklang, Kritawan Siriboon « Segmentation d'images IRM anatomiques par inférence bayésienne multimodale et détection de lésions » World Academy of Science, Engineering and Technology International Journal of Computer, Information, Systems and Control Engineering Vol:1 No:5, 2007
- [14] F. Cesarini, M. Gori, S. Marinai, G. Soda « Structured Document Segmentation and Representation by the Modified X-Y tree » OSI-Università di Firenze Via S.Marta ,3 - 50139 Firenze –Italy (1999)
- [15] Sukhvir Kaur, Palvinder Singh Mann « Improved XY cut Page Segmentation Algorithm for Border Noise » Sukhvir Kaur et al | IJCSET | May 2013 | Vol 3, Issue 5, 149-151
- [16] Xavier Philippeau, « segmentation en regions » document issu de <http://www.developpez.com> <http://www.developpez.com> (2008)
- [17] S.L. Horowitz and T. Pavlidis. Picture segmentation by a directed split and merge procedure. In CMetImAly77, pages 101.11, 1977.
- [18] Jean-Christophe Baillie « Segmentation », ES322 - Traitement d'Image et Vision Artificielle, ENSTA, Paris (2003)
- [19] Aliouche, W et Djadi, N.A « Reconnaissance tridimensionnelle à partir de coupes sériées : Application aux images médicales » Institut I.N.I Oued-Smar, Alger.(2008)
- [20] Cyril Meurie « Segmentation d'images couleur par classification pixellaire et hiérarchie de partitions » thèse de doctorat, université de CAEN/BA (2005)

- [21] Samia Bensalem « Segmentation automatique de tumeur cerebrale par contours actifs contraints par des relations spatiales » mémoire de magister en électronique, université de Blida 1(2015)
- [22] Dan S. Bloomberg « Multiresolution Morphological Approach to Document Image Analysis » Xerox Palo Alto Research Center 3333 Coyote Hill Road Palo Alto, CA 94304, USA. Presented at ICDAR, Saint-Malo, France, Oct, 1991
- [23] Syed Saqib Bukhari, Faisal Shafait, and Thomas M. Breuel « Improved Document Image Segmentation Algorithm using Multiresolution Morphology» Technical university of Kaiserslautern, Kaiserslautern, Germany. German Research Center for Arti_cial Intelligence (DFKI), Kaiserslautern, Germany (2010)
- [24] Hung-Ming Sun « Enhanced Constrained Run-Length Algorithm for complex Layout Document Processing» International Journal of Applied Science and Engineering 2006.4, 3: 297-309, Department of Information Management, Kainan University, No. 1 Kainan Road, Luchu, Taoyuan County, 33857, Taiwan, R.O.C
- [25] Asma OUJI, « Segmentation et classification dans les images de documents numérisés» Thèse de doctorat. École doctorale : InfoMaths Laboratoire d'Informatique en Image et Systèmes d'information (LIRIS), University Saad Dahleb de Blida I (2012)
- [26] Thibault Lelore « Segmentation d'image Application aux documents anciens» Mémoire de master Laboratoire des sciences de l'information et des systèmes, école polytechnique, université de Nantes (2007)
- [27] Jayat Maxime, Garcin David, Jolie Nicolas, «EyeLiner » Rapport de Soutenance 1, NightWare(2012).
- [28] Stefano Ferilli, Fabio Leuzzi, Fulvio Rotella and Floriana Esposito« A Run Length smoothing-Based Algorithm For Non-Manhattan Document Segmentation » Dipartimento di Informatica – Universit,,a di Bari Via E. Orabona, 4 70126 Bari (2012)
- [29] Ferilli, Stefano, Biba, Marenglen, Esposito, Floriana, and Basile, Teresa M.A. (2009). A distance-based technique for non-manhattan layout analysis. In Proceedings of the 10th International Conference on Document Analysis and Recognition (ICDAR-2009), volume I, pages 231–235. IEEE Computer Society.

- [30] J.Y. Ramel, S. Leriche « Segmentation et analyse interactives de documents anciens imprimés » Lab. d'Informatique, École Polytechnique de l'Université de Tours, 64, avenue Jean Portalis 37200 Tours (2004)
- [31] Robert M. Haralick, « Document Image Understanding: Geometric and Logical Layout» Electrical Engineering FT-10, University of Washington, Seattle, WA 98115.(1994)
- [32] J. M . Chassery; M. Melkemi « Diagramme de Voronoï appliqué à la segmentation d'images et à la détection d'événements en imagerie multi-sources» projet Segmentation de l'action Images du GDR 134 TDSI du CNRS. (12 décembre 1990)
- [33] Thai V. Hoang, Salvatore Tabbone, Ngoc-Yen Pham « Extraction of Nom Text Regions from Stele Images Using Area Voronoi Diagram » MICA Research Center, Hanoi University of Technology, Hanoi, Vietnam, Université Nancy 2, LORIA, UMR 7503, 54506 Vandoeuvre-lès-Nancy, France (2009)
- [34] K. Kise, Akinori Sato, Motoi Iwata. «Segmentation of page images using the area Voronoi diagram », Computer Vision and Image Understanding archive, Volume 70 (3), Special issue on document image understanding and retrieval, 1998, p. 370-382.
- [35] Koichi Kise, Motoi Iwata and Keinosuke Matsumoto « On the Application of Voronoi Diagrams to Page Segmentation» Department of Computer and Systems Sciences, College of Engineering, Osaka Prefecture University 1-1 Gakuencho, Sakai, Osaka 599-8531, Japan (1999)
- [36] Sylvain Chevalier « Reconnaissance d'écriture manuscrite par des techniques markoviennes : une approche bidimensionnelle et générique» Université René Descartes - Paris 5 Centre Universitaire des Saints - Pères UFR de mathématiques et informatique (2004).
- [37] Maitine Bergounioux « Quelques Methodes de Filtrage en Traitement d'Image » Université d'Orléans, Département de mathématique, BP 6759, F-45067 Orléans Cedex 2 (2010).
- [38] G, Bartholomew « Operateur de traitement d'image », 1920

- [39] S.D.Hassen « Classification de Blocs de documents», Université Saad Dahleb de Blida 1 (2012).
- [40] Marquegnies Julien « Eye Liner » Rapport de soutenance remis le 11 février 2009
- [41] Nghiem Anh Tuan « Reconnaissance d'écriture manuscrite» (2005)
- [42] Gilson Braviano. « Logique floue en segmentation d'images: seuillage par entropie et structures pyramidales irrégulières », thèse de doctorat à l'Université Joseph Fourier-Grenoble 1. Octobre 1995
- [43] Nguyen Thi Oanh « Binarisation d'images de documents graphiques » rapport de stage de fin d'études, Université de Nancy 2, Institut de la Francophonie pour l'Informatique, Laboratoire Lorrain de Recherche en Informatique et ses Applications
- [44] Azizi Rebiai « Une Approche Hybride Pour La Reconnaissance d'Écriture Arabe Manuscrite» Thèse de Magistère, Université de Constantine Faculté des Sciences de l'Ingénieur Département d'Électronique
- [45] Philippe Lefèvre « Reconnaissance de l'imprimé» Centre français d'exploitation, Techniques de l'Ingénieur, traité Informatique (1998)
- [46] N. Otsu, « , A threshold selection method from grey scale histogram » IEEE Trans. on Syst. Man and Cyber., vol 1, pp 62-66, 1979
- [47] T.Gadi, R. Benslimane. «Fuzzy hierarchical segmentation ». Traitement du signal, vol.7, n°1, 2000
- [48] « PJE-Projet Encadré IVI Semaine 5 : Analyse en Composantes Connexes » ; Université de Lille 1.
- [49] R. Gonzalez, R. Woods. « Digital Image Processing », Pearson Prentice Hall, First impression, 2009, pp.665, 682.

- [50] Mokhtar M. Hasan, Pramod K. Mishra « Improving Morphology Operation for 2D Hole Filling Algorithm» International Journal of Image Processing (IJIP), Volume (6) : Issue (1) : 2012
- [51] « Morphological operations on binary images », Technical University of Cluj-Napoca, Computer Science Departmen, Image Processing – Laboratory 7: Morphological operations on binary images (2002).
- [52] Yves Usson « Analyse d’images numériques en microscopie » Reconnaissance et Microscopie Quantitative, Laboratoire TIMC MR5525 CNRS Institut d’Ingénierie et d’Information de Santé (IN3S), La Tronche (2006)
- [53] M.Boutaunte, M.Fakir et B.Bouikhalene « Contribution à la reconnaissance des documents Tifinagh » Equipe de Traitement de l’Information et Télécommunication, Département d’Informatique, Facultés des Sciences et Techniques Béni Mellal, USMS, Maroc (2012)
- [54] Laurence Likforman-Sulem « Apport du traitement des images la numérisation des documents manuscrits anciens » Ecole Nationale Supérieure des Télécommunications, Département TSI, Paris (2003)
- [55] Rangachar Kasturi, Lawrence O’Gormane and Venu Govindaraju « Document image analysis: A primer » *Sadhana* Vol. 27, Part 1, February 2002, pp. 3–22. © Printed in India
- [56] Mohamed Ben Jlaiel, Slim Kanoun, Adel M. Alimi « Une méthode de segmentation d’Images de Documents Composites» Université de Sfax, Ecole Nationale d’Ingénieurs de Sfax, BP. W-3038 - Sfax – Tunisie (2006)
- [57] Y.Kabir « Programmer En langage C » Université de Blida Faculté des sciences de l’ingénieur Département de génie électrique (2011/2012).