

THÈSE

Contribution à la Restauration des Images de Documents Anciens

présentée devant
L'INSTITUT NATIONAL DES SCIENCES APPLIQUÉES DE LYON

pour obtenir
Le grade de docteur

École Doctorale Informatique et Information pour la Société (EDIIS)
Spécialité : Informatique

par
Fadoua DRIRA

Soutenue le – décembre 2007 devant la commission d'examen

COMPOSITION DU JURY

Examineurs	M. Christian JUTTEN	PROFESSEUR, U.J.F DE GRENOBLE
	M. Adel ALIMI	PROFESSEUR, ENI-SFAX-TUNISIE
Rapporteurs	M. Abdel BELAID	PROFESSEUR, UNIVERSITÉ DE NANCY2
	M. Christian OLIVIER	PROFESSEUR, UNIVERSITÉ DE POITIERS
Directeurs	M. Hubert EMPTOZ	PROFESSEUR, I.N.S.A DE LYON
	M. Frank LEBOURGEOIS	MAÎTRE DE CONFÉRENCES, I.N.S.A DE LYON

Résumé

La numérisation massive de documents patrimoniaux ouvre de nouvelles perspectives de Recherche comme la restauration des documents dégradés. Ces dégradations sont dues au mauvais état de conservation et à la numérisation. De telles altérations sur des images aussi complexes ne peuvent pas être traitées par une simple transposition des outils traditionnels de restauration. Cette thèse traite donc de la restauration des dégradations des images des documents anciens, et plus particulièrement des dégradations du support papier et des formes des lettres car elles nuisent à la lisibilité des documents numérisés et empêchent tout traitement automatisé.

Pour les dégradations du support papier, assimilées à un problème de superposition de couches, nous proposons deux méthodes. La première est une méthode de segmentation récursive issue de l'algorithme des nuées dynamiques appliquée sur les données décorréélées par l'analyse en composantes principales. La deuxième méthode est une amélioration de l'algorithme du Mean-Shift qui réduit sa complexité de calcul.

Nous proposons, ensuite, de traiter les dégradations des formes des lettres par des méthodes de diffusion. Tout d'abord, nous présentons les modèles existants les plus adaptés à la restauration des images de document. Ensuite, nous proposons un modèle de diffusion capable de filtrer les contours et de renforcer la continuité des traits tout en préservant les singularités.

Pour concevoir une approche capable de restaurer à la fois les dégradations du support papier et des formes des lettres, nous proposons une combinaison entre la segmentation et la diffusion. Cette combinaison permet de réaliser une segmentation en couches des images de documents tout en préservant la forme des contours des lettres.

Mots—Clés : Images de documents dégradés - restauration d'image - traitement par EDP - séparation recto/verso - segmentation colorimétrique - paradigme local/global.

Abstract

The massive digitization of heritage documents raised new prospects for Research like the restoration of the degraded documents. These degradations are due to the bad conditions of conservation and even to the digitization process. Images of old and degraded documents cannot be the retored directly by classical approaches. Hence, we propose in this thesis to develop and analyze document image restoration algorithms. We are mainly interested in foreground/background degradations, since they harm the legibility of the digitized documents and limit the processing of these images.

For background degradations, considered as a problem of the superposition of layers, we propose two-based segmentation methods. The first is a recursive approach that relies on the k-means clustering algorithm and the principal component analysis. The second method is an improvement of the initial algorithm of MeanShift in an attempt to reduce its complexity.

For foreground degradations, we propose to tackle the problem with PDE-based diffusion approaches. This solution has many useful features that are relevant for use in character restoration. Our comparative study of existing methods allows us to select the best approaches well adapted to our problem. We propose also a new diffusion method preserving singularities and edges while smoothing.

Our previously proposed solutions, the diffusion and the Mean-Shift algorithms, are used with success in a joint iterative framework to solve foreground and background degradations. This framework generates segmented images with more reduced artefacts on the edges and on the background than those obtained in separate application of each method.

Key–Words : Degraded Document images - image restoration - PDE based approaches- background/foreground separation - colorimetric segmentation - local/global paradigm.

Table des matières

1	Introduction Générale	3
1.1	La numérisation des documents du patrimoine	4
1.2	La restauration des images dégradées	7
1.3	Notre démarche	8
1.4	Organisation du mémoire	10
I	Dégradations/Restauration des Images de Documents	13
2	Dégradations des Images de Documents	15
2.1	Qu'est ce qu'une dégradation ?	16
2.2	Dégradations physiques du document	16
2.2.1	Dégradations chimiques	16
2.2.2	Dégradations biologiques	16
2.2.3	Dégradations physiques d'origine humaine	17
2.3	Dégradations dues à la capture	18
2.3.1	Dégradations matérielles	18
2.3.2	Dégradations numériques	21
2.4	Proposition d'une classification non-causale des dégradations	24
2.4.1	Dégradations uniformes	24
2.4.2	Dégradations non-uniformes	24
2.5	Conclusion	26
3	Restauration des Images de Documents	27
3.1	Préambule	28
3.2	Méthodes de restauration supervisées	28
3.2.1	Séparation recto/verso par recalage	28
3.2.2	Séparation recto/verso par analyse de l'orientation	30
3.2.3	Séparation recto/verso par analyse des formes	30
3.2.4	Séparation recto/verso par analyse de la couleur	30
3.2.5	Restauration des formes par modèle	31
3.2.6	Restauration des formes par prototypage	33
3.2.7	Restauration géométrique de la courbure	34

3.3	Méthodes de restauration non-supervisées	36
3.3.1	Restauration du recto/verso par séparation de sources	36
3.3.2	Restauration des formes par morphologie mathématique	36
3.3.3	Restauration Colorimétrique	37
3.4	Conclusions et discussions	38
II Dégradations du Fond : Restauration par Segmentation		41
4	Segmentation : Réflexion Méthodologique	43
4.1	Introduction	44
4.2	Approches globales	44
4.3	Approches locales	46
4.4	Approches duales	47
4.5	Conclusions et discussions	48
5	Restauration par Segmentation Globale	49
5.1	Introduction	50
5.2	Représentation des données	50
5.2.1	Principe de l'analyse en composantes principales	51
5.2.2	Décorrélation des couleurs par l'analyse en composantes principales	51
5.3	Limite d'une segmentation linéaire	53
5.3.1	Principe de l'algorithme des K -moyennes	53
5.3.2	Évaluation d'une segmentation linéaire	54
5.4	Proposition d'une classification en cascade	55
5.4.1	Justification de l'approche proposée	55
5.4.2	Critère de sélection pour une segmentation non-supervisée	56
5.5	Étude expérimentale de l'approche proposée	57
5.5.1	Étude des performances	57
5.5.2	Étude des limites	58
5.6	Conclusions et perspectives	58
6	Restauration par Segmentation Duale	65
6.1	Introduction	66
6.2	Le principe du Mean-Shift	66
6.2.1	Procédure du Mean-Shift global	67
6.2.2	Procédure du Mean-Shift spatial	68
6.3	Notre Proposition : le Mean-Shift local-global	70
6.3.1	Principe du Mean-Shift local-global	70
6.3.2	Processus de classification par le Mean-Shift local-global	71
6.4	Étude expérimentale de l'approche proposée	72
6.4.1	Étude de l'influence du choix des paramètres	72
6.4.2	Étude de la qualité du résultat	75
6.4.3	Étude du temps de calcul	75
6.4.4	Étude de la complexité	78
6.5	Conclusions et perspectives	79

III	Dégradations des Formes : Restauration par Diffusion	81
7	Restauration des Formes par Diffusion	83
7.1	Introduction	84
7.2	Bruit : Lissage par EDP	84
7.2.1	Diffusion linéaire isotrope	85
7.2.2	Diffusion non-linéaire	86
7.3	Flou : Rehaussement	96
7.3.1	Rehaussement sous l'angle d'une opération inverse de celle de la diffusion	96
7.3.2	Renforcement des contours : Filtre de choc	96
7.4	Expérimentation sur les images de documents	102
7.4.1	Evaluation des méthodes de diffusion	102
7.4.2	Évaluation par OCR	105
7.5	Conclusion	112
8	Reconstruction par Inpainting	117
8.1	Introduction	118
8.2	Principe de l'inpainting	118
8.2.1	La diffusion des intensités de l'image dans la direction des isophotes .	120
8.2.2	La diffusion des intensités suivant le modèle de la totale variation (TV)	120
8.2.3	La diffusion des intensités dirigée par la courbure (CDD)	121
8.3	Application aux documents	122
8.4	Expérimentation et évaluation	123
8.4.1	Expérimentation sur des images de synthèses	123
8.4.2	Expérimentation sur des images réelles de documents	126
8.5	Discussion sur les méthodes de reconstruction par inpainting	129
8.6	Conclusion	131
9	Proposition d'un Nouveau Modèle de Diffusion	133
9.1	Les limites des modèles existants	134
9.2	Proposition d'un nouveau modèle de diffusion	135
9.2.1	Vers un modèle de diffusion qui préserve les singularités	135
9.2.2	Fonctionnement de notre modèle de diffusion	137
9.3	Notre modèle versus les autres modèles de diffusion	139
9.4	Expérimentations et paramétrages	140
9.4.1	Résultats sur les images de documents	141
9.4.2	Application aux autres images	144
9.5	Conclusion	146
IV	Dégradations Non-Uniformes : Restauration par une Approche Unificatrice	153
10	Segmentation/Diffusion :Vers un Cadre Commun	155
10.1	Introduction	156
10.2	Étude comparative entre le Mean-Shift global et la diffusion	156

10.2.1	Formulation discrète de la diffusion scalaire de Perona-Malik	156
10.2.2	Formulation discrète simplifiée du Mean-Shift global	157
10.3	Proposition d'une combinaison coopérative	158
10.3.1	Principe de notre proposition	158
10.3.2	Implémentation et discrétisation	159
10.4	Combinaison coopérative versus combinaison séquentielle	159
10.5	Conclusion	162
11	Segmentation/Diffusion : Évaluation des Performances	167
11.1	Introduction	168
11.2	Étude du choix des paramètres	168
11.2.1	Cas de (σ_R, K)	168
11.2.2	Cas de (α, β)	169
11.3	Étude de la convergence	170
11.4	Mise en œuvre pratique	171
11.4.1	Application sur des images de documents	171
11.4.2	Application sur des images naturelles	179
11.5	Conclusion	179
12	Conclusion Générale	183
12.1	Résumé de nos contributions	183
12.2	Perspectives de recherche	185
A	EDP : Définitions et notations	187
A.1	Définition des EDP	187
A.2	Application aux images discrètes	187
A.3	Dérivées successives d'une image et leurs significations	189
A.4	Le tenseur de dérivées premières et son interprétation géométrique	193
A.5	L'application des EDP aux images couleurs	195
	Bibliographie	197

Chapitre 1

Introduction Générale

« Il faut toujours connaître les limites du possible. Pas pour s'arrêter, mais pour tenter l'impossible dans les meilleures conditions. »

Romain Gary

Résumé

Depuis quelques années, on numérise quotidiennement des quantités importantes de documents issus du patrimoine. Cette révolution technologique a ouvert de nouveaux domaines de recherche comme la restauration des images qui constitue le cadre de notre travail. Nous introduisons ainsi dans ce chapitre le principe de la numérisation des documents, ses principaux avantages et limites. Nous enchaînons ensuite par la présentation du cadre de nos travaux de recherche et nous terminons par l'organisation du mémoire et l'articulation entre les parties.

Sommaire

1.1	La numérisation des documents du patrimoine	4
1.2	La restauration des images dégradées	7
1.3	Notre démarche	8
1.4	Organisation du mémoire	10

1.1 La numérisation des documents du patrimoine

Le document joue un rôle important dans le développement des civilisations et de leurs cultures ; il permet de conserver et de transmettre les connaissances d'une génération à l'autre. Au fil des générations, le support physique des documents écrits a beaucoup évolué, allant des supports rigides (le bois, la pierre, le marbre, les galets, les tablettes de cire ou d'argile) aux supports souples (le papyrus, le parchemin, le cuir, le papier). Le papier, moins coûteux et plus commode, s'est généralisé dans le monde comme le principal support, mais, le papier est un support fragile dont la conservation est difficile. L'invention de l'imprimerie moderne au XV^{ème} siècle a marqué le début d'une nouvelle ère de production et de diffusion massive du savoir ; l'imprimerie n'est pas le résultat d'une seule technique mais d'un ensemble de techniques parvenues à maturité et correctement agencées.

Le récent développement des technologies numériques de l'information et de la communication a révolutionné la transmission des connaissances et remis en cause la suprématie du papier comme support unique de diffusion des savoirs. Le monde de l'édition physique (sur support papier) est complètement bouleversé par le monde du numérique. Le document électronique est aujourd'hui le vecteur de diffusion le plus utilisé dans le monde. Le document électronique possède de nombreux avantages par rapport aux documents papier.

La numérisation est la technologie qui permet de convertir des documents physiques en images électroniques. Nous définissons la numérisation au sens large du terme comme l'ensemble des opérations nécessaires pour capturer les images et les transformer en données exploitables informatiquement. La numérisation fournit les images et les méta-données associées ainsi que les informations nécessaires au fonctionnement des services offerts aux usagers. Cette définition de la numérisation est encore plus large que celle de la rétro-conversion qui consiste à convertir les documents papier dans un format logique pivot les rendant de nouveau rééditables et donc réutilisables. La rétro-conversion concerne la transformation complète d'un document papier en un document électronique proche de l'original alors que la numérisation, au sens large, s'intéresse aussi à l'accessibilité et aux services que l'on peut offrir aux usagers. Au sens strict du terme, la numérisation consiste seulement à effectuer la capture des images des documents, on parlera alors de capture d'images.

Les données numériques ne remplaceront jamais les supports physiques traditionnels, elles ne sont qu'une représentation plus ou moins fidèle des documents originaux. Par conséquent, la «copie numérique» ne constitue pas une alternative pérenne à la conservation du patrimoine culturel. Cependant, la numérisation présente de nombreuses propriétés très intéressantes :

- **Facilité de reproduction** : Les données numériques peuvent être dupliquées très rapidement sans aucune perte d'information. En outre, ces données peuvent être échangées, stockées, manipulées, interrogées et recopiées avec des coûts très réduits. Une dissémination très rapide des connaissances dans le monde entier est donc possible notamment avec le développement du World-Wide-Web qui rend les ouvrages de notre patrimoine accessibles au monde entier, réduisant un peu plus la fracture entre les pays du nord et les pays du sud.
- **Indépendance du support physique** : Les données numériques peuvent être stockées sur des supports variés de façon permanente ou bien itinérante sur des disques

d'ordinateurs ou des supports amovibles. Ces données peuvent être reproduites sur de nombreux supports (papiers, écrans, papiers électroniques...) et même être consultables n'importe où et par n'importe quel moyen informatique fixe ou mobile (PC, PDA, ebook, téléphone portable...). La numérisation va donc révolutionner l'accès à l'information et à la connaissance. De nombreux chercheurs peuvent travailler simultanément à distance sur les mêmes documents ou bien consulter des documents rares autrefois peu accessibles.

- **Facilité de consultation** : La consultation de documents numériques offre un confort visuel sans précédent, une navigation plus performante et une recherche d'information plus efficace en comparaison avec les formes traditionnelles de consultation des livres ou des microfilms. En effet, les moteurs de recherche permettent de retrouver efficacement et de comparer rapidement les informations, dans plusieurs bibliothèques numériques. Néanmoins, ces moteurs de recherche fonctionnent uniquement sur des textes électroniques et ne peuvent pas indexer l'information directement à partir des images. Une solution à ce problème consiste soit à décrire manuellement leurs contenus, soit à développer des systèmes de reconnaissance capables d'extraire automatiquement des informations dans les images.
- **Préservation et conservation des documents originaux** : La numérisation permet de garder une copie des documents originaux risquant de partir en poussière. Nous rappelons, à titre d'exemple, l'incendie qui a eu lieu le 12 juin 1999 à la Bibliothèque Universitaire des Lettres et Sciences Humaines et Sociales de Lyon. Cet incendie a permis de faire prendre conscience du caractère éphémère des ouvrages anciens et de l'intérêt d'une conservation alternative. Les copies numériques réduisent les consultations excessives des lecteurs qui contribuent à l'usure des documents originaux fragiles. C'est la raison qui pousse les gestionnaires des fonds patrimoniaux à financer de coûteux projets de numérisation. L'UNESCO, responsable de la protection du patrimoine culturel mondial, a lancé en 1992 le programme «Mémoire du monde», pour sauvegarder et promouvoir ce patrimoine par la numérisation [STE-95]. Depuis, les services des archives, tant nationales que départementales, lancent d'importantes campagnes de numérisation. En France, le projet de numérisation de la Bibliothèque Nationale (Gallica) a permis de donner vie à l'une des plus importantes bibliothèques numériques, avec plus de 80000 documents accessibles en ligne.

Mais la numérisation seule ne suffit plus [AND-99]. Il faut absolument qu'elle cohabite avec le développement d'outils informatiques destinés à améliorer les conditions d'accès et de recherche.

L'analyse d'images de documents est un domaine de recherche actif à la frontière de la reconnaissance des formes et de l'analyse d'images. L'analyse d'images de documents permet aujourd'hui de retrouver des mots dans les manuscrits médiévaux ou bien de reconnaître les mots si on dispose de dictionnaires ; il convient de rappeler que les traitements automatiques ne fonctionnent correctement que si les images possèdent les qualités suffisantes. En effet, la numérisation permet de réaliser une «copie numérique» des documents ; mais, pour un grand nombre de ces documents, la numérisation arrive trop tard et beaucoup de documents ont été numérisés en mauvais état. En plus de la détérioration physique des documents

anciens, les images montrent des dégradations dues au processus même de numérisation. En effet, le développement du marché des capteurs pour un usage quotidien (scanner de bureau, appareil photo numérique, Webcam, caméra embarquée sur les téléphones mobiles, les caméras numériques DV...), ont poussé les constructeurs à développer des technologies à bas coût en réduisant la qualité des capteurs, de l'électronique, de la mécanique et des optiques. Toutes ces technologies sont incompatibles avec une numérisation de qualité. C'est ce qui explique pourquoi les numériseurs de qualité, spécialement dédiés, sont des machines complexes très coûteuses. Ces numériseurs à haute performance réduisent les contacts avec les ouvrages précieux en évitant de les presser contre une vitre grâce à une caméra montée sur un portique. Ils évitent d'imposer des contraintes physiques aux reliures avec des plateaux spéciaux. De plus, ils utilisent des lumières froides pour éviter d'élever la température des ouvrages fragiles (Fig.1.1). En cherchant à préserver les documents originaux, ces numériseurs produisent des défauts géométriques et photométriques qui nécessitent des prétraitements.



Figure 1.1 : *Trois types de scanners ; respectivement de gauche à droite : Scanners sans vitre, Scanner avec lumière froide, Scanner avec plateau pour préserver la reliure du livre*

En plus du coût élevé des numériseurs, les coûts engendrés par les opérations manuelles de manipulation, de suivi de la qualité des images, de stockage et d'indexation rendent les projets de numérisation très onéreux pour les bibliothèques. Face aux coûts excessifs d'une numérisation de qualité et à la grande quantité de documents à numériser, les responsables des fonds patrimoniaux privilégient la quantité au détriment de la qualité des images. Ils favorisent les procédés rapides de numérisation comme la numérisation des microfilms. Pour réduire les coûts de stockage, ils choisissent des formats destructeurs de compression comme JPEG ou bien conservent les images monochromes ou bitonales, rendant difficile leurs traitements futurs. Pour les mêmes raisons économiques, ils privilégient la numérisation avec une résolution moyenne ou faible, tout juste adaptée à la lecture à l'écran mais impropre à l'analyse d'images. Enfin, même si les responsables choisissent d'utiliser des numériseurs de qualité avec des paramètres optimaux, il n'est pas pour autant facile de mettre en œuvre leurs performances.

La restauration des images est donc très souvent utilisée comme prétraitement pour rendre les images exploitables informatiquement. La restauration est un domaine clé qui doit être privilégié et traité en priorité car la performance des systèmes de reconnaissance est fortement limitée par la qualité d'images.

1.2 La restauration des images dégradées

Il existe actuellement des images de documents numérisés de différentes origines, de qualité très variable ; elles présentent des dégradations dues au processus de numérisation et montrent des documents dégradés par l'usure du temps. Ce sont ces deux sources de dégradation qui sont au cœur des problèmes actuels et qui rendent difficile l'exploitation de ces images. La restauration numérique des images constitue alors une solution pour retrouver la représentation la plus fidèle possible du document d'origine. Nous définirons plusieurs types de restauration :

- **La restauration physique** : Elle a pour principal objectif la conservation et la sauvegarde des documents anciens ou fragiles. En effet, ce patrimoine écrit dont la valeur historique ou artistique est souvent inestimable n'échappe pas, en dépit des précautions prises, à l'usure du temps (vieillesse, modification des couleurs, porosité du papier, acidité de l'encre, humidité et variation de la température de stockage, présence d'insectes, sinistres...). La restauration physique consiste à redonner au document sa forme originale. Ce sont des méthodes de restauration physico-chimiques qui permettent de ralentir le vieillissement des documents et de renforcer les supports.
- **La restauration numérique** : Elle regroupe l'ensemble des opérations de traitements informatiques qui ont pour objectif de retrouver la qualité de l'image originale avant sa dégradation par le processus de numérisation. Le traitement de l'image permet de corriger les défauts qui apparaissent lors du processus de capture d'images. Ces défauts sont notamment des défauts géométriques (inclinaison du document, courbure apparente le long de la reliure...), des défauts optiques (éclairage non-homogène, netteté de l'image...), des défauts électroniques (bruit et qualité du capteur), des défauts informatiques (traitements destructifs de l'image sans sauvegarde des originaux, compression avec perte...) ou même des défauts extérieurs dus aux conditions de la prise d'image (présence d'éléments non-désirés, extérieurs au document). Dans ce cas précis, le traitement numérique doit être le dernier recours si on ne peut ni refaire une autre capture d'image ni améliorer les conditions de la numérisation ; il peut rarement retrouver l'information déformée par le processus de numérisation.
- **La restauration virtuelle** : Elle cherche à retrouver la forme originale des documents par des méthodes de traitement des images. Elle est donc complémentaire à la restauration physique. Elle permet par exemple de retrouver la couleur d'origine, d'atténuer le verso quand il traverse le recto, de remplir les trous, d'enlever les taches ou de corriger la courbure après la restauration de la reliure. Il existe aujourd'hui une polémique sur l'intérêt de restaurer «virtuellement» les documents numérisés. Certains pensent qu'il est important de retrouver le document tel qu'il était avant les dégradations du temps, d'autres estiment qu'il faut conserver la représentation la plus fidèle possible de l'état actuel du document avec toutes les dégradations subies avec le temps ; tout dépend de l'usage que l'on fait de ces documents.

La restauration virtuelle a été appliquée à un nombre très réduit de domaines comme la restauration des films à l'INA¹, des bandes sonores, des peintures et des fresques. La restauration des documents écrits du patrimoine est donc un domaine émergent et qui constitue le cadre de notre travail.

1.3 Notre démarche

Nous avons pour objectif de restaurer des images dégradées de documents. La restauration (numérique ou virtuelle) doit permettre de répondre à un maximum de besoins et redonner une nouvelle vie aux documents de notre patrimoine.

Nous disposons actuellement de nombreux documents anciens, du Moyen-Âge jusqu'à nos jours. Ces documents possèdent des contenus très variés, ils peuvent être manuscrits ou imprimés, écrits avec des alphabets différents (Latin, Arabe, Grec...) et avoir des formes différentes de mise en page. Nous avons étudié les nombreuses bases d'images que le LIRIS a accumulé, depuis plusieurs années, à partir de différents projets de numérisation. Les principaux sont (Fig.1.2) :

- Les manuscrits de Tombouctou au Mali (XI^{ème}-XVI^{ème}).
- Les imprimés de la Bibliothèques Municipale de Lyon (livre de Nostradamus, XVI^{ème}).
- Les premiers imprimés de la Renaissance (fond du projet Européen DEBORA XV^{ème}-XVI^{ème}).
- Les manuscrits médiévaux de l'IRHT (Institut de Recherche et d'Histoire des Textes, XV^{ème}-XVI^{ème}).
- Les archives de Châtillon-sur-Chalaronne dans l'Ain (XII^{ème}-XVIII^{ème}).
- Les archives départementales de Savoie (début XX^{ème}).
- Les gazettes de Leyde (XVIII^{ème}-XIX^{ème}).
- Le dictionnaire de Trevoux (XVIII^{ème}).

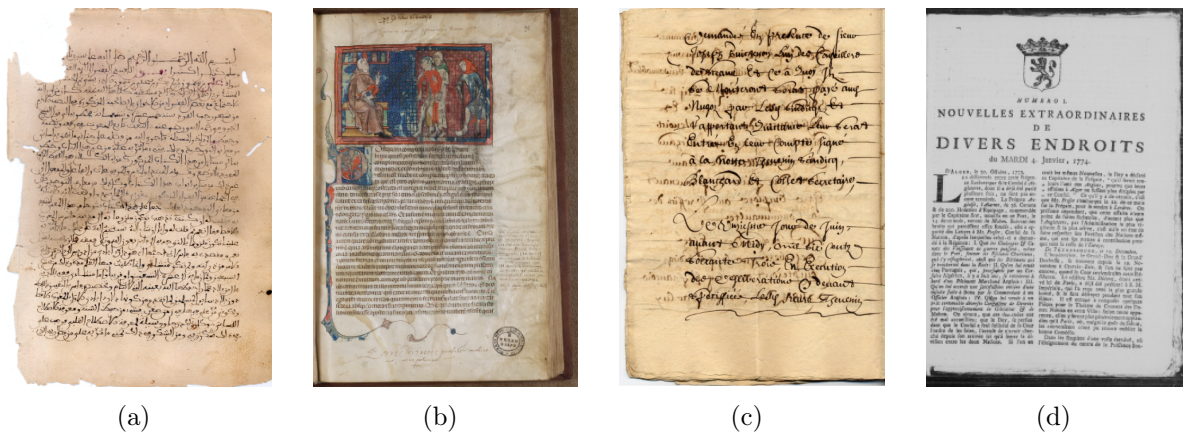


Figure 1.2 : (a) Manuscrits de Tombouctou, (b) Images de documents de l'IRHT (XIV^{ème}), (c) Archives de Chatillon sur Chalaronne (XVI^{ème}), (d) Gazettes de Leyde

¹Projet Européen Aurora (AUtomedated Restoration of ORiginal film and video Archives)

C'est à travers ces bases d'images et de l'expérience acquise par le LIRIS dans le domaine de la numérisation que nous avons pu identifier et analyser les dégradations les plus courantes. Ces bases de plusieurs dizaines de milliers d'images couvrent une large période de l'Histoire et nous ont permis de mesurer la grande variété des contenus et des problèmes liés à la qualité de ces images. On ne connaît actuellement aucune méthode de restauration qui peut s'adapter à une telle variété de documents. Il faut donc trouver les traitements les plus génériques possibles qui soient indépendants des contenus des images.

Les techniques développées pour la restauration des images de documents ne sont pas la simple transposition des techniques existantes dédiées aux images naturelles. L'explication se trouve tout simplement dans les différences qui existent entre les images naturelles (Fig.1.3) et les images de documents (Fig.1.4). Les images de documents sont des images particulières qui possèdent les spécificités suivantes :

- Ce sont des images de traits qui forment des structures complexes qui doivent être préservées comme les lettres.
- Les informations sont concentrées sur des parties très réduites de l'image. La simple perte d'une partie infinitésimale peut rendre un texte illisible.
- Il existe une certaine continuité le long des lignes de mêmes intensités.
- La distribution statistique des couleurs montre des classes rarement séparables linéairement.
- La proximité colorimétrique entre les différentes couches d'information rend particulièrement difficile leurs séparations quand elles sont mélangées comme la séparation entre le recto et le verso par transparence.
- Les dégradations sont aléatoires et donc rarement modélisables et prédictibles. Les méthodes à base d'apprentissage ou de modélisation sont généralement limitées au traitement d'une seule dégradation sur des documents particuliers.
- Les dégradations peuvent se superposer de façon non-linéaire rendant très difficiles leurs séparations.

Nous avons établi une classification non-causale des dégradations des documents qui permet de traiter un ensemble de dégradations par une même méthode. Ainsi, nous différencions deux familles de dégradations :

- **Les dégradations du fond** : Souvent liées à la détérioration du support papier, elles font apparaître des taches d'humidité et laisse les textes du verso passer au recto par transparence. Nous considérons ces défauts comme un problème de séparation de couches d'information. Nous traiterons ce problème par la segmentation couleur.
- **Les dégradations des formes** : Elles effacent une partie des traits et déforment leurs contours. Ces dégradations rendent les textes illisibles ou difficiles à segmenter. Nous avons choisi des méthodes à base de diffusion pour restaurer ou reconstruire les formes des traits et des caractères.

Cette classification des dégradations et des traitements associés à chaque classe constitue le fil d'Ariane de notre travail autour duquel nous avons construit le plan de ce mémoire.

1.4 Organisation du mémoire

Ce mémoire est construit en quatre parties. La première partie resitue nos travaux dans son contexte et présente les dégradations des images de documents et les méthodes actuelles de restauration. La deuxième partie décrit les travaux relevant de la restauration du fond par analyse de la couleur. La troisième partie étudie l'adaptation des approches par diffusion et leurs performances pour la restauration des formes. La quatrième partie s'inspire des précédentes pour dégager une approche générique capable de restaurer efficacement les dégradations locales du fond et des formes.

Partie I : Dégradations et restaurations des images de documents

La première partie de ce manuscrit définit les dégradations et les méthodes actuelles pour les traiter. Elle est décomposée en deux chapitres :

Le *chapitre 2* présente une typologie causale des dégradations suivant leurs origines. Puis nous proposons une classification fonctionnelle orientée traitements d'image, utile pour définir les outils nécessaires à la restauration des différentes dégradations.

Le *chapitre 3* traite des méthodes existantes de restauration des images dégradées de documents.

Partie II : Dégradations du fond : Restauration par segmentation

La deuxième partie s'intéresse à l'apport de la segmentation couleur pour la restauration des dégradations du fond pour séparer le texte du fond quand ils sont mélangés colorimétriquement. Elle se scinde en trois chapitres.

Le *chapitre 4* présente une réflexion méthodologique sur la segmentation d'images couleur. Nous proposons de différencier méthodes locales et globales de segmentation car c'est autour de cette dualité que se situent les apports essentiels de cette thèse.

Le *chapitre 5* propose une méthode de segmentation globale basée sur une segmentation récursive à base de l'algorithme des nuées dynamiques.

Le *chapitre 6* introduit une méthode de segmentation locale/globale basée sur une adaptation de l'algorithme du Mean-Shift et qui permet de réduire la complexité de calcul sur des images de grandes tailles comme les images de documents.

Partie III : Dégradations des formes : Restauration par diffusion

La troisième partie du document étudie les apports de la diffusion pour la restauration des dégradations des traits et des formes. Cette partie comporte trois chapitres.

Le *chapitre 7* décrit les performances et les limites des méthodes de diffusion appliquées à la restauration des images de documents. Nous prouverons que les méthodes de diffusion permettent d'améliorer les performances des OCR et de restaurer visuellement les textes.

Le *chapitre 8* décrit et évalue les méthodes de reconstruction virtuelle de parties manquantes d'une image de documents par la diffusion (inpainting).

Le *chapitre 9* propose un nouveau modèle de diffusion adapté aux images de documents qui préserve les singularités et renforce les contours.

Partie IV : Dégradations fond/formes : Restauration par une approche unificatrice

Cette dernière partie propose de combiner les méthodes locales de diffusion pour réparer les dégradations des formes et les méthodes globales de segmentation couleur pour séparer les formes du fond. C'est une réponse possible au paradigme local/global et à la correction simultanée du fond et des formes. Cette dernière partie est organisée en deux chapitres :

Le *chapitre 10* introduit une approche générique convergente par une combinaison coopérative entre la segmentation colorimétrique et la diffusion anisotrope. Nous comparons notamment les performances de cette combinaison avec celles données par un enchaînement séquentiel des méthodes.

Le *chapitre 11* présente une évaluation des performances de l'approche unificatrice proposée sur des images de documents et des images naturelles.



Figure 1.3 : Exemples d'images naturelles



Figure 1.4 : Exemples d'images de documents

Première partie

Dégradations/Restauration des Images
de Documents

Chapitre 2

Dégradations des Images de Documents

« La vraie philosophie est de voir les choses telles qu'elles sont. »

George Louis Buffon

Résumé

Les images de documents numérisés laissent apparaître souvent de nombreux défauts. Certains dépendent de l'état de conservation du document (la qualité du papier, l'acidité de l'encre, l'humidité du lieu de stockage, vieillissement du support...) ou des modifications apportées par l'Homme (annotations, soulignements, mauvaise restauration physique...). D'autres proviennent du processus même de numérisation (résolution insuffisante, courbure apparente, restitution non fidèle des couleurs...). La plupart des travaux, traitant les images de documents anciens, font référence à une classification des dégradations en fonction de leurs causes. Nous rappelons ce système de classification et nous proposons ensuite un système de classification mieux adapté aux traitements d'images et donc à nos travaux sur la restauration.

Sommaire

2.1	Qu'est ce qu'une dégradation ?	16
2.2	Dégradations physiques du document	16
2.3	Dégradations dues à la capture	18
2.4	Proposition d'une classification non-causale des dégradations .	24
2.5	Conclusion	26

2.1 Qu'est ce qu'une dégradation ?

La notion de dégradation des images tient une place centrale dans le présent mémoire. Il convient donc de la définir. On appelle dégradation tous les effets indésirables cumulés nuisant à la lisibilité, au traitement ou à la conservation des images. Les dégradations ont plusieurs origines et l'accumulation des défauts rend difficile leurs séparations. Les dégradations ont des effets plus destructeurs sur les images de documents que sur toutes autres images dites «naturelles».

2.2 Dégradations physiques du document

2.2.1 Dégradations chimiques

Le papier est un moyen de stockage fragile et éphémère. Il est constitué par l'essentiel de matériaux naturels ou organiques qui sont chimiquement peu stables, les réactions chimiques, qui peuvent avoir lieu, peuvent provoquer des modifications irréversibles dans la structure des matériaux (Fig.2.1). Les processus chimiques ont des chances de s'intensifier avec l'accroissement de la température, l'humidité, la luminosité et la pollution atmosphérique. Toutes ces réactions chimiques peuvent se traduire, d'une part, par le jaunissement du papier, et de l'autre, par une décoloration des encres et des pigments. Une destruction complète du papier peut avoir lieu pour des réactions photochimiques au contact de l'oxygène. L'utilisation des matériaux telles que la colle ou les feuilles de protection en contact direct avec le support papier aide à propager des substances nuisibles qui peuvent ainsi provoquer des dommages.



Figure 2.1 : Exemples des dégradations chimiques

2.2.2 Dégradations biologiques

Les dégradations biologiques des documents (Fig.2.2) sont causées par des organismes vivants qui représentent «les parasites de la civilisation». Les dommages d'origine biologique diffèrent des précédents quant à la leurs effets. Ils peuvent être causés par différents agents :

- **Les micro-organismes** (bactéries, moisissures) provoquent des taches dont la couleur varie avec l'organisme pathogène. Ces taches envahissent divers types de matériaux et provoquent des altérations irréversibles. L'humidité et la chaleur favorisent le développement de ces micro-organismes.
- **Les insectes** laissent des traces de morsure et d'excoriations. La concentration de matériaux pouvant servir de nourritures dans les bibliothèques, favorise la prolifération des insectes.
- **Les rongeurs** nuisent en particulier à la partie extérieure des ouvrages. Ils grignotent les couvertures, les reliures et les bords des livres.



Figure 2.2 : Exemples de dégradations biologiques

2.2.3 Dégradations physiques d'origine humaine

L'homme peut contribuer à la dégradation progressive du patrimoine documentaire. Outre les désastres inévitables comme les accidents ou les guerres, certaines dégradations qui sont imputables aux manques dans les règles élémentaires de conservation, peuvent être évitées. Par manque de moyens financiers, les bibliothécaires ne peuvent pas toujours investir dans des systèmes de climatisation pour réduire l'humidité. Certaines n'ont ni sécurité efficace contre les incendies ni alarmes appropriées.

Toute manipulation physique, par les conservateurs eux-mêmes, détériore les documents fragiles. Citons à titre d'exemple, le cas d'un livre qui tombe par maladresse d'un chariot en cours de transport, un mauvais rangement des microfilms provoquant des rayures. En outre, les bibliothécaires n'hésitent pas à apposer systématiquement des tampons à toutes les pages et même sur les parties écrites les rendant ainsi illisibles (Fig.2.3). Les annotations (numéros de pages, dates, explications, traductions...) et les rayures (corrections, censures...) effectuées par les chercheurs et les conservateurs contribuent à «l'histoire» de l'ouvrage mais aussi à sa dégradation. Les erreurs de restauration constituent le cas le plus extrême de dégradations humaines pouvant être facilement évitées. La restauration de la reliure, des iconographies ou la consolidation des supports papier peut être mal réalisée et détériore encore plus le document. Enfin, la consultation physique des ouvrages est aussi une des

principales sources de destruction des documents fragiles. C'est la raison pour laquelle, les bibliothécaires limitent l'accès aux ouvrages les plus précieux.



Figure 2.3 : Exemples de dégradations d'origine humaine

2.3 Dégradations dues à la capture

La numérisation, même de qualité, n'est qu'une représentation imparfaite, à un moment donné, d'un objet physique. Tout dépend des conditions de la prise d'image (placement, éclairage, réglages...) et de la qualité du numériseur (capteur linéaire ou matriciel, bruit du capteur, précision de l'optique et de la mécanique, algorithmes de prétraitement, codage interne des couleurs...). Les choix techniques effectués (choix des paramètres du matériel, des prétraitements et des formats de fichiers) par les acteurs de la numérisation ont plus d'incidence sur la qualité des images que les problèmes précédemment cités.

2.3.1 Dégradations matérielles

Nous distinguons principalement deux façon de numériser : la numérisation directe du document original et la numérisation indirecte faite à partir d'une copie du document original (microfilms, photographies, négatifs, reproductions imprimées ou photocopies). La numérisation de copies est plus simple à mettre en œuvre et à automatiser mais la qualité des images produites est très inférieure à celle obtenue par la numérisation directe du document original.

a. Les limites technologiques de la numérisation de documents

La numérisation est un procédé de codification numérique des rayonnements visibles réfléchés par un objet éclairé. Elle réalise deux opérations principales [BEN-05] :

- **L'échantillonnage spatial** : Les rayonnements lumineux sont discrétisés spatialement sur une grille dont la finesse dépend du nombre d'éléments photosensibles du capteur. Chaque élément de la grille constitue une information élémentaire appelée pixel qui

est décrit par ses coordonnées spatiales (x,y) et sa valeur dans l'espace des couleurs $I(x,y)$. Le processus d'échantillonnage peut ainsi être vu comme la multiplication de l'intensité par un réseau de distributions de Dirac.

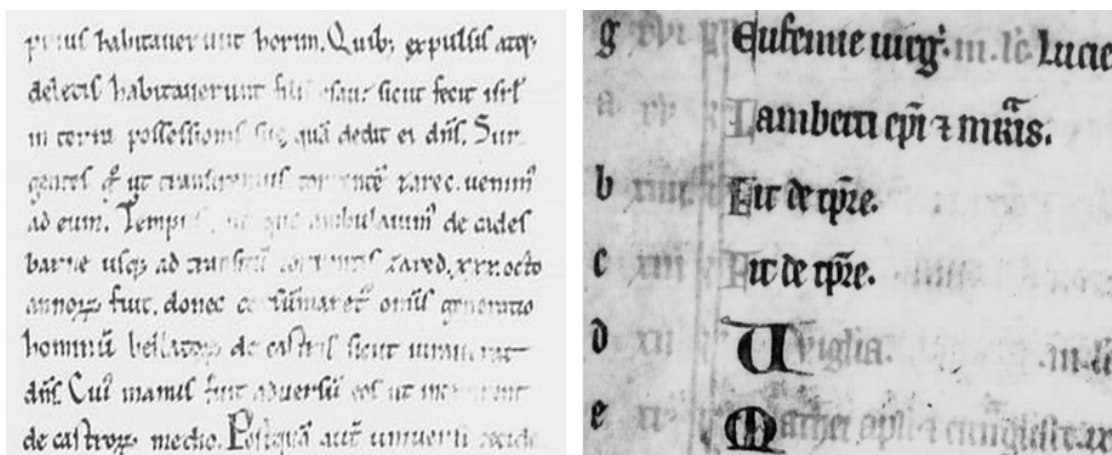
- **Une quantification des intensités lumineuses** : La valeur de chaque pixel $I(x,y)$ mesure le rayonnement quantifié sur cette surface élémentaire par le capteur. Les couleurs sont codées par des nombres entiers dont la précision dépend du type de capteur utilisé et de la qualité du codage des couleurs. Généralement, le codage est effectué sur n bits autorisant 2^n valeurs. Pour l'acquisition en couleur, chaque pixel $I(x,y)$ est alors codé par un vecteur à trois composantes (R, V, B) pouvant chacune prendre 2^n valeurs. Par conséquent, les couleurs ne sont pas toutes représentées lors de la phase de quantification.

Par définition même du passage du continu au discret, la numérisation discrétise et donc appauvrit l'information spatiale et colorimétrique. La réduction de l'information spatiale va générer des effets de pixellisation (lignes en escalier, image en mosaïque) visibles en agrandissant l'image autour des traits fins. La perte d'information colorimétrique, moins perceptible, gêne les traitements informatiques. En outre, l'éclairage est crucial pour la fidélité de la couleur. Ainsi, la présence d'éclairage, comme la lumière du jour ou un mauvais réglage de la position de l'éclairage, produit des effets indésirables sur l'image finale. Étant donné que le capteur enregistre une lumière réfléctie par l'objet, celui-ci doit être positionné perpendiculairement à l'axe de la caméra. Dans le cas contraire, la lumière peut être mal réfléctie, ce qui réduit la luminosité et affecte les couleurs. Suivant la qualité de l'optique, la distance focale doit être constante pour ne pas perdre la netteté sur toute la surface du document. Ce sont ces phénomènes qui expliquent les problèmes liés à la courbure des livres autour de la reliure. Les supports physiques n'ont ni la même texture ni la même qualité de réflexion. Les supports anciens en peaux, par exemple, peuvent être trop brillants par rapport à d'autres supports comme le papier ou le parchemin qui absorbent plus la lumière. Par conséquent, les réglages de la caméra et des éclairages sont très critiques et ne doivent absolument pas changer lors de la numérisation d'un même ouvrage. En plus, la qualité de la caméra, de l'optique, de la mécanique et de l'électronique est aussi très importante dans la restitution fidèle de l'image.

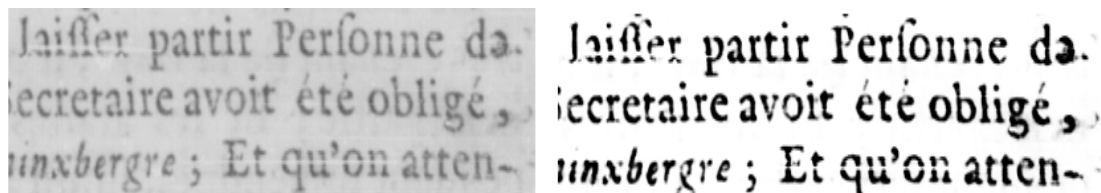
b. Les limites de la numérisation des microfilms

Les responsables de fonds patrimoniaux ont, depuis plusieurs décennies, microfilmé une grande quantité de documents dont les originaux sont, pour certains, disparus ou détruits. Pour des raisons économiques, les responsables de fonds patrimoniaux ont longtemps préféré la numérisation des microfilms à la numérisation directe des originaux ; aussi, il existe actuellement une grande quantité d'images de documents issues de la numérisation des microfilms. Il convient donc de s'intéresser aux dégradations causées par le processus de microfilmage, de conservation des microfilms et de leur numérisation.

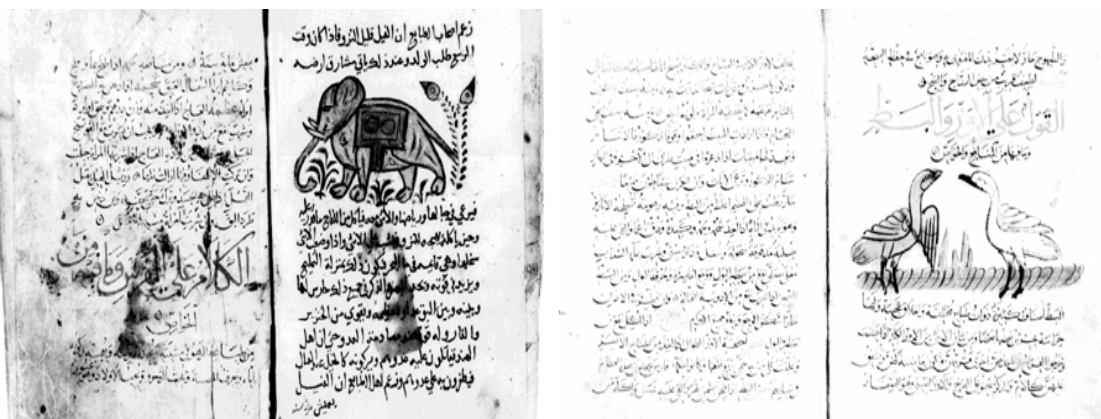
Le microfilmage est une opération photographique visant à stocker un grand nombre d'images sur une bobine de faible taille. Pour réduire l'image sur un microfilm, il faut tout d'abord perdre l'information couleur et augmenter le contraste de façon à conserver la lisibilité des lettres. Par conséquent, le microfilm est une copie qui a perdu définitivement toute



(a) Perte d'information couleur sur les images numérisées des microfilms (source : IRHT)



(b) Rayures sur les images numérisées des microfilms (source : Gazette de Leyde Lyon II)



(c) Variation du temps d'exposition sur les images issues de la numérisation de la même bobine de microfilm et d'un même ouvrage (source BNF)

Figure 2.4 : Limites de la numérisation des microfilms

l'information colorimétrique et une grande partie de l'information sur la luminosité. La numérisation de ces microfilms ne permet pas de restituer l'information perdue. Les images numériques ainsi générées utilisent seulement quelques niveaux de gris. Les traits les moins contrastés, à l'origine en couleurs claires, sont irrémédiablement effacés (Fig.2.4(a)).

Le microfilm est aussi un support fragile qui s'use après de nombreuses manipulations lors du visionnage. Les rayures laissées par les frottements et les manipulations mécaniques sont généralement visibles (Fig.2.4(b)).

Sous la forme de bobines, les microfilms sont numérisés à très grande vitesse avec des appareils dédiés. Pour augmenter la vitesse de numérisation d'une bobine de films, les fabricants réduisent les temps d'exposition et poussent les limites optiques et mécaniques des

matériels produisant ainsi des images de qualité très moyenne avec une résolution souvent insuffisante. De plus, on observe fréquemment des temps d'exposition très différents d'une image à l'autre sur la même bobine produisant périodiquement des images surexposées et donc trop claires suivies d'images trop sombres (Fig.2.4(c)). Compte tenu de ce qui vient d'être rappelé, la qualité des images issues de la numérisation de microfilms s'apparente à celle des images binaires.

2.3.2 Dégradations numériques

Les dégradations numériques sont les conséquences des choix technologiques délibérés ou accidentels comme le paramétrage du matériel, les prétraitements irréversibles et la compression avec perte des images. Les bibliothécaires ont très souvent privilégié la numérisation rapide de masse sur une numérisation de qualité. Pour réduire les temps d'acquisition et diminuer les coûts de la numérisation, ils ont fait des choix techniques qui auront des conséquences beaucoup plus graves sur la qualité des images. La réduction délibérée de la résolution spatiale ou colorimétrique et l'utilisation de formats de compression avec perte sont encore très largement utilisées pour sauver l'espace disque et réduire les coûts de stockage des images.

Les dégradations numériques sont aussi causées par une mauvaise utilisation du matériel. Citons, à titre d'exemples, l'application d'une méthode de compression d'images couleur sur une image binaire, le changement d'un format de fichier à un autre, l'application de traitements destructifs comme la correction du contraste ou le rehaussement de la netteté. Tous ces traitements réalisés modifient définitivement l'information. Il convient donc d'étudier les conséquences de ces pertes d'information sur les traitements et la lisibilité.

a. L'inadéquation de la résolution à la taille des textes

La résolution représente la finesse de la grille utilisée pour discrétiser les rayonnements visibles. Elle s'exprime en nombre de points par pouce (DPI : Dot Per Inch) ou en nombre de pixels par pouce (PPI : Pixels Per Inch). Plus la finesse de la grille diminue, moins les détails sont préservés (Fig.2.5). Une faible résolution s'accompagne généralement d'une perte de la topologie de l'objet transformant par exemple la lettre «e» en lettre «c» ou supprimant les espaces entre les lettres adjacentes (Fig.2.6). Une bonne résolution aide à mieux distinguer les détails fins, mais elle requiert la réservation d'un grand espace de stockage. Pour réduire les coûts tout en générant des images numériques exploitables, il faut trouver un compromis entre la taille des images et leur qualité. En outre, lors de la numérisation d'une collection, la résolution est fixée à l'avance indépendamment des contenus des textes. La figure 2.7 montre un exemple d'une image illisible à cause d'une résolution insuffisante sur des textes de très petite taille.

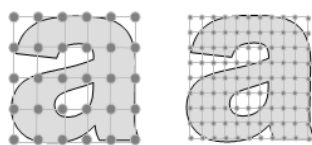


Figure 2.5 : Finesse de la grille pour une faible et une forte résolution.

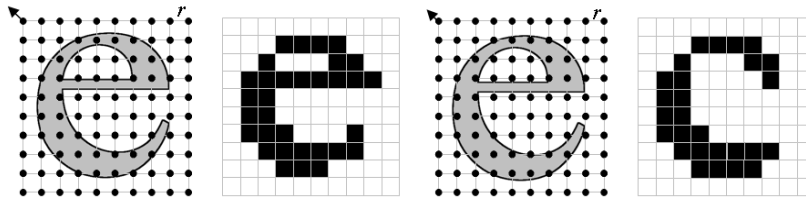


Figure 2.6 : Perte de la topologie avec le déplacement aléatoire d'une grille de résolution.

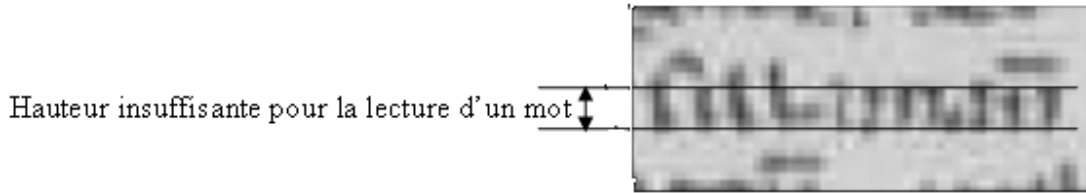


Figure 2.7 : Problème d'une faible résolution sur une image de trait

b. Choix de la compression avec perte

La manipulation des images de documents requiert un temps et un espace disque qui augmentent en fonction de leur taille ; d'où l'intérêt de les compresser afin de réduire leur taille. Le choix des formats de compression est très critique car il peut entraîner une altération irréversible des données. Toute information abandonnée ne sera pas récupérée à la décompression, si il s'agit d'une compression avec perte. L'exemple le plus célèbre est certainement l'encodage JPEG des images.



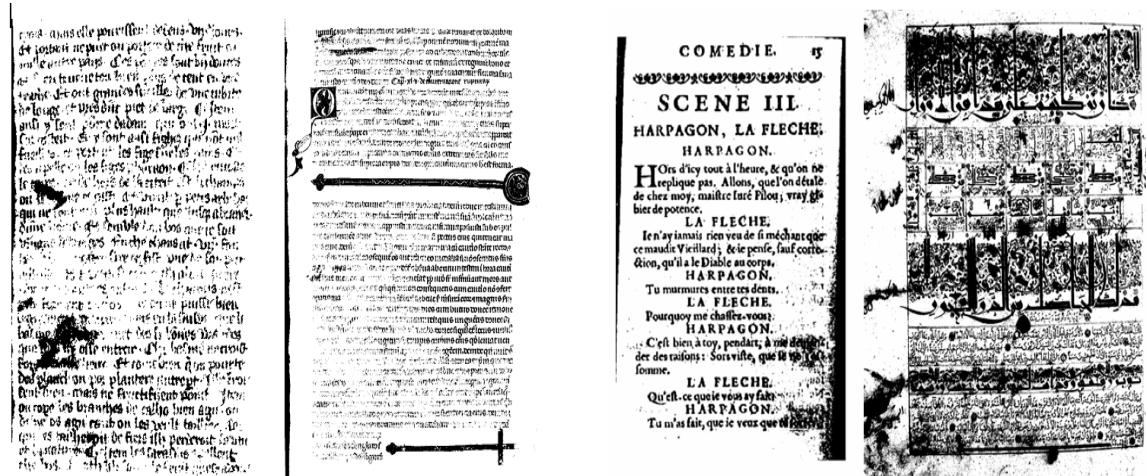
Figure 2.8 : Aspect destructif de la compression JPEG ; de gauche à droite respectivement : l'image originale et ses versions compressées avec un facteur de qualité 50 % et 10 %.

Cet encodage introduit une analyse fréquentielle sur des blocs de taille réduite (généralement de taille 8x8 pixels). Il a pour principe de privilégier un fort lissage sur les zones uniformes et un faible lissage sur les zones très contrastées. La quantification indépendante de chaque bloc 8x8 de l'image provoque des ruptures des traits. Nous percevons sur l'exemple (Fig.2.8) la présence d'artéfacts sur les frontières des carrés qui sont de plus en plus flagrants si on autorise à supprimer de nombreuses fréquences. On peut perdre jusqu'à 50% de l'information pour un facteur de qualité variant entre 100% et 90% [EMP-03]. Le format de compression JPEG dégrade ainsi les formes des caractères ; dégradation d'autant plus sensible que la taille de la police est petite ou que la résolution est très faible. Cela s'explique par l'augmentation de la complexité des frontières des lettres à l'intérieur des macro-blocs

que le filtrage fréquentiel ne peut plus restituer. Malgré le développement de formats de compression moins destructifs comme JPEG2000 ou comme les méthodes de compression adaptées aux documents comme le format DjVu [BOT-98] ou le format développé pour le projet Debora [BOU-07], la compression JPEG est le format de compression le plus utilisé pour la sauvegarde et la diffusion.

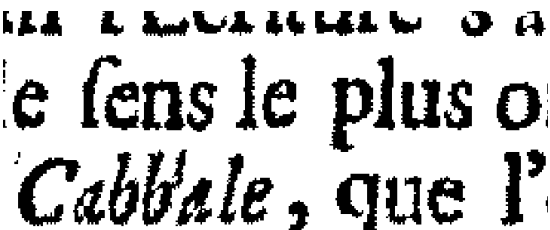
c. Le problème de la numérisation en binaire

La binarisation est un traitement irréversible qui permet de transformer une image en niveaux de gris ou en couleur en une image noir et blanc en fonction d'un seuil à définir. Toute valeur de pixel supérieure à un seuil fixe aura la valeur 255 (blanc), 0 (noir) sinon. Le choix du seuil de binarisation est donc crucial. Des algorithmes de binarisation à seuils adaptatifs, comme ceux proposés par Sauvola ou Niblack, conduisent à des résultats intéressants, mais la qualité des images binaires produites dépend du document original. Un document original dégradé aboutit généralement à une image binaire de mauvaise qualité.



(a) Binarisation de documents en couleurs non homogènes (IRHT)

(b) Binarisation de documents dont le support est dégradé (BNF)



(c) Agrandissement des défauts visibles du bruit sur le contour des caractères dans une image binaire (Trévoux)



(d) Effet d'un manque de résolution : déformation des caractères, collage de caractères proches et division des caractères en morceaux

Figure 2.9 : Limites de la numérisation en binaire

La binarisation échoue systématiquement sur des documents faiblement contrastés, ou lorsque les couleurs des encres et du papier ne sont pas uniformes (Fig.2.9(a)). Les dégradations du support, comme l'apparition du verso par transparence sur le recto et la présence de taches d'humidité, sont conservées dans l'image binaire (Fig.2.9(b)). En outre, le bruit

du capteur étant plus élevé sur les zones de fortes variations d'intensité, la binarisation fait systématiquement apparaître des irrégularités aléatoires sur les contours des caractères (Fig.2.9(c)). Elle modifie la topologie des formes en fonction de la résolution limitant ainsi les possibilités de segmentation des caractères et des mots (Fig.2.9(d)). Les images en couleur ou en niveaux de gris sont beaucoup plus lisibles et plus facilement interprétables par une machine que les images binaires, les niveaux de gris ou la couleur apportent une information supplémentaire qui compense le manque de résolution.

2.4 Proposition d'une classification non-causale des dégradations

Nous proposons une nouvelle classification orientée traitements d'image qui regroupe les dégradations décrites précédemment en fonction des méthodes de restauration possibles.

2.4.1 Dégradations uniformes

Ce sont les dégradations qui affectent tout le document de façon uniforme. Elles englobent les dégradations colorimétriques ou géométriques comme l'inclinaison, les variations uniformes des couleurs ou de la luminosité; elles regroupent les défauts issus d'une transformation globale de l'image qui peut agir uniformément soit sur la localisation des pixels (inclinaison, courbure) soit sur leurs valeurs (transformation de la couleur ou correction de la luminosité). La restauration revient, dans le cas présent, à définir la transformée inverse qui permettra de se rapprocher de l'information telle qu'elle l'était à l'origine. La restauration sera donc possible avec des méthodes globales qui appliquent les mêmes traitements sur toute l'image.

2.4.2 Dégradations non-uniformes

Par dégradations non-uniformes on entend toutes altérations qui varient suivant la position dans l'image. Il sera nécessaire d'utiliser des traitements adaptatifs. Nous avons défini deux sous-classes en fonction de la localisation de l'information par rapport au papier ou au texte : dégradations du fond et dégradations des formes.

a. Les dégradations du fond

Les dégradations du fond (Fig.2.10) se rapportent aux défauts du support papier comme les taches dues à l'humidité, au passage en transparence du verso, aux annotations et aux ajouts indésirables. Ces artefacts représentent une couche supplémentaire de bruit qui vient se superposer sur le texte original. Restaurer revient dans ce cas à séparer les différentes couches pour en extraire celles qui appartiennent au document original. Une dégradation du fond peut aussi contribuer à la dégradation du texte. Par exemple, le passage du verso sur le recto, les trous et les taches du support papier, rendent la lecture du texte très difficile.



Figure 2.10 : Dégradations du fond sur une image (source IRHT)

b. Les dégradations des formes

Nous appelons dégradations de formes celles qui modifient la continuité des traits et la topologie des objets. Différentes dégradations peuvent apparaître sur les caractères : rupture des traits, caractères tronqués, vide dans les traits, fusion de caractères, effet d'escalier... En outre, l'apparition des effets de transparence ou de trous risque de rendre certains caractères illisibles. Le mot de la figure 2.11 peut représenter les mots «parts», «paris» ou «parle». Ces dégradations sont très pénalisantes pour la lisibilité des textes et pour la reconnaissance automatique des formes.

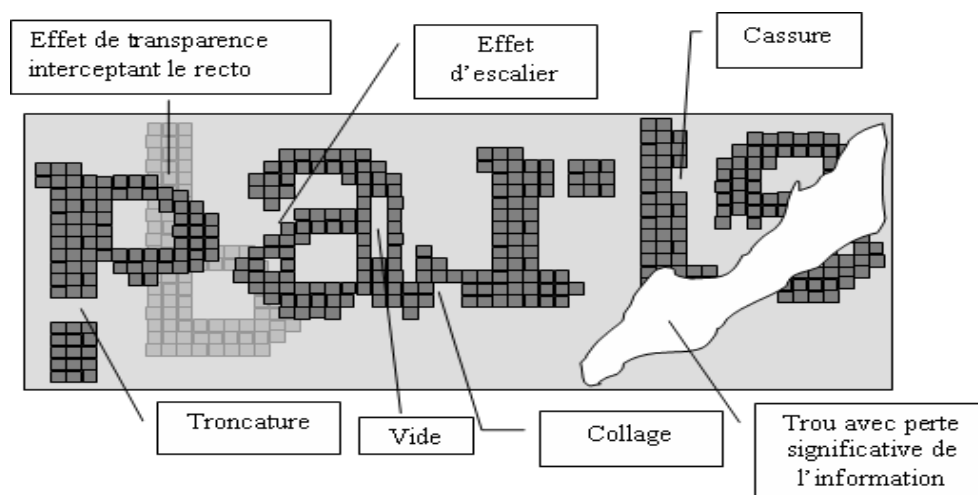


Figure 2.11 : Dégradations des formes sur une image de synthèse

2.5 Conclusion

Nous venons de voir que les dégradations peuvent provenir de causes multiples. Nous avons montré qu'elles affectent la qualité visuelle des images et peuvent engendrer une perte significative de l'information. Nous avons mis en évidence les limites de quelques traitements (binarisation, compression) et l'impact d'un mauvais paramétrage (résolution) sur les images de documents. Compte tenu de la variété des dégradations et de leurs causes, nous avons proposé une classification qui regroupe les dégradations qui peuvent être traitées avec les mêmes méthodes.

Chapitre 3

Restauration des Images de Documents

« L'art, c'est le moyen de corriger la vie. »

Robert Choquette

Résumé

Le chapitre précédent a démontré que les images de documents présentent de nombreux défauts physiques ou numériques que l'on doit corriger avec une restauration virtuelle ou numérique. Dans ce chapitre, nous dressons un bilan des méthodes de restauration existantes dans la littérature. Nous les présentons suivant l'utilisation ou non des connaissances a priori.

Sommaire

3.1	Préambule	28
3.2	Méthodes de restauration supervisées	28
3.3	Méthodes de restauration non-supervisées	36
3.4	Conclusions et discussions	38

3.1 Préambule

Nous avons précédemment défini la restauration numérique qui a pour objectif de corriger les défauts causés par le processus de numérisation et la restauration virtuelle cherchant à réparer les dégradations physiques des documents. Qu'elle soit numérique ou virtuelle, la restauration essaie toujours de retrouver l'image la plus fidèle possible. Un autre domaine très proche de la restauration est le domaine de l'amélioration des images. La frontière entre ces deux domaines demeure floue. En effet, l'amélioration et la restauration partagent des méthodes communes mais leur finalité reste différente. L'amélioration cherche à augmenter la qualité visuelle et non à retrouver l'image la plus fidèle possible. Une image fidèle au document original, nettoyée de toutes les dégradations physiques et numériques, n'est pas forcément une belle image qui offre un confort de lecture optimal. Qualité et beauté sont donc deux choses différentes.

Le nombre de travaux spécifiques à la restauration des images de documents est très réduit. Les articles scientifiques portent sur les corrections de dégradations bien particulières. À notre connaissance, il n'existe ni travaux abordant plusieurs dégradations ni méthodes génériques susceptibles de traiter des documents aux contenus différents.

Nous avons fait le choix de classer les travaux sur la restauration suivant l'utilisation ou non de connaissances *a priori*. Cette classification orientée traitement est un peu moins redondante que celle basée sur la classification des types de défauts et de leur localisation sur le fond ou la forme.

3.2 Méthodes de restauration supervisées

Les méthodes dites «non-aveugles» ou supervisées sont des méthodes de restauration basées sur des connaissances *a priori* du contenu des documents. Elles utilisent des informations externes, fournies par l'utilisateur, qui ne peuvent pas être extraites de l'image elle-même. Les méthodes supervisées sont très efficaces mais elles nécessitent toujours des données extérieures qui ne sont pas toujours disponibles. Par exemple, la séparation entre le recto et le verso par recalage requiert deux prises de vue séparées du recto et du verso de la même page. Ces méthodes de restauration ne sont pas suffisamment génériques pour traiter tous les types de document. En outre, la restauration des formes de caractères par l'observation de toutes les occurrences du même caractère dans la page ne fonctionne que sur des documents imprimés où les caractères peuvent être segmentés et comparés.

Parmi les travaux de recherche sur la restauration des images de documents, nous trouvons beaucoup d'articles qui abordent la séparation entre le recto et le verso.

3.2.1 Séparation recto/verso par recalage

Le recalage a été plus particulièrement appliqué pour la suppression des marques en transparence avec la présence des deux faces du document (recto/verso associé). Le principe consiste à mettre en correspondance les deux faces d'une même page pour en extraire le texte original associé à chaque face. L'approche de Sharma [SHA-00] applique un filtre linéaire adaptatif sur les deux faces du document simultanément. Le cas de passage en transparence de l'écrit n'a, semble-t-il, pas été traité et nous donnons dans la figure 3.1 le seul résultat

publié par l'auteur ; il faut remarquer que les deux parties à séparer sont de natures différentes (texte et image).

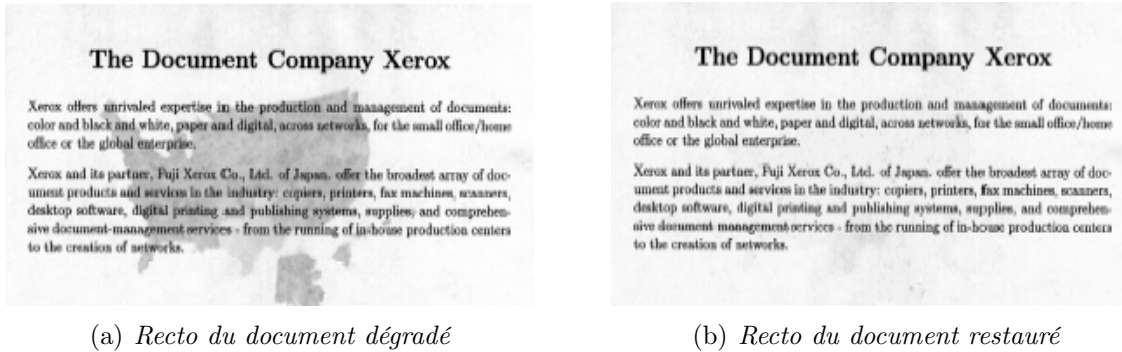


Figure 3.1 : Suppression des taches en transparence par recalage [SHA-00]

Une autre méthode, proposée par E. Dubois et A. Pathak [DUB-01], consiste à traiter en même temps les deux faces (recto/verso) du document et à comparer les intensités des pixels du recto avec son correspondant du verso. Si l'intensité du pixel du recto est moins élevée que celle du verso, sa valeur est alors conservée. La figure 3.2 montre un exemple d'application de cette méthode.

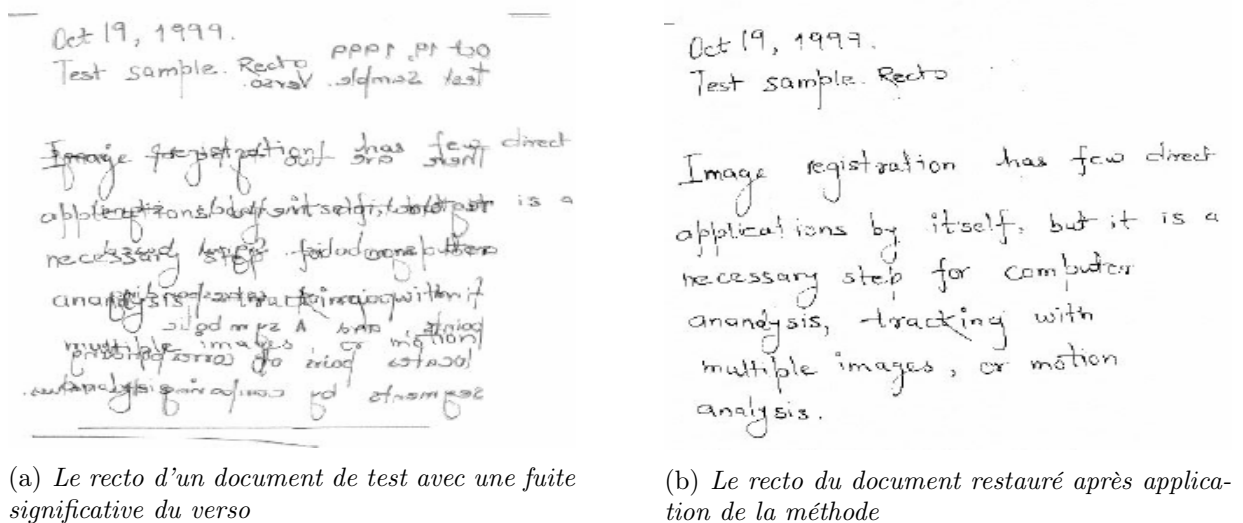


Figure 3.2 : Suppression des taches en transparence par recalage [DUB-01]

Certes les méthodes à base de recalage donnent de bons résultats mais leur dépendance vis-à-vis des deux faces du document limite énormément leur application ; il est nécessaire d'effectuer une mise en correspondance parfaite des deux faces ce qui n'est pas évident à mettre en œuvre quand l'image du recto est géométriquement déformée par rapport à celle du verso. Nous concevons donc l'intérêt des méthodes aveugles de séparation recto/verso qui n'utilisent qu'une seule face à la fois.

3.2.2 Séparation recto/verso par analyse de l'orientation

Certains documents sont caractérisés par des styles d'écriture très orientés. L'analyse de l'orientation s'avère donc efficace pour distinguer le texte du recto de celui du verso. Q. Wang et al. [WAN-03] proposent une analyse de l'orientation à base de la transformée en ondelettes. Le principe consiste à définir une transformée en ondelettes capable de séparer les traits suivant les directions 45° et 135° . Ainsi, le recto va être obtenu dans l'image filtrée dans la direction 45° et le verso dans l'image filtrée dans la direction orthogonale à la première. Néanmoins, certains détails du recto peuvent se projeter sur l'image du verso cassant ainsi les caractères. Les auteurs proposent d'extraire du document original les parties manquantes sur un certain voisinage de taille 7×7 (Fig.3.3). Nous classons donc cette approche dans les méthodes de restauration supervisées puisqu'elle nécessite la connaissance *a priori* des orientations des textes du recto et du verso.

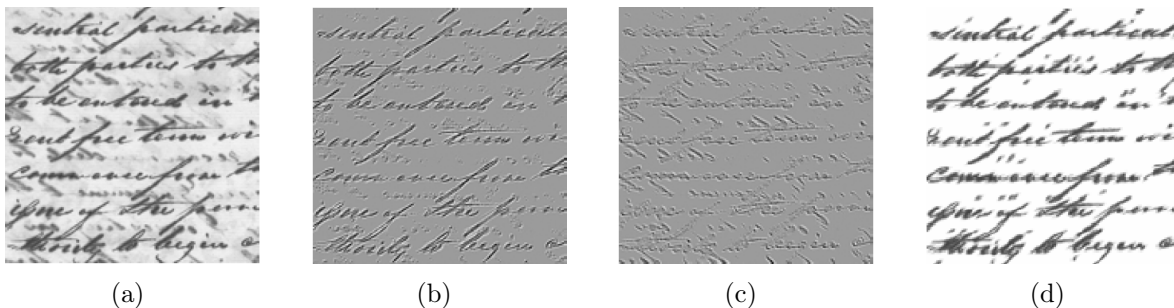


Figure 3.3 : Suppression des taches en transparence par analyse de l'orientation [WAN-03] : (a) image de test, (b) image associée à l'orientation 135° , (c) image associée à l'orientation 45° , (d) image restaurée de (a)

3.2.3 Séparation recto/verso par analyse des formes

[WOL-06] propose une méthode aveugle de séparation entre le recto et le verso basée sur une régularisation de la segmentation par les champs de Markov. La segmentation utilise une approche Bayésienne probabiliste pour classer les pixels du verso qui appartiennent aussi au recto. Cette séparation entre le recto et le verso n'utilise pas d'information sur la couleur ou sur les niveaux de gris. Les champs de Markov capturent la distribution spatiale des pixels et prédisent la configuration la plus probable à partir d'une configuration initiale donnée par la segmentation. Ils tendent à régulariser la segmentation pour préserver la continuité des traits du recto qui croisent ceux du verso (Fig.3.4).

3.2.4 Séparation recto/verso par analyse de la couleur

E. Smigiel [SMI-04] utilise une carte de Kohonen pour résoudre le problème de suppression des taches en transparence. Il détermine une classification des pixels en quatre classes (fond, texte, texte coloré et texte du verso). Cette approche nécessite une fastidieuse phase d'apprentissage sur une portion représentative d'une page. La classification ne marche pas si la luminosité varie sur une même page ou d'une page à l'autre. Les performances de la séparation par classification dépendent donc de l'apprentissage.

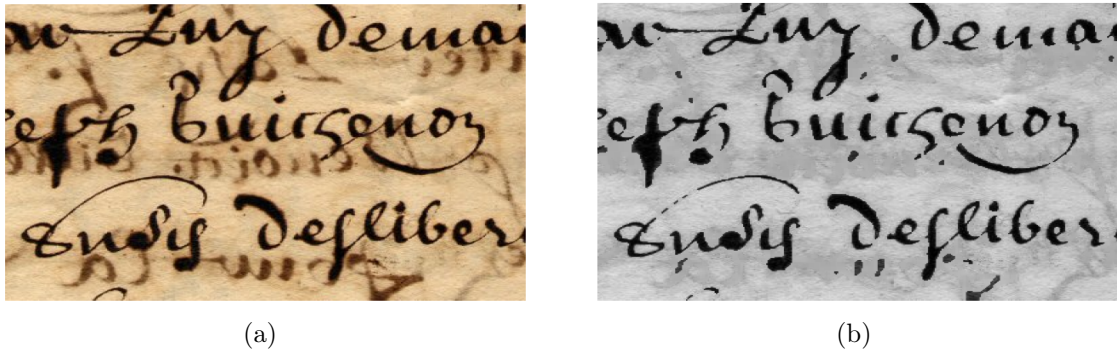


Figure 3.4 : (a) Image d'origine, (b) Image finale de recto séparée du verso [WOL-06]

U. Garain [GAR-06] propose de séparer le recto du verso par une approche classique de segmentation en régions de couleur qui sont ensuite agrégées suivant des règles prédéfinies. Le résultat montre un effet de mosaïque autour des régions agrégées et présente de nombreuses fragmentations des traits. De plus, cette approche échoue sur notre image de référence (Fig.3.5).

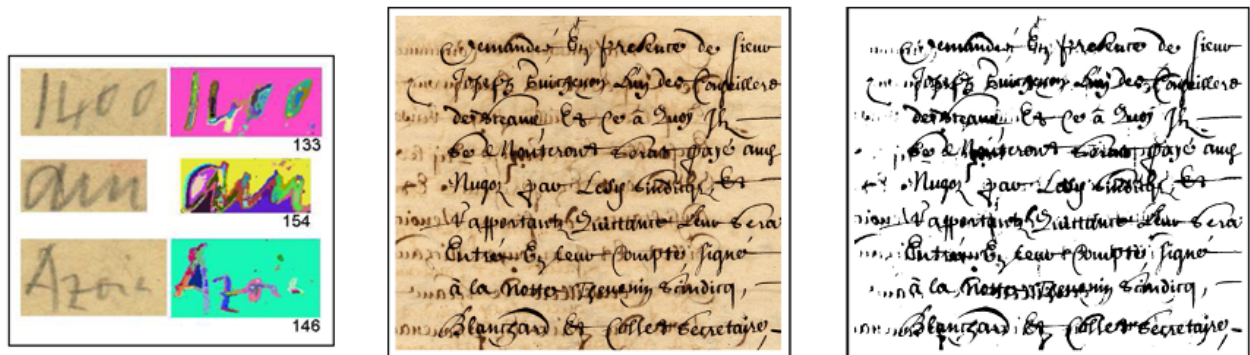


Figure 3.5 : Séparation recto/verso par segmentation et agrégation de régions couleurs : Résultats de la sur-segmentation en régions, image dégradée et restaurée par [GAR-06]

La méthode de segmentation d'images couleurs de Y. Leydier [LEY-04] est basée sur la sérialisation de l'algorithme des nuées dynamiques ; elle donne de bons résultats pour la séparation recto/verso (Fig.3.6). Il s'agit d'une classification automatique des couleurs sur une fenêtre glissante parcourant toute l'image. Durant le déplacement de la fenêtre, l'algorithme adapte automatiquement les centres des classes au contexte local. Cette approche permet donc de combiner les avantages d'une segmentation par classification et les méthodes de segmentation adaptatives classiques que l'on trouve dans le domaine du document. Son principal inconvénient réside dans la détermination manuelle du nombre de classes et des couleurs pour chaque classe. C'est donc une approche efficace mais entièrement supervisée qui nécessite l'intervention de l'opérateur pour chaque document.

3.2.5 Restauration des formes par modèle

Certaines contributions pour la restauration des formes se sont orientées vers la définition de modèles reflétant plus ou moins fidèlement ces dégradations. Cependant, au-

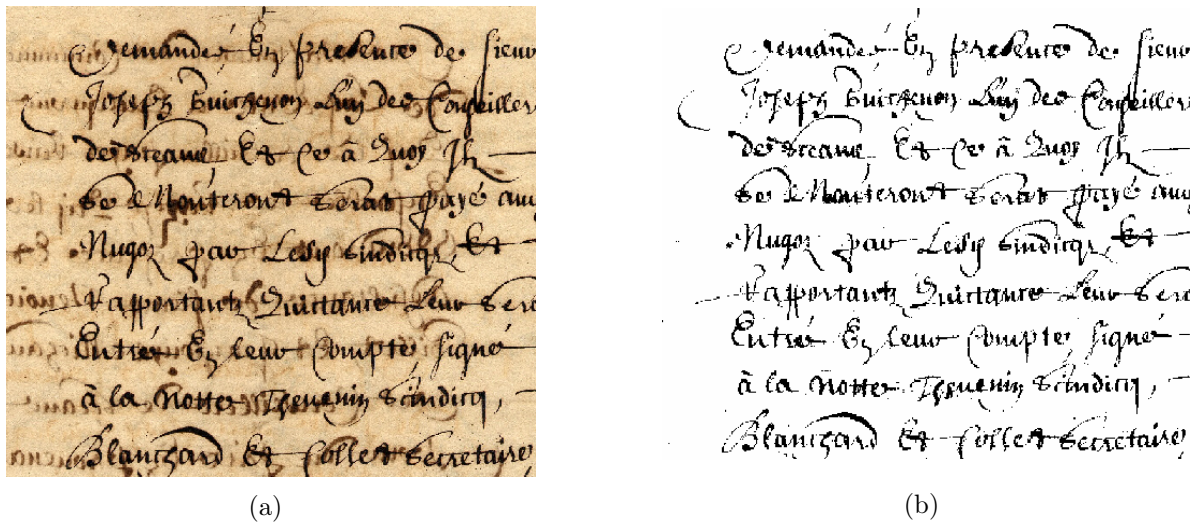


Figure 3.6 : Application de l'approche de [LEY-04] : (a) Image originale (b) Résultat pour $K=2$, le nombre de classes

cun modèle suffisamment générique pour décrire toutes les distorsions possibles n'a jamais été publié [BAI-00]. On peut citer, à titre d'exemple, le modèle de défaut proposé dans [BAI-93, BAI-92], modèle qui inclut un certain nombre de paramètres, à savoir, la taille du texte, la résolution, la translation, la rotation, le changement d'échelle, le défaut de jet d'encre. Les différents paramètres pris en compte sont essentiellement reliés aux conditions de numérisation de la page. Un autre modèle de dégradation, proposé par Kanungo [KAN-94, KAN-00], suppose que les probabilités d'altération des formes ne dépendent que de la distance à leur frontière et suivent une loi normale. Ils simulent certains défauts, comme le flou, dus à l'impression et la numérisation, par des opérations de fermeture morphologique. Zhang et Kanungo [ZHA-01] proposent d'utiliser un modèle de dégradation défini à partir de la numérisation d'une image idéale de caractères. Ce modèle est basé ainsi sur les paramètres extraits de l'image dégradée produite. Baird élabore dans [BAI-92] un modèle de dégradation utilisé pour entraîner un logiciel de reconnaissance optique de caractères et donc améliorer ses performances.

La restauration par modèle conduit en général à des résultats intéressants. Mais, elle exige des hypothèses de déformation homogène sur l'ensemble de la collection. Ces hypothèses figées ne restent applicables que pour des conditions très particulières d'acquisition. Une simple modification des conditions prises en compte (changement de la résolution ou de la taille du texte, mauvais positionnement de la page sur le scanner) met en échec tout le modèle proposé.

[ALL-03] propose la restauration des caractères cassés par contour actif (snake). Cette restauration consiste à faire évoluer un contour actif, guidé par le modèle d'une lettre entière, sur la lettre cassée (Fig.3.7). Cette réparation ne peut avoir lieu que si l'on dispose du modèle de lettre à réparer et si la lettre possède une topologie simple. Ces limites rendent cette approche difficilement exploitable en conditions réelles. De plus, cette méthode nécessite une segmentation en caractères et une recherche de la forme entière à partir de la forme cassée.

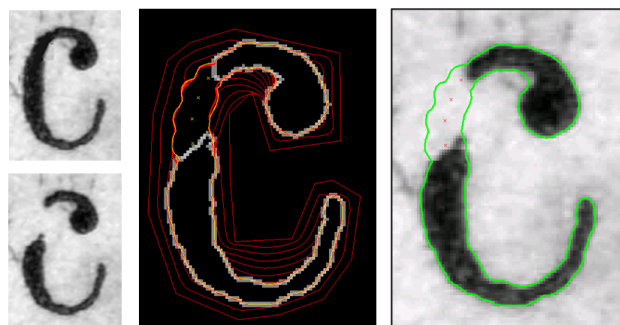


Figure 3.7 : Principe de la restauration d'une lettre cassée par les contours actifs [ALL-03]

3.2.6 Restauration des formes par prototypage

Le prototypage est l'opération qui consiste à collectionner toutes les formes similaires d'une page. On l'applique généralement aux caractères d'un document imprimé. À partir de la superposition de toutes les formes dégradées d'une même lettre, on construit le modèle de la forme idéale de la lettre. La figure 3.8 montre différentes occurrences dégradées du caractère «e». Malgré les différences générées par les dégradations partielles, pour les trois premières occurrences, la structure est bien préservée. Toutefois, la structure n'est plus préservée avec les autres occurrences.

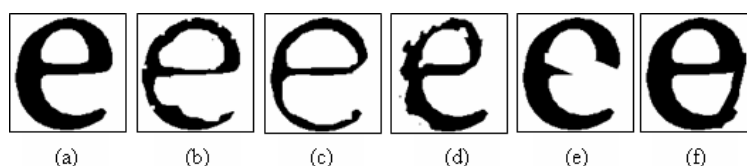


Figure 3.8 : Cas de dégradations de la lettre «e» : (a) caractère idéal ; (b),(c) et (d) caractères dégradés avec structure du «e» conservée, (e) et (f) caractères dégradés avec structure confuse : (e) «e» ou «c», (f) «e» ou «o»

Hobby et Baird [HOB-97] proposent une technique de restauration basée sur l'interpolation du caractère idéal en utilisant le principe du prototypage. Le principe consiste à superposer plusieurs occurrences du caractère dégradé et à moyenner l'image résultante pour obtenir une approximation de l'image idéale du caractère. Chaque type de caractère dans le document sera ensuite substitué par son prototype (Fig.3.9).

So, we find special
term rewriting for

So, we find special
term rewriting for

Figure 3.9 : Restauration par prototypage des formes dégradées de lettres

Cette méthode reste confrontée aux problèmes de la bonne définition d'une similarité de formes pour le prototypage et de l'application de la segmentation. Elle ne reste efficace que pour des documents imprimés qui contiennent de faibles dégradations n'affectant pas

la structure générale des formes des caractères ; elles peuvent alors améliorer les taux de reconnaissance des logiciels d'OCR. Mais, cette méthode ne peut pas s'appliquer aux documents manuscrits. En outre, elle ne permet pas de retrouver les parties irrégulières comme les formes anguleuses (Fig.3.10).



Figure 3.10 : Limite d'une restauration par prototypage

3.2.7 Restauration géométrique de la courbure

La numérisation de livres volumineux reliés fait souvent apparaître des défauts géométriques, notamment l'apparition de la courbure à proximité de la reliure du livre : un traitement est alors nécessaire pour rendre le document lisible (Fig.3.11). Plusieurs solutions ont été proposées pour réaliser la mise à plat de l'image du document après les corrections géométriques. Certaines méthodes utilisent une analyse en deux dimensions de l'orientation des traits ou des lignes de texte, d'autres sont basées sur une analyse en trois dimensions permettant de construire le relief à partir de l'analyse de la réflexion de la lumière.

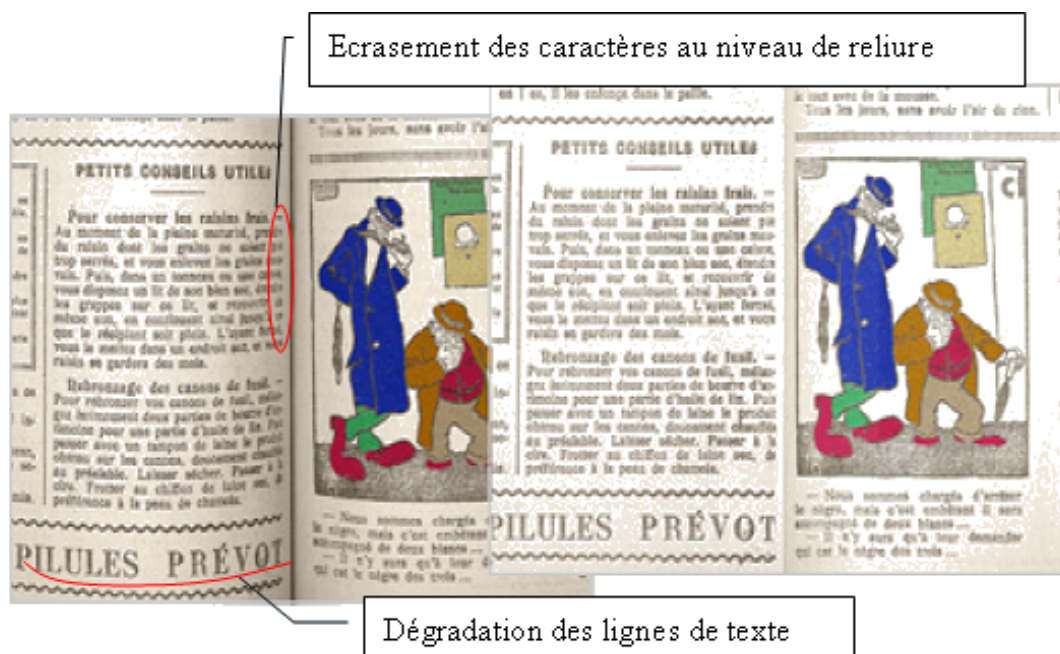


Figure 3.11 : Défauts géométriques de la photographie d'un livre ouvert à gauche et sa version restaurée à droite (source I2S)

a. Restauration par un modèle 2D

Les méthodes de restauration de la courbure basées sur un modèle 2D procèdent en deux étapes : calcul de la courbure et traitement géométrique pour la corriger. Le calcul de la courbure peut être fait à partir des lignes de textes [WU-02, CAO-03, ZHA-03] ou par extrapolation des bords de la page [TAK-01, HIR-01]. Q. Zheng [ZHE-01] propose une correction basée sur une analyse des lignes et des mots de l'image. E. Trinh [TRI-03] combine l'analyse de la courbure des lignes de textes et des bords des pages et obtient un résultat final de meilleure qualité. Sa méthode a contribué à l'élaboration du logiciel Book Restorer de la société I2S.

La modélisation 2D utilise les propriétés d'un document textuel (parallélisme et espacement vertical constant des lignes de texte au sein d'un même paragraphe) afin d'estimer la mise à plat du document. Elle sera limitée pour la segmentation des lignes de textes des documents dégradés [BOU-06] ou des manuscrits pour lesquels les alignements ne sont pas naturellement horizontaux. En outre, les dégradations du fond comme la présence de taches ou le passage du verso sur le recto complexifient les opérations de segmentation des lignes de texte.

b. Restauration par un modèle 3D

Les méthodes de restauration, basées sur un modèle 3D, peuvent être réalisées avec une prise de vue soit normale (un appareil photographique [WAD-95, WAD-97, COU-04, KAN-93, CAO-03]), soit particulière (deux appareils photographiques ou deux photographies du document prises [YAM-04] sous des éclairages différents [CHO-97]).

Pour une prise de vue normale, les déformations des feuilles sont généralement modélisées par une surface cylindrique aux abords de la reliure [CAO-03, KAN-93]. Cette modélisation implique que tous les points de la surface ayant la même altitude forment des droites parallèles. La surface est définie par son rayon de courbure. Les lignes de textes sont ensuite corrigées le long de l'axe des abscisses par déplacement horizontal des pixels. Une alternative à cette modélisation est la reconstruction de relief des documents scannés à partir des niveaux de gris des pixels de l'image. Il s'agit de l'application de la méthode «Shape from shading» qui est une technique bien connue de reconstruction tridimensionnelle à partir des variations de lumière sur une surface de texture homogène. Une fois le relief reconstruit, il suffit d'opérer par la mise à plat sur l'axe des abscisses en fonction de l'altitude du point [WAD-95, WAD-97, COU-04].

Pour une prise de vue particulière avec l'utilisation de plusieurs vues sous des angles ou des éclairages différents d'un même document, une mise en correspondance est calculée entre les images pour reconstruire le relief de la page. Ce calcul du relief s'effectue soit par triangulation [YAM-04] soit par «Shape from shading» [CHO-97]. La mise à plat d'un document dégradé par une modélisation 3D offre l'avantage d'utiliser les informations directement liées à la déformation 3D. Néanmoins, sa dépendance vis-à-vis de nombreuses hypothèses (texture, forme *a priori* de la courbure...) et la complexité du système de capture d'image limitent fortement son usage.

Les méthodes de restauration géométrique de la courbure doivent être suivies d'une restauration des formes. Mais quelle que soit la méthode utilisée, l'image issue de cette correction de courbure sera une image interprétée car les caractères situés le long d'une forte courbure

ne pourront pas retrouver leurs formes d'origine.

3.3 Méthodes de restauration non-supervisées

3.3.1 Restauration du recto/verso par séparation de sources

Les taches, qui viennent se superposer sur le texte original du recto, ont généralement une couleur différente qui les distingue du texte original. Nous pouvons ainsi modéliser ces taches par des couches supplémentaires. La séparation de sources permet de séparer ces différentes couches d'information.

A. Tonazzini et al. [TON-04] proposent, dans ce contexte, une approche utilisant l'Analyse en Composantes Indépendantes (ACI) [HYV-99A] et plus particulièrement l'algorithme de FastICA [HYV-99B] pour faire une séparation aveugle du recto et du verso. Cette approche ne nécessite ni l'information du verso, ni un modèle, ni un apprentissage. C'est donc une approche très intéressante du point de vue de l'utilisateur. Cette méthode résout le problème de transparence pour quelques images de documents dégradés (s'agissant le plus souvent d'images de synthèse). Cependant, les résultats associés à des images réelles de document restent bien modestes (Fig.3.12).

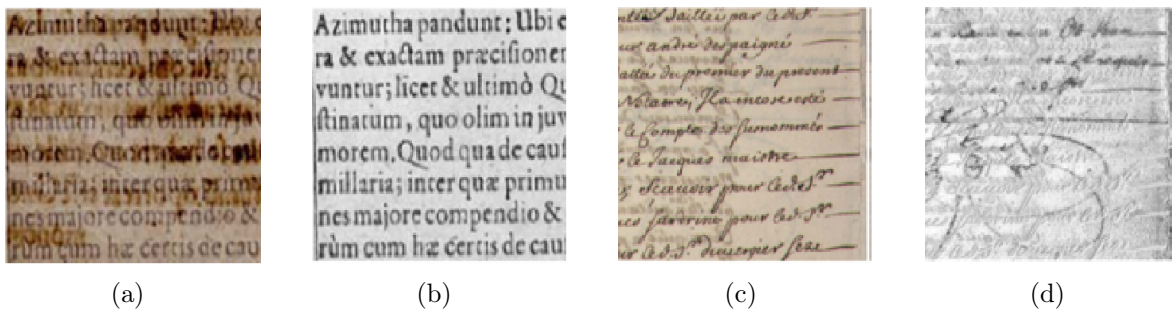


Figure 3.12 : Suppression des taches en transparence par analyse de la couleur [TON-04] : (a) image de synthèse, (b) image restaurée de (a), (c) image réelle, (d) image restaurée de (c)

3.3.2 Restauration des formes par morphologie mathématique

La restauration des formes par morphologie mathématique est possible pour les petites dégradations qui concernent quelques pixels seulement. Les opérations de dilatations et d'érosions conditionnelles permettent de réparer les ruptures de traits et de combler les trous [BRE-03] (Fig.3.13).

Wichello et Yan [WHI-96] combinent la méthode de restauration par morphologie avec un suivi de contours. L'objectif consiste essentiellement à rétablir une certaine continuité des contours dégradés. Le suivi de contours, fait à l'aide de masques de tailles variables, autorise la connexion des parties du caractère dans une direction privilégiée ; en revanche, il risque de relier les éléments qui ne devraient pas l'être s'ils ont une distance inférieure à la taille du masque. Ce problème a été résolu par Yu et Yan [YU-01] qui ajoutent l'information sur la distance entre deux composantes connexes. Par ailleurs, cette proposition, testée pour

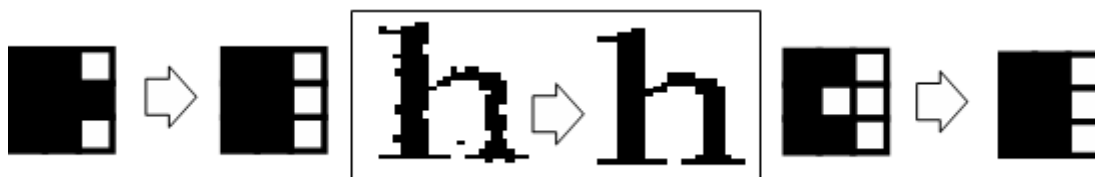


Figure 3.13 : Effets des opérateurs morphologiques sur les images binaires

des chiffres manuscrits cassés, fait appel à des critères de décision figés qui limitent son caractère générique. La définition de critères de décision a été également proposée par Shi et Govindaraju [SHI-96] pour la restauration des chiffres imprimés et manuscrits. Le principe consiste à voir pour un pixel donné la possibilité de le faire appartenir à un contour ou à un segment noir proche. Cette méthode présente l'avantage de ne pas modifier l'épaisseur des traits. Toutefois, les améliorations apportées dépendent de la taille de la fenêtre utilisée.

La restauration par morphologie mathématique permet donc de remplir les vides et/ou de corriger les ruptures des traits ; cependant, elle n'est applicable que sur un petit voisinage du caractère et ne prend pas en compte des informations structurelles ou globales de la forme. En outre, la morphologie mathématique, par sa définition même, ne présente aucune précision sous-pixellique. La taille de la fenêtre est toujours un nombre entier, ce qui n'est pas toujours le cas lors de la manipulation de traits plus ou moins fins. Enfin, les méthodes de restauration proposées restent très modestes et manquent même d'une évaluation objective des performances.

3.3.3 Restauration Colorimétrique

Toute chaîne de numérisation est exposée à deux types de dégradations que ce soit d'une source externe ou interne [TSI-03]. Le support à numériser peut influencer sur la réflexion de la lumière. Les contraintes physiques impliquées par la reliure engendrent des dégradations photométriques. Nous observons alors une diminution des intensités lumineuses autour des courbures. E. Trinh [TRI-03] propose une méthode de correction des modifications de la couleur et de la luminosité près de la reliure à partir du calcul du profil de l'éclairement. Ce profil utilise l'histogramme de chrominance et de luminosité de toute l'image et des colonnes.

Dans le cas d'un changement d'illumination interne ou externe à la source émettrice, il s'avère primordial d'appliquer un prétraitement global identique sur tous les pixels de l'image. Changement de la valeur de luminance, correction gamma et recalage d'histogramme peuvent être des solutions pour résoudre ce problème et améliorer les couleurs de l'image. Certaines dégradations photométriques ont tendance à accentuer un ton par rapport à un autre. Pour remédier à ce problème, il suffit de supprimer la tonalité dominante et de mettre en place un équilibrage pour les couleurs [GAS-03]. Toutefois, ce traitement doit conserver la balance des couleurs de l'ensemble de l'image pour ne pas en perdre les détails. Une solution simple pour résoudre ce problème consiste à accentuer les contrastes en général. Le contraste est une propriété intrinsèque à une image qui permet de quantifier la capacité de distinguer deux régions distinctes (Fig.3.14).

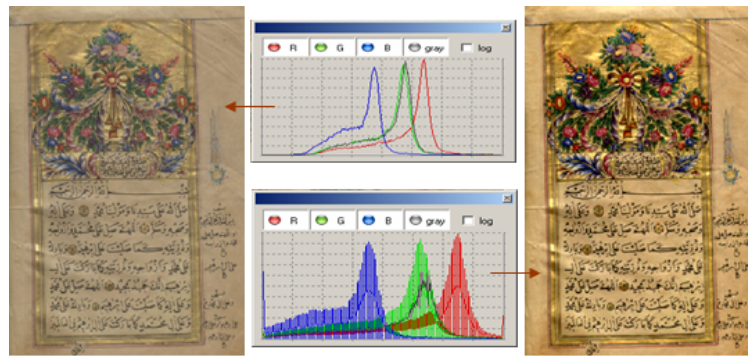


Figure 3.14 : Restauration de la dégradation colorimétrique par ajustement du contraste

3.4 Conclusions et discussions

Dans ce chapitre, nous nous sommes intéressés à différentes approches de restauration proposées dans la littérature. Nous avons dégagé différentes techniques selon le problème étudié : dégradation non-uniforme de fond ou de forme, dégradation uniforme colorimétrique ou géométrique. Nous avons pu retenir que les dégradations non-uniformes posent encore de vrais problèmes qui ne sont pas encore résolus. D'ailleurs, les travaux traitant les dégradations géométriques prennent comme hypothèse de départ l'uniformité de la dégradation. En outre, toute restauration géométrique doit être suivie par une restauration des formes. Toutes ces contraintes nous incitent à approfondir l'analyse des dégradations non-uniformes. Nous ne traiterons pas donc les dégradations uniformes. Des algorithmes de restauration des couleurs et de correction de la courbure apparente de l'image ont été déjà développés par [TRI-03] au LIRIS et intégrés dans le logiciel commercial BookRestorer de traitement d'image.

Nous avons aussi constaté que les méthodes de restauration non-supervisées étaient beaucoup plus rares que les méthodes supervisées reposant sur un modèle ou sur des connaissances *a priori*. Toutefois, les dégradations sont généralement des phénomènes aléatoires difficilement prévisibles. C'est la raison pour laquelle nous croyons difficilement à la généralisation des approches de restauration basées sur l'apprentissage ou sur une modélisation du processus de dégradation [BAI-92]. De plus, le cumul des dégradations rend très difficile la séparation de dégradations mélangées ou superposées.

Notre objectif est de développer des méthodes de restauration capables d'améliorer une large gamme de documents souffrant d'une dégradation non-uniforme. Or, les méthodes non-supervisées ont l'avantage de ne pas dépendre du contenu de l'image et ne nécessitent pas d'autres informations supplémentaires que celles déjà contenues dans l'image. Nous proposons ainsi de traiter ces dégradations par des méthodes non-supervisées de restauration.

Pour les dégradations du fond, le problème majeur est celui du passage en transparence de l'encre, problème dont souffre la plupart de nos images ; nous ne disposons pas nécessairement des deux faces du document (recto/verso). Nous avons donc choisi de nous orienter vers des méthodes de restauration aveugle, cas d'une restauration sans verso. L'analyse de la couleur et plus particulièrement l'application des techniques de segmentation/classification nous paraît la piste la plus adaptée à notre problème.

Dans le cadre des dégradations des formes, la restauration par morphologie mathématique est à notre avis une piste intéressante qui reste limitée car elle exige une taille de

masque entière et ne tient pas compte des structures de l'image. Nous nous sommes rendus compte que la morphologie mathématique discrète peut être avantageusement remplacée par la morphologie continue à base d'équation aux dérivées partielles. Ce domaine nous a attirés notamment par sa nature locale, sa base mathématique solide et rigoureuse, sa capacité de modéliser les images dans le domaine continu et encore la variété de schémas qui peuvent être implémentés. Ce domaine a été évalué pour la restauration des images naturelles, mais aucune évaluation n'existe, à notre connaissance, pour la restauration des images de documents. L'adaptation de ces techniques à notre contexte de restauration des formes dégradées des lettres nous semble la piste la plus prometteuse.

Pour mesurer la performance de la restauration, il est possible de mesurer certaines dégradations de façon objective si on possède l'image non-dégradée. Mais le plus souvent, nous ne disposons pas de l'image «optimale», la plus fidèle au document original. Par conséquent, nous ne pouvons évaluer que visuellement et de façon subjective la correction de certaines dégradations. Pour évaluer la restitution des formes, il est possible de mesurer objectivement l'augmentation des performances des traitements automatiques comme l'OCR sur les documents imprimés. Dans le cas de documents manuscrits sur lesquels l'OCR ne peut pas fonctionner, nous pouvons mesurer qualitativement la restauration à partir de l'amélioration de la lisibilité par exemple. C'est pour cela que nous aurons parfois recours à des images de synthèses reproduisant artificiellement des défauts de façon à comparer l'image initiale et l'image restaurée. Mais la synthèse artificielle des dégradations telle qu'elle a été faite par IBM pour l'apprentissage des OCRs [BAI-93] repose sur un modèle imparfait éloigné de la réalité. L'évaluation des performances de la restauration des images de documents reste un problème ouvert.

Deuxième partie

Dégradations du Fond : Restauration par Segmentation

Chapitre 4

Segmentation : Réflexion Méthodologique

« La Philosophie c'est la réflexion aboutissant à reconnaître sa propre insuffisance et la nécessité d'une action absolue partant du dedans. »

Jules Lagneau

Résumé

Nous définissons les dégradations du fond par l'ensemble des défauts du support papier, comme le passage en transparence du verso sur le recto ou la présence de taches qui rendent le document difficile à lire. Nous proposerons de séparer les couches d'information par une classification des couleurs. Dans ce chapitre, nous présentons une réflexion méthodologique sur les approches de classification locale ou globale des couleurs car nous retrouverons cette problématique tout au long de notre étude.

Sommaire

4.1	Introduction	44
4.2	Approches globales	44
4.3	Approches locales	46
4.4	Approches duales	47
4.5	Conclusions et discussions	48

4.1 Introduction

Les dégradations du fond forment des couches supplémentaires qui viennent se superposer sur l'image originale du recto. Les techniques de segmentation/classification des couleurs nous paraissent la piste la plus adaptée à ce problème. Restaurer le document dégradé revient alors à utiliser les approches de segmentation par classification des couleurs.

La segmentation est en effet un des problèmes fondamentaux de l'analyse d'images et reste une étape presque incontournable en vision par ordinateur. Elle s'inspire du système de perception visuel humain pour générer une interprétation d'une image [COC-95]. La segmentation cherche à extraire, de façon aussi exacte que possible, des régions homogènes et disjointes les unes des autres dont chacune caractérise un des objets présents dans cette image. Dans le cas d'une sur-segmentation, un objet peut être fragmenté en plusieurs régions. Inversement, une sous-segmentation réunit plusieurs objets dans une même région de l'image. Dans le cas idéal, tous les objets sont représentés par une seule et unique région de l'image séparée des autres objets.

La littérature recense un très grand nombre de techniques utilisées pour la segmentation couleur. Les sections suivantes ne cherchent pas à faire une étude exhaustive de ces différentes techniques. Le lecteur intéressé par une telle étude pourra se référer par exemple à [COC-95, TRE-04]. Mais, nous nous limitons dans cette étude aux méthodes de segmentation qui sont uniquement basées sur la classification des couleurs. Les méthodes basées sur les caractéristiques de texture ou de formes ne nous paraissent pas adaptées à notre problématique. On classe, en général, les méthodes de segmentation selon une approche région ou une approche contour. Nous proposons d'aborder les méthodes de segmentation par classification des couleurs suivant leur caractère local, global ou dual. Les approches globales utilisent l'information couleur de l'image toute entière alors que les approches locales ne prennent en considération que les informations couleurs dans un voisinage de chaque point de l'image. Les approches locales ont l'avantage de s'adapter au contexte et aux variations des couleurs mais elles provoquent plus facilement une sur-segmentation à cause du bruit présent dans l'image. Inversement, les approches globales, plus résistantes au bruit, conduisent le plus souvent à une sous-segmentation de l'image. Enfin, les approches duales, qui utilisent simultanément l'information locale et globale permettent de tirer avantage des deux approches complémentaires. C'est autour de cette dualité que nous conduisons notre réflexion méthodologique.

4.2 Approches globales

Les approches globales, dont le principe est illustré par la figure 4.2, ne sont basées que sur les attributs couleurs de tous les pixels de l'image. Effectuer une segmentation avec une approche globale revient à extraire les régions qui correspondent à une partition des couleurs dans l'espace RVB. Ces régions extraites ont les mêmes caractéristiques colorimétriques. L'appartenance à une région homogène est définie par une distance entre deux couleurs. Généralement, on utilise des espaces couleurs comme le $L^*a^*b^*$ afin de mesurer la distance entre deux couleurs avec la distance Euclidienne [TRE-04]. Nous distinguons les approches supervisées des approches non-supervisées.

L'algorithme des K -moyennes [DUD-98] est un algorithme très populaire, trop souvent considéré comme une approche non-supervisée bien qu'elle nécessite la définition du nombre



Figure 4.1 : Images dégradées de document par passage en transparence de l'encre

de classes et la donnée de leur centre initial. Les pixels sont affectés à la classe dont le centre de gravité est le plus proche au sens de la distance colorimétrique considérée. Après affectation de l'ensemble des couleurs de l'image, les centres de gravité sont remis à jour

par calcul de la moyenne des couleurs associées à la classe. L'algorithme prend fin lorsque les centres de gravité de toutes les classes sont suffisamment stabilisés. Cette classification globale des couleurs se base sur l'hypothèse que les classes de couleurs sont homogènes.

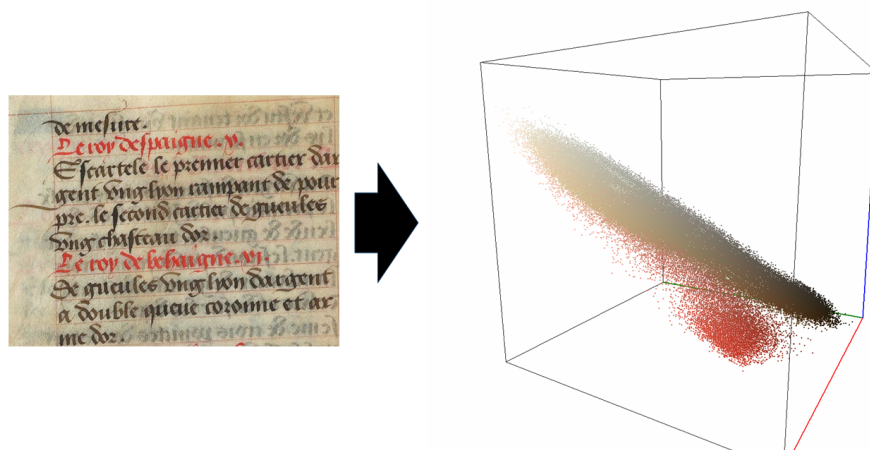


Figure 4.2 : Principe de classification globale des couleurs

Parmi les approches non-supervisées, nous pouvons citer la méthode de classification basée sur les densités de la distribution des couleurs de l'image (Mean-Shift). Cette méthode ne nécessite ni la fixation du nombre de classes ni la définition des couleurs représentatives de chaque classe.

4.3 Approches locales

Les approches globales sont incapables de s'adapter aux variations locales de l'image. Pour contourner cette difficulté et améliorer la qualité de la segmentation, il convient d'introduire l'information contextuelle autour de chaque point de l'image [TOR-03, KUM-03]. Rappelons que la perception humaine utilise de manière courante l'information contextuelle lorsque les conditions de vision sont mauvaises [SER-04]. Dans des situations où la vision est de qualité médiocre (bruitée, lointaine, non statique, entrelacée...), le contexte semble avoir un rôle important à jouer afin d'augmenter la fiabilité de la reconnaissance. Les approches locales tiennent compte des relations spatiales entre les couleurs dans le voisinage de chaque point (Fig.4.3).

Elles n'introduisent que des propriétés calculées dans un sous-espace de l'image et ne tiennent pas compte des informations colorimétriques de toute l'image. Les approches locales de classification des couleurs apparaissent principalement dans le domaine de la segmentation des images de documents ; ce sont des méthodes adaptées à la binarisation des images de traits. Le premier algorithme de classification locale a été élaboré par Niblack qui a proposé de calculer un seuil sur la luminosité pour chaque pixel en fonction de son voisinage sans utiliser aucune information globale liée à l'image entière. Les algorithmes adaptatifs de binarisation comme le Niblack et ses variantes comme le Sauvola remportent un grand succès dans la segmentation des images de documents et sont aujourd'hui intégrés dans les scanners professionnels. Ces méthodes concernent uniquement les images en niveaux de gris et il

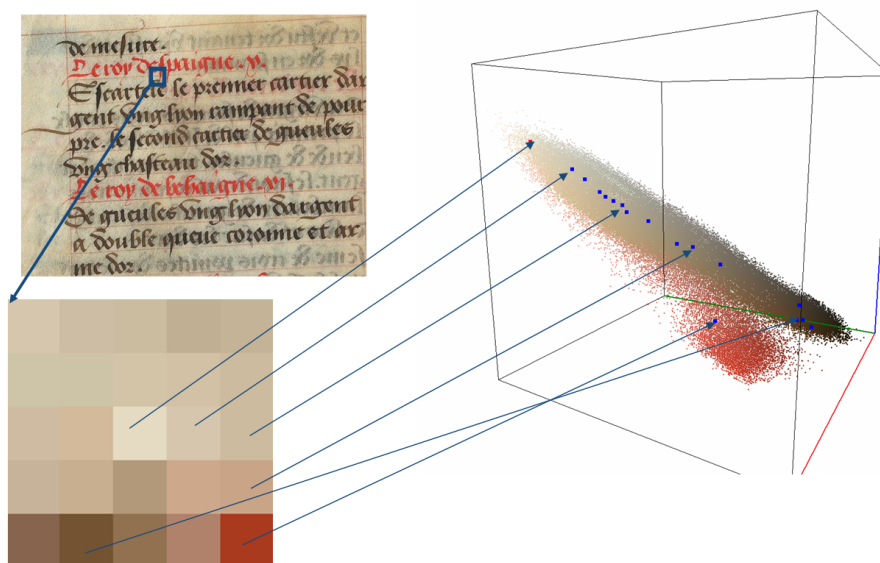


Figure 4.3 : Principe de classification locale des couleurs

n'existe pas d'adaptations de ces méthodes aux images couleurs. C'était pour répondre à ce manque que le LIRIS avait développé la méthode du K -means sérialisé [LEY-04]. Cette approche reste une approche supervisée et requiert en particulier l'intervention de l'utilisateur pour chaque image différente ou pour un lot d'images ayant les mêmes caractéristiques.

Pour les images naturelles, la classification des pixels de l'image dans un espace joint combinant la position et les caractéristiques colorimétriques locales a été proposée par C. Tomasi et al. [TOM-98] qui ont introduit le filtrage bilatéral. L'information spatiale est ajoutée comme une pondération dans le calcul de la distance entre les pixels. De même [COM-02] introduit l'information spatiale dans le calcul de la distance colorimétrique pour la classification des couleurs. Ces deux approches produisent une sur-segmentation en micro-régions très similaire à celle obtenue avec la morphologie mathématique, en particulier avec la méthode des bassins d'attraction. Nous détaillerons ces méthodes dans le chapitre 6.

Ces approches locales conduisent souvent à une sur-segmentation de l'image en petites régions disjointes car elles sont sensibles au bruit. Il est nécessaire de procéder à une difficile étape de fusion des régions en fonction d'un critère d'homogénéité couleur. Si une région satisfait le critère d'homogénéité avec ses régions adjacentes, alors elles sont agglomérées jusqu'à aboutir de proche en proche à une image correctement segmentée. Cette délicate phase de fusion de régions utilise généralement des critères complexes et des heuristiques très difficiles à mettre en œuvre. Cette approche a déjà été utilisée par [GAR-06] pour séparer le recto du verso avec des résultats peu convaincants.

4.4 Approches duales

Nous avons pu remarquer que les approches locales et globales s'appuient sur des informations différentes et complémentaires. Il existe des approches hybrides avec une combinaison séquentielle en série des méthodes locales/globales d'analyse. En général, une approche globale est suivie d'une approche locale pour affiner les résultats de la première analyse.

La relaxation probabiliste [HSI-89, COH-95, BON-01] est un exemple de méthode qui utilise le résultat d'une analyse globale et qui cherche ensuite à affiner le résultat par une approche locale. L'estimation des probabilités initiales basée sur la distribution des couleurs de l'image s'apparente à une segmentation globale. Ces probabilités initiales sont ensuite affinées progressivement en analysant les distributions locales dans un voisinage de chaque point de l'image.

La combinaison d'un filtrage pour réduire le bruit, suivie d'une segmentation constitue un autre exemple de combinaison local/global. Le filtrage repose sur une analyse locale qui est suivie d'une méthode de segmentation à partir d'informations globales. Dans les chapitres 10 et 11 nous étudierons l'apport d'une combinaison séquentielle en série d'une diffusion pour réduire le bruit et d'une segmentation couleur basée sur une analyse globale. Nous démontrerons que l'utilisation simultanée des informations locales et globales dans une combinaison non séquentielle donne de meilleurs résultats.

4.5 Conclusions et discussions

La segmentation par classification des couleurs devrait *a priori* permettre de séparer le recto du verso ou des taches superposées pour une grande majorité des documents. Choisir la méthode de classification la plus adaptée à notre problématique, revient à répondre à plusieurs questions : autorise-t-on l'intervention de l'utilisateur ? doit-on opter pour une approche globale, locale ou duale ?

Nous choisissons d'orienter nos travaux sur des approches essentiellement non-supervisées qui ne nécessitent pas la détermination du nombre et des attributs des classes à extraire. Les limites des approches supervisées ont été clairement soulevées par [LEY-04] dont la méthode nécessite la surveillance constante de l'utilisateur et qui empêche le traitement d'images hétérogènes.

Nous avons décidé de laisser de côté les approches purement locales et de restreindre notre choix entre les approches globales et duales. En effet, les approches purement locales conduisent à une sur-segmentation de l'image qui nécessite un post-traitement de fusion de régions. Ce post-traitement exige la définition des critères de proximité colorimétrique et spatiale qui sont loin d'être évidents à mettre en œuvre notamment pour le cas de traitement de documents. Les approches duales nous paraissent très intéressantes car elles introduisent l'information contextuelle tout en conservant l'information globale sur l'image entière. Tout au long de notre étude, nous retrouvons la dualité entre les traitements locaux et globaux de l'information. Le chapitre 5 présentera une segmentation globale des couleurs par une classification récursive des informations décorréées par une Analyse en Composantes Principales. Le chapitre 6 introduit une méthode qui combine des informations locales et globales pour réduire la complexité de calcul d'une classification basée sur les densités de la distribution des couleurs. Dans la partie IV, nous présenterons une méthode unificatrice entre les approches de diffusion décrits dans la partie III et la classification des couleurs basée sur les densités décrites dans le chapitre 6. Cette combinaison permet à la fois de réduire le bruit, restaurer la continuité des traits et de séparer les couches d'informations par classification des couleurs pour supprimer les dégradations du support papier.

Chapitre 5

Première Méthode : Restauration par Segmentation Globale

« Mieux vaut allumer des petites lumières que de se plaindre de l'obscurité. »

Lao Tseu

Résumé

Dans ce chapitre, nous exposons notre première contribution pour la restauration des dégradations du fond. En particulier, nous abordons le problème de la séparation entre le recto et le verso. L'objectif est de nettoyer le fond des différentes taches et des traces du verso qui peuvent nuire à la lisibilité d'une part et à la performance des traitements numériques d'autre part. Pour résoudre ce problème, nous proposons une méthode de classification récursive des couleurs à partir de l'algorithme des K-moyennes sur des données décorréélées.

Sommaire

5.1	Introduction	50
5.2	Représentation des données	50
5.3	Limite d'une segmentation linéaire	53
5.4	Proposition d'une classification en cascade	55
5.5	Étude expérimentale de l'approche proposée	57
5.6	Conclusions et perspectives	58

5.1 Introduction

Dans une première étape, nous voulons tester une méthode de classification globale des couleurs avec des méthodes simples pour vérifier si la séparation entre le recto et le verso est un problème trivial ou non. Nous allons utiliser une méthode de classification statistique comme la méthode des K -moyennes pour classer les couleurs de l'image entière. Mais la méthode des K -moyennes nécessite la fixation *a priori* du nombre de classes de couleurs et des centres initiaux des classes. Comme le nombre de classes de couleurs varie d'un document à l'autre, nous proposons donc de modifier la méthode précédente de façon à éviter la détermination du nombre de classes par l'utilisateur. Notre modification de l'algorithme de classification des couleurs porte sur deux points essentiels :

- Nous proposons de réaliser la classification des couleurs dans un espace où l'information est décorrélée par une analyse en composantes principales (ACP). Cette analyse permet d'accentuer la séparation entre les classes de couleurs.
- Pour éviter de déterminer le nombre K de classes de couleurs lors de l'application de l'algorithme des K -moyennes, nous proposons de classer récursivement les couleurs de l'image avec l'algorithme des 2-moyennes. L'application récursive d'une classification en deux classes permet de réaliser une classification non-linéaire des couleurs.

5.2 Représentation des données

La représentation ou plus exactement le choix d'un espace de représentation des données en couleur est la première étape de toute méthode de segmentation. Il s'agit de transformer l'image de l'espace RVB (Rouge-Vert-Bleu) à un espace adapté à la segmentation. En pratique, l'espace RVB, composé des trois couleurs fondamentales, est l'espace le plus couramment utilisé pour coder les images. Il associe à chaque couleur trois composantes, qui correspondent aux intensités respectives de trois couleurs primaires de la synthèse additive. La figure 5.1 illustre l'espace de couleur RVB où toutes les couleurs sont représentées par un point appartenant au cube RVB. Cet espace de couleur est un espace linéaire qui a pour inconvénient d'avoir des composantes chromatiques corrélées. Or la segmentation qui implique souvent l'utilisation d'un critère d'homogénéité est inefficace si l'information couleur est corrélée. D'où l'idée de procéder par un changement d'espace de couleur plus adéquat et aidant à mieux séparer les éléments d'une image.

L'Analyse en Composantes Indépendantes (ICA), issue des travaux sur la séparation de sources, est un outil d'analyse de données permettant de trouver une nouvelle base de représentation dans laquelle les observations apparaissent indépendantes. Cette approche paraissait la mieux adaptée à notre problème, mais elle a déjà fait l'objet de nombreuses études par [TON-04] dans le cadre de la séparation entre le recto et le verso dans les images de documents (voir chapitre 3). Il nous a semblé plus judicieux de commencer par une décorrélation des données par l'Analyse en Composantes Principales (ACP) plus simple à mettre en œuvre, contrairement à l'analyse en composantes indépendantes qui ne possède pas d'une méthode unique de calcul. En proposant une classification récursive en cascade des

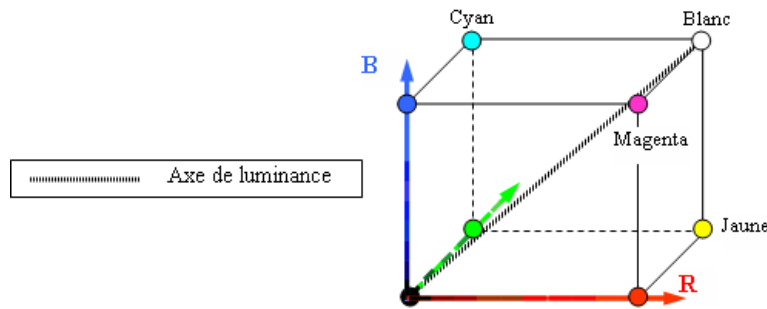


Figure 5.1 : Espace des couleurs RVB

données, nous pensons réaliser une séparation non linéaire des informations colorimétriques.

5.2.1 Principe de l'analyse en composantes principales

L'objectif principal de l'Analyse en Composantes Principales (ACP)¹ en traitement de signal est de transformer un tableau de données comprenant un certain nombre d'observations, pour pouvoir en donner une meilleure interprétation. Pour traiter une image couleur décomposée en RVB avec l'ACP, il suffit de considérer l'image (N pixels) comme un tableau de données ayant trois caractéristiques (R, V et B) et N observations. En diagonalisant la matrice 3×3 de covariance calculée à partir du tableau des observations sur les trois canaux couleurs, on obtient une matrice passage de trois vecteurs propres représentant les nouveaux axes qui permettent de projeter les points de l'image dans le nouvel espace et trois valeurs propres ($\lambda_1 > \lambda_2 > \lambda_3$) qui indiquent la quantité de variance expliquée de chaque axe. La transformation va réaliser un changement de base dans laquelle l'information contenue sur chaque axe est distribuée de façon optimale suivant la variance expliquée des informations. L'ACP permet donc de représenter les données d'origine dans un espace de dimension inférieure ou égale à l'espace de départ où les classes sont mieux séparées du point de vue de la variance. L'ACP possède donc deux propriétés intéressantes : la première est la décorrélation des informations, la seconde propriété est qu'elle assure une réduction de la dimension des données tout en conservant le maximum d'information possible.

5.2.2 Décorrélation des couleurs par l'analyse en composantes principales

Dans cette section, nous allons étudier l'apport d'une décorrélation de l'information couleur par l'analyse en composantes principales. Sur la figure 5.2 nous remarquons que le vecteur propre (u_1), associé au premier axe factoriel de la projection à base d'ACP, correspond à la valeur propre qui porte 99.8% de la variance totale dans l'image. La projection suivant le vecteur (u_3) met en évidence une combinaison bruitée du texte du recto et du verso en transparence (Fig.5.2).

¹L'ACP fut introduite par Karl Pearson en 1901 et intégrée à la statistique mathématique par Harold Hotelling en 1933. L'ACP est aussi connue sous le nom de transformée Karhunen-Loève à partir des années 50

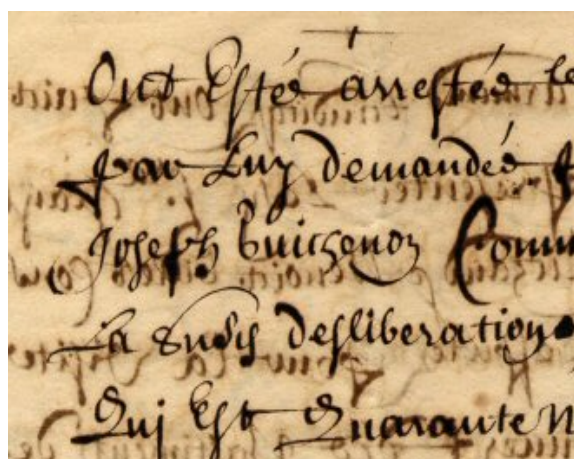
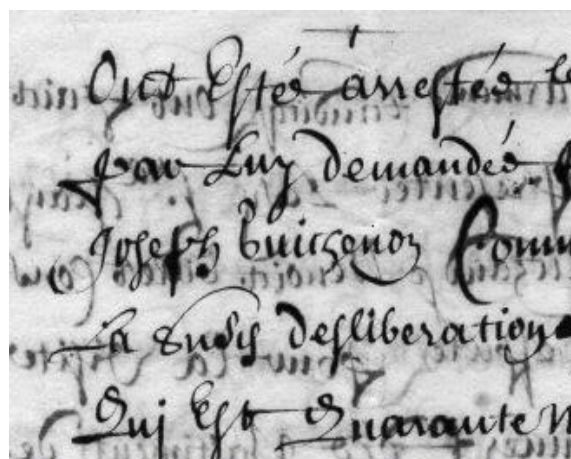
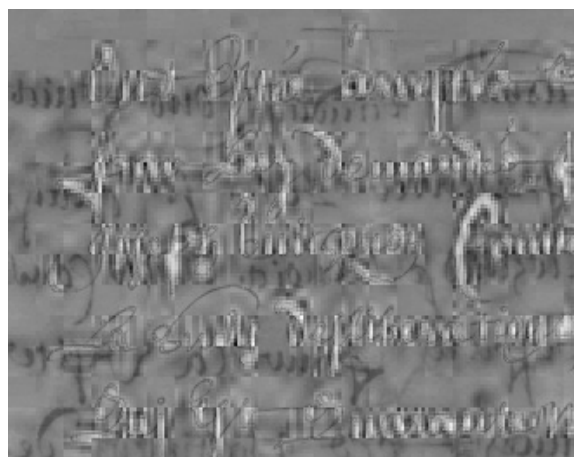


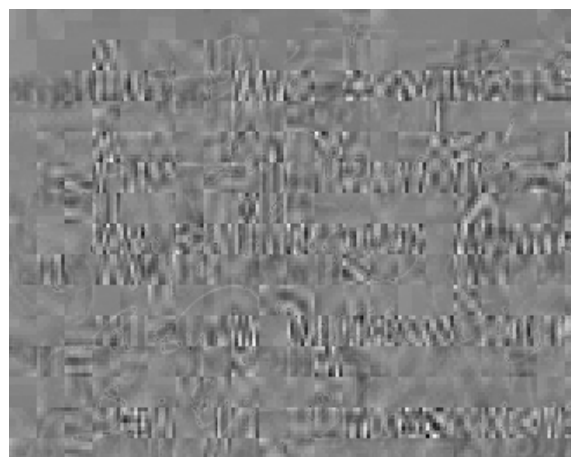
Image originale B



B suivant $u_1 = (-0.6, -0.58, -0.43)$
 associé à la valeur propre λ_1 représente
 99.8% de la variance



B suivant $u_2 = (0.6, -0.18, -0.75)$ associé
 à la valeur propre λ_2 représente 0.19% de
 la variance



B suivant $u_3 = (0.35, -0.79, 0.49)$ associé
 à la valeur propre λ_3 représente 0.01%
 de la variance

Figure 5.2 : La projection de l'image B sur les trois axes factoriels générés par l'ACP

Le premier axe factoriel correspond approximativement à une combinaison linéaire équivalente à celle de la luminance. L'axe de luminance est, par définition, l'axe des niveaux de gris où les composantes rouges, vertes et bleues sont égales. Il passe par les sommets Noir et Blanc. En d'autres termes, l'axe de luminance est la bissectrice associée au tétraèdre formé par les axes ; rouge, vert et bleu et qui passe par les points d'intersection de ces trois axes.

Existe-t-il ainsi une différence entre l'axe de luminance et le premier axe factoriel de l'analyse en composantes principales ? Des tests effectués sur des exemples d'images de document ont apporté les éléments de réponse (Fig.5.3). En effet, la distribution des triplets RVB des pixels d'une image de document se présente sous la forme d'un tube. Ce tube démarre à l'origine (noir) et suit une courbe approximativement parallèle à l'axe de luminance. Cette

distribution des couleurs dans le cube RVB est une caractéristique représentative des images de document que n'ont pas les images de scènes naturelles. Ceci explique l'importance de l'axe de luminance qui apporte une information discriminante dans les images de documents. Ce fait ne signifie pas pour autant que la numérisation des documents en niveaux de gris soit suffisante pour toutes les applications en analyse d'images. Une distribution des couleurs en tube dans le cube RVB n'est forcément ni linéaire ni parallèle à l'axe de luminance. La figure 5.3 montre que, dans l'espace RVB d'origine, les classes associées au «recto», «verso en transparence» et «papier» forment des amas qui sont imbriqués. Ce qui nous emmène à constater que :

- L'axe de luminance n'est pas le meilleur axe sur lequel les pixels projetés montrent une meilleure séparation. Seul le premier axe factoriel de l'ACP est susceptible d'ajuster les amas de façon à maximiser la séparation des pixels projetés sur cet axe.
- Les classes semblent suivre une distribution linéaire, qui peut résulter de la juxtaposition de plusieurs distributions gaussiennes.

L'analyse en composantes principales peut représenter une étape préliminaire à la classification. Elle permet d'appréhender la qualité de la classification à partir de la réduction des nombres de canaux et par élimination de certaines associations entre les bandes. Nous allons utiliser, dans le processus de segmentation, les deux premiers axes factoriels, les plus porteurs de l'information, le troisième axe pouvant être négligé puisqu'il correspond à la mise en évidence du bruit de l'image et des artefacts de la compression JPEG. Nous allons étudier, dans les sections suivantes, la faisabilité de la classification des couleurs dans cet espace.

5.3 Limite d'une segmentation linéaire

5.3.1 Principe de l'algorithme des K -moyennes

Le principe de la méthode des K -moyennes consiste au début, à choisir un représentant ou un centre pour chaque classe. Ces centres initiaux sont généralement choisis au hasard. Une procédure itérative permet de les mettre à jour jusqu'à stabilisation. Une telle procédure a pour objectif de minimiser la variance inter-classe définie par :

$$I = \sum_{j=1}^K \sum_{i=1}^{n_j} D_E(x_{ij}, c_j) \quad (5.1)$$

où K est le nombre de classes, c_j le représentant de la classe j , n_j le nombre des points x_{ij} de l'image appartenant à la classe j et D_E la distance Euclidienne.

La minimisation de ce critère est en fait un problème d'optimisation par descente de gradient. Essayons de mieux expliciter le principe pour le cas particulier d'une dimension ou d'un nombre d'attributs pour chaque point de l'image égal à 1. Étant donnée une partition initiale formée de K classes, le problème se formule en tant que calcul des centres c_j des différentes classes qui minimisent le critère donné par l'équation 5.1. Ces centres sont définis

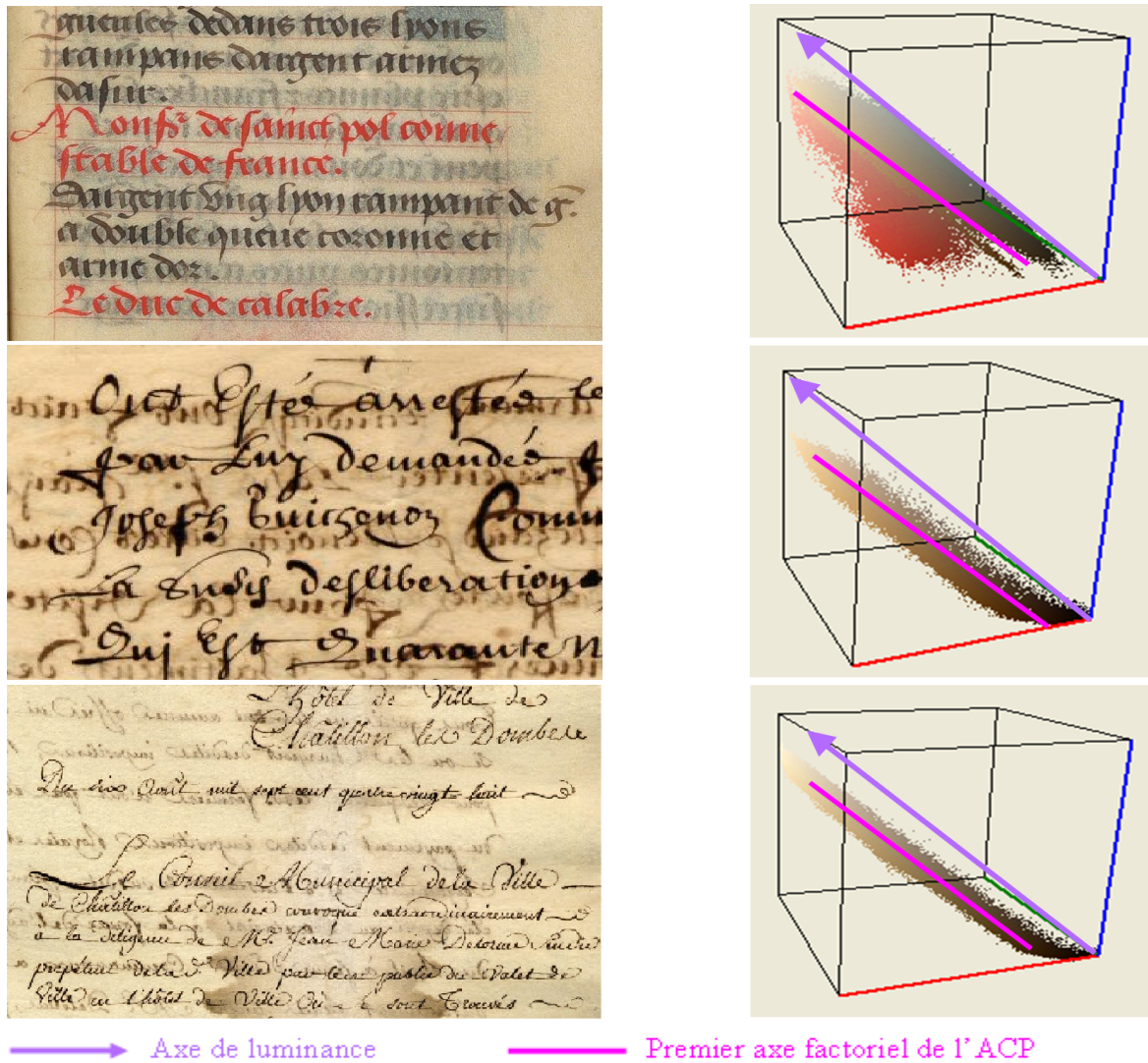


Figure 5.3 : Images de document et distribution des couleurs associées dans le cube RVB

par :

$$c_j = \frac{1}{n_j} \sum_{i=1}^{n_j} x_{ij} \quad \forall j = 1, \dots, K \quad (5.2)$$

L'équation 5.2 prouve que les centres c_j ne sont autres que les centres de gravités des différentes classes. La nouvelle classification des couleurs est obtenue en affectant à chaque point de l'image le centre de gravité qui lui est le plus proche. Si la nouvelle distribution est égale à la précédente, alors on atteint une situation de convergence où les centres sont fixes et l'algorithme s'arrête. En cas de différence entre les deux partitions, il faut reprendre le calcul des centres c_j .

5.3.2 Évaluation d'une segmentation linéaire

Nous avons soumis à l'algorithme de classification automatique du type K -moyennes l'image entière ayant comme caractéristiques les composantes RVB. Nous avons donné à

K la valeur 3 (puisque nous avons trois classes; le «recto», le «verso en transparence» et le «fond»). Comme nous le pensions, l'application d'une classification linéaire directe ne permet pas de séparer correctement le recto du verso. Soit la figure 5.4 illustrant une partie représentative d'un document possédant ce type de dégradation. La figure 5.5 montre les résultats obtenus suite à une classification automatique appliquée sur l'image donnée par la figure 5.4. Nous remarquons que la classe n°1, correspondant au recto, contient toujours des informations reliées au verso même après avoir effectué cette classification. Cette constatation confirme que la séparation entre le recto et le verso est un problème non trivial qui ne peut pas être résolu avec des techniques linéaires simples. C'est la raison pour laquelle nous proposons une adaptation de l'algorithme des K -moyennes, dans la section suivante.

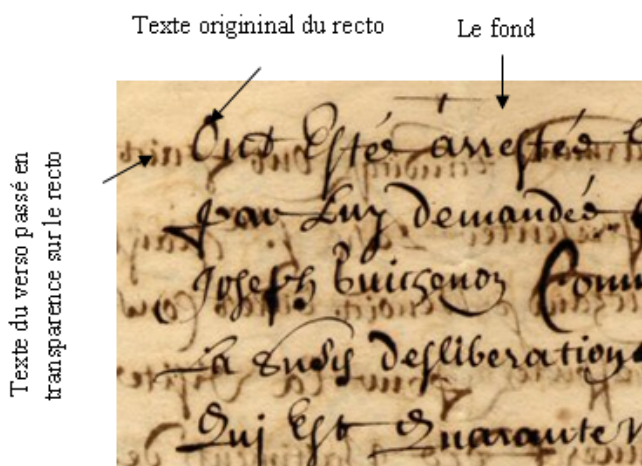


Figure 5.4 : Extrait d'une image dégradée de document

5.4 Proposition d'une classification en cascade

5.4.1 Justification de l'approche proposée

Nous allons procéder par une analyse dichotomique des couleurs de l'image. Cette idée revient à appliquer récursivement la décorrélation de l'information par l'ACP suivie d'une classification en deux classes. Le choix d'une telle approche de segmentation s'explique par le fait que :

1. aucune connaissance *a priori* du nombre total des classes n'est exigée. Il s'agit d'une procédure par une approche de segmentation non-supervisée.
2. une classification dichotomique en deux classes est plus performante qu'une classification directe en K classes (ce qui sera mieux illustré par la suite dans les résultats).

La première récursion réalise généralement une séparation fond /recto+verso. Par conséquent, pour les autres récursions, on ne soumet à l'analyse de la variance qu'un problème à deux classes (recto/verso). En effet, les pixels du support papier sont statistiquement plus fréquents dans l'image que les pixels du recto ou du verso et ils dirigent le premier axe factoriel. En appliquant l'analyse en composantes principales sur l'image sans le support papier,

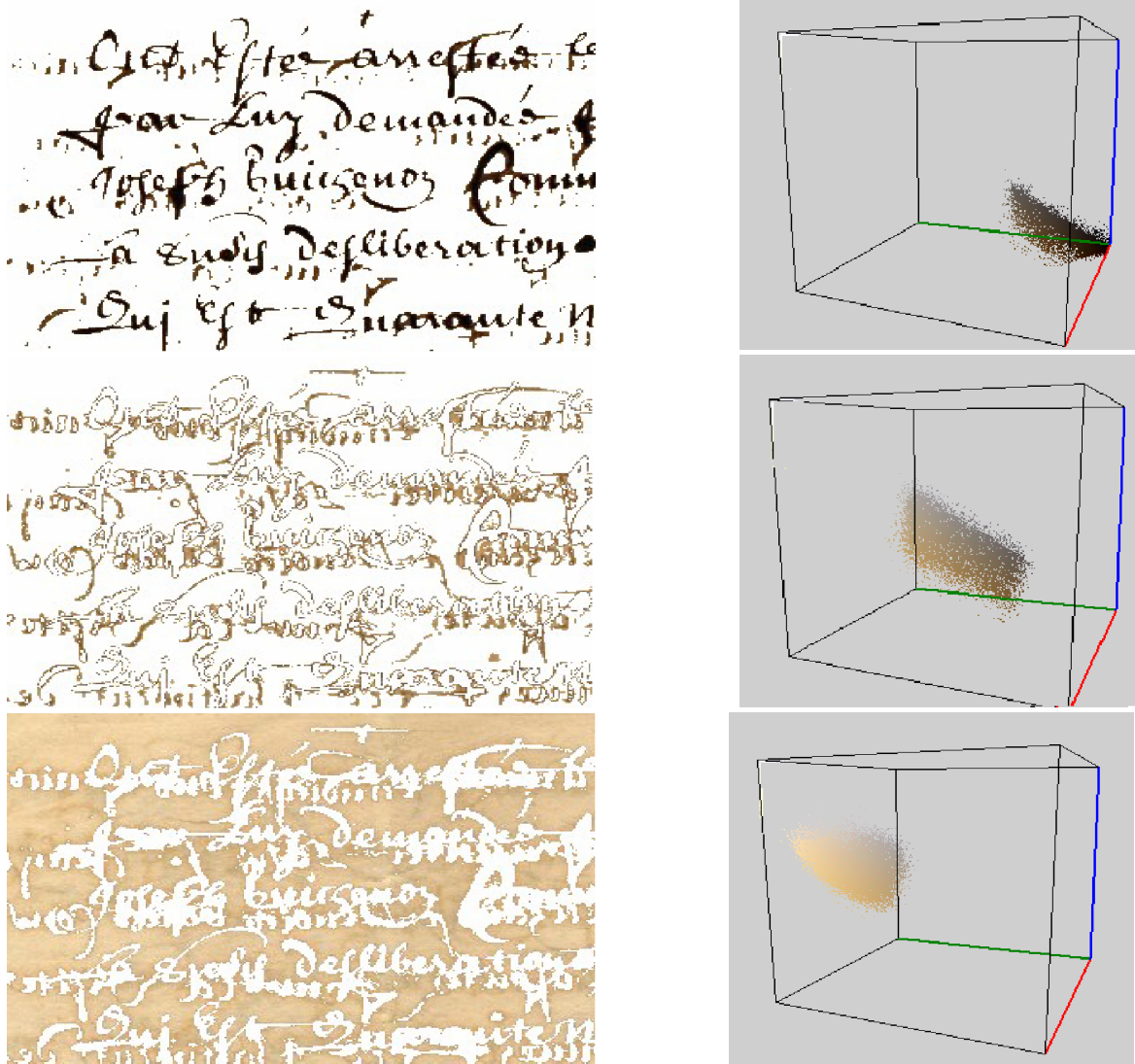


Figure 5.5 : Résultat d'une classification automatique par K -moyennes ($K=3$) sur l'image donnée sur la figure 5.4 et affichage de la distribution des couleurs des classes images générées dans le cube RVB

le résultat attendu serait un premier axe factoriel dont la projection des pixels devrait séparer au mieux le recto du verso. De même en appliquant plusieurs récursions, nous pourrions séparer les différentes classes des couleurs d'encre. Le résultat final est une arborescence dont la base représente l'image initiale et les feuilles décrivent les images contenant les différentes classes de couleurs. Le problème qui se pose à ce niveau est la sélection de la feuille qui correspond au texte original du recto. L'intervention de l'utilisateur est nécessaire à ce niveau, d'où l'idée d'introduire un critère de sélection.

5.4.2 Critère de sélection pour une segmentation non-supervisée

Notre critère de sélection fait appel à l'histogramme logarithmique. Un histogramme est défini comme l'outil statistique par excellence qui permet d'estimer la distribution des pixels

de l'image indépendamment de leur position. Si on note $Histo$ l'histogramme de l'image, alors la détermination de ses valeurs $i_{0 < i < 255}$ est telle que :

$$Histo(i) = Card \{Pixel(x, y) = i\} \quad (5.3)$$

avec $Card$ le cardinal ou la fréquence d'apparition d'une valeur de pixel.

Nous rappelons que l'histogramme logarithmique est un histogramme à échelle logarithmique. Il permet une appréciation plus globale du contenu de l'image en mettant en valeur les faibles fréquences. En partant de l'hypothèse que le texte original du recto est le texte le plus foncé, l'analyse de l'histogramme logarithmique de l'image peut être une solution adéquate à notre problème. Il permet, en effet, de ne favoriser que le traitement des classes d'images possédant un contenu. Ainsi, la classe image correspondant à l'histogramme logarithmique ayant le plus grand nombre de valeurs d'intensités faibles est la classe qui sera subdivisée en deux autres classes. La figure 5.6 donne une idée des résultats obtenus suite à l'application de notre méthode sur une image de test. Ces résultats montrent que notre approche dichotomique est plus performante qu'une simple classification globale en K classes appliquée à toute l'image (Fig.5.4). Au fur et à mesure qu'on descend dans l'arbre, nous arrivons à mieux discerner le texte original du recto. Au bout d'un nombre d'itérations (égal dans ce cas à 3), l'algorithme converge vers la classe recherchée.

5.5 Étude expérimentale de l'approche proposée

Pour résumer le principe de notre proposition, nous donnons le processus général de notre méthode dans la figure 5.7. Les étapes suivantes sont exécutées successivement :

1. Analyse en composantes principales (ACP) de l'image en cours,
2. Segmentation en deux classes par l'application des K -moyennes sur les données décorréelées,
3. Génération d'une image par classe en conservant les couleurs d'origine,
4. Calcul de l'histogramme logarithmique de chaque classe image. L'histogramme ayant le plus grand nombre de pixels sombres est associé à la classe image qui doit être ensuite traitée. Cette image va subir tous les traitements précédents en commençant par l'étape (1).

5.5.1 Étude des performances

Pour évaluer la méthode proposée, nous avons fait des tests sur des exemples d'images dégradées fournies par les archives de Chatillon-sur-Chalaronne. Les résultats obtenus (Fig.5.8) montrent l'efficacité de cette méthode qui réussit à extraire le texte original du recto du support dégradé. Les différentes marques du verso en transparence sont remplacées par la moyenne des pixels du fond.

La figure 5.9 montre la différence, pour une image de test donnée, entre l'image seuillée par la méthode de Sauvola et l'image obtenue après application de notre méthode.

La méthode proposée ne requiert :

1. **aucun traitement numérique de l'autre face du document dégradé.** Elle est capable de supprimer les marques, conséquence du passage en transparence du texte du verso sur le recto, à l'aide d'une seule face du document. Notre méthode se classe donc parmi les approches de restauration aveugles.
2. **aucun apprentissage.** Cette méthode est différente des méthodes d'apprentissage par réseau de neurones : les Self-Organizing Maps (ou cartes auto-organisatrices) [SMI-04] où un apprentissage doit avoir lieu sur des parties du document à traiter.
3. **aucun paramètre d'entrée.** Cette méthode est différente de l'approche de [LEY-04] qui a comme paramètres d'entrée le nombre de classes et les échantillons pour initialiser les nuées.

5.5.2 Étude des limites

L'approche dichotomique proposée s'est révélée efficace pour la restauration aveugle des images dégradées de documents et notamment pour la séparation entre le recto et le verso ; cette approche est conçue pour le cas où le recto ne correspond qu'à une seule couleur (Fig.5.8), ce qui n'est pas toujours vrai dans la réalité. Par exemple, l'image de la figure 5.10.a. affiche deux couleurs d'encre. Un traitement dichotomique de cette image nous permet de retrouver toutes les classes 2 à 2 dans un ordre probablement différent. Il n'est pas évident, dans ce cas, de sélectionner automatiquement les feuilles correspondantes aux versions restaurées. Nous avons laissé de côté pour ce traitement l'histogramme logarithmique, utile que pour le cas d'images contenant une couleur unique du texte original du recto. Nous avons dû choisir manuellement les deux sous-classes représentant respectivement le texte rouge et le texte noir. Ces sous-classes peuvent servir de masque pour créer une image artificielle du document et le restaurer partiellement. Le fond est remplacé par sa moyenne et les couches rouge et noire sont copiées telle quelles (Fig.5.10).

Notre proposition permet de résoudre le cas du recto monocolore dégradé et elle est mal adaptée aux textes des rectos multicolores. Si nous avons réussi à automatiser le choix de la classe image correspondant au texte original du recto monocolore parmi toutes celles présentes dans l'arborescence générée, il nous a été impossible de l'adapter au cas de texte multicolore. L'intervention de l'utilisateur est nécessaire, cependant, il serait judicieux et utile de rendre cette sélection automatique.

D'autres inconvénients de notre proposition sont hérités de l'algorithme des K -moyennes. Cet algorithme est très sensible au choix des conditions initiales. Le résultat final n'est qu'une solution parmi d'autres qui ne minimise le critère donné par l'équation (5.1) que localement. L'algorithme des K -moyennes est une procédure non déterministe puisque des conditions initiales différentes mènent à des résultats différents.

5.6 Conclusions et perspectives

Dans ce chapitre, nous avons mis en évidence que la restauration des dégradations du fond et plus particulièrement la séparation entre le recto et le verso n'est pas un problème trivial qui peut être résolu par une simple classification directe des couleurs en K classes. Notre première démarche consiste à appliquer une segmentation récursive utilisant à la fois l'algorithme de segmentation par K -moyennes et l'analyse en composantes principales pour

décorrélér l'information. Cette proposition a permis de séparer correctement le texte du support papier sur certaines images dont le recto est représenté par une seule couleur. En revanche, nous échouons à traiter les autres documents à rectos multicolores. Étant données les limites de cette approche et son manque de généralité, nous chercherons d'autres voies plus sûres et plus génériques dans les chapitres suivants.



Figure 5.6 : Extrait de l'arbre g n r  suite   l'application de notre m thode sur une image de test (a); (a1), (a2), (a11), (a12), (a121) et (a122) repr sentent les diff rentes classes images g n r es; (LH1), (LH2), (LH11), (LH12), (LH121) et (LH122) repr sentent les histogrammes logarithmiques respectifs de (a1), (a2), (a11), (a12), (a121) et (a122).

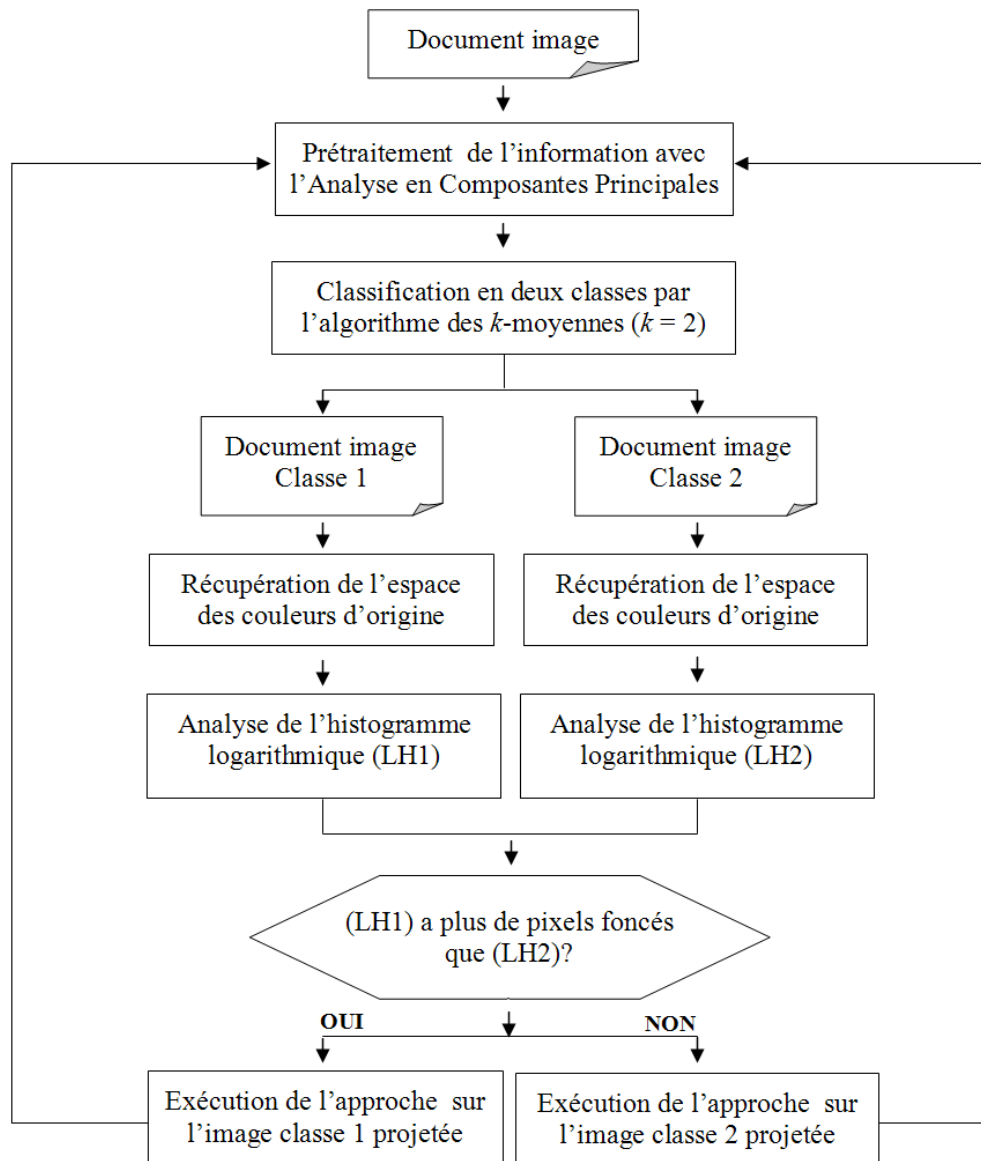


Figure 5.7 : Le synopsis de la méthode proposée



Figure 5.8 : De gauche à droite respectivement : l'image d'origine et sa version restaurée

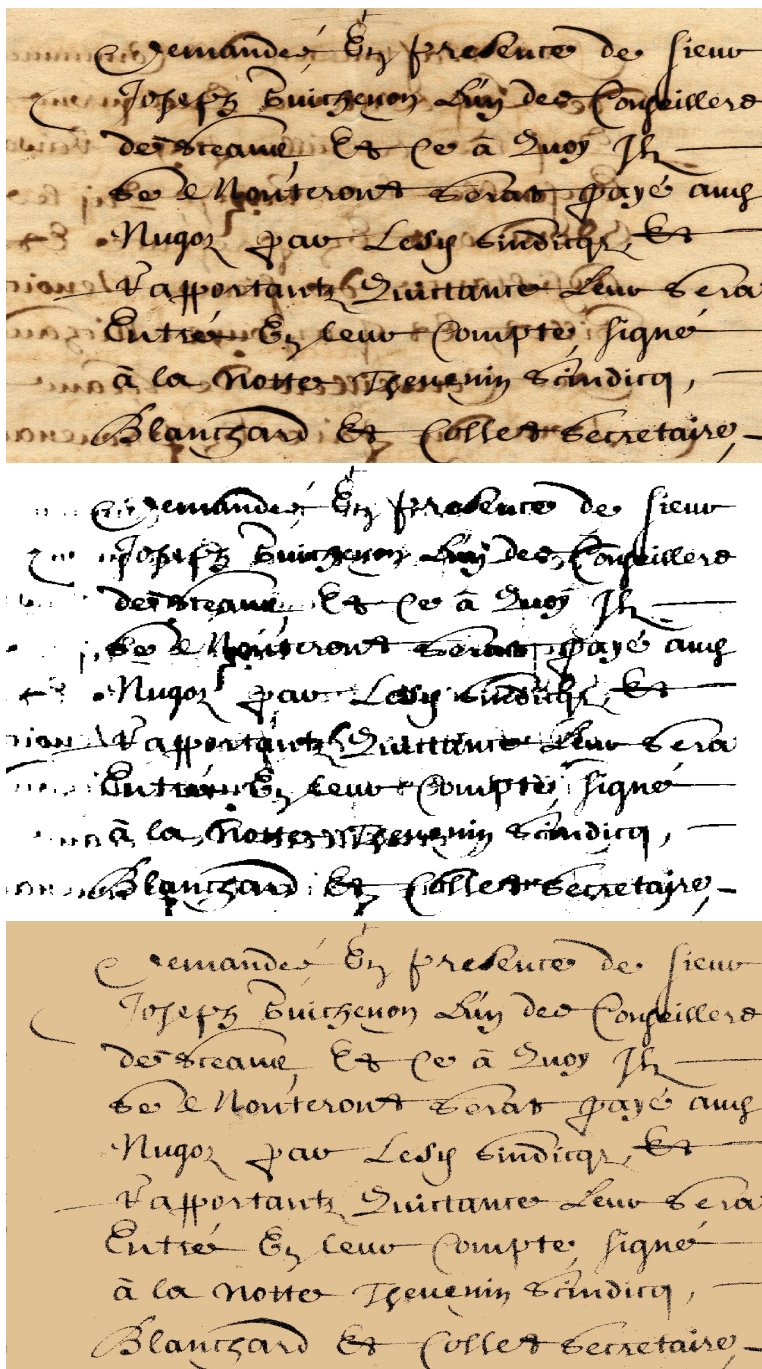
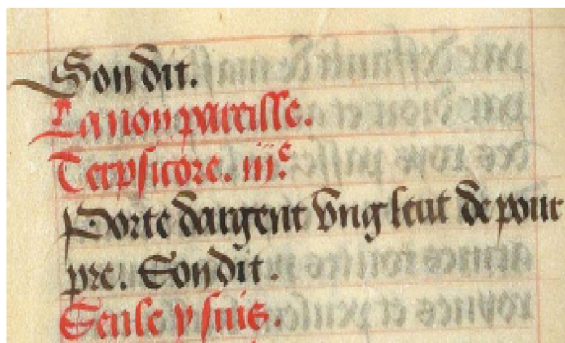
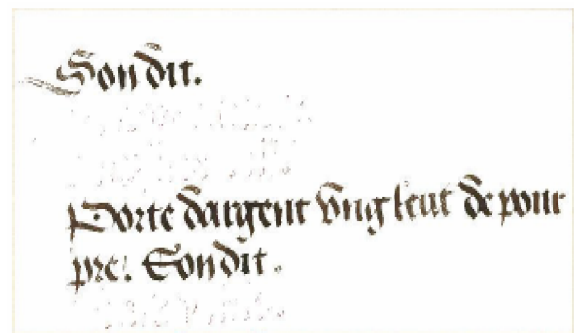


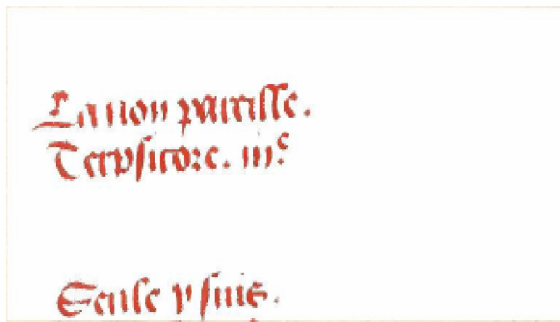
Figure 5.9 : De haut en bas : Image dégradée de document, résultat d'un seuillage par Sauvola, résultat obtenu par la classification récursive



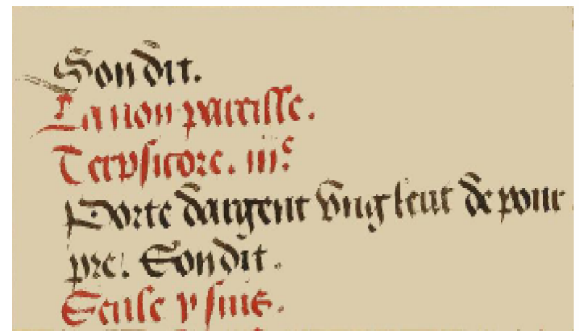
(a) *L'image dégradée*



(b) *La classe contenant le texte noir*



(c) *La classe contenant le texte rouge*



(d) *La version restaurée résultante*

Figure 5.10 : Application de la classification récursive sur une image dégradée multicolore

Chapitre 6

Deuxième Méthode : Restauration par une approche duale

« Sous le masque de la complexité, la simplicité se questionne. »

Monique Keurentjes

Résumé

Dans ce chapitre, nous proposons une autre approche globale de classification des couleurs pour nettoyer les dégradations du support papier. Il s'agit d'une adaptation du Mean-Shift qui a déjà fait ses preuves dans le domaine de la segmentation des images naturelles mais qui n'a encore jamais été testée sur des images de documents. Nous montrons que la densité des classes de couleur est un bon critère pour séparer le texte du support papier bruité. Cette méthode de segmentation est suffisamment générique pour permettre de traiter une très grande variété de documents, de différentes origines, sans nécessiter l'intervention de l'utilisateur. Mais le Mean-Shift souffre d'une forte complexité de calcul qui limite son application dans le traitement d'images de documents de grandes tailles. Nous proposons, une approche duale qui introduit, en plus des informations globales sur la distribution des couleurs, des informations locales pour réduire la complexité du Mean-Shift.

Sommaire

6.1	Introduction	66
6.2	Le principe du Mean-Shift	66
6.3	Notre Proposition : le Mean-Shift local-global	70
6.4	Étude expérimentale de l'approche proposée	72
6.5	Conclusions et perspectives	79

6.1 Introduction

Les méthodes de classification par recherche des modes sont des méthodes performantes qui ne requièrent aucune intervention de l'utilisateur. Leur principe consiste à regrouper les points qui convergent vers le même maximum local de la densité des couleurs de l'image, appelé mode. Chaque point peut donc être associé à un mode de cette densité et chaque mode est associé à une classe. Le nombre total de modes représente ainsi le nombre de classes à extraire. En effet, ces méthodes de classification reposent sur l'estimation d'une fonction de densité des couleurs de l'image. Il existe deux types de méthodes d'estimation de la densité de données multidimensionnelles :

- **les méthodes paramétriques** : Elles supposent que les données (les couleurs) sont distribuées suivant un modèle connu. L'objectif de ces méthodes est donc d'estimer les paramètres du modèle à partir des observations. Dans notre cas, nous ne pouvons pas faire l'hypothèse que les distributions des couleurs suivent une loi Gaussienne ou bien une autre loi quelconque. Les méthodes paramétriques ne garantissent pas la généralité de la segmentation car les hypothèses de départ ne leur permettent de fonctionner que sur des images particulières de documents.
- **les méthodes non-paramétriques** : Elles ne font aucune hypothèse sur la distribution des données et permettent de définir des classes qui peuvent avoir des formes quelconques. La densité des couleurs peut être estimée dans une fenêtre de Parzen [PARZ-61]. Cette catégorie de méthodes non-paramétriques est adaptée au traitement générique des documents de différentes origines. Un algorithme qui appartient à la catégorie des méthodes non-paramétriques est l'algorithme du Mean-Shift.

Nous proposons donc, dans ce chapitre, de résoudre les problèmes de séparation recto/verso et de restauration des dégradations du fond par l'algorithme du Mean-Shift. Il existe dans la littérature deux procédures de cet algorithme : le Mean-Shift global et le Mean-Shift spatial ; la différence entre les deux se situe au niveau de la prise en compte ou non, durant la classification, des informations spatiales des couleurs. Mais quelle que soit la différence, le principe reste le même : il s'agit d'une procédure itérative basée sur une estimation non-paramétrique de la densité locale des couleurs de l'image. Nous verrons que le Mean-Shift global possède de bonnes propriétés de robustesse et de généralité pour différencier les couleurs du texte original du recto des différentes autres couleurs présentes sur le support papier (couleurs du verso en plus des couleurs du fond). Toutefois, le Mean-Shift global entraîne des calculs complexes qui limitent son application directe au traitement des images de document en haute résolution. Nous proposerons ainsi une troisième procédure de Mean-Shift que nous avons appelée le Mean-Shift local-global ; elle possède les performances de la version d'origine sans la complexité de calcul. Elle est donc bien adaptée à la séparation entre le recto et le verso dans les images de document de grandes tailles.

6.2 Le principe du Mean-Shift

Le principe du Mean-Shift est inspiré du gradient de densité introduit par Fukunaga en 1975 [FUK-75, FUK-90], puis généralisé par Cheng en 1995 [CHE-95] et mis en pratique,

notamment pour la segmentation des images, à partir de 1997 avec les travaux de Comaniciu et Meer [COM-97, COM-99]. Le principe du Mean-Shift consiste à associer à chaque point de l'image le mode (mean) de la densité des couleurs de l'image dont elle est issue. Une fenêtre de Parzen [PARZ-61] permet d'estimer la fonction de densité des classes de couleurs correspondant aux différentes régions de l'image. Cette approche repose sur l'hypothèse selon laquelle les régions homogènes de l'image sont caractérisées par une densité de probabilité monomodale de la distribution des couleurs. Trouver ces modes revient à trouver les points pour lesquels le gradient de la fonction de densité est nul. C'est donc la méthode itérative de descente de gradient qui permet de déplacer (Shift) les couleurs de l'image vers les modes (mean) dans la direction où la densité des couleurs est la plus forte.

6.2.1 Procédure du Mean-Shift global

La procédure du Mean-Shift global est une procédure itérative de recherche de modes ou de maxima locaux. Pour tout point $\{x_i\}_{i=1\dots N}$ de l'image, la recherche d'un mode est décrite par l'algorithme suivant :

1. Choisir le paramètre h associé au facteur d'échelle,
2. Initialiser le nombre d'itérations $p = 1$ et le centre de la fenêtre $y^1 = x_i$
3. Calculer le vecteur Mean-Shift $m_{h,G}(y^p)$, avec pour g le noyau uniforme ou Gaussien

$$m_{h,G}(x) = \frac{\sum_{i=1}^K x_i g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}{\sum_{i=1}^K g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)} - x .$$

pour les K points x_i dans la fenêtre de Parzen $N_c(x)$ centrée en x :

$$N_c(x) = \{x_i / d_{couleur}(x, x_i) < h\}$$

4. Translater le point y^p vers la nouvelle position y^{p+1} tel que $y^{p+1} = y^p + m_{h,G}(y^p)$,
5. Incrémenter le nombre d'itérations $p = p + 1$,
6. Reprendre l'algorithme en 3 jusqu'à converger vers le point y où $\|m_{h,G}(y)\| < \epsilon$.

Le principe de cette méthode consiste donc à itérer le déplacement d'un point y par le vecteur Mean-Shift. À chaque itération, le point de l'image y est translaté vers la moyenne des couleurs pondérée de l'ensemble des points x_i trouvés dans la fenêtre de Parzen centrée en x . La procédure s'arrête dès qu'on aboutit à un point stationnaire caractérisé par un vecteur gradient nul. Ce point de convergence est, par définition, un mode de la distribution sous-jacente de la classe des couleurs associée. Les points ayant une distance colorimétrique en dessous du seuil h se voient converger vers un même mode. La figure 6.1 donne un exemple de trajectoires de deux points différents vers les centres de leurs classes respectives. Cette figure montre cette convergence dans le plan de la chrominance (U^*, V^*) en deux dimensions alors que cette convergence se fait dans un espace couleur à trois dimensions incluant aussi la luminance (Y) ou directement dans l'espace (R, G, B). La norme du vecteur Mean-Shift est toujours plus petite que h et décroît dans les zones plus denses pour devenir nul à la position du mode d'une classe couleur.

La localisation et le nombre de modes dépendent du paramètre h qui correspond à la taille de la fenêtre de Parzen. Ce paramètre peut s'apparenter également à un paramètre de lissage; pour une valeur de h très faible, la fonction de densité est calculée à un niveau de détail très fin, ce qui conduit à un nombre élevé de modes et donc à une sur-segmentation

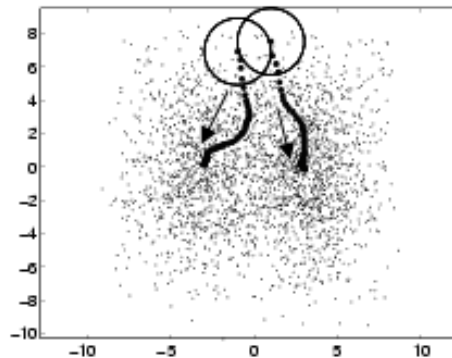


Figure 6.1 : Convergence de la procédure Mean-Shift. Les points noirs représentent les trajectoires des points par translations successives selon le vecteur Mean-Shift à chaque itération

de l'image. Par contre, une valeur élevée de h risque d'augmenter le lissage des données, de regrouper les maxima locaux proches de faibles niveaux en maxima plus élevés et de réduire le nombre de modes conduisant à une sous-segmentation de l'image. La figure 6.2 met en évidence l'influence du paramètre h sur l'image de test donnée. Le choix de la valeur h dépend de l'application envisagée. Dans notre application, la valeur de $h=7$ donne une meilleure approximation du texte original du recto par rapport aux autres taches.

Les performances du Mean-Shift global ont été prouvées lors de leur application dans le cadre de la segmentation d'images couleurs [COM-97]. Mais cette procédure présente un gros inconvénient : son temps de calcul. Sa complexité est de l'ordre de $O(N^2)$ où N est le nombre total des points de l'image. En effet, le calcul du vecteur Mean-Shift en chaque point de l'image consiste à trouver, parmi les N points de l'image, ceux qui sont distants à moins de h du point étudié. Sans optimisation, cette recherche nécessite N^2 distances pondérées par le noyau g . De plus, nous ne tenons pas compte du nombre aléatoire d'itérations nécessaires pour translater tous les points de l'image vers leurs modes respectifs. L'optimisation basée sur un partitionnement des données dans un KD-tree permet de réduire la complexité en $O(N \log N)$ au détriment d'un encombrement mémoire trop volumineux pour les images de grandes tailles. Or les images de documents sont généralement numérisées avec des résolutions élevées et possèdent des tailles plus grandes que les photos, les images du Web ou les images issues des vidéos. À titre d'exemple, l'extrait d'image de la figure 6.2.a, de taille 100×129 , a nécessité plus de 20 *minutes* de calcul pour trouver tous les modes de la fonction de densité des 12900 pixels de cette image. Le traitement d'une image d'un document numérisé en 400dpi, de taille 3000×4000 , nécessite au moins une journée de calcul par image. Pour traiter ces images dans un temps raisonnable, il convient de réduire la complexité de l'algorithme.

6.2.2 Procédure du Mean-Shift spatial

Comaniciu et Meer [COM-02] proposent une méthode connue sous le nom du Mean-Shift spatial qui consiste à appliquer le principe général du Mean-Shift à un vecteur de caractéristiques plus grand auquel on ajoute les informations spatiales (x,y) des points. Elle mélange ainsi les informations spatiales (position des pixels) et les informations colorimétriques (valeurs des pixels) pour segmenter les images dans des délais raisonnables. Les régions segmen-



Figure 6.2 : Application du Mean-Shift global pour différentes valeurs de h

tées remplissent à la fois les critères de proximité spatiale et d'homogénéité colorimétrique des pixels qui les composent. La simplicité, la robustesse et la vitesse d'exécution mises en avant par Comaniciu et Meer ont fait du Mean-Shift spatial une méthode particulièrement attractive pour de nombreux problèmes en segmentation d'images. Cependant, cette méthode

non-paramétrique provoque une sur-segmentation des images naturelles en micro-régions très similaires à celles obtenues avec les bassins d'attraction par la morphologie mathématique [BRE-03]. L'application du Mean-Shift spatial sur les documents dégradés présente le même inconvénient de sur-segmenter les traits et le support papier (Fig.6.3). L'étape du Mean-Shift spatial ne suffit pas dans le cas de taches étendues spatialement, c'est pourquoi elle est complétée par la fusion de proche en proche des modes voisins pour corriger la sur-segmentation. L'utilisation de deux méthodes différentes de classification (par détection de mode puis regroupement basé sur la distance) est alors nécessaire. Une fois les modes obtenus, ils sont regroupés en fonction d'un seuil arbitraire sur leurs distances respectives dans l'espace des caractéristiques. La phase de fusion des régions est cruciale et la performance d'une telle phase ne garantit pas toujours l'obtention d'une image de bonne qualité. Elle donne un aspect de «mosaïque» le long des traits ainsi que des ruptures de continuité comme nous l'avons déjà constaté dans [GAR-06]. De plus, la fusion de certaines zones uniformes impose la fixation de certains critères pas du tout évidents à mettre en œuvre.

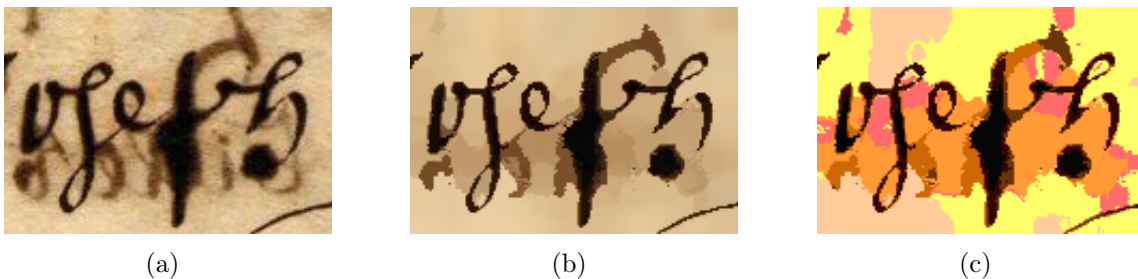


Figure 6.3 : (a) Image dégradée, (b) Résultat du Mean-Shift spatial, (c) Sur-segmentation en régions affichées en pseudo couleurs

6.3 Notre Proposition : le Mean-Shift local-global

Nous allons introduire la méthode dite de Mean-Shift local-global qui permet de diminuer la complexité du Mean-Shift global sans pour autant dégrader la qualité du résultat final obtenu. Son concept repose sur l'extension de la méthode originale du Mean-Shift global dans un espace qui prend en compte les informations colorimétriques locales et globales.

6.3.1 Principe du Mean-Shift local-global

L'algorithme du Mean-Shift global définit, pour chaque point x de l'image, un voisinage dans l'espace des caractéristiques liées à l'information couleur. Ce voisinage inclut tout pixel appartenant à une fenêtre de rayon arbitraire h et de centre le pixel courant x de l'image. Il représente l'information clé pour l'estimation globale de la densité. La recherche de tous les pixels de l'image dont la distance par rapport à x est inférieure à h nécessite le calcul de $(N - 1)$ distances. Pour ne pas calculer toutes les distances, nous faisons l'hypothèse que localement, autour de x , il existe R couleurs proches de x dont la distance est inférieure à h . Ce qui revient à rechercher, localement autour du point x , le voisinage qui regroupe l'ensemble des R points les plus proches localement et colorimétriquement de x . Nous notons ce nouveau voisinage par le voisinage $R - NNC$ du point courant x défini par :

$$R - NNC(x) = \left\{ \begin{array}{l} y/d_{couleur}(x, y) < h \text{ avec } y \text{ un point parmi les} \\ R \text{ plus proches voisins spatialement de } x \end{array} \right\}$$

Le voisinage $R - NNC$ est extrait à l'aide d'une fenêtre glissante, centrée sur x et dont la taille s'agrandit jusqu'à obtenir les R points localement et colorimétriquement proches de x (Fig.6.4). Le principe consiste à commencer par un voisinage 3x3 dont on extrait les points ayant une distance colorimétrique inférieure à h . Tant qu'on n'a pas atteint le nombre R de points colorimétriquement proches de x suivant le seuil h , on incrémente la taille de la fenêtre jusqu'à une taille de fenêtre limite 16x16. La contrainte spatiale, introduite par le voisinage $R - NNC$, réduit considérablement le coût d'une recherche sur l'ensemble des pixels de l'image. La valeur R détermine la qualité de l'estimation du gradient de densité des couleurs. Plus la valeur R est grande, plus le nombre de points intervenant dans le calcul du gradient de densité est élevé et plus l'estimation du vecteur de Mean-Shift est fiable.

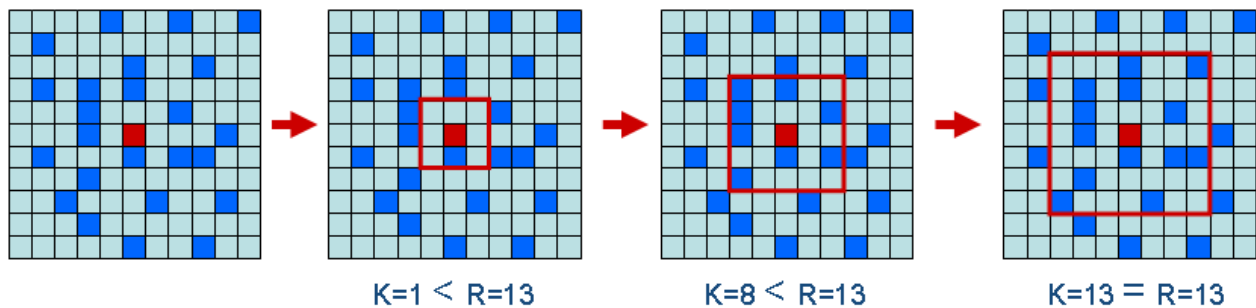


Figure 6.4 : Principe de la fenêtre adaptative pour la sélection des voisins : le pixel rouge représente le point courant de l'image, les points bleus sont les points ayant une distance colorimétrique inférieure à un seuil h , R est le nombre de voisins à extraire, K le nombre de voisins extraits pour la fenêtre en cours.

La convergence du Mean-Shift local est prouvée car notre proposition n'est autre qu'un Mean-Shift global avec un voisinage colorimétrique réduit aux points situés dans un voisinage autour du point x . Un exemple de convergence est donné sur la figure 6.5 qui montre, pour trois points choisis aléatoirement et appartenant respectivement au «fond», au «verso en transparence» et au «recto», que la procédure converge vers le mode du nuage quel que soit l'emplacement du point. Le Mean Shift local permet de trouver un nombre réduit de modes locaux qui serviront de points de départ pour la deuxième phase. Cette dernière consiste à appliquer le Mean-Shift global classique sur un nombre réduit de modes locaux, détectés au cours de la première phase. Plus la valeur de R est élevée, plus le nombre de modes locaux sera réduit et plus ils seront proches des modes globaux calculés sur l'image entière. Il est facile de constater que si $R \rightarrow \infty$ alors notre Mean-Shift local-global, qui est une combinaison séquentielle d'un Mean-Shift local suivi d'un Mean-Shift global, devient un Mean-Shift entièrement global.

6.3.2 Processus de classification par le Mean-Shift local-global

Nous pouvons résumer les principales étapes de l'algorithme comme suit (Fig.6.6) :

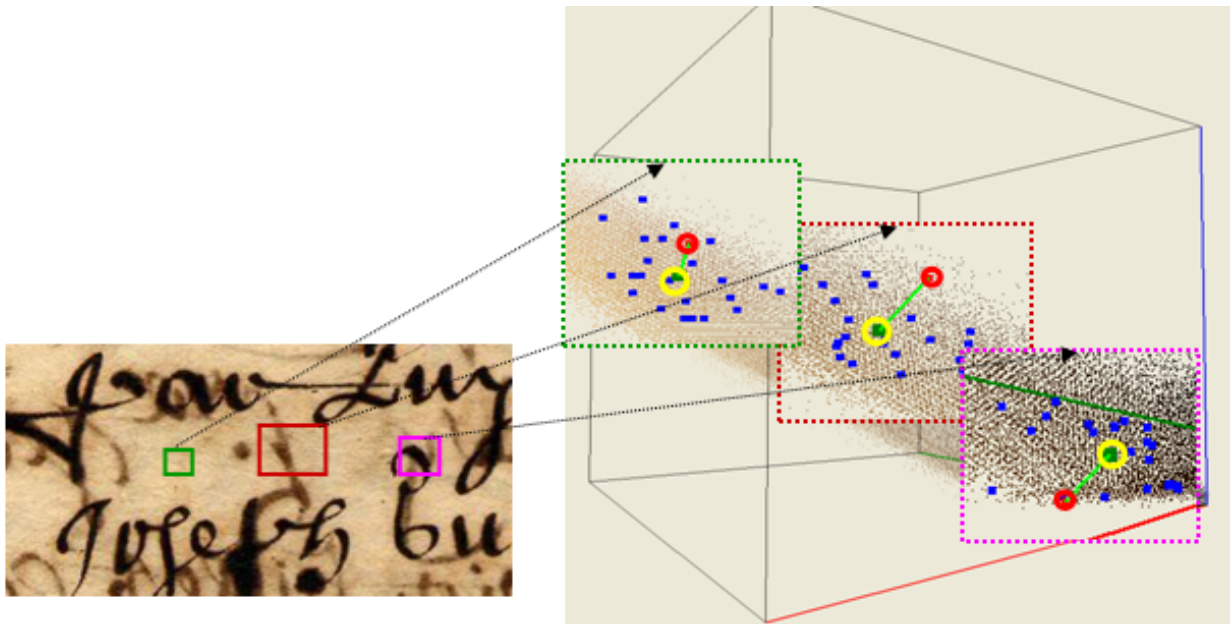


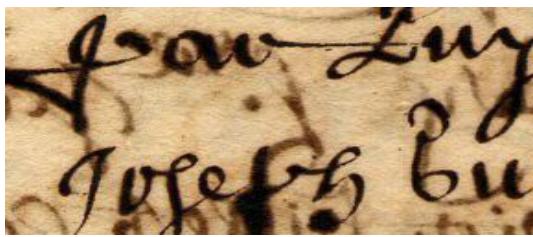
Figure 6.5 : Extrait d'une image de document avec un zoom sur des fenêtres locales dans le cube RVB après application du Mean-Shift local. Les points bleus forment le voisinage $R - NCC$ de chaque point en cours (présenté par un cercle rouge), les cercles jaunes sont les modes locaux extraits.

1. Appliquer la procédure du Mean-Shift local pour tout point de l'image. Cette procédure ne tient en compte que le voisinage $R - NCC$ d'un pixel donné. Quand tous les pixels de l'image sont traités, nous aboutissons à l'extraction de divers modes locaux.
2. Appliquer la procédure itérative du Mean-Shift global aux différents modes locaux extraits à l'étape précédente par rapport à tous les pixels de l'image originale. Quand tous les modes locaux sont traités, nous aboutissons à l'extraction de divers modes globaux.
3. Affecter à chaque pixel de l'image originale la valeur du mode global trouvé. Le nombre total de classes extraites est donné par le nombre de modes globaux détectés.

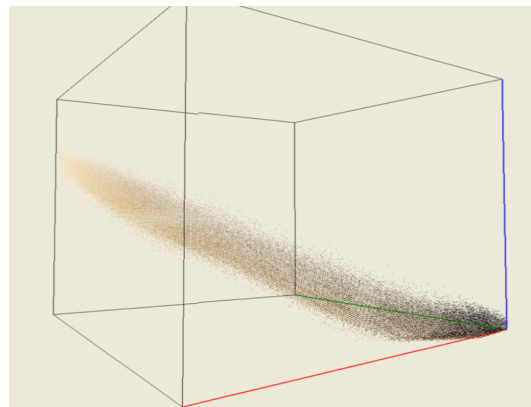
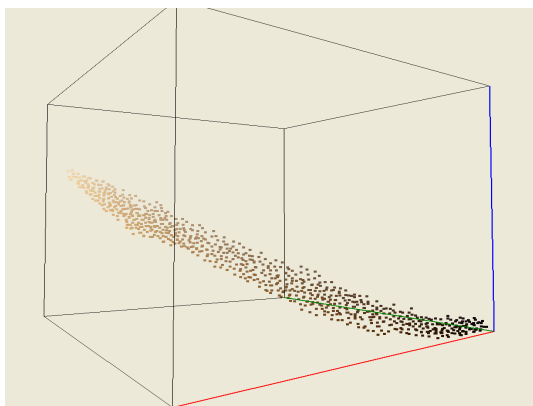
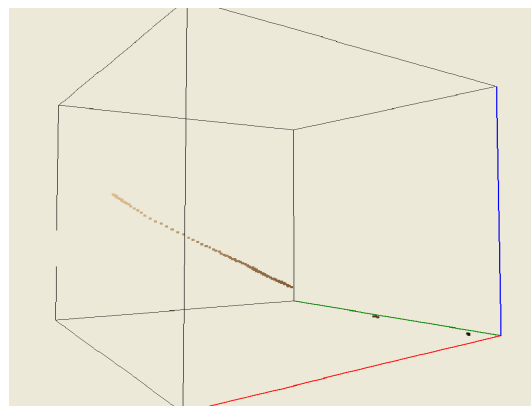
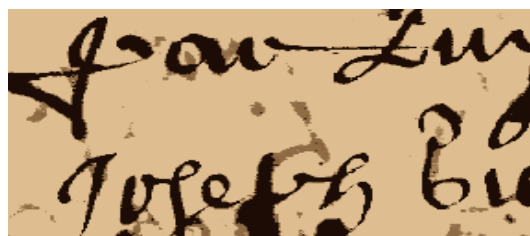
6.4 Étude expérimentale de l'approche proposée

6.4.1 Étude de l'influence du choix des paramètres

L'approche proposée se déroule en deux étapes et fait appel à deux paramètres : le nombre R du $R-NNC$ et le paramètre d'échelle h . Ces deux paramètres sont fixés au préalable par l'utilisateur. Étudier les performances de notre méthode pour différentes valeurs de R et h permettra de voir de plus près l'impact de la variation de ces paramètres sur le nombre de classes détectées et donc sur la qualité du résultat final. Pour cela, nous avons effectué plusieurs exécutions en faisant varier R et h sur un extrait d'une image de document (Fig.6.6.a).



(a) Extrait d'une image de document

(b) Distribution des N couleurs de l'image (a) dans l'espace RVB(c) Distribution des modes locaux (ML) après application du Mean-Shift local sur tous les points de l'image (a)
($R = 25 \Rightarrow ML = 870$)(d) Distribution des modes globaux (MG) après application du Mean-Shift global sur les modes locaux détectés (b)
($ML = 870 \Rightarrow MG = 23$)

(d) Image segmentée après application du Mean-Shift local-global

Figure 6.6 : Application du Mean-Shift local-global sur un extrait d'une image de document

de taille 118×101 . Les valeurs considérées sont respectivement 5, 10, 25, 50 et 100 pour R et 2, 3, 4, 5, 6 et 7 pour h . La figure 6.7 et la figure 6.8 montrent l'évolution du nombre total d'itérations sur tous les pixels de l'image en fonction de h et pour différentes valeurs

de R . Nous remarquons que le nombre d'itérations et donc le temps de calcul est fortement dépendant de la valeur de h . Ce résultat est prévisible étant donné qu'une augmentation de la valeur de h entraîne un champ de recherche plus grand autour du point à traiter.

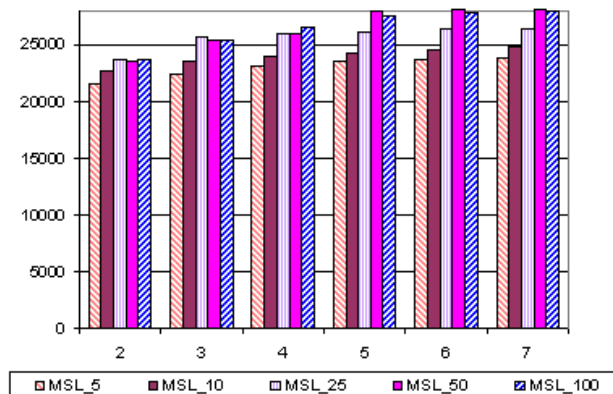


Figure 6.7 : Cas de la première étape du Mean-Shift local-global : Évolution du nombre total d'itérations en fonction de h et R

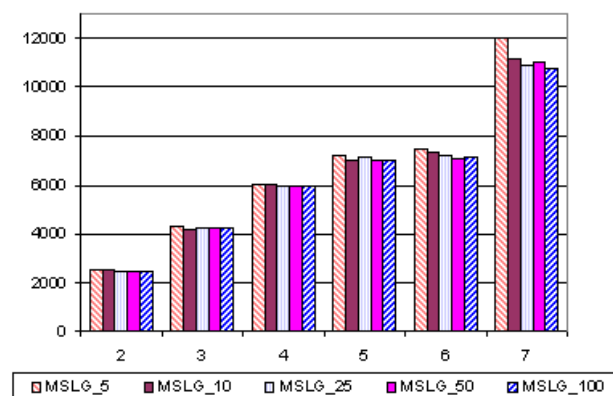


Figure 6.8 : Cas de la deuxième étape du Mean-Shift local-global : Évolution du nombre total d'itérations en fonction de h et R

Dans le cas du Mean-Shift local, le nombre R relié au voisinage R -NNC interagit avec la valeur de h sélectionnée. En effet, une faible valeur de h engendre un nombre R de voisins dont la propriété colorimétrique est très voisine du point courant. La convergence vers un mode local est assurée au bout d'un nombre réduit d'itérations. En revanche, une forte valeur de h rend la sélection des R voisins moins sélective. Plus particulièrement, les voisins les plus proches dans le domaine spatial du point courant peuvent satisfaire cette distance colorimétrique sans pour autant avoir besoin d'augmenter énormément la taille de la fenêtre. Pour avoir une meilleure convergence vers un mode local, nous avons choisi, pour nos tests, une valeur de R égale à 100. Une grande valeur de R entraîne un coût plus élevé de calcul durant la première phase. En contrepartie, cette valeur permet une meilleure détermination du gradient de densité et des modes locaux proches des modes globaux. Comme la deuxième phase est la plus coûteuse en temps de calcul puisqu'elle correspond à l'application du Mean-

Shift global par rapport à tous les points de l'image, nous obtenons une convergence plus rapide avec des grandes valeurs de R .

Lors de la deuxième étape, le champ de recherche des voisins aux modes locaux précédemment calculés s'étend sur tous les pixels de l'image. Plus la valeur de h augmente, plus un nombre important de voisins sera extrait et plus le nombre total d'itérations s'élève. La figure 6.9 montre une nette baisse du nombre des modes locaux et des modes globaux avec l'augmentation de la valeur du paramètre d'échelle h . Ce paramètre influe ainsi sur le nombre de modes détectés : plus il est grand, plus le nombre de maxima ou modes détectés est petit et inversement.

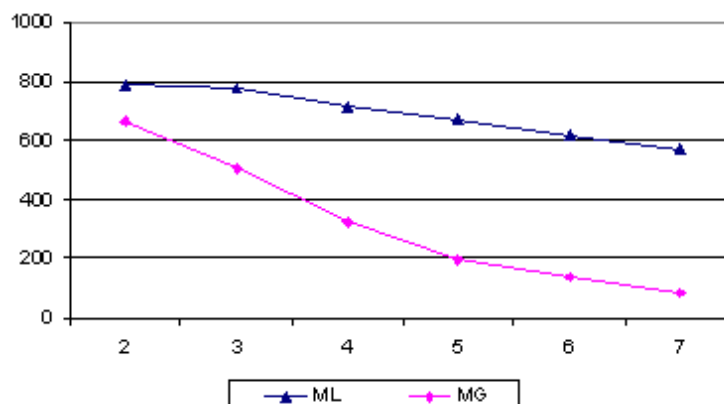


Figure 6.9 : Évolution du nombre des modes locaux (ML) et des modes globaux (MG) en fonction de h

6.4.2 Étude de la qualité du résultat

Des tests sur des images dégradées de documents (Fig.6.10 et Fig.6.11) montrent l'aptitude de notre algorithme du Mean-Shift local-global à extraire correctement le texte original du recto du support dégradé, sans aucune connaissance *a priori* sur les caractéristiques de l'image.

Notre proposition est une version modifiée et optimisée en vitesse de l'algorithme du Mean-Shift global. Comme nous le souhaitons, le résultat final obtenu avec notre Mean-Shift local-global est très proche de celui du Mean-Shift global. La figure 6.12 montre les différences entre les résultats obtenus par ces deux algorithmes. Nous notons sur cet exemple (Fig.6.12) que notre amélioration porte surtout sur les temps de calculs.

6.4.3 Étude du temps de calcul

Pour étudier le temps de réponse, nous proposons d'analyser le nombre total d'itérations mesuré en moyenne sur tous les pixels d'une image donnée. L'analyse a pour objet de comparer les temps de réponse respectivement associés à l'algorithme du Mean-Shift global et à l'algorithme du Mean-Shift local-global. La figure 6.13 montre les résultats obtenus suite à l'application de ces deux algorithmes sur une image de test (Fig.6.6.a). Nous notons que quel



Figure 6.10 : Application du Mean-Shift local-global sur des extraits d'images dégradées de documents, à gauche les images de test, à droite les images restaurées pour respectivement de haut en bas $h=7$, $h=6$ et $h=7$

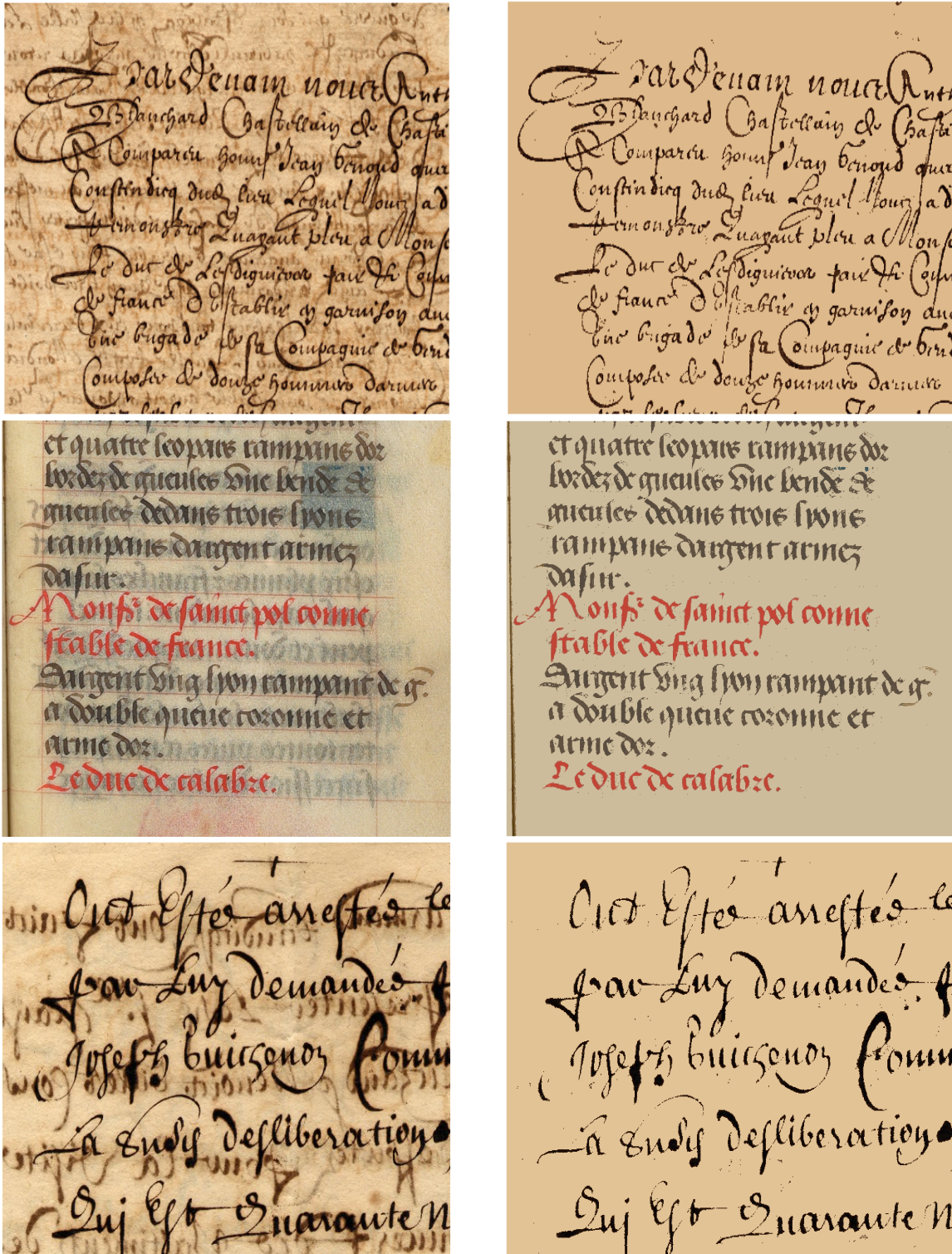


Figure 6.11 : Application du Mean-Shift local-global sur des extraits d'images dégradées de documents, à gauche les images de test, à droite les images restaurées pour respectivement de haut en bas $h=5$, $h=3$ et $h=3$

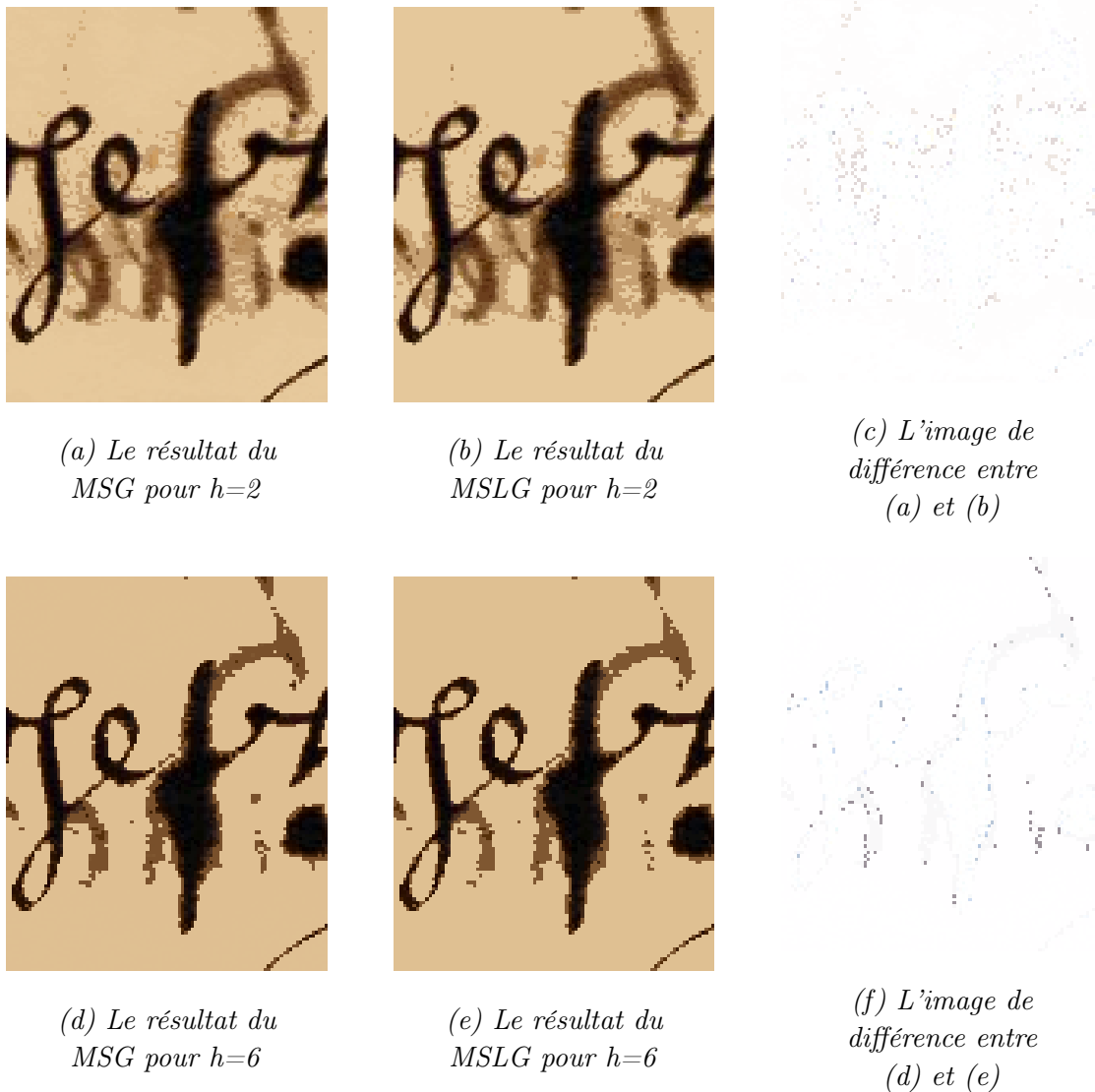


Figure 6.12 : Étude comparative de la qualité du résultat final pour le Mean-Shift global (MSG) et le Mean-Shift local-global (MSLG). L'image de test est donnée sur la figure 6.6.a

que soit la valeur du facteur d'échelle h , le nombre d'itérations nécessaire pour la convergence est nettement plus petit pour notre algorithme qu'avec celui du Mean-Shift global. Ce résultat nous incite à comparer la complexité de ces deux algorithmes.

6.4.4 Étude de la complexité

L'application de l'algorithme du Mean-Shift local, comme première étape, réduit considérablement le temps de calcul et donc la complexité de l'algorithme du Mean-Shift local-global. En effet, la complexité de l'algorithme initial est estimée, sans optimisation, à $O(N \times N)$, où N est le nombre de pixels de l'image. L'opération la plus coûteuse dans cet algorithme est la recherche des voisins les plus proches du point courant dans l'espace couleur. En utilisant les structures les plus populaires, le *KD-tree*, la complexité peut être réduite à $O(N \log N)$ opé-

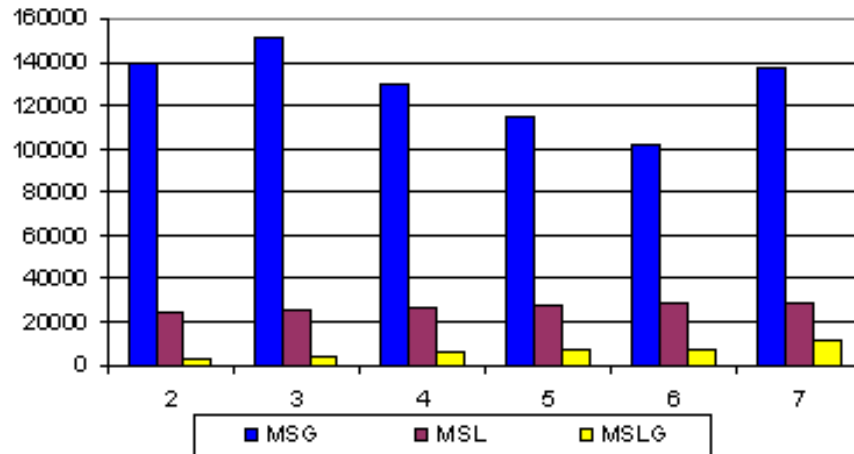


Figure 6.13 : Évolution du nombre total d'itérations en fonction de h respectivement pour le Mean-Shift global (MSG), la première étape du Mean-Shift local-global (MSL) et sa deuxième étape (MSLG)

rations. Mais une telle structure demande un espace mémoire considérable pour les images de document de très grande taille. La proposition du Mean-Shift local-global réduit la complexité, dans le cas idéal, à $O(N \times (R + M))$, où R est la taille du R -NNC et M le nombre de modes locaux extraits suite à la première étape. Bien évidemment, $(R + M) < \log N < N$ et par conséquent le Mean-Shift local-global nécessite des temps de calcul plus faibles comparés à un Mean-Shift global même optimisé et pour un résultat quasiment similaire.

6.5 Conclusions et perspectives

Pour élaborer une amélioration de l'algorithme du Mean-Shift global, nous avons proposé de réduire la complexité de l'algorithme original en limitant la recherche des points colorimétriquement proches aux points locaux situés dans un voisinage autour de chaque point de l'image. Le Mean-Shift local-global, proposé comme solution, est réalisé en deux étapes. La première étape consiste à appliquer un Mean-Shift sur chaque point de l'image par rapport aux R points les plus proches spatialement du point courant et colorimétriquement similaires. Cette première phase donne un nombre réduit M de modes locaux. Dans une deuxième étape, nous appliquons le Mean-Shift global aux M modes locaux trouvés précédemment par rapport à tous les points de l'image. L'existence des R points localement et colorimétriquement proches de x trouve sa justification avec la propriété de continuité des images de documents.

D'après la propriété de la continuité des couleurs dans les images de document, il doit presque toujours exister, localement autour d'un point x donné, des points colorimétriquement proches. Mais, il est possible que l'on ne trouve aucun point ou bien un nombre trop faible de points colorimétriquement proches de x pour estimer le gradient de la fonction de densité. Dans ce cas, le point x représente une couleur statistiquement trop peu fréquente dans l'image, ce qui est la caractéristique d'un bruit ponctuel dans l'image. On retrouve le même problème avec le Mean-Shift global, lorsque l'on ne trouve, dans l'image entière,

aucune couleur située à une distance colorimétrique inférieure à h par rapport au point x . Le gradient de densité ne peut pas alors être estimé et le vecteur de Mean-Shift est remplacé par le vecteur nul. Ces points particuliers de l'image ne se déplacent pas et forment donc le centre d'un nouveau mode dès la première itération. Ainsi, les couleurs aléatoires, provenant du bruit de l'image, forment de nouvelles classes couleurs caractérisées par des modes centrés sur un seul point de l'image dans l'espace couleur. Aucun auteur ne parle de ce problème de détermination du gradient de densité lorsqu'il n'est pas défini. En effet, les images naturelles modérément bruitées sont suffisamment riches en couleurs pour éviter le problème de l'indétermination du gradient de densité. Pour les images de documents, le problème est plus perceptible que pour les images naturelles. Les images de documents sont caractérisées par un nombre très faible de couleurs comme les couleurs du support papier et les couleurs des encres utilisées. Si l'image est suffisamment bruitée pour faire apparaître des couleurs marginales (comme par exemple avec la compression JPEG), alors elles seront toutes segmentées en nouvelles classes. Par conséquent, le Mean-Shift local-global ne devrait pas permettre de traiter des images fortement texturées ou moirées.

Le Mean-Shift local-global permet de séparer le recto du verso et de restaurer les dégradations du support papier pour une grande majorité des cas. Cette procédure est entièrement automatique et non-supervisée. Elle ne nécessite pas l'intervention de l'utilisateur. Cependant, la classification des pixels se fait sans tenir compte des informations locales des couleurs. Même si notre approche tient compte de l'information locale pour accélérer les calculs, la seconde étape correspond à un Mean-Shift global qui ne tient compte que des densités des classes de couleurs de l'image entière. Ainsi, la suppression des marques en transparence peut endommager les traits du texte original du recto si ces derniers interceptent les marques en transparence. De même, des pixels isolés peuvent être classés différemment de leurs voisins. L'irrégularité de la couleur de l'encre et la variation de la densité de l'encre le long d'un trait peuvent conduire à des ruptures de tracés ou des trous dans les traits. Nous allons nous focaliser dans la partie suivante à ce problème de restauration des caractères dégradés ou cassés. Nous faisons appel à des méthodes de diffusion.

Troisième partie

Dégradations des Formes : Restauration par Diffusion

Restauration des Formes par Diffusion

« Dans le bonheur d'autrui, je cherche mon bonheur. »

Pierre Corneille

Résumé

L'objectif de ce chapitre est d'étudier les méthodes de diffusion dans le cadre de la restauration des images de documents dégradés. Nous analysons les différents modèles de diffusion pour réduire le bruit, lisser les contours, restaurer la continuité des traits ou rehausser la netteté des formes. Nous évaluons visuellement chaque modèle sur des images de synthèse ainsi que sur des images réelles de documents. Nous présentons également une étude de faisabilité sur l'amélioration des performances des OCR à l'aide de prétraitements de l'image par des modèles de diffusion.

Sommaire

7.1	Introduction	84
7.2	Bruit : Lissage par EDP	84
7.3	Flou : Rehaussement	96
7.4	Expérimentation sur les images de documents	102
7.5	Conclusion	112

7.1 Introduction

L'application des Équations aux Dérivées Partielles (EDP) en traitement des images est un domaine de recherche récent qui permet de traiter un grand nombre de problèmes comme la restauration, la segmentation, le suivi d'objets, l'évolution des courbes planes, les contours déformables, l'estimation du flot optique, la stéréovision, la morphologie continue, etc.

Notre objectif étant la restauration, nous allons donc nous focaliser sur la réduction du bruit, le rehaussement des informations utiles, le renforcement de la continuité des traits et la reconstruction des parties manquantes dans l'image. Nous allons montrer que les traitements par EDP sont adaptés aux images de documents et à leurs contenus. En effet, ces traitements reposent principalement sur les dérivées de l'image qui sont beaucoup mieux définies dans les images de documents que dans les images naturelles : les images de documents sont composées majoritairement de traits autour desquels les dérivées peuvent être précisément calculées. Ainsi, les traitements par EDP pourront facilement réduire le bruit tout en préservant la netteté des contours des traits. Un autre avantage important des méthodes à base d'EDP est qu'elles ne traitent pas les images sous leur forme discrète mais qu'elles considèrent les images comme des fonctions continues et différentiables. Les traitements à base d'EDP sont par nature non-limités par les frontières des pixels et peuvent traiter des images avec de très faibles résolutions spatiales. Ces approches à base d'EDP sont aussi dites «variationnelles» car elles reposent sur les infimes variations de l'image et réalisent des traitements par petites corrections des valeurs d'intensités. Les traitements sont donc plus fins qu'avec des traitements classiques dans le domaine discret. Ces méthodes variationnelles nous semblent *a priori* très bien adaptées au traitement d'images, en faible ou moyenne résolution, de traits dont l'épaisseur est généralement de un à quelques pixels seulement.

Nous allons effectuer tout au long de ce chapitre une étude des méthodes types associées à la réduction du bruit et au rehaussement des détails des formes dégradées. La reconstruction des parties manquantes de l'image sera introduite dans le chapitre suivant. Pour chaque type de problème, nous expliquerons le traitement par EDP le plus adapté et le principe de son fonctionnement. Nous renvoyons le lecteur à [ALV-93, GUI-95, WEI-96, WEI-97A, WEI-98] pour une description plus détaillée d'un grand nombre d'approches par EDP.

Les définitions et les notations mathématiques nécessaires à la compréhension de cette partie sont données en annexe.

7.2 Bruit : Lissage par EDP

Le débruitage d'image est une tâche qui a suscité beaucoup d'intérêt dans le domaine du traitement d'image. La diffusion par EDP est une des solutions adéquates à ce type de problème. C'est historiquement une des premières applications des EDP pour résoudre le problème du filtrage du bruit dans une image. Nous distinguons fondamentalement deux types de diffusion :

- *La diffusion linéaire* : le traitement est appliqué de façon identique indépendamment

de la position (x, y) dans l'image.

- **La diffusion non-linéaire** : le traitement est appliqué différemment suivant la localisation (x, y) des points dans l'image. Généralement, le traitement est fonction des gradients de l'image et de sa magnitude qui indiquent la présence ou non d'un bruit ou du contour d'un objet.

Nous différencions aussi les diffusions isotropes des diffusions anisotropes :

- **La diffusion isotrope** : le traitement ne présente aucune direction privilégiée.
- **La diffusion anisotrope** : le traitement présente une orientation privilégiée et diffuse plus dans une direction que dans les autres directions.

Si les méthodes linéaires n'ont pas un réel intérêt, les méthodes non-linéaires et plus particulièrement les méthodes non-linéaires et anisotropes ouvrent la voie à de nombreux traitements possibles qui peuvent s'adapter aux contenus des images.

7.2.1 Diffusion linéaire isotrope

Si on considère le bruit comme l'addition d'un signal à haute fréquence et de faible amplitude, alors le filtrage du bruit consiste à minimiser les variations de l'image estimées par la magnitude du gradient $\|\nabla I\|$. Tikhonov [Tik-77] formalise le problème en proposant de trouver l'image I qui minimise la somme de toutes les variations de l'image pour tous les points (x, y) de Ω (Eq. 7.1).

$$I_{restaurée} = \underbrace{\text{Min}}_{I:\Omega \rightarrow \mathbb{R}} E(I) = \int_{\Omega} \|\nabla I\|^2 dx dy \quad (7.1)$$

En appliquant le schéma de résolution d'Euler Lagrange, l'équation de Tikhonov revient à résoudre l'EDP Eq.7.2, connue sous le nom de l'équation de la chaleur.

$$\begin{cases} I_t = \Delta I = \text{div}(\nabla I) \\ I_0 = I_{bruitée} \end{cases} \quad (7.2)$$

Koenderink [KOE-84] est le premier à remarquer que l'évolution de cette équation est équivalente à un lissage gaussien de variance $\sigma = \sqrt{2t}$ (Eq.7.3).

$$\frac{\partial I}{\partial t} = \text{div}(\nabla I) \Leftrightarrow I_t = I_0 \otimes G_{\sigma=\sqrt{2t}} \quad (7.3)$$

Nous pouvons observer le résultat de cette évolution au fil des itérations sur la figure 7.1. L'équation de la chaleur réduit progressivement toutes les variations de l'image au cours du temps t ce qui correspond également à un changement d'échelle σ de l'image, c'est-à-dire à un lissage par un noyau Gaussien d'écart-type σ . L'équation de la chaleur est une opération linéaire dans le sens où le traitement s'effectue indistinctement sur tous les pixels de l'image. Elle minimise progressivement toutes les variations de l'image en commençant par les variations de faibles intensités correspondant au bruit dans l'image pour finir de réduire aussi les variations issues des contours des traits, ce qui rend l'image plus floue. Cette EDP ne privilégie aucune direction en particulier et la diffusion s'effectue de façon identique dans toutes les directions et donc de façon isotrope.

L'équation originale de Tikhonov ajoute un terme supplémentaire sur la différence entre l'image restaurée I et l'image initiale I_0 afin de trouver une solution pas trop différente de l'image d'origine (Eq. 7.4).

$$I_{restaurée} = \underset{I:\Omega \rightarrow \mathbb{R}}{\text{Min}} E(I) = \int_{\Omega} \|\nabla I\|^2 + \frac{\lambda}{2} \|I_0 - I\|^2 dx dy \quad (7.4)$$

La résolution de cette équation par Euler revient à l'EDP 7.5 avec un terme supplémentaire $|I - I_0|$ aussi appelé «attachement aux données» et qui permet de converger vers une solution pas trop éloignée de l'image initiale I_0 .

$$\begin{cases} I_t = \text{div}(\nabla I) + \lambda |I - I_0| \\ I_0 = I_{bruitée} \end{cases} \quad (7.5)$$

Malgré le terme d'attachement aux données qui stoppe la diffusion avant son terme, la diffusion linéaire isotrope ne fait pas la distinction entre les variations de l'image dues au bruit des variations d'intensités des contours des traits qu'il faut préserver. Les approches de diffusion non-linéaires et anisotropes permettent de pallier à cet inconvénient inhérent à la diffusion isotrope et donc de préserver les contours et les structures. Leur principe consiste à adapter la diffusion aux contenus de l'image de manière locale.



Figure 7.1 : De gauche à droite respectivement : l'image bruitée, les images traitées par un filtrage isotrope linéaire (200 itérations) sans et avec le terme d'attache aux données ($\lambda = 0.2$)

7.2.2 Diffusion non-linéaire

La motivation essentielle des modèles basés sur la diffusion non-linéaire est la construction d'un opérateur de diffusion dépendant des propriétés locales de l'image. Le principal objectif consiste à réduire le bruit tout en préservant les contours des traits. Les EDP restent basées sur l'équation de la chaleur et la différence entre les EDP réside dans l'introduction d'un coefficient de diffusivité C calculé localement en chaque point (x, y) de Ω dans le terme de la divergence. On distingue deux types d'opérateurs de diffusivité :

- **Un coefficient de diffusivité de type scalaire** $I_t = \text{div}(c \times \nabla I)$, avec c un scalaire, généralement fonction de la magnitude du gradient et qui adapte la diffusion aux propriétés locales du signal en diminuant ou en augmentant localement la diffusion.
- **Un coefficient de diffusivité de type matricielle** $I_t = \text{div}(D \times \nabla I)$, avec D une matrice qui représente la structure locale de l'image et impose une direction à la diffusion en chaque point (x, y) de Ω .

a. Diffusion scalaire

Les premiers travaux qui se situent dans ce contexte sont ceux de Perona et Malik [PER-90, PER-94]. Ils ont proposé d'effectuer une diffusion adaptative dont le comportement est fonction de la norme du gradient (Eq.7.6). Ils introduisent une fonction qui dépend de la norme du gradient dans l'équation de la chaleur afin de contrôler la diffusion. Cette fonction dite de diffusivité $C(\|\nabla I\|)$ doit prendre des valeurs entre $[0,1]$. Si le coefficient de diffusivité C est nul, la diffusion sera alors stoppée. Inversement, lorsque C est égale à «1», on retrouve une diffusion isotrope par l'équation de la chaleur.

$$\begin{cases} I_t = \operatorname{div}(C(\|\nabla I\|) \times \nabla I) \\ I_0 = I_{\text{bruitée}} \end{cases} \quad (7.6)$$

En prenant une fonction C positive et décroissante en fonction de la magnitude du gradient, avec $C(0) = 1$ et $\lim_{v \rightarrow \infty} C(v) = 0$, on atténue la diffusion dans les régions à fort gradient (lissage faible voire absent pour préserver les contours) et on la maintient dans les zones à faible gradient (lissage fort sur les zones homogènes mais bruitées). Perona et Malik proposent les fonctions suivantes :

$$C(\|\nabla I\|) = \frac{1}{1 + \left(\frac{\|\nabla I\|}{k}\right)^2} \quad \text{ou} \quad C(\|\nabla I\|) = \exp\left(-\frac{\|\nabla I\|}{k}\right)^2 \quad (7.7)$$

La valeur du paramètre k permet de distinguer les zones à faible gradient de celles à fort gradient. Les zones de l'image où $\|\nabla I\| < k$ sont assimilées à des régions uniformes où le bruit doit être réduit. La fonction de diffusivité $C(\|\nabla I\|)$ prend alors des valeurs élevées proche de «1» et l'équation 7.6 est identique à l'équation de la chaleur (Eq.7.2) qui réalise une forte diffusion dans les zones éloignées des contours. Inversement, les parties de l'image où $\|\nabla I\| \geq k$ sont considérées comme étant des variations d'intensités dues aux contours des traits qu'il faut préserver. La fonction de diffusivité $C(\|\nabla I\|)$ prend alors des valeurs proches de «0» bloquant ainsi la diffusion pour préserver les variations d'intensité dues aux contours. La figure 7.2 montre, sur une image de document imprimé, les différences entre le coefficient de diffusivité constant de l'équation de la chaleur (Eq.7.2) et le coefficient de diffusivité variable dans le filtrage non-linéaire de Perona et Malik (Eq.7.6). On constate que le filtrage non-linéaire réduit la diffusion lorsque l'on s'approche des contours ce qui explique leurs préservations.

Afin de mieux comprendre le comportement diffusif de ce modèle, nous allons le réécrire sous la forme des dérivées secondes dans la direction parallèle ou orthogonale au gradient [CHA-97, KOR-97A] :

$$I_t = \operatorname{div}(C(\|\nabla I\|) \times \nabla I) \Leftrightarrow I_t = c_\xi I_{\xi\xi} + c_\eta I_{\eta\eta} \quad (7.8)$$

avec les coefficients de diffusivités $c_\xi = C(\|\nabla I\|)$ et $c_\eta = (C(\|\nabla I\|) \times \|\nabla I\|)'$. Lorsque $c_\xi > c_\eta$, la diffusion devient anisotropique et diffuse plus largement dans la direction des isophotes ξ que dans celle du gradient η , préservant ainsi les contours des traits.

La figure 7.3 montre que lorsque le coefficient de diffusion dépend de la magnitude des gradients de l'image originale I_0 ($I_t = \operatorname{div}(C(\|\nabla I_0\|) \times \nabla I)$), le bruit de niveau supérieur

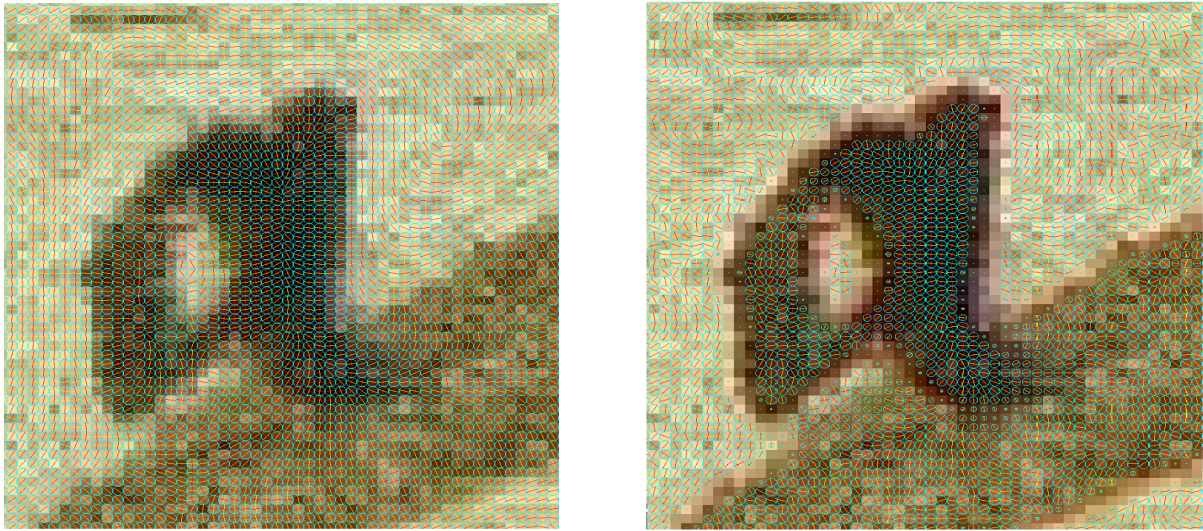


Figure 7.2 : Différence entre, respectivement de gauche à droite, une diffusion linéaire isotrope et une diffusion non-linéaire scalaire

$\|\nabla I_0\|$ sera conservé. En revanche, lorsque le coefficient de diffusion dépend de la magnitude des gradients de l'image en cours de diffusion ($I_t = \text{div}(C(\|\nabla I\|) \times \nabla I)$), le résultat final montre une image où les objets sont presque segmentés lorsque $\|\nabla I\| \geq k$ et les zones de l'image où $\|\nabla I\| < k$ seront entièrement uniformisées en intensité.



Figure 7.3 : Résultat de Perona-Malik sur l'image de la figure 7.1 pour $k=40$: de gauche à droite respectivement ; l'image bruitée, résultat d'une diffusion fonction des gradients de l'image originale et des gradients de l'image courante

Cette démarche introduite par Perona et Malik a fait l'objet de plusieurs études critiques [WEI-97B, CAT-92]. Elle présente encore des lacunes telles que son inefficacité dans les zones où le bruit présente de grosses discontinuités et même l'absence de solutions dans certains cas. De nombreuses variantes ont ensuite été proposées à partir de cette idée. Les premières ont pour objectif d'améliorer le modèle de Perona-Malik qui pose des problèmes théoriques de stabilité et n'admet pas de solution régulière. Catté [CAT-92] propose d'appliquer un filtrage Gaussien sur les gradients lors du calcul du coefficient de diffusion pour rendre le problème mieux posé (Eq.7.9).

$$\begin{cases} I_t = \text{div}(C(\|G_\sigma \otimes \nabla I\|) \times \nabla I) \\ I_0 = I_{\text{bruitée}} \end{cases} \quad (7.9)$$

La figure 7.4 montre que les gradients ∇I d'une image possèdent des directions désordonnées car ils sont calculés localement, chacun de façon indépendante. En appliquant un filtrage Gaussien sur l'image des gradients, on renforce la cohérence des directions de chaque vecteur en tenant compte des vecteurs voisins. Le champ de gradient régularisé $G_\sigma \otimes \nabla I$

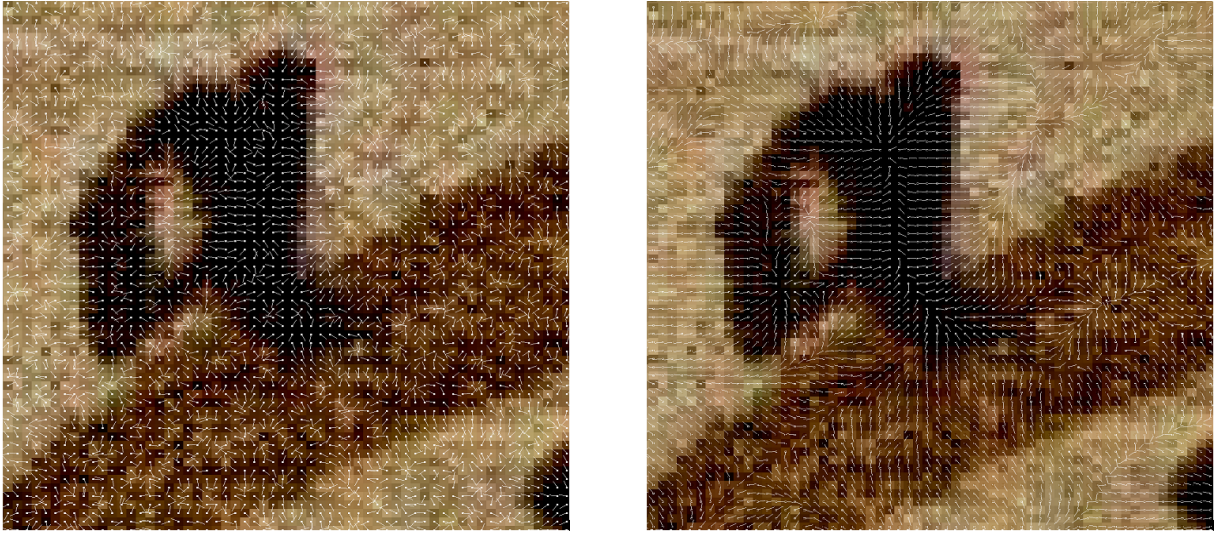


Figure 7.4 : Comparaison entre le champ de gradient ∇I et le champ régularisé $G_\sigma \otimes \nabla I$

montre des vecteurs correctement orientés le long des contours (autour des fortes variations) comme dans les zones uniformes de l'image avec de faibles variations.

Contrairement à la figure 7.3, l'image de synthèse de la figure 7.5 est déformée par un bruit poivre et sel. L'image n'est pas suffisamment régulière pour que le filtrage de Perona et Malik puisse marcher et l'image n'a pratiquement pas évolué après 400 itérations. Le filtrage des gradients dans le calcul du coefficient de diffusivité rend le problème mieux posé et le filtrage redevient opérationnel. En calculant le coefficient de diffusivité sur l'image en cours de filtrage avec le modèle de Catté, le filtrage devient très efficace et les formes sont beaucoup mieux restaurées.



Figure 7.5 : De gauche à droite respectivement : La diffusion de Perona-Malik ($k = 40$), Perona-Malik-Catté ($k = 20$) et l'image traitée par une diffusion Perona-Malik-Catté fonction de l'image originale ($k = 20$)

b. Diffusion matricielle

La diffusion non-linéaire matricielle utilise une matrice de diffusion D dans le terme de la divergence de l'équation de la chaleur [WEI-94, WEI-98].

$$\begin{cases} I_t = \operatorname{div}(D\nabla I) \\ I_0 = I_{\text{bruitée}} \end{cases} \quad (7.10)$$

Si la matrice 2×2 de diffusion D est symétrique et définie positive, alors elle représente l'équation d'une ellipse qui guide la diffusion en lui imposant une direction privilégiée.

Contrairement à la diffusion scalaire qui adapte le niveau de diffusion, la diffusion matricielle impose une orientation différente suivant le contenu de l'image. Une représentation du comportement de cette forme de diffusion matricielle est donnée dans la figure 7.6. Nous pouvons noter que cette forme de diffusion se comporte de la même manière qu'une diffusion isotropique sur une zone homogène.

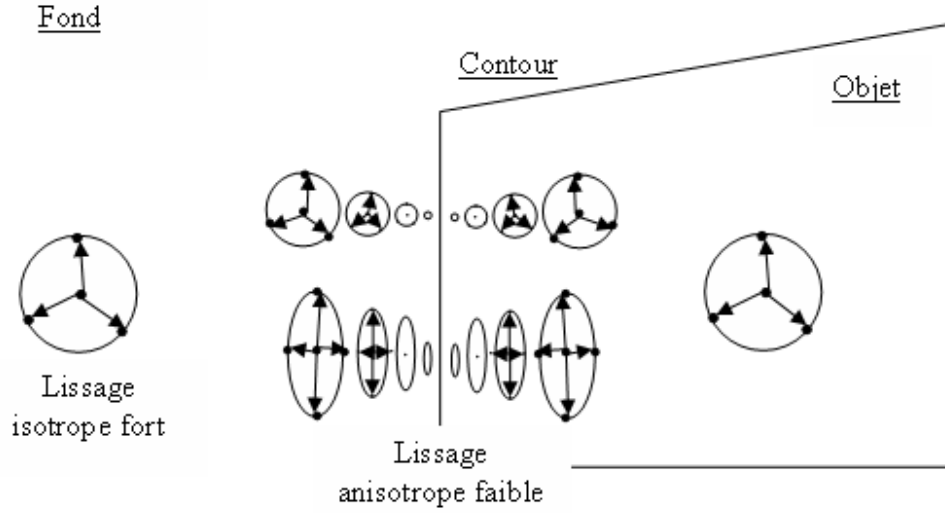


Figure 7.6 : Comportement d'une diffusion anisotrope à proximité des contours et isotrope partout ailleurs

Si $D = C(\|\nabla I\|) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, on retrouve une EDP non-linéaire de diffusion scalaire de type Perona et Malik.

Kœndring a montré que la diffusion matricielle fait intervenir une partie anisotropique et une partie isotropique. Pour séparer les deux parties, il suffit de réécrire la matrice de diffusion $D = \alpha G + \beta Id$ comme une combinaison linéaire $\alpha \beta$ d'une matrice G représentant la forme anisotropique et la matrice identité Id correspondant à la partie isotrope :

$$\begin{aligned} I_t &= \text{div}(D\nabla I) = \text{div}((\alpha G + \beta Id)\nabla I) = \alpha \text{div}(G\nabla I) + \beta \text{div}(\nabla I) \\ &= \alpha(c_\eta I_{\eta\eta} + c_\xi I_{\xi\xi}) + \beta(c_{isotrope}\Delta I) \end{aligned} \quad (7.11)$$

On retrouve un filtrage anisotropique sous la forme d'une combinaison linéaire des dérivées secondes $c_\eta I_{\eta\eta} + c_\xi I_{\xi\xi}$ avec les coefficients de diffusivité c_η dans la direction du gradient η et c_ξ dans la direction de la tangente aux isophotes ξ , auquel s'ajoute une diffusion isotrope classique identique à celle de l'équation de la chaleur $c_{isotrope}\Delta I$.

Weickert propose d'utiliser le tenseur T des dérivées premières (Eq.A.16) pour construire une matrice de diffusion D .

$$T = \nabla I \times \nabla I^t = \begin{pmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{pmatrix} \quad (7.12)$$

Nous savons que le tenseur peut se décomposer dans une base orthonormale formée par les vecteurs propres θ_+ et θ_- associés respectivement aux valeurs propres λ_+ et λ_- .

$$T = \lambda_- \times \Theta_- \Theta_-^t + \lambda_+ \times \Theta_+ \Theta_+^t \quad (7.13)$$

Le vecteur propre Θ_+ indique l'orientation principale de la structure locale des traits et les valeurs propres λ_+ et λ_- indiquent la forme de la structure (isotrope ou anisotrope) dans les directions principales. La matrice de diffusion D utilise les mêmes vecteurs propres Θ_+ et Θ_- du tenseur T . Mais Weickert utilise de nouvelles valeurs $\mu_{+/-} = f_{+/-}(\lambda_+, \lambda_-)$ en fonction des valeurs propres λ_+ et λ_- pour modifier à volonté le comportement de la diffusion le long des axes Θ_+ et Θ_- [WEI-01] :

$$D = f_+(\lambda_+, \lambda_-) \times \Theta_- \Theta_-^t + f_-(\lambda_+, \lambda_-) \times \Theta_+ \Theta_+^t \quad (7.14)$$

Ainsi les axes $\Theta_{+/-}$ indiquent la direction de la diffusion et les nouvelles valeurs $\mu_{+/-} = f_{+/-}(\lambda_+, \lambda_-)$ indiquent le comportement isotropique ou anisotropique du filtrage en chaque point de l'image. La figure 7.7 montre que la matrice de diffusivité possède la même orientation que celle du tenseur et montre que les coefficients de diffusion changent suivant une fonction $f_{+/-}$ le long des axes, modulant ainsi la diffusion aux contenus de l'image.

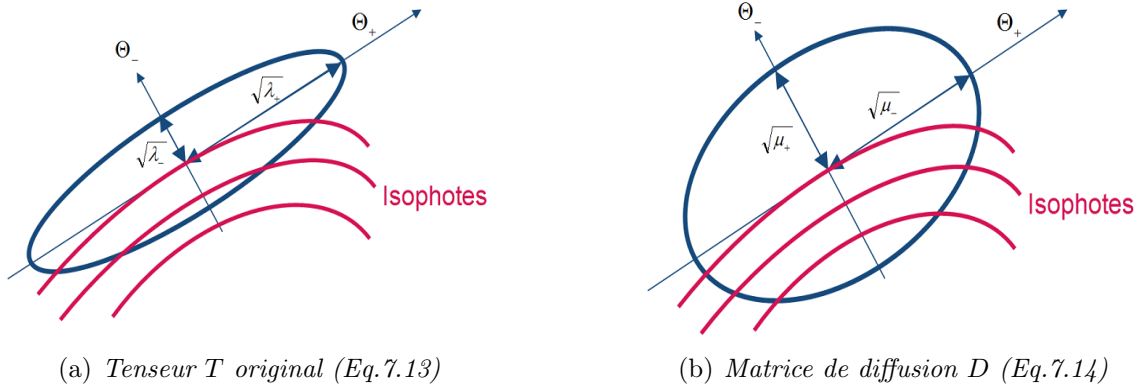


Figure 7.7 : Relation entre le tenseur original T et la matrice de diffusion D construite

En pratique, les gradients doivent être calculés sur une image légèrement lissée pour éviter les points singuliers de l'image. Comme dans la régularisation proposée par Catté [CAT-92], il faut lisser l'image avec une Gaussienne d'écart-type σ pour calculer le tenseur T (Eq. 7.15).

$$T_\sigma = \nabla (G_\sigma \otimes I) \nabla (G_\sigma \otimes I)^t \quad (7.15)$$

De plus, Weickert propose de renforcer la cohérence locale des directions principales Θ_+ des tenseurs en tenant compte des orientations des tenseurs voisins. Pour cela, il propose de convoluer la carte des tenseurs par un noyau Gaussien d'écart-type ρ . Pour renforcer la direction générale des orientations de l'image dans une image particulièrement bruitée, Weickert propose de définir ρ par $\rho = 3 \times \sigma$, mais le choix des paramètres σ et ρ dépend de la nature de l'image, du niveau du bruit et de la perte de la continuité des traits.

$$T_\sigma^\rho = G_\rho \otimes (\nabla (G_\sigma \otimes I) \nabla (G_\sigma \otimes I)^t) \quad (7.16)$$

La figure 7.8 affiche la matrice de diffusivité $D = 0.25 \times \Theta_- \Theta_-^t + 1 \times \Theta_+ \Theta_+^t$ avec $\Theta_{+/-}$ les vecteurs propres du tenseur T_σ^ρ suivant les paramètres σ et ρ . Pour $\rho < 2$, les ellipses montrent

des orientations différentes qui suivent le bruit et les ruptures des traits car elles sont calculées de façon indépendante. À partir de $\rho > 3$, le lissage Gaussien donne une meilleure cohérence dans les orientations des ellipses qui suivent la direction des traits indépendamment des ruptures des tracés. La régularisation de la carte des tenseurs permet aux matrices d'être toujours orientées dans la direction majoritaire des traits et de diffuser de façon à réduire les discontinuités.

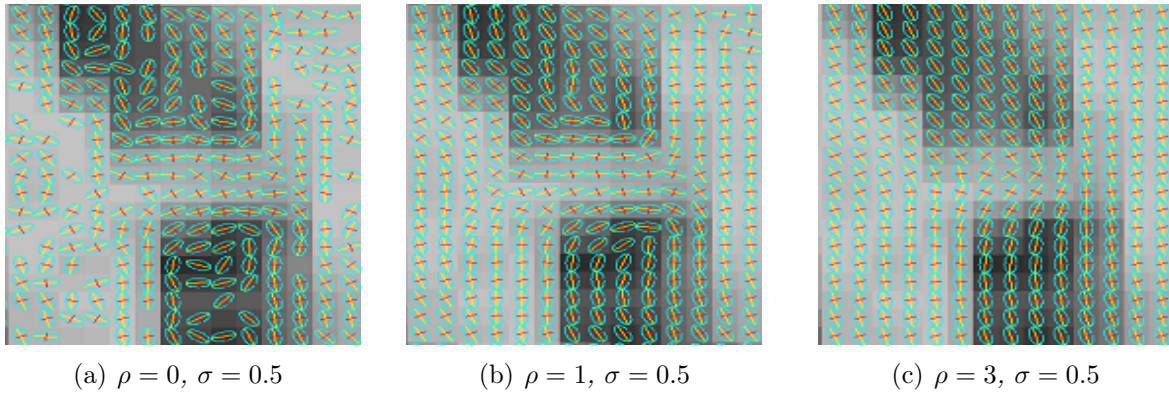


Figure 7.8 : Matrice de diffusivité $D = 0.25 \times \Theta_- \Theta_-^t + 1 \times \Theta_+ \Theta_+^t$ pour T_σ^ρ suivant (ρ, σ)

Parmi les méthodes définies dans la littérature et appartenant à la catégorie des approches de diffusion matricielle, nous avons noté l'approche de Weickert [WEI-96, WEI-97B, WEI-98], et l'approche de R. Kimmel et al. [SOC-96, SOC-98, KIM-00] connue sous le nom du flux de Beltrami ou Beltrami Flow. Le principe est le même mais les valeurs de $\mu_{+/-}$ diffèrent d'une approche à l'autre (Tab.7.1). De façon synthétique, les EDP à diffusion matricielle s'écrivent donc toutes sous la forme suivante :

$$\begin{cases} I_t = \operatorname{div}(D\nabla I) \\ I_0 = I_{\text{bruitée}} \end{cases} \text{ avec } \begin{cases} D = \mu_+ \times \Theta_- \Theta_-^t + \mu_- \times \Theta_+ \Theta_+^t \\ \Theta_{+/-}, \lambda_{+/-} \text{ vecteurs et valeurs propres de } T_\sigma^\rho \end{cases}$$

La figure 7.9 montre, sur l'image d'une lettre, les matrices de diffusivité D , associées à chaque auteur et ayant la forme d'une ellipse orientée par Θ_+ et de longueur d'axe $\sqrt{\mu_{+/-}}$. Comme les matrices de diffusion changent au cours du temps et que chaque modèle transforme les images différemment, ces cartes ont été donc calculées lors de la première itération (pour t proche de 0) pour les rendre comparables. En revanche, pour la première itération, le bruit est encore trop présent et les variations de l'image empêchent de voir clairement la forme des matrices de diffusion. C'est pour ces raisons que nous avons calculé ces cartes sur une image légèrement diffusée par Perona-Malik et les ellipses ont été ensuite superposées à l'image d'origine. Ce mode d'affichage permet d'éviter l'influence de la diffusion de chaque modèle sur la forme des matrices de diffusivité D . Les différentes valeurs de $\sqrt{\mu_{+/-}}$ modifient considérablement le comportement de la diffusion au voisinage des contours des traits. Avec $\alpha = 0.001$, les ellipses de diffusion de Weickert sont très orientées le long de l'axe principal et le rayon peut atteindre 1 le long de l'axe principal alors qu'il reste constant et égal à α sur l'axe secondaire. La diffusivité de Weickert est très forte dans la direction de l'axe principal du tenseur pour renforcer la structure des traits. Avec un lissage des tenseurs T_σ^ρ ($\rho = 3$,

Perona-Malik	$\mu_+(\lambda_+, \lambda_-) = \mu_-(\lambda_+, \lambda_-) = \exp - \left(\frac{\ G_\sigma \otimes \nabla I\ }{K} \right)^2 = \exp - \left(\frac{\lambda_+ + \lambda_-}{K} \right)^2$
Weickert	$\begin{cases} \mu_+(\lambda_+, \lambda_-) = \alpha + (1 - \alpha) \exp \left(\frac{-C}{(\lambda_+ - \lambda_-)^2} \right) & \text{si } \lambda_+ \neq \lambda_- \\ = \alpha & \text{sinon} \\ \mu_- = \alpha = 0.001 \end{cases}$
Beltrami	$\begin{cases} \mu_+ = \sqrt{\frac{1+\lambda_+}{1+\lambda_-}} \\ \mu_- = \sqrt{\frac{1+\lambda_-}{1+\lambda_+}} \end{cases} \quad I_t = \frac{1}{\sqrt{(1+\lambda_-)(1+\lambda_+)}} \operatorname{div}(D\nabla I)$
Tschumperlé	$\begin{cases} \mu_+(\lambda_+, \lambda_-) = \frac{1}{\sqrt{1+\lambda_++\lambda_-}} \\ \mu_-(\lambda_+, \lambda_-) = \frac{1}{1+\lambda_++\lambda_-} \end{cases}$

Tableau 7.1 : Affichage des valeurs de $\mu_{+/-}$ pour respectivement Perona-Malik, Weickert, Beltrami et Tschumperlé

$\sigma = 1$), on assiste à un effet de «fibre» qui explique la déformation de la structure des formes complexes comme celles des caractères. Les matrices de diffusivité de Tschumperlé comme celles de Beltrami sont isotropes dans les zones de l'image contenant de faibles variations et deviennent anisotropes à l'approche d'un contour ou d'une forte variation des intensités de l'image. Les matrices de diffusivité de Tschumperlé deviennent nulles à l'approche d'une forte variation lorsque λ_+ et λ_- atteignent de grandes valeurs alors que les matrices de diffusivité de Beltrami possèdent toujours un rayon non-nul.

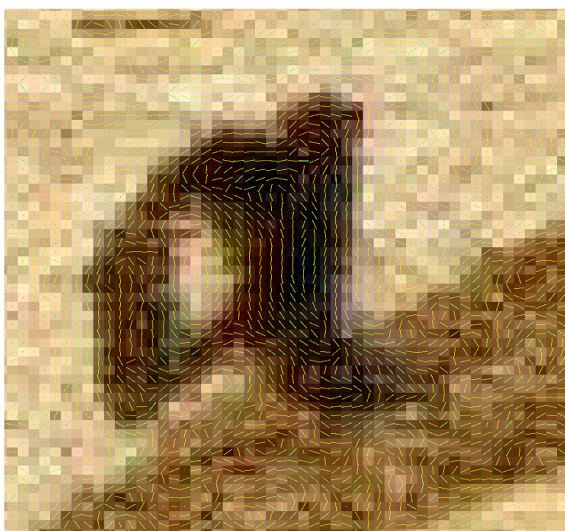
D. Tschumperlé introduit une formulation différente de la diffusion matricielle en substituant la divergence du gradient par la trace de la Hessienne plus stable numériquement [TSC-01, TSC-02A, TSC-02B, TSC-05].

$$\begin{cases} I_t = \operatorname{trace}(DH) \\ I_0 = I_{\text{bruitée}} \end{cases} \quad \text{avec} \quad \begin{cases} D = \frac{1}{\sqrt{1+\lambda_++\lambda_-}} \times \Theta_- \Theta_-^t + \frac{1}{1+\lambda_++\lambda_-} \times \Theta_+ \Theta_+^t \\ \Theta_{+/-}, \lambda_{+/-} \text{ vecteurs et valeurs propres de } I_\sigma^t \end{cases}$$

[TSC-02A] montre que la différence entre les EDP basées sur la divergence du gradient $I_t = \operatorname{div}(D\nabla I)$ et les EDP basées sur la trace de la Hessienne $I_t = \operatorname{trace}(DH)$ est la suivante :

$$\operatorname{div}(D\nabla I) = \operatorname{trace}(DH) + \nabla I^t \cdot \operatorname{div}(D) \quad (7.17)$$

En proposant une équation basée sur la trace de la Hessienne, D. Tschumperlé rend

(a) *Résultat de Weickert $\rho = 0, \sigma = 0$* (b) *Résultat de Weickert $\rho = 3, \sigma = 1$* (c) *Résultat de Tschumperlé*(d) *Résultat de Beltrami***Figure 7.9** : *Matrices de diffusivité D suivant les auteurs*

l'interprétation plus facile en développant les EDP directement sous la forme de combinaison de dérivées secondes $I_t = c_u I_{uu} + c_v I_{vv}$.

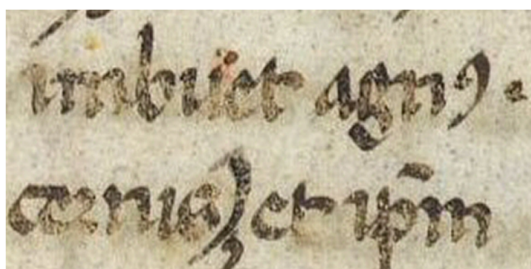
Le comportement numérique des EDP de D. Tschumperlé semble aussi plus stable que celles qui sont basées sur la divergence. La figure 7.10 montre les résultats des différentes méthodes de diffusion matricielle. Les méthodes de diffusion matricielle ont tendance à arrondir les angles mais, en contrepartie, elles lissent les contours.

Après plusieurs essais sur des documents différents contenant des dégradations variées (Fig.7.11), nous avons constaté que la diffusion matricielle de Tschumperlé et la diffusion scalaire de Perona-Malik modifiée par Catté donnaient de très bons résultats pour atténuer le bruit dans les images de documents sans dégrader les contours des traits. En revanche,

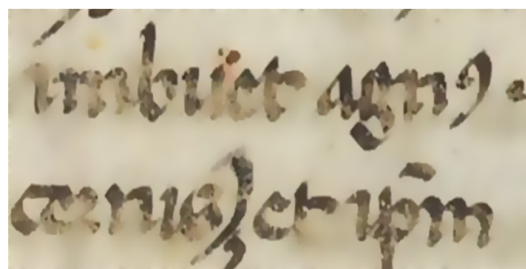


Figure 7.10 : Comparaison entre les différentes méthodes de diffusion non-linéaires anisotropiques et matricielles sur l'image de la figure 7.1

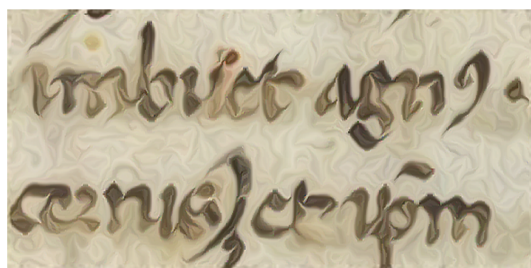
lorsque le bruit a provoqué un certain nombre de ruptures dans la continuité des traits, la diffusion matricielle de Beltrami ou de Tschumperlé avec des valeurs ρ élevées de T_ρ^p permet de renforcer la continuité des caractères.



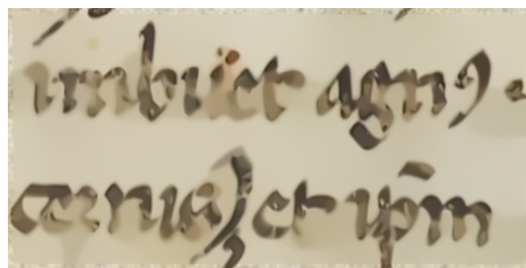
(a) Image originale



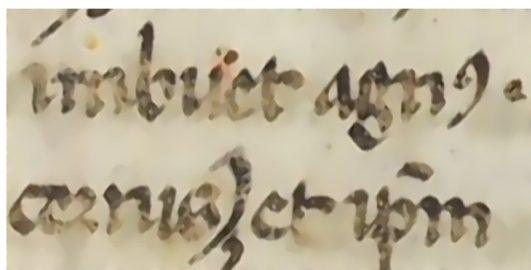
(b) Résultat de Perona-Malik-Catté $k = 10$



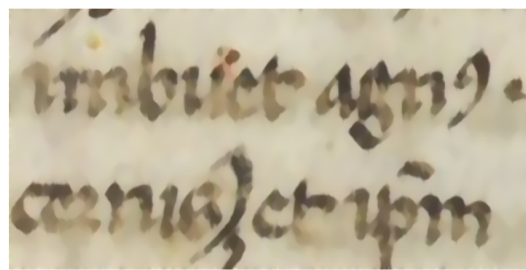
(c) Résultat de Weickert ($\rho = 1.5, \sigma = 0.5$)



(d) Résultat de Beltrami ($\rho = 1.5, \sigma = 0.5$)



(e) Tschumperlé $\rho = 0$



(f) Résultat de Tschumperlé $\rho = 1.0$

Figure 7.11 : Comparaison des méthodes de diffusion sur une image réelle

7.3 Flou : Rehaussement

Le rehaussement consiste à renforcer les variations utiles des contours des traits sans renforcer le bruit de fond de l'image. Visuellement, l'image devient plus nette et les détails de l'image apparaissent plus clairement au lecteur.

7.3.1 Rehaussement sous l'angle d'une opération inverse de celle de la diffusion

Dans l'équation 7.18, si $+F$ est le terme de régularisation pour réduire les variations de l'image alors l'EDP inverse $-F$ rehausse les variations. Par conséquent, pour rehausser les détails de l'image et corriger le flou, il faut appliquer les équations inverses des EDP utilisées pour la diffusion.

$$\begin{cases} I_t = +F(I, \nabla I, \dots) \\ I_0 = I_{bruitée} \end{cases} \quad \begin{cases} I_t = -F(I, \nabla I, \dots) \\ I_0 = I_{floue} \end{cases} \quad (7.18)$$

Par exemple, l'équation inverse de la chaleur renforce toutes les variations, celles issues des contours comme celles dues aux bruits de l'image (Eq. 7.19).

$$\begin{cases} I_t = -\Delta I = -div(\nabla I) \\ I_0 = I_{floue} \end{cases} \quad (7.19)$$

En 1960, Gabor avait déjà proposé d'atténuer le bruit d'une image en réalisant une soustraction avec son Laplacien. Sans le savoir, il avait trouvé la forme inverse de l'équation de la chaleur bien avant que les bases théoriques des EDP en traitement des images aient été fondées. Cependant, les opérations inverses des filtres non-linéaires de diffusion scalaire ou matricielle vont plutôt renforcer les variations de l'image issues du bruit de l'image que les variations des contours. Pour renforcer exclusivement les contours sans rehausser le bruit de l'image, il faut appliquer des filtres spéciaux aussi appelé «filtres de choc».

7.3.2 Renforcement des contours : Filtre de choc

Les filtres de choc sont des équations basées sur les EDP morphologiques que nous allons brièvement introduire; ces définissent la morphologie continue par opposition à la morphologie discrète.

a. EDP Morphologiques

La morphologie mathématique introduite par Serra est le champ d'application de la théorie des ensembles aux images discrètes, elle permet de réaliser des opérations ensemblistes pour comparer les formes. Ces opérations peuvent être reproduites par des EDP de la forme suivante :

$$I_t = c \|\nabla I\| \quad (7.20)$$

Elles permettent les opérations morphologiques de base comme l'érosion et la dilatation comme suit :

$$\begin{aligned} \text{Dilatation} & : I_t = \|\nabla I\| \\ \text{Érosion} & : I_t = -\|\nabla I\| \end{aligned}$$

Lorsque l'on discrétise les EDP morphologiques, on comprend mieux les opérations morphologiques d'érosion et de dilatation à partir d'une image initiale I^0 .

$$\begin{aligned} \text{Dilatation} & : I^{n+1}(x, y) = I^n(x, y) + dt \|\nabla I^n(x, y)\| \\ \text{Érosion} & : I^{n+1}(x, y) = I^n(x, y) - dt \|\nabla I^n(x, y)\| \end{aligned}$$

Ainsi, pour respectivement éroder ou dilater une image, il suffit à chaque itération d'enlever ou d'ajouter à l'image une fraction dt des intensités des contours de l'image. Le principe de la morphologie continue repose sur le fait que la frontière (ou contour) de la forme transformée est le résultat d'un processus de déformation du contour de la forme originale. Or, cette transformation d'une courbe à une autre est le résultat d'une séquence des déformations infinitésimales qui ne dépendent que localement de la forme originale, à condition d'utiliser un élément structurant convexe. La morphologie continue basée sur les EDP présente de nombreux avantages par rapport à la morphologie discrète :

- **Meilleure modélisation mathématique.** La morphologie continue n'a pas à tenir compte de toutes les difficultés du passage d'un domaine continu à un domaine discret.
- **Précision sous-pixélique.** Nous obtenons la possibilité d'utiliser les transformations de distance les plus précises avec une précision qui n'est pas limitée aux pixels de l'image. La morphologie continue permet d'effectuer des opérations avec des éléments structurants de taille λ réelle et non-entière. Ainsi, il est possible de réaliser de nombreuses érosions d'objets d'épaisseur de 1 à 2 pixels avec un élément structurant de taille λ infinitésimale, opération qui est impossible en morphologie discrète. La morphologie continue est donc adaptée au traitement des images de documents en faible résolution. La figure 7.16 montre la différence entre une érosion discrète avec le plus petit rayon possible que l'on puisse obtenir en morphologie discrète, c'est-à-dire $\lambda = 1$ et une érosion par EDP avec un rayon très faible $\lambda = 0.2$.
- **Indépendance à la géométrie de la grille discrète.** La morphologie continue permet d'utiliser des éléments structurants dont la forme ne dépend ni de la géométrie ni de la finesse de la grille de discrétisation. Nous ne sommes plus obligés de faire un choix entre une géométrie discrète carrée ou hexagonale.
- **Simplification des traitements morphologiques non-linéaires.** Les EDP morphologiques peuvent facilement réaliser des traitements non-linéaires basés sur les informations locales de l'image et ainsi créer des opérateurs morphologiques déformables. Une EDP morphologique déformable se présente sous la forme suivante (Eq.7.21) :

$$I_t = C(I) \|\nabla I\| \quad (7.21)$$

Ces EDP déformables modifient localement la forme et le rayon de l'élément structurant suivant une fonction de diffusivité $C(I)$ qui est calculée localement en chaque

point (x, y) de Ω . La morphologie déformable peut sélectionner l'érosion et la dilatation suivant le signe de $C(I)$ et adapter la géométrie et le rayon de l'élément structurant suivant les valeurs de $C(I)$ pour chaque point de l'image. Les EDP morphologiques non-linéaires sont à la base des filtres de choc.



Figure 7.12 : De gauche à droite respectivement : l'image originale, le résultat d'une érosion continue par EDP de rayon $\lambda = 0.2$ et le résultat d'une érosion discrète de rayon $\lambda = 1$ avec un masque 3×3

b. Filtres de choc

À la différence des filtres de rehaussement classique qui accentuent indistinctement toutes les variations de l'image, y compris le bruit, les filtres de choc renforcent uniquement les contours des traits; leur emploi va donc réduire les effets de flou autour des traits des caractères.

Osher et Rudin [OSH-88B] ont été les premiers à introduire dans le domaine du traitement d'images les filtres de choc jusque là utilisés en mécanique des fluides. Ces filtres sont des opérateurs dont l'application permet d'accentuer les transitions en augmentant localement leur amplitude. Les chocs sont plus particulièrement développés sur les points d'inflexion (passage par zéro des dérivées secondes) et les extrema locaux restent inchangés dans le temps (Fig.7.13). Aucun extremum local erroné n'est créé. Ce schéma préserve donc la totale variation.

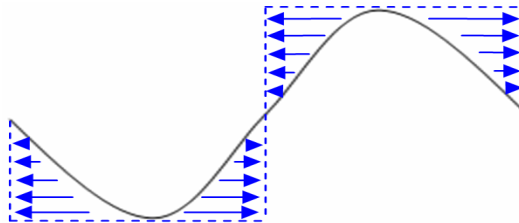


Figure 7.13 : Principe de rehaussement par filtre de choc dans le cas 1-D

L'équation proposée par Rudin (Eq. 7.22) montre que c'est bien une EDP morphologique [JAC-96] qui dilate ou érode les contours des objets suivant le signe du Laplacien de l'image.

$$\text{Filtre de choc de Rudin} : \begin{cases} I_t = -\text{sign}(\Delta I) \|\nabla I\| \\ I_0 = I_{\text{floue}} \end{cases} \quad (7.22)$$

La figure 7.14 explique avec un signal 1D le fonctionnement du filtre ce choc de Rudin. Les variations de l'image autour des traits sont très progressives et donnent une impression

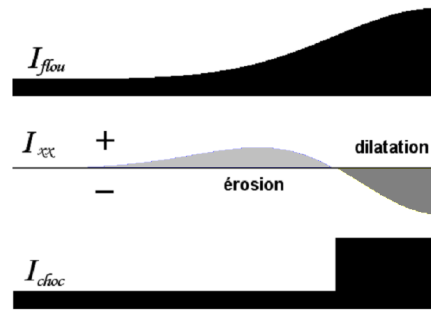


Figure 7.14 : Principe du filtre de choc de Rudin dans le cas 1-D.

de manque de netteté des contours. La dérivée seconde est positive au début de la variation, passe par zéro au point d'inflexion de la variation et devient négative jusqu'à la fin de la variation des intensités. Comme le Laplacien ΔI d'un signal 1D est égal à la dérivée seconde I_{xx} de ce signal, alors le filtre de Rudin érode la partie positive puis dilate la partie négative de la dérivée seconde jusqu'au point d'inflexion. Le résultat du filtrage I_{choc} revient à redresser la pente progressive de la variation exactement au point du changement d'inflexion du signal d'origine. Les contours apparaissent alors plus nets.

Le filtre de choc de Rudin permet de rehausser les contours mais il reste très sensible au bruit qui sera aussi amplifié ; des améliorations ont vu ainsi le jour. Osher et Rudin proposent de remplacer le Laplacien par la dérivée seconde directionnelle $I_{\eta\eta}$ parallèle au gradient et qui est connue pour être un meilleur détecteur de contour que le passage par zéro du Laplacien [OSH-88B]. Pour rendre le filtre encore plus robuste face aux faibles variations du signal, Alvarez et Mazorra [ALV-94] remplacent le détecteur de contour $I_{\eta\eta}$ par sa version lissée $I_{\eta\eta}^\sigma$ avec un noyau Gaussien de paramètre σ . La nouvelle formule du filtre de choc s'écrit comme suit :

$$\text{Filtre de choc d'Alvarez} : \begin{cases} I_t = -\text{sign}(I_{\eta\eta}^\sigma) \|\nabla I\| \\ I_0 = I_{floue} \end{cases} \quad (7.23)$$

Nous devons vérifier si les filtres de chocs sur des images 2D respectent la topologie des caractères et la continuité des formes complexes, notamment si ils peuvent restaurer les formes complexes et préserver les angles droits autour des intersections des traits et des empattements des lettres (car c'est l'information la plus importante pour les systèmes de reconnaissance de caractères ou OCR). La figure 7.15 montre, sur une image de synthèse, que le filtre d'Alvarez est plus stable que celui de Rudin sur une image bruitée.

c. Filtre de renforcement de la cohérence des structures

Le filtre de choc pour le renforcement de la cohérence des traits (coherence enhancing shock filter) est basé sur la dérivée seconde $I_{\Theta_+\Theta_+}^\sigma$ dans la direction principale Θ_+ des tenseurs T_σ calculée sur une image lissée par une Gaussienne d'écart-type σ (Eq.7.15). Cette EDP [WEI-03] est donnée par :

$$\text{Filtre de choc de Weickert} : \begin{cases} I_t = -\text{sign}(I_{\Theta_+\Theta_+}^\sigma) \|\nabla I\| \\ I_0 = I_{bruitée} \end{cases} \quad (7.24)$$

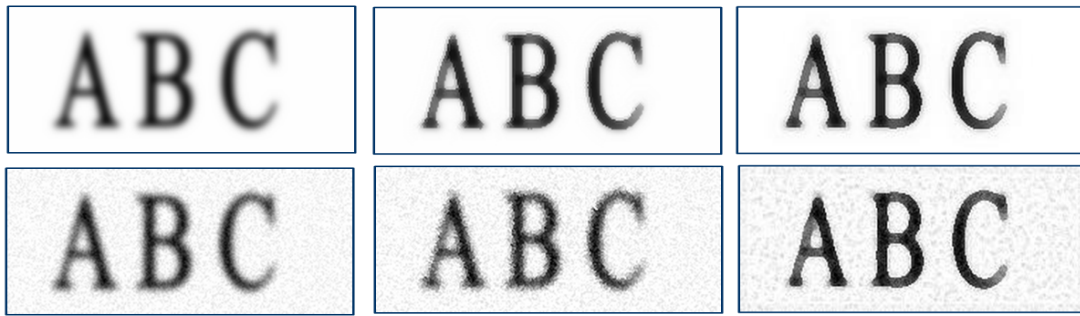


Figure 7.15 : De gauche à droite respectivement : une image floue (en haut) et une image floue bruitée (en bas), l'image rehaussée par le filtre de Rudin et l'image rehaussée par le filtre d'Alvarez.

Ce processus, comme son nom l'indique, dilate les formes le long des structures cohérentes. La direction du vecteur propre Θ_+ garantit que ce processus de diffusion crée des chocs orthogonaux à la direction du flux malgré la discontinuité des traits. Ces filtres permettent de renforcer la continuité des traits discontinus et de lisser le bruit sur les contours des traits. Weickert prouve la performance de son filtre de choc pour la reconstruction des images d'empreintes digitales dégradées comme l'indique la figure 7.16.



Figure 7.16 : De gauche à droite respectivement : l'image originale , le résultat pour un choc d'Alvarez $\sigma = 1.5$ et le résultat pour un choc de Weickert ($\rho = 5, \sigma = 1.5$)

L'image de l'empreinte digitale est intéressante car elle contient majoritairement des traits parallèles avec très peu d'intersections en angle droit. Sur des images de caractères composées de nombreuses intersections formant des structures complexes, le filtre de choc de Weickert tend à déformer les caractères et à générer des structures allongées et parallèles. Le filtre de choc de Weickert renforce les traits parallèles mais tend à diverger sur des images de caractères, si la diffusion n'est pas contrôlée (Fig.7.17).

Pour contrôler le filtre de Weickert, on peut ajouter un terme d'attachement aux données $|I - I_0|$ ou bien on peut réduire la dérive de l'EDP en calculant les tenseurs sur l'image d'origine I_0 et non plus sur l'image en cours d'évolution I .

d. La combinaison Choc/diffusion

[ALV-94, KOR-97B, COU-00] ont proposé de coupler le choc et la diffusion dans la même EDP pour simultanément rehausser la netteté des contours et diffuser les zones de faible



Figure 7.17 : Rehaussement par le filtre de choc de Weickert ($\sigma = 1.5, \rho = 5$) ; de gauche à droite respectivement : les résultats obtenus après 5, 10, 20 et 40 itérations

variation. Il suffit d'ajouter le terme correspondant au filtre de choc avec un terme de diffusion dans la même équation ; le filtre de Rudin peut ainsi être associé avec le filtrage linéaire et isotropique de l'équation de la chaleur (Eq.7.25).

$$\begin{cases} I_t &= -\text{sign}(\Delta I) \|\nabla I\| + \lambda \Delta I \\ I_0 &= I_{\text{bruitée}} \end{cases} \quad (7.25)$$

Plus généralement, on peut écrire une combinaison de choc et de diffusion avec l'EDP suivante (Eq.7.26) qui regroupe toutes les formes d'EDP vues précédemment.

$$\begin{cases} I_t &= -\text{sign}(I_{uu}^\sigma) \|\nabla I\| + \lambda \text{div}(D \nabla I) \\ I_0 &= I_{\text{bruitée}} \end{cases} \quad (7.26)$$

Avec $\vec{u} = \vec{\eta}$ pour le filtre de choc d'Alvarez, $\vec{u} = \vec{\Theta}_+$ pour le filtre de choc de Weickert. Le terme de la diffusion peut être associé à une diffusion matricielle ou scalaire avec $D = c \times Id$.

La figure 7.18 montre les résultats obtenus, pour une image de test, avec les filtres combinant choc/diffusion pour les différents modèles présentés précédemment.

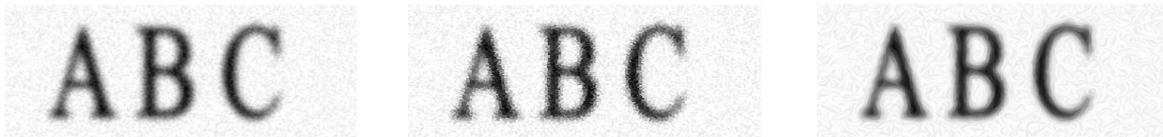


Figure 7.18 : De gauche à droite respectivement, image de synthèse floue avec un bruit gaussien additif, résultat avec un filtre de choc de Rudin seul (Eq.7.22) et le résultat pour le filtre de choc de Weickert couplé avec la diffusion matricielle (Eq.7.26)

e. Mean Curvature Motion

Le Mean Curvature Motion (MCM) est une technique de filtrage des irrégularités des traits. Cette technique, proposée dans [KIM-92, SET-85, SET-87, OSH-88A, FED-03], est

basée sur la théorie de l'évolution des courbes planes sous l'effet d'une équation de chaleur intrinsèque. Elle s'écrit sous la forme suivante (Eq.7.27) :

$$\begin{cases} I_t = \operatorname{div} \left(\frac{\nabla I}{\|\nabla I\|} \right) \|\nabla I\| \\ I_0 = I_{\text{bruitée}} \end{cases} \quad \text{avec} \quad \operatorname{div} \left(\frac{\nabla I}{\|\nabla I\|} \right) = I_{\xi\xi} = \kappa \|\nabla I\| \quad (7.27)$$

Il s'agit d'une diffusion non-linéaire qui réduit progressivement les courbures des contours de l'image. Cette EDP revient à faire évoluer les lignes de niveaux de l'image plus ou moins rapidement selon leurs courbures. Les lignes de niveaux de fortes courbures sont plus rapidement dilatées ou érodées que les lignes droites. En fonction du signe de la courbure, les convexités et les concavités des lignes de niveaux s'estompent et se lissent. Les lignes de niveaux perdent progressivement leurs singularités et les angles s'arrondissent (Fig. 7.19). Ce modèle de diffusion permet de corriger les effets d'escalier dus à une perte de résolution ou de réduire le bruit localisé le long des contours des traits.

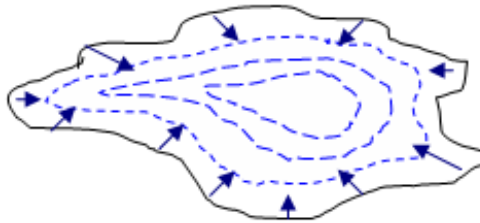


Figure 7.19 : Évolution d'une image de synthèse avec le processus de l'évolution des courbes

Mais, ce processus finit par converger vers des formes circulaires qui s'éloignent de l'information initiale. En effet, en réduisant toujours les courbures des contours, les formes se simplifient et deviennent convexes (Fig.7.20). Le MCM semble adapté au filtrage des images binaires de caractères où les traits montrent des effets d'escalier ainsi que sur les images couleurs comprimées en JPEG qui laisse apparaître des artefacts aux frontières des macro-blocs.

7.4 Expérimentation sur les images de documents

7.4.1 Évaluation des méthodes de diffusion

Nous avons testé les différents modèles de diffusion sur des images de documents. Nous distinguons deux familles de modèles de diffusion :

- *Les modèles qui ne préservent pas la topologie des formes des caractères* comme, par exemple, le Mean Curvature Motion (MCM) ou le filtre de renforcement de la cohérence des traits de Weickert. Sans un terme d'attachement aux données, ces filtres dégradent trop rapidement les formes des objets et ne conviennent donc pas au traitement des images de documents.
- *Les modèles qui préservent la topologie et les structures des lettres* comme les modèles de Perona-Malik, Beltrami ou de Tschumperlé. Ces modèles de diffusion

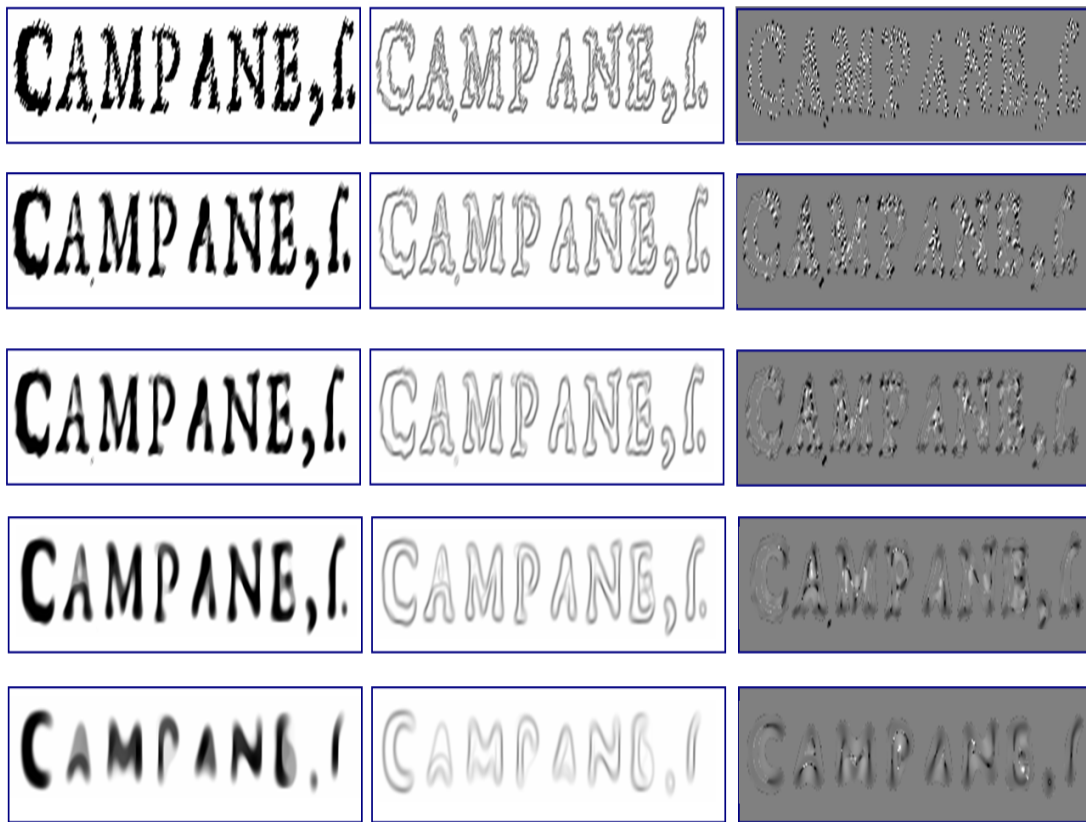


Figure 7.20 : Application du processus du Mean Curvature Motion sur une image du dictionnaire de Trévoux : de gauche à droite : respectivement l'image reconstruite à l'itération N , l'image de contour et l'image de courbure correspondante, les itérations N sont définies, respectivement de haut en bas, par 0, 5, 10, 30 et 100.

préservent suffisamment longtemps les contours et les structures pour pouvoir être appliqués au traitement des documents.

Sur une image de manuscrit médiéval de l'IRHT, le modèle de diffusion MCM et le filtre de renforcement de la cohérence de Weickert dégradent très rapidement (après quelques itérations) les formes. Les lettres finissent par converger vers des formes stylisées pour le filtre de Weickert ou vers des formes qui se convexifient pour le filtre MCM (Fig.7.21).

Les modèles de Perona-Malik, Beltrami et Tschumperlé préservent les structures même après plusieurs centaines d'itérations (Fig.7.22). Ces filtres sont très efficaces pour supprimer le bruit tout en préservant les contours. Le modèle de diffusion de Perona-Malik modifié par Catté donne de très bons résultats pour la préservation de la topologie des formes des caractères. En revanche, ce modèle ne lisse pas les contours perturbés par le bruit et ne renforce pas la continuité des traits. Les modèles de Beltrami et de Tschumperlé donnent de meilleurs résultats sur les images de documents car ils ont la propriété de réduire le bruit, de lisser les contours et de préserver les structures. En revanche, ils renforcent très modérément la continuité des traits.

Cependant, après un nombre trop élevé d'itérations, les modèles tensoriels de Tschumperlé et de Beltrami finissent pas perdre les structures complexes (Fig.7.23). Pour tous ces



Figure 7.21 : Les filtres de diffusion qui ne préservent pas la topologie des formes. De gauche à droite respectivement : l'image originale, l'image traitée par le filtre de Mean Curvature Motion (500 itérations) et l'image traitée par le filtre de Weickert (300 itérations)



Figure 7.22 : Les filtres de diffusion qui préservent la topologie des formes : De gauche à droite respectivement, l'image traitée par le filtre de Perona-Malik modifié Catté (500 itérations, $k = 10$, $\sigma = 0.5$), le filtre de Beltrami (1000 itérations) et le filtre de Tschumperlé (500 itérations)

modèles, la préservation des structures s'explique facilement par le fait que les coefficients de diffusivités $\mu_{+/-}$ tendent vers zéro lorsque l'on s'approche d'une forte variation. Mais, comme ces coefficients ne sont jamais totalement nuls, une infime diffusion est toujours réalisée à chaque itération, expliquant ainsi la perte progressive des structures après plus de 1000 itérations.

Après de nombreux tests sur des documents d'origines diverses, imprimés ou manuscrits, nous avons pu mettre en évidence les avantages et les inconvénients des trois modèles de diffusion qui sont les plus adaptés au traitement de documents (Tab.7.2). Le modèle de Perona-Malik se distingue de celui de Beltrami et de Tschumperlé par une convergence rapide vers une image pratiquement segmentée. Cela s'explique par le coefficient de diffusivité $c(\|\nabla I\|)$ qui dépend du gradient ∇I calculé sur l'image itérée I et non sur l'image d'origine I_0 . Pour les modèles tensoriels de Beltrami ou de Tschumperlé, ils sont très complets car ils réduisent tous les types de bruits, y compris les bruits localisés sur les contours. En revanche, ils restaurent difficilement la continuité des traits et finissent par rendre les contours flous. En outre, les modèles tensoriels de diffusion dépendent de la résolution de l'image. Plus l'image est numérisée avec une faible résolution, moins les détails fins seront préservés.



Figure 7.23 : Les filtres de diffusion qui finissent par perdre la topologie après un nombre trop élevé d'itérations : De gauche à droite respectivement, l'image traitée par le filtre de Perona-Malik modifié Catté (1000 itérations, $k = 10$, $\sigma = 0.5$), le filtre de Beltrami (2000 itérations) et le filtre de Tschumperlé (1000 itérations)

	Avantages	Inconvénients
Perona-Malik	<ul style="list-style-type: none"> - Réduit les bruits additifs ou ponctuels. - Préserve les contours et les structures. - Modèle générique adapté à tous les types de documents. 	<ul style="list-style-type: none"> - Paramètre k très critique. - Converge vers une image segmentée. - Ne lisse pas les contours. - Ne renforce pas les traits.
Beltrami ou Tschumperlé	<ul style="list-style-type: none"> - Pas de paramètres. - Modèle générique adapté à tous les types de documents. - Préserve les contours et les structures. - Réduit tous les types de bruits et particulièrement les bruits localisés sur les contours. 	<ul style="list-style-type: none"> - Nécessité de contrôler les excès de la diffusion avec un terme d'attachement aux données. - Renforce très modérément la continuité des traits. -Dépendance à la résolution de l'image. - Les contours deviennent flous et un filtre de choc peut être utile pendant ou après le traitement.

Tableau 7.2 : Avantages et inconvénients des modèles de diffusion les plus adaptés au traitement des images de documents

Cette évaluation des modèles de diffusion a été réalisée de façon visuelle. Dans la partie suivante, nous allons tester objectivement la restauration de caractères imprimés à l'aide de logiciels de reconnaissance automatique de caractères.

7.4.2 Évaluation par OCR

Notre toute première expérimentation sur la diffusion a consisté à appliquer directement des modèles existants et de les évaluer le plus objectivement possible par un système de reconnaissance de caractères ou OCR (abréviation du terme anglais Optical Character Re-

cognition). Son but est de convertir un texte imprimé en langage interprétable par machine pour des traitements informatiques. Nous avons ici deux objectifs, d’abord démontrer qu’un prétraitement basé sur la restauration des images permet d’améliorer les performances d’un OCR. Ensuite, nous voulons réaliser une validation objective de la restauration par diffusion plutôt qu’une évaluation subjective visuelle.

a. Le fonctionnement d’un OCR

Un système de lecture automatique de caractères localise et identifie les caractères dans une image puis les transforme en texte éditable qui pourra être exporté vers un traitement de texte. L’OCR comporte essentiellement trois modules de traitements [NAG-92, BEL-06] définis suivant un ordre chronologique d’exécution par :

1. **La segmentation.** L’objectif de cette étape est d’identifier les blocs de texte des autres composants de l’image. Elle utilise des mesures sur les espacements entre les lignes et les caractères pour reconnaître le texte dans un ordre convenable. Ces mesures génèrent divers types d’erreurs étant donnée la diversité des mises en formes des textes. Si le document est particulièrement dégradé et fait apparaître des caractères cassés ou collés, la segmentation des caractères devient alors une tâche difficile. De nombreuses erreurs de l’OCR sont donc expliquées par les problèmes de segmentation des mots et des caractères.
2. **La reconnaissance.** Les formes de lettres sont caractérisées par des descripteurs qui sont analysés puis comparés à ceux stockés dans une base de connaissances. Malgré la présence des descripteurs de formes robustes, les dégradations des caractères localisés à des endroits importants comme, par exemple, la barre du «t» ou la queue du «Q» peuvent tromper le système de reconnaissance.
3. **La correction.** C’est une étape de post-traitement qui met en évidence les caractères non-reconnus et les mots qui n’ont pas été trouvés dans le dictionnaire. Cette étape propose à l’utilisateur des corrections lexicales et syntaxiques à l’aide des dictionnaires intégrés au logiciel pour améliorer le taux de reconnaissance.

Cette description ne fait pas référence à l’étape importante de prétraitement des images que tous les logiciels d’OCR actuels effectuent pour adapter l’image aux conditions de sa reconnaissance. À notre connaissance, ces prétraitements ne concernent que la normalisation de l’échelle si la résolution est inadaptée à celle prévue par le programme d’analyse ou la binarisation si l’image est en niveaux de gris ou en couleur. Le développement d’outils avancés de restauration des caractères pour améliorer les performances des logiciels de reconnaissance de caractères est une piste encore largement inexplorée.

b. L’évaluation des performances d’un OCR

Comme les logiciels d’OCR sont basés sur des systèmes experts à bases de règles dont le taux d’erreur ne peut pas être prévisible à l’avance de manière théorique, les performances d’un logiciel d’OCR sont mesurées statistiquement en comptant le nombre n_e de caractères erronés par rapport au nombre n_c de caractères analysés [BEL-06] :

$$\Gamma_{err} = \frac{n_e}{n_c} \quad (7.28)$$

Une erreur est signalée si le caractère reconnu ne correspond pas au caractère attendu (substitution), ou bien s'il n'a pas été reconnu par le système (rejet). Du point de vue de l'utilisateur, le taux de substitution est plus critique que le taux de rejet. Mais, pour notre application, nous ne faisons pas la différence entre les deux types d'erreurs. Le taux de reconnaissance Γ_{rec} représente le pourcentage de caractères correctement reconnus. Il est exprimé par le rapport suivant :

$$\Gamma_{rec} = \frac{n_c - n_e}{n_c} = 1 - \Gamma_{err} \quad (7.29)$$

On trouve désormais sur le marché des logiciels de reconnaissance de caractères avec de bonnes performances. Ces performances restent très dépendantes de la qualité du texte imprimé à reconnaître et chutent considérablement sur les documents anciens dégradés.

c. Choix des documents

Nous avons choisi d'évaluer la restauration par diffusion sur les images du dictionnaire de Trévoux du XVIII^{ème} siècle et qui constitue le tout premier dictionnaire complet de la langue française. Le dictionnaire est un document imprimé qui peut être traité par OCR. Cet ouvrage de plusieurs dizaines de milliers de pages a été numérisé et prêté par l'ATILF¹ à des fins de recherches pour l'évaluation des OCR [BEL-06] et à des fins de la restauration [DRI-07]. Il a été numérisé en noir et blanc avec un bruit important le long des contours des caractères (Fig 8.4) qui perturbe les systèmes de reconnaissance de formes. C'est donc un document idéal pour tester les modèles de diffusion.

d. Choix des modèles de diffusions adaptés

Comme nous sommes en présence d'images binaires dont les dégradations se portent essentiellement sur les contours des lettres, nous allons nous concentrer uniquement sur les modèles qui lissent les contours. Nous avons retenu le modèle de diffusion matricielle de Tschumperlé et le filtre de choc de Alvarez pour lisser les traits et celui de Weickert pour renforcer la continuité des traits tout en lissant leurs contours. Comme nous l'avons vu dans la section précédente, les modèles de diffusion finissent par perdre la topologie des formes des lettres. Nous devons donc ajouter aux EDP un terme d'attachement aux données où I est l'image traitée et I_0 est l'image initiale bruitée. Pour trouver le bon paramètre λ devant le terme $\lambda |I - I_0|$ d'attachement aux données, nous proposons comme solution de regarder l'évolution des résultats de la reconnaissance des caractères sur des portions représentatives de chaque document. Pour le meilleur taux de reconnaissance atteint, nous mémorisons le coefficient λ . Les images restaurées par les approches d'Alvarez, de Weickert et de Tschumperlé, qui donne le meilleur taux de reconnaissance, sont représentées respectivement par les figures 7.26 , 7.27 et 7.28.

Les modèles de diffusion ont été choisis pour les propriétés de lissage des contours. Ils ne permettent donc pas de réparer les ruptures trop larges des traits. Sur la figure 7.29, la

¹ATILF : Analyse et Traitement Informatique de la Langue Française, UMR 7118 Nancy, Images prêtées par Pr. Isabelle Turcan.

LE C. DE GABALIS. La *Cabbale* est une science sérieuse, & il n'y a que les mélancoliques qui s'y adonnent. A. B. DE VILLARS. La *Cabbale* est une de ces chimères qu'on autorise quand on les combat gravement, & qu'on ne doit entreprendre de détruire qu'en se jouant. Ib. Robert Flud Anglois, en a fait d'amples Traitez & Apologies dans ses neuf grands Volumes.

CABBALÉ, se dit aussi de la Sècte des Juifs, qui suit & pratique la *Cabbale*, qui interprètent l'écriture selon l'art de la *Cabbale*, prise au second sens que nous avons expliqué; car les Juifs font

Figure 7.24 : Extrait de l'image originale du dictionnaire du Trévoux.

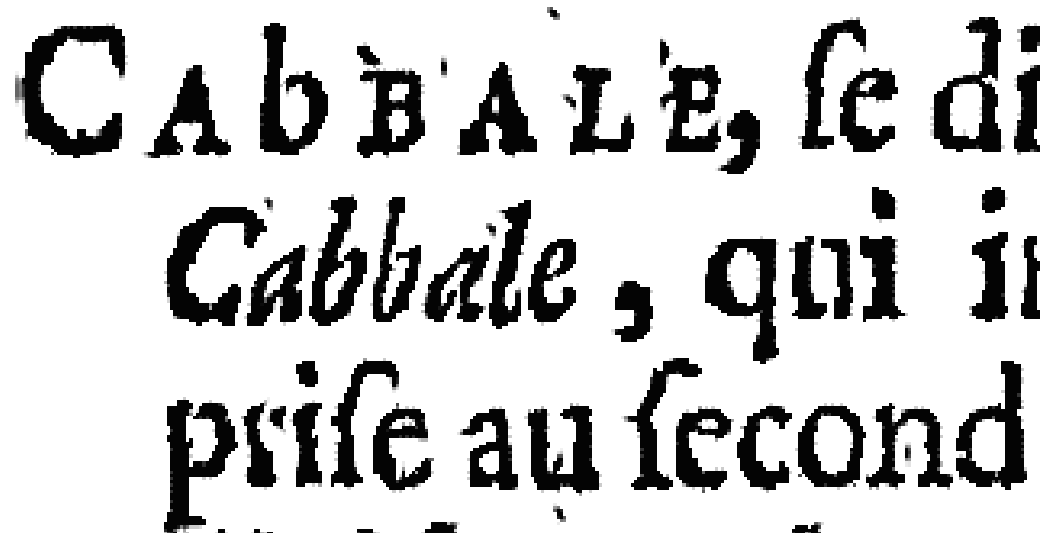
C A B B A L É, se di
Cabbale, qui i
 prise au second

Figure 7.25 : Agrandissement de l'image originale.

diffusion par l'approche de Tschumperlé échoue à réparer la lettre «h» coupée en deux. Les matrices de diffusivités D suivent les contours de la rupture du tracé et ne parviennent pas à s'orienter dans le prolongement de la continuité des traits.

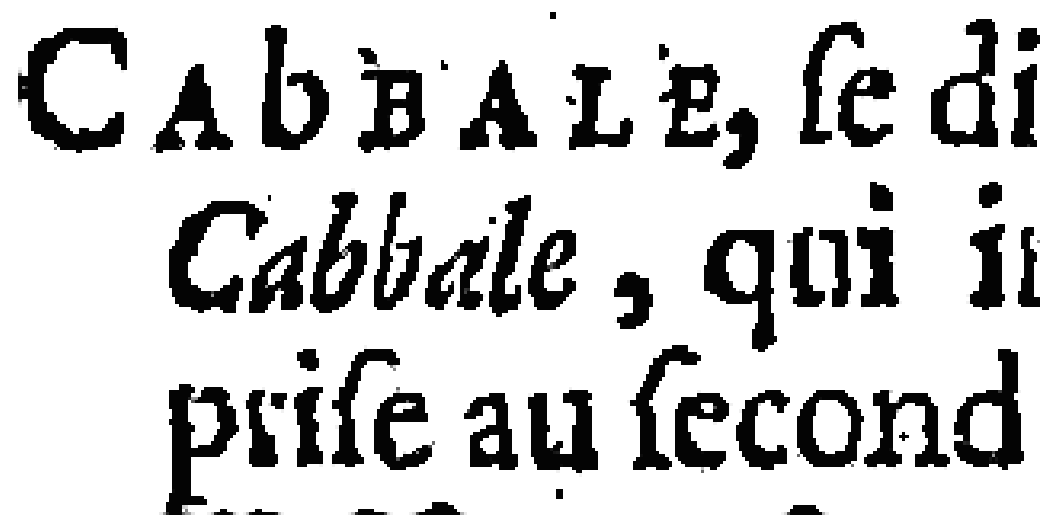
e. Résultats des logiciels d'OCR sur les images restaurées

Nous allons comparer les erreurs commises par deux logiciels d'OCR : ABBY FineReader 7.01 et OmniPagePro 15.02; ces deux logiciels se comportent, dans notre étude, comme



C A B B A L E, se di
Cabbale, qui il
prise au second

Figure 7.26 : La meilleure version restaurée par l'approche de Weickert de l'image 7.25



C A B B A L E, se di
Cabbale, qui il
prise au second

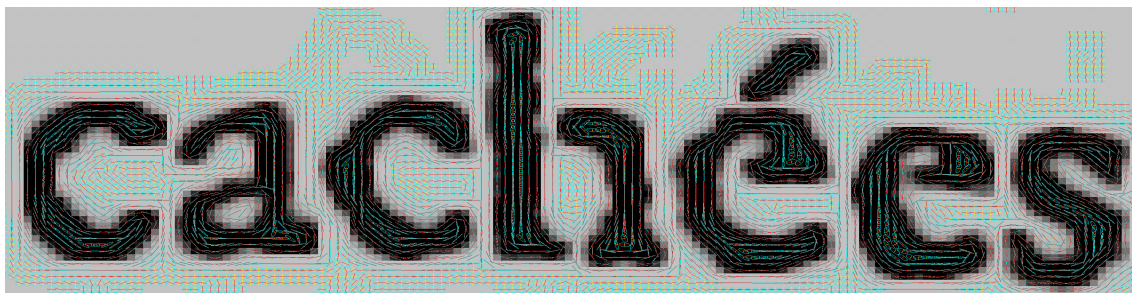
Figure 7.27 : La meilleure version restaurée par l'approche d'Alvarez de l'image 7.25

CABBALE, se di
Cabbale, qui it
prise au second

Figure 7.28 : La meilleure version restaurée par l'approche de Tschumperlé de l'image 7.25

cachées

(a) Mot extrait d'une page du dictionnaire de Trévoux



(b) Affichage des tenseurs de D . Tschumperlé à l'itération 200

Figure 7.29 : Image originale du dictionnaire du Trévoux et l'image en cours de restauration avec les matrices de diffusivités de Tschumperlé

une boîte noire. Nous ne nous intéressons qu'au résultat de la reconnaissance de texte et nous supposons que celui de la segmentation est correct. Pour évaluer les résultats de la reconnaissance avant et après restauration, nous avons comparé les erreurs détectées dans le texte. En général, un logiciel d'OCR peut faire trois types d'erreurs sur la reconnaissance

des caractères :

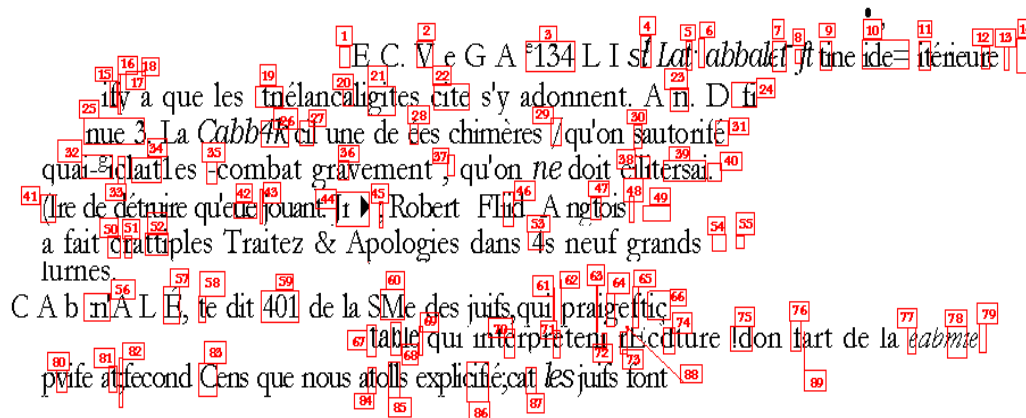
- *Une substitution*, en remplaçant un caractère par un autre ;
- *Une suppression*, en ignorant un caractère à cause d'un rejet ou de la substitution de plusieurs lettres par une lettre de forme similaire («n» à la place de «ri») ;
- *Un ajout*, en décomposant un caractère par deux autres dont l'apparence peut être proche de la lettre non reconnue («rn» à la place de «m»).

Prenons l'exemple de la reconnaissance de l'image initiale par l'OCR OmniPagePro (Fig.7.30), nous notons à titre d'exemple que le logiciel échoue à reconnaître l'image du mot «VILLARS» et rejette le mot en entier. Pour la reconnaissance du même mot par FineReader (Fig.7.32), seules quelques lettres sont reconnues et les autres sont substituées par d'autres lettres, la séquence «ILL» est remplacée par «ut» et la lettre «R» est remplacée par «a».

Les figures 7.30 et 7.31(respectivement 7.32 et 7.33) montrent les résultats donnés par OmniPagePro (respectivement FineReader), avant et après la restauration par le modèle de Tschumperlé. Les tableaux expliquent comment nous avons manuellement calculé le taux d'erreur en tenant compte à la fois du rejet, de la substitution, du dédoublement ou de la suppression de caractères et également de la ponctuation. Quel que soit l'OCR utilisé, après une phase de restauration, on peut constater une nette diminution des erreurs de l'OCR.

Le tableau 7.3 récapitule les taux d'erreurs Γ_{err} calculés en tenant compte ou non des erreurs sur la ponctuation et les caractères accentués pour chaque modèle de restauration et suivant le logiciel OCR utilisé. Quel que soit le modèle de diffusion choisi, le taux de reconnaissance est toujours plus élevé après la restauration. Nous observons aussi que le modèle de Tschumperlé restaure beaucoup mieux les caractères que celui d'Alvarez qui lui-même donne de meilleurs résultats que le modèle de Weickert. Nous avons observé le meilleur taux de reconnaissance pour le logiciel FineReader associé au modèle de diffusion de D. Tschumperlé. La restauration des caractères dégradés par ce modèle a permis de gagner en moyenne un point de reconnaissance en plus sur les résultats de l'OCR. Une telle performance peut s'expliquer par le fait qu'il réalise un lissage des contours par une diffusion alors que les modèles d'Alvarez et de Weickert sont des filtres de chocs qui réalisent un lissage implicite des contours par des érosions ou des dilatations des aspérités le long des traits. Enfin, le modèle de Weickert adapté au renforcement de la continuité des traits, finit par modifier de façon trop importante la topologie de certaines lettres et pose des problèmes à l'OCR. Par conséquent, le meilleur modèle de restauration des imprimés aux contours bruités est celui de la diffusion par un modèle tensorielle.

Nous avons réalisé alors un test sur un nombre plus élevé de pages du dictionnaire de Trévoux en prenant en compte uniquement la méthode de restauration de D. Tschumperlé et le logiciel de reconnaissance FineReader sélectionnés pour leurs performances. Cette série de tests évalue la reconnaissance en ignorant les problèmes de ponctuation et de diacritique sur des extraits de cinq pages du dictionnaire de Trévoux. La variabilité de la qualité des dégradations des formes de traits d'une image à l'autre influence considérablement les performances du logiciel de reconnaissance. Le tableau 7.4 synthétise les résultats de la reconnaissance par OCR et donne les taux d'erreurs avant et après la restauration pour chaque page extraite du dictionnaire du Trévoux. Sur certaines pages, ce taux d'erreur chute considérablement après une phase de restauration.



Numéro	Correct	Généré	Numéro	Correct	Généré	Numéro	Correct	Généré
1	L	<rien>	31	e	é	61	fait	<rien>
2	D	V	32	nd	i-fi	62	&	<rien>
3	BA	134	33	on	c	63	t	<rien>
4	.	t	34	<espace>	lait	64	qu	g
5	<espace>	t	35	<espace>	<espace>	65	<espace>	f
6	C	<rien>	36	à	a	66	là	tic
7	e	et	37	&	<rien>	67	C	t
8	e	<rien>	38	en	cih	68	bba	b
9	u	ti	39	repre	ersai	69	,	<rien>
10	fcience	ide=	40	-	.	70	è	e
11	f	it	41	d	l	71	é	e
12	f	r	42	n	ue	72	t	i
13	,	<rien>	43	fe	<rien>	73	l	r
14	&	<rien>	44	ID	Jr>	74	ri	d
15	l	f	45	.	<rien>	75	fel	ld
16	<espace>	<rien>	46	u	ii	76	l	t
17	n	<rien>	47	l	t	77	c	e
18	'	<rien>	48	,	<rien>	78	bal	mi
19	m	tn	49	en	<rien>	79	,	<espace>
20	o	a	50	d	c	80	r	v
21	qu	git	51	'	r	81	u	T
22	qui	cite	52	m	tii	82	<espace>	;
23	B	n	53	fe	4	83	f	C
24	E	fi	54	Vo	<rien>	84	v	T
25	VILLARS	<rien>	55	-	<rien>	85	n	Ll
26	ale	4k	56	B	n'	86	qu	Cif
27	eft	cil	57	E	E	87	r	T
28	c	é	58	f	t	88	'	<rien>
29	<espace>	<espace>/	59	auffi	401	89	'	<rien>
30	<espace>	<espace>s	60	ect	M	90		

Figure 7.30 : La traduction par OmnipagePro de l'image initiale et affichage des erreurs détectées

7.5 Conclusion

La diffusion permet bien de restaurer visuellement les formes dégradées de caractères et augmente de façon très nette le taux de reconnaissance des OCR. Nos tests démontrent que la diffusion est bien un outil puissant de restauration des formes dégradées dans les

L E C. DE GA L I t s. *La Cabbale* etc L i n e f i c i e n c e f é r i e u C e , &
 il n'y a que les mélancoliques qui s'y adonnent. A B. D t
 V I L L A R s. *La Cabbale* est une de ces chimères qu'on autorise
 quand on les combat gravement, & qu'on ne doit entrepren-
 dre de détruire qu'en le jouant. b. Robert Flud Anglois, c a
 a fait d'amples Traitez & Apologies dans Ces neuf grands Vol.
 l u t e s
 C A B A L I s t e a e t h de la S e t e de s J u i f s , q u i f u i t 6 e p h a t i q u e
 C a b b a l e q u i i n t e r p r e t e n t l e t r u e d o n t p a r t d e l a C a b b e a l e
 F i f e a u f e c o n d f e u s q u e n o u s a v o n s e x p l i c i t e s t a r l e s j u i f s f o n t

Numéro	Correct	Généré	Numéro	Correct	Généré	Numéro	Correct	Généré
1	BA	<espace>	19	f	C	37	E	t
2	I	t	20	-	..	38	fel	c
3	<espace>	<rien>	21	m	t	39	l	F
4	C	t	22	B	<rien>	40	'	<espace>
5	ft	tc	23	E	<rien>	41	l	.l
6	u	Li	24	,	<rien>	42	C	IC
7	f	C	25	f	r	43	a	ea
8	E	t	26	uffi	eth	44	,	<espace>
9	a	à	27	c	<rien>	45	pr	F
10	e	é	28	&	6°	46	u	i
11	n	is	29	r	h	47	<espace>	f
12	â	a	30	là	<espace>	48	n	u
13	en	ca	31	<espace>	<rien>	49	o	ô
14	<espace>	<rien>	32	è	e	50	p	y
15	f	4	33	é	e	51	qu	cp
16	ID	,b	34	t	r	52	c	st
17	u	ii	35	l	l	53		
18	en	cà	36	'	<rien>	54		

Figure 7.31 : Le résultat de la reconnaissance par OmnipagePro de l'image restaurée et affichage des erreurs détectées après une restauration par l'approche de Tschumperlé

images de documents. Nous avons aussi prouvé que l'on pouvait gagner plus d'un point de reconnaissance aux OCR en développant des méthodes de prétraitement des images. Ces premiers résultats encourageants ouvrent la voie sur des nouvelles pistes de recherche comme l'amélioration des performances des OCR à l'aide de prétraitements adaptés et l'application des modèles de diffusion pour la restauration des formes dégradées de caractères. Dans le chapitre suivant, nous allons montrer que les modèles de diffusion peuvent également servir dans la reconstruction des parties manquantes des images.

LE C. CE G ABAUS. L & Cébale(k une factice férieufe, 6C
 il n'y a que les mélancoliques qui s'y adonnent. Ais
 V ut A la s= Lz' Cbaktâ une < les chimères qui on aurife
 quatioâ les combat grâvemenc, & quon ne doit entrepren
 dre de ddiure qa'pa' e jouant. Il >. Robert FJud Anglois >> ca
 a fait d'amples Traitez ôc Apologies dans Les neuf grands Voy
 limes >>
 C A b f c A L % fe dit aujfi ôc la Sede des Juifs qui fuit Se pratique là
 Cabale qui interprètent l'Ecriture félon l'art de Ja fô^&ffefc
 prife au fécond fens que nous a ons expk quév car les Juifs onc

Numéro	Correct	Généré	Numéro	Correct	Généré	Numéro	Correct	Généré
1	.	,	31	e	è	61	i	ï
2	D	C	32	n	ñ	62	u	ü
3	LI	U	33	<espace>	<rien>	63	f	fi
4	a	&'	34	n	â	64	de	âc
5	ab	é	35	t	c	65	l	i
6	eft	t(k	36	,	?	66	ct	d
7	cien	acti	37	<espace>	:	67	i	i
8	e	è	38	'	<rien>	68	s	e
9	&	6C	39	t	c	69	,	<rien>
10	l	i	40	étr	di	70	&	Se
11	'	3	41	i	t	71	l	ï
12	l	i	42	u	a	72	bb	b
13	u	y	43	en	pa	73	,	^
14	'	*	44	<espace>	^	74	e	e
15	B	is	45	f	<rien>	75	l	T
16	DE	<espace>	46	<espace>	<espace>f	76	'	<rien>
17	ILL	ut	47	D	ÿ>	77	e	é
18	R	a	48	lu	Jù	78	l	F
19	.	-	49	i	i	79	'	<espace>
20	a	z'	50	,	>>	80	l	J
21	ab	d	51	en	ca	81	Cabbale	Fô&ffefc
22	le	ktâ	52	m	in	82	e	é
23	eft	<rien>	53	&	ôc	83	vo	^ô
24	d	<J	54	f	T	84	li	K
25	<espace>	<rien>	55	-	v.	85	;	v
26	<espace>	<espace>'	56	u	û	86	<espace>	<rien>
27	u	ii	57	.	>>	87	o	b
28	o	ô	58	B	fc	88	t	c
29	a	â	59	E	%	89	<espace>	<rien>
30	t	r	60	,	<rien>	90	,	<espace>

Figure 7.32 : La traduction par FineReader de l'image initiale et affichage des erreurs détectées

LE C. DE G ABALIS. *Li Cabbale élk* une science jerieufe, & il n'y a que les mélancoliques qui s'y adonnent. As DE VI 11 A R s. La *Cabbak* est une de ces chimères qu'on aurorife quand on les combat gravement, & qu'on ne doit entreprendre de d&unre qu'ea fe jouant. I». Robert F Jbd Anglois, cu a fait d'amples Traitez & Apologies dans fes neuf grands Vo lumes.

C A B B A L fe, fe dit au (ii de la Sède des Juifs, <jui fu<> & pratique là *Cabbak*, qui interprètent l'Escriture félon l'art de *h ftibudlet* prife au fécond fens que nous avons expliqué, car les Juifs fbfc

Numéro	Correct	Généré	Numéro	Correct	Généré	Numéro	Correct	Généré
1	a	ii	13	e	ê	25	q	<j
2	<espace>	<rien>	14	â	a	26	it	<<
3	e	é	15	étr	&i	27	le	k
4	ft	lk	16	i	u	28	è	e
5	f	/	17	n	à	29	e	é
6	f	j	18	D	>>	30	la	h
7	é	e	19	lu	Jb	31	Ca	fii
8	e	è	20	cù	en	32	b	ù
9	B	s	21	è	e	33	,	t
10	le	k	22	E	fe	34	e	é
11	a	â	23	ff	Gi	35	;	;
12	t	r	24	ct	â	36	ont	bfic

Figure 7.33 : La traduction par FineReader de l'image restaurée et affichage des erreurs détectées après une restauration par l'approche de Tschumperlé

	Logiciel de reconnaissance	Reconnaissance d'une page : Affichage du taux d'erreur.			
		Avec ponctuation		Sans ponctuation	
		Avec diacritique	Sans diacritique	Avec diacritique	Sans diacritique
Image initiale non-traitée	OmnipagePro	0.280	0.270	0.262	0.251
	FineReader	0.243	0.215	0.211	0.181
Image restaurée (Tschumperlé)	OmnipagePro	0.140	0.127	0.128	0.114
	FineReader	0.109	0.087	0.109	0.08
Image restaurée (Weickert)	OmnipagePro	0.219	0.212	0.197	0.190
	FineReader	0.171	0.142	0.156	0.125
Image restaurée (Alvarez)	OmnipagePro	0.174	0.162	0.173	0.162
	FineReader	0.123	0.120	0.118	0.102

Tableau 7.3 : Reconnaissance d'une page du dictionnaire du Trévoux avant et après restauration et affichage du taux d'erreur.

	Nombre total de caractères	Reconnaissance par FineReader : Affichage du taux d'erreur.	
		Avec restauration	Après restauration
Extrait N°1	448	0.140	0.093
Extrait N°2	788	0.101	0.087
Extrait N°3	981	0.218	0.191
Extrait N°4	778	0.079	0.077
Extrait N°5	606	0.178	0.176

Tableau 7.4 : Résultats récapitulatifs sur cinq extraits du dictionnaire du Trévoux.

Reconstruction par Inpainting

« L'expérience est une lanterne que l'on accroche dans le dos
et qui éclaire le chemin déjà parcouru. »

Confucius

Résumé

La présence de symétries et de redondances d'information sur les formes des lettres permet la reconstruction des parties manquantes dans les images de documents. Nous proposons d'étudier dans ce chapitre l'apport de l'inpainting pour la reconstruction des parties détruites des documents.

Sommaire

8.1	Introduction	118
8.2	Principe de l'inpainting	118
8.3	Application aux documents	122
8.4	Expérimentation et évaluation	123
8.5	Discussion sur les méthodes de reconstruction par inpainting .	129
8.6	Conclusion	131

8.1 Introduction

La reconstruction des parties manquantes est un problème récurrent en traitement d'images, en particulier dans les images de documents. Trois approches de reconstruction sont envisageables :

- **Reconstruction par apprentissage** : il s'agit de réaliser un apprentissage des configurations possibles des formes pour prédire les pixels de la zone manquante. Cette méthode est adaptée aux images de documents qui contiennent des formes continues et répétitives. Elle possède des limites telles que la complexité de calcul liée au nombre des configurations possibles et la taille de la base d'apprentissage. Les méthodes par apprentissage sont, par nature, plutôt limitées aux images de documents imprimés. Parmi les méthodes de reconstruction par apprentissage, on retrouve les champs de Markov (Markov Random Field) pour réparer les images binaires de caractères [WOL-06] ou l'approche par les contours actifs guidée par un modèle [ALL-03].
- **Reconstruction par copie de formes similaires** : La méthode PMS consiste à trouver (Pattern Matching) et à recopier (Substitution) la forme la plus similaire à la place de celle qui est dégradée. Cette méthode est très difficile à mettre en œuvre car elle nécessite une segmentation précise du contenu des images et une mesure de similarité efficace pour trouver les formes identiques. Pour les images de documents, elle est actuellement en cours d'étude pour réparer les caractères imprimés des gazettes [EYN-07]. Cette solution n'est pas adaptée aux documents manuscrits, difficiles à segmenter et dont les contenus sont trop variables.
- **Reconstruction par diffusion** : Cette méthode introduite par Bertalmio [BER-00], aussi appelée «digital inpainting», utilise les EDP pour reconstruire les parties manquantes. Elle a déjà servi à réparer les photos anciennes déchirées, les rayures sur les films anciens ou les craquelures sur les images numériques de peinture. À notre connaissance, elle n'a pas été étudiée en restauration d'images de documents.

La reconstruction de parties manquantes des images par diffusion nous paraît très intéressante à tester dans le cadre de la restauration des images de documents. Elle ne nécessite ni apprentissage, ni segmentation ni interprétation préalable des contenus des images ; elle est totalement générique car elle peut traiter des documents aux formes complexes et très variables comme les manuscrits anciens. En effet, la variabilité des manuscrits rend les approches par apprentissage ou les méthodes par copie complètement inutilisables. Nous étudierons les limites des méthodes de diffusion pour la reconstruction des formes complexes des caractères, la conservation de leurs structures et la préservation de leurs topologies.

8.2 Principe de l'inpainting

L'inpainting (il n'existe pas d'équivalent dans la langue française) fait référence aux artistes qui restaurent les tableaux de peinture en repeignant les parties dégradées. L'inpainting numérique [BAL-00, BER-00, CHA-00] est l'opération équivalente pour les images numé-

riques. C'est donc une méthode de reconstruction des parties manquantes de l'image à partir des informations situées sur les bords de la zone à reconstruire ; elle utilise les EDP pour diffuser l'information des bords vers l'intérieur de la zone à reconstruire. L'inpainting consiste à joindre les lignes de même niveau (isophotes) de part et d'autre des bords du domaine à remplir et à poursuivre les contours dans leurs directions initiales (Fig. 8.1). Intuitivement, la meilleure façon de reconstruire la partie manquante consiste à connecter les contours et à conserver des lignes de niveaux les plus courtes et les plus lisses possible.

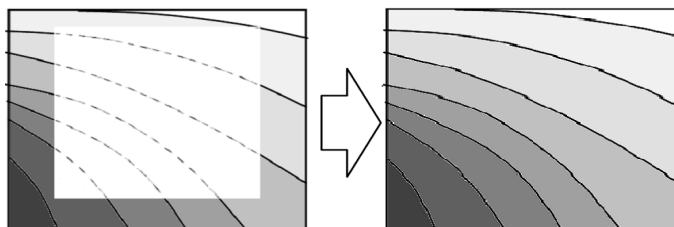


Figure 8.1 : Principe de reconstruction par jonction des isophotes

La méthode ne permet pas de reconstruire l'image de l'objet original, mais une image plausible, compatible avec les lois de la perception (Fig.8.2.b) ; elle ne peut pas reconstruire des formes complètement inattendues ou improbables comme celles présentées à la figure 8.2.c.

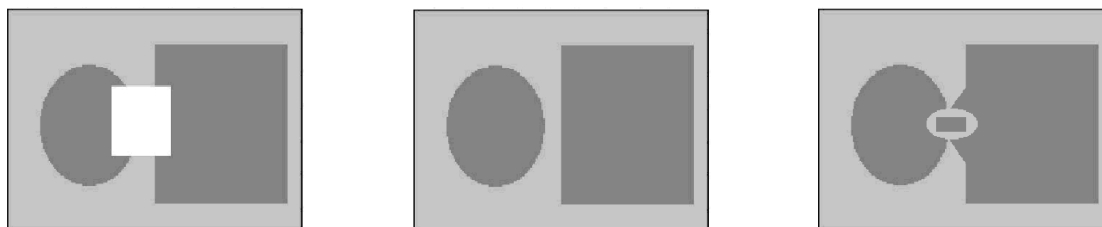


Figure 8.2 : De gauche à droite respectivement : l'image à reconstruire, l'image la plus probable visuellement et l'image visuellement improbable ou inattendue

L'inpainting nécessite deux étapes :

1. **Repérer l'information manquante et créer un masque qui délimite de façon précise les zones à reconstruire.** On définit donc un masque binaire $M : \Omega \rightarrow \{0, 1\}$ avec $M(x, y) = 1$ si le pixel (x, y) se situe dans la zone à reconstruire et $M(x, y) = 0$ sinon. On note ∂M le bord de la zone à reconstruire.
2. **Reconstruire la partie manquante de l'image en transportant les valeurs des pixels des bords ∂M à l'intérieur de la zone M à reconstruire.** La méthode synthétise artificiellement une image plausible, compatible avec les lois de la perception. La prolongation des bords est notamment dirigée par les lignes de niveaux de l'image et cela peut conduire à deux approches différentes :
 - (a) Diffuser l'information dans la direction des isophotes. C'est l'approche de Bertalmio.

- (b) Diffuser l'information perpendiculairement aux isophotes dans la direction du gradient soit en minimisant le critère de la variation totale variation (TV) soit en diffusant suivant la courbure CDD (Curvature-driven Diffusion).

8.2.1 La diffusion des intensités de l'image dans la direction des isophotes

Bertalmio [BER-00] définit l'inpainting comme une propagation des informations L du contour du domaine ∂M dans le domaine M suivant la direction N des isophotes (Fig.8.3).

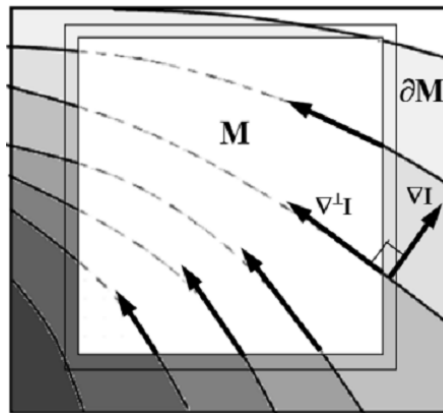


Figure 8.3 : Direction de la propagation des informations situées le long du contour ∂M à l'intérieur du domaine M dans la direction des isophotes $\vec{N} = \nabla^\perp I$

L'équation $I_t = \nabla L \cdot \vec{N}$ converge et se stabilise lorsque $\nabla L \cdot \vec{N} = 0$; c'est-à-dire lorsque l'information L a été complètement propagée dans la direction N . La meilleure solution consiste à prendre pour N la direction des isophotes. Comme on souhaite un remplissage «lisse», on propage un estimateur de lissage de l'image, par exemple le Laplacien $L = \Delta I = I_{xx} + I_{yy}$. L'EDP se stabilisera donc quand le lissage sera constant le long des isophotes. La forme générale de l'équation adaptée à ce processus s'écrit sous la forme suivante (Eq.8.1) :

$$\begin{cases} I_t = \nabla(\Delta I) \cdot \nabla^\perp I & \text{si } (x, y) \in M \\ I_t = 0 & \text{sinon} \end{cases} \quad (8.1)$$

Au niveau de l'implémentation, le modèle de Bertalmio nécessite des aménagements; si on ne fait que diffuser dans le prolongement des isophotes, les lignes de niveaux risquent de se mélanger et de devenir incohérentes. Il faut donc appliquer une diffusion non-linéaire entre chaque itération de l'équation 8.1.

8.2.2 La diffusion des intensités suivant le modèle de la totale variation (TV)

Le modèle de reconstruction par une diffusion des intensités suivant le modèle de la totale variation (TV) a été introduit par Chan et Shen [CHA-02]. Il repose sur la minimisation des

variations totales (TV) dans le domaine M à remplir (Eq. 8.2). Le critère de la variation totale est proche de celui de Tikhonov (Eq. 7.1) sauf que la variation des intensités d'une image est mesurée par la norme du gradient et non par son carré.

$$I_{restaurée} = \underbrace{Min}_{I:M \rightarrow \mathbb{R}} TV(I) = \int_M \|\nabla I\| dx dy \quad (8.2)$$

En appliquant le schéma de résolution d'Euler Lagrange, la minimisation de la totale variation donne une EDP équivalente à celle de l'équation 8.3. Cette EDP est guidée par la courbure κ des isophotes (Eq. A.14).

$$\begin{cases} I_t = \operatorname{div} \left(\frac{\nabla I}{\|\nabla I\|} \right) & \text{si } (x, y) \in M \\ I_t = 0 & \text{sinon} \end{cases} \Leftrightarrow \begin{cases} I_t = \frac{I_{\xi\xi}}{\|\nabla I\|} & \text{si } (x, y) \in M \\ I_t = 0 & \text{sinon} \end{cases} \quad (8.3)$$

On retrouve l'équation non-linéaire de Perona et Malik avec $C(\|\nabla I\|) = \frac{1}{\|\nabla I\|}$ dans l'équation 7.6. Cette EDP basée sur la totale variation ne conserve pas la courbure (Fig. 8.4) et ne peut pas prolonger les isophotes lorsque la zone à reconstruire possède une largeur l supérieure à l'épaisseur d du trait (Fig. 8.4).



(a) la diffusion par la totale variation ne préserve pas la courbure des isophotes

(b) la reconstruction est impossible sur une zone plus large que celle des traits

Figure 8.4 : Limite d'une reconstruction par la totale variation

8.2.3 La diffusion des intensités dirigée par la courbure (CDD)

Pour remédier aux inconvénients du modèle précédent, [CHA-02] propose d'ajouter l'information sur la courbure des isophotes dans l'équation de la diffusion. Cette astuce permet d'encourager la reconstruction de larges zones même si celles-ci ont une dimension supérieure à celle des traits. La CDD (Curvature Driven Diffusion) ou diffusion guidée par la courbure des isophotes (Eq. 8.4) prolonge donc les parties de fortes courbures aux extrémités des traits pour renforcer la reconstruction dans des zones larges (Fig.8.5).

$$\begin{cases} I_t = \operatorname{div} \left(\kappa^p \frac{\nabla I}{\|\nabla I\|} \right) & \text{si } (x, y) \in M \\ I_t = 0 & \text{sinon} \end{cases} \text{ avec } \begin{cases} p \geq 1 \\ \kappa = \operatorname{div} \left(\frac{\nabla I}{\|\nabla I\|} \right) \end{cases} \quad (8.4)$$

Mais la CDD ne préserve toujours pas la courbure car le prolongement des isophotes s'effectue toujours linéairement. C'est avec l'élastique d'Euler [CHA-02] que le modèle d'inpainting permet de reconstruire des courbures (Eq.8.5).

$$\begin{cases} I_t = \operatorname{div} \left((a + bk^2) \frac{\nabla I}{\|\nabla I\|} \right) & \text{si } (x, y) \in M \\ I_t = 0 & \text{sinon} \end{cases} \quad (8.5)$$

Ce nouveau modèle permet aussi de réunir la formulation de la totale variation avec $a = 1$ et $b = 0$ ainsi que la diffusion guidée par la courbure (CDD) avec $a = 0$ et $b = 1$. Cette EDP peut aussi s'écrire sous la forme $I_t = \nabla L \cdot \nabla^\perp I$ généralisant aussi le modèle de diffusion de Beltramio. Ce modèle permet de reconstruire correctement les courbures des formes (Fig.8.5).

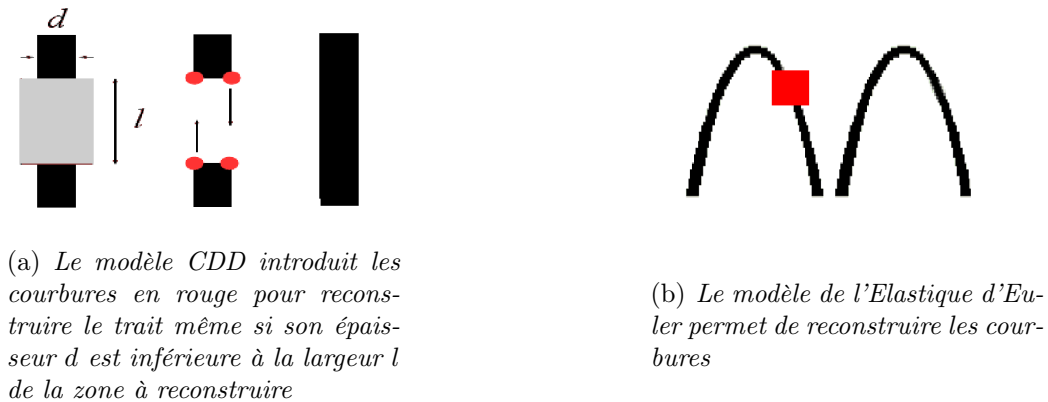


Figure 8.5 : Les avantages des modèles de diffusion dirigée par la courbure (CDD) et de l'élastique d'Euler par rapport à la méthode de diffusion basée sur la totale variation (TV).

8.3 Application aux documents

L'inpainting a d'abord été appliqué pour effacer les craquelures ou remplir les traits des textes incrustés dans les images ou les vidéos. L'information sur tout le contour du masque est utilisée pour remplir la zone à reconstruire. Dans le cas des images de documents, seule l'information autour des traits représentant une infime partie de l'information sur le contour du masque est utilisable pour la reconstruction. L'application de l'inpainting pour la reconstruction des lettres dans les images de documents donne des résultats beaucoup moins impressionnants que pour les autres applications où l'inpainting a généralement été utilisé. En effet, dans le cas de photos ou de peintures, les défauts de reconstruction sont difficilement perceptibles. En revanche, pour la reconstruction de formes de caractères, le moindre défaut de reconstruction devient très visible. La comparaison entre l'image d'origine et l'image reconstruite pourrait se faire avec les mesures habituelles en terme de SNR (Signal to Noise Ratio) ou bien en terme de MSE (Mean Square Error). Mais nous trouvons que ces mesures, généralement utilisées en traitement du signal, ne sont pas adaptées à l'appréciation des dégradations dans les images de documents. En effet ces mesures globales sont essentiellement quantitatives et non qualitatives. Elles ne tiennent pas compte de l'importance de l'information le long des traits des caractères non plus des petits changements dans la structure des formes qui modifieraient la reconnaissance de certaines lettres. Malgré une recherche active dans le domaine de l'évaluation de la qualité visuelle des images de documents

[LU-02], nous ne disposons toujours pas de réelles méthodes objectives de comparaison de la qualité d'images de documents en terme de lisibilité. Par conséquent, nous nous limiterons à mettre en évidence les différences entre l'image originale et l'image reconstruite superposée au contour de l'image d'origine. Les figures 8.6 et 8.7 montrent respectivement l'application du modèle d'inpainting de Beltramio et de celui de Chan (CDD) sur une image de synthèse de façon à mettre en évidence les défauts de réparation par soustraction. Le modèle de diffusion dirigée par la courbure semble plus adapté à la réparation des courbures des lettres par rapport au modèle de Beltramio.



Figure 8.6 : (a) Image de synthèse avec en rouge les zones à reconstruire, (b) Résultat en cours d'évolution du modèle d'inpainting de Beltramio, (c) Résultat final du modèle d'inpainting de Beltramio, (d) Différence entre l'image reconstruite et l'image originale



Figure 8.7 : (a) Image de synthèse avec en rouge les zones à reconstruire, (b) Résultat en cours d'évolution du modèle d'inpainting par CDD, (c) Résultat final du modèle d'inpainting par CDD, (d) Différence entre l'image reconstruite et l'image originale

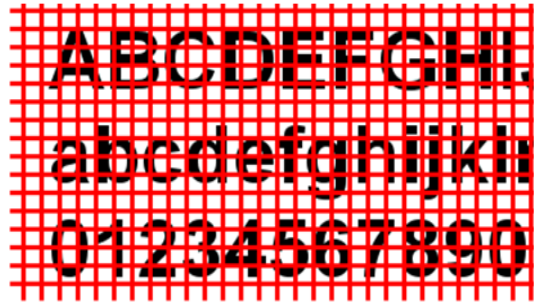
8.4 Expérimentation et évaluation

8.4.1 Expérimentation sur des images de synthèses

Pour évaluer les performances de la reconstruction, nous avons créé plusieurs images contenant des caractères imprimés de polices très différentes avec ou sans empattements, avec différents styles typographiques en italique ou en gras. Nous avons synthétisé des masques artificiels à partir de grilles aléatoires de différente taille. Nous avons augmenté graduellement les pertes d'informations en augmentant la finesse de la grille et l'épaisseur des traits jusqu'à ce que la reconstruction devienne difficile. Avec la perte d'information la plus élevée possible, le texte reconstruit par inpainting reste lisible. Cela s'explique par le fait qu'il existe dans l'alphabet Latin de nombreuses redondances d'information et des symétries qui permettent de conserver la lisibilité malgré de nombreuses altérations.

Des différents tests réalisés (Fig.8.8, Fig.8.9, Fig.8.10 et Fig.8.11), nous constatons que le modèle de reconstruction dirigé par la courbure de Chan (CDD) et plus particulièrement la généralisation avec l'élastique d'Euler (CDD2) donne globalement de meilleurs résultats que l'approche de Beltramio. Cela s'explique par la présence de nombreuses courbures dans

**ABCDEFGHIJ
 abcdefghijklr
 01234567890**

(a) *Image originale*(b) *Image avec le masque rouge*

**ABCDEFGHIJ
 abcdefghijklr
 01234567890**

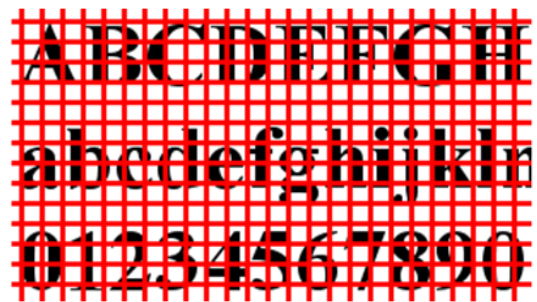
(c) *Image reconstruite par Beltramio*

**ABCDEFGHIJ
 abcdefghijklr
 01234567890**

(d) *Image reconstruite par CDD2*

Figure 8.8 : Comparaison des modèles d'inpainting sur une police Arial.

**ABCDEFGH
 abcdefghijkln
 01234567890**

(a) *Image originale*(b) *Image avec le masque rouge*

**ABCDEFGH
 abcdefghijkln
 01234567890**

(c) *Image reconstruite par Beltramio*

**ABCDEFGH
 abcdefghijkln
 01234567890**

(d) *Image reconstruite par CDD2*

Figure 8.9 : Comparaison des modèles d'inpainting sur une police Times New Roman gras.

l'alphabet Latin que le modèle de Beltramio ne peut pas reconstruire ; le modèle de Chan est, en revanche, adapté à la reconstruction des courbures. Il est clair que le taux de reconstruction dépend de la graisse de la police de caractères ; les polices de caractères imprimés en gras sont

**ABCDEFGHIJ.
abcdefghijklm
01234567890**

(a) *Image originale*(b) *Image avec le masque rouge*

**ABCDEFGHIJ.
abcdefghijklm
01234567890**

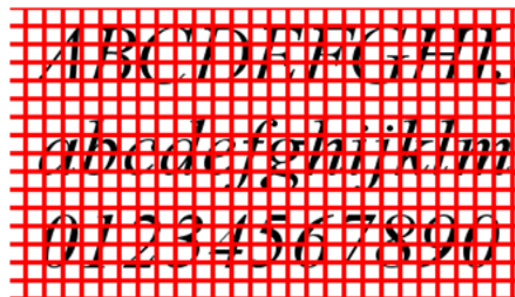
(c) *Image reconstruite par Beltramio*

**ABCDEFGHIJ.
abcdefghijklm
01234567890**

(d) *Image reconstruite par CDD2*

Figure 8.10 : Comparaison des modèles d'inpainting sur une police Times Italique gras.

**ABCDEFGHIJ.
abcdefghijklm
01234567890**

(a) *Image originale*(b) *Image avec le masque rouge*

**ABCDEFGHIJ.
abcdefghijklm
01234567890**

(c) *Image reconstruite par Beltramio*

**ABCDEFGHIJ.
abcdefghijklm
01234567890**

(d) *Image reconstruite par CDD2*

Figure 8.11 : Comparaison des modèles d'inpainting sur une police Times Italique maigre.

plus faciles à restaurer car en général l'épaisseur des traits est supérieure à celle de la zone à reconstruire (Fig.8.8, Fig.8.9, Fig.8.10). Quand la police de caractère est fine (Fig.8.11), la reconstruction devient plus difficile. Les formes plus complexes des polices de caractères

avec empattements (Fig.8.9, Fig.8.10 et Fig.8.11) sont plus difficiles à reconstruire que celles des polices sans empattement (Fig.8.8). Enfin, les polices imprimées en italiques (Fig.8.10 et Fig.8.11) posent plus de problèmes à cause de l'inclinaison des traits.

8.4.2 Expérimentation sur des images réelles de documents

Nous avons appliqué la même expérimentation sur des images couleurs de manuscrits du Moyen-Âge en augmentant progressivement la finesse de la grille et l'épaisseur des traits. La figure 8.12 montre la grille obtenue en blanc superposée à l'image originale. Malgré les pertes considérables d'information, la reconstruction par les modèles de Beltramio (Fig.8.13) et de Chan (Fig.8.14) rendent les images à nouveau lisibles. Étant donné que nous ne pouvons pas mesurer objectivement la reconstruction des images de manuscrits par OCR, nous avons superposé l'image reconstruite en gris et l'image originale représentée par ses contours. Comme nous l'avons déjà observé précédemment, le modèle de reconstruction dirigé par la courbure reconstruit beaucoup mieux les images réelles de documents que l'approche de Beltramio.



Figure 8.12 : Image originale (IRHT) et superposition du masque en blanc.

La figure 8.15 montre un exemple concret de reconstruction de document dégradé par les trous de ver qui peuvent être facilement localisés par un simple seuillage. Pour s'assurer que les bords des zones décrivent bien les courbes de niveaux de l'image originale, nous avons convexifié, lissé et enfin dilaté les zones détectées (Fig.8.16). En appliquant directement le modèle de diffusion dirigé par la courbure (CDD), toutes les zones simples ont été reconstruites. Seules les zones complexes contenant des intersections de traits ou des traits fins ne sont pas correctement reconstruites (Fig.8.17).

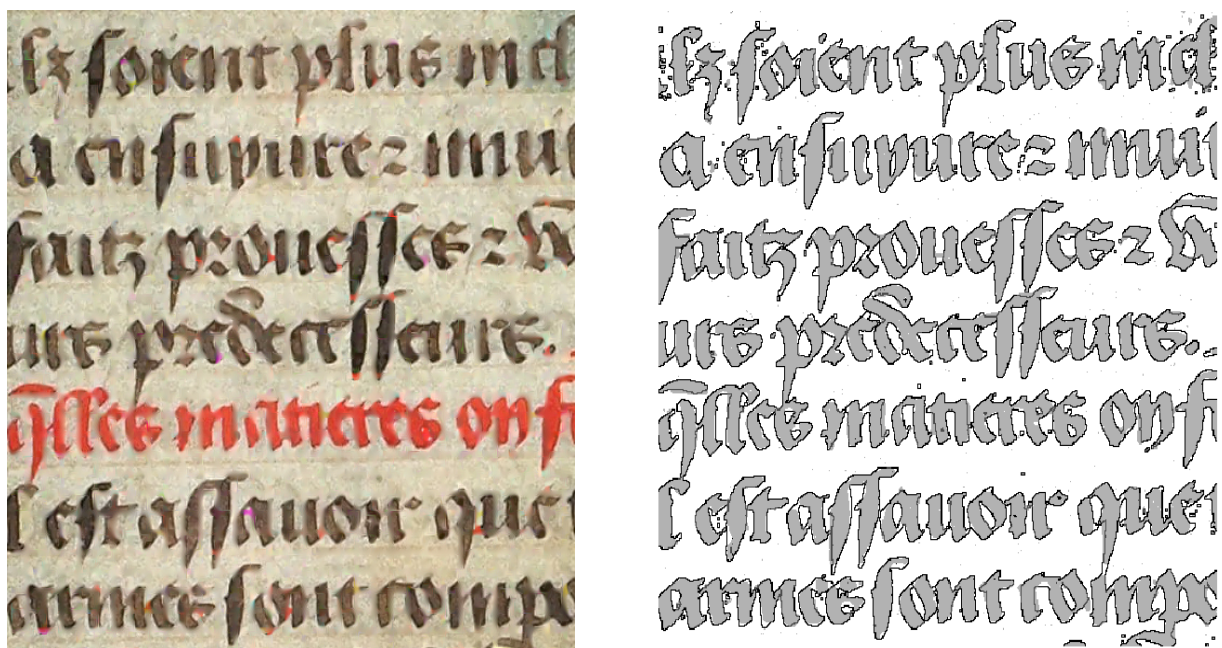


Figure 8.13 : Reconstruction par inpainting avec le modèle de Beltramio : De gauche à droite respectivement, l'image reconstruite par Beltramio et l'image reconstruite (en gris) superposée aux contours de l'image originale.

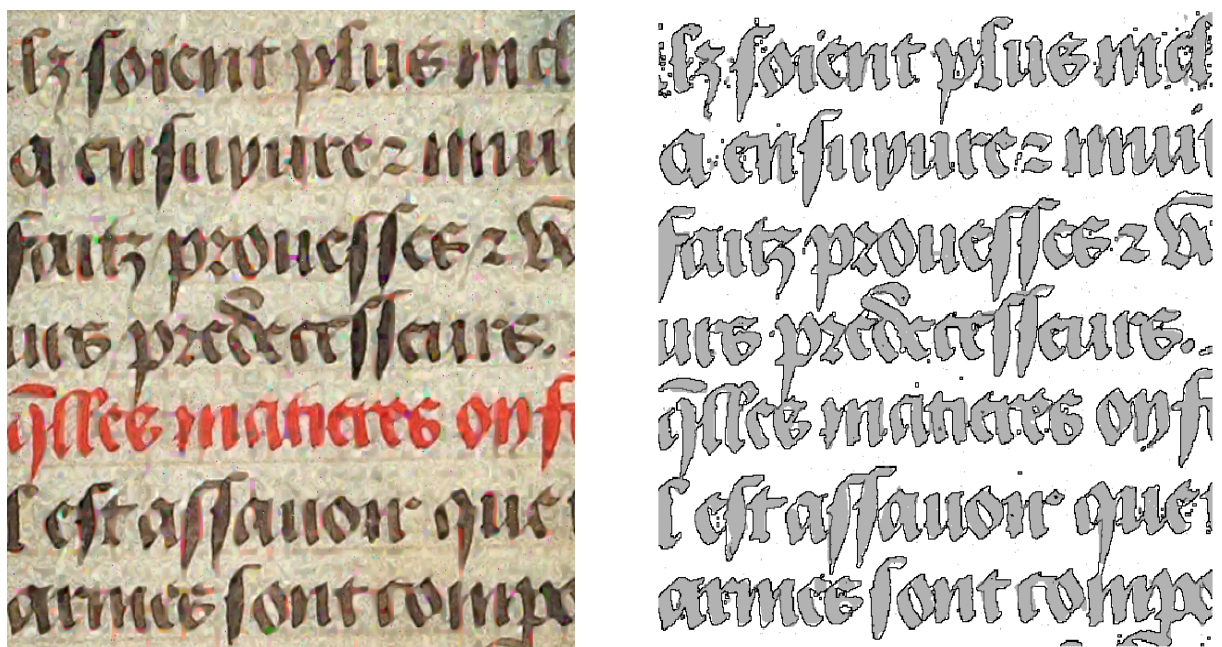


Figure 8.14 : Reconstruction par inpainting avec le modèle de Chan et l'élastique d'Euler : De gauche à droite respectivement, l'image reconstruite par le modèle CDD2 et l'image reconstruite (en gris) superposée aux contours de l'image originale.

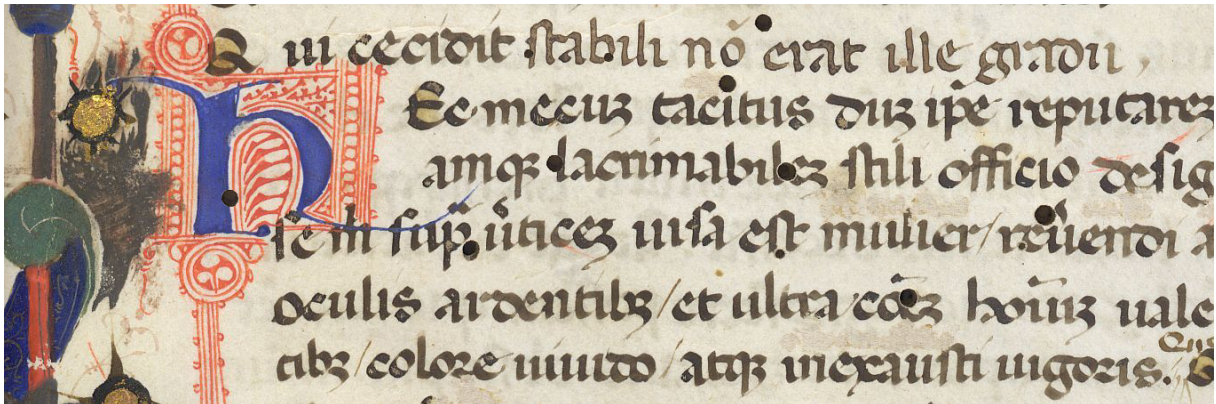


Figure 8.15 : Image dégradée par les trous de ver

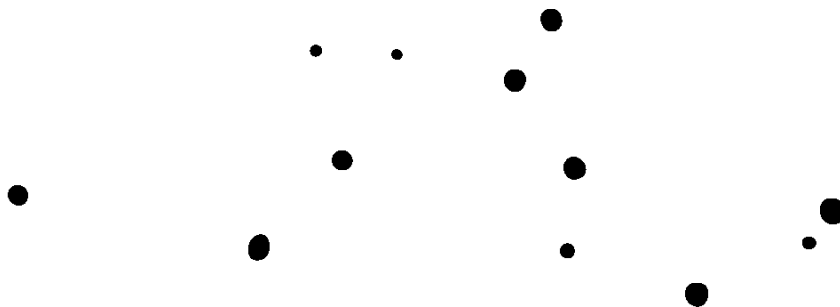


Figure 8.16 : Masque des zones à reconstruire

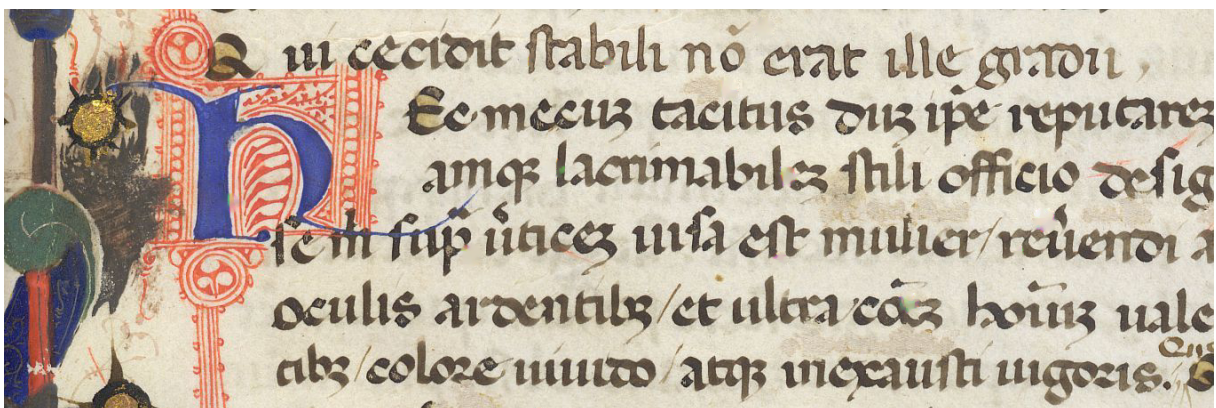


Figure 8.17 : Image reconstruite par le modèle de diffusion dirigé par la courbure (CDD).

8.5 Discussion sur les méthodes de reconstruction par inpainting

L'inpainting par EDP est applicable aux images de documents pour deux raisons :

L'inpainting classique ne marche pas sur les images texturées, ce qui est presque le cas pour les images de documents. En effet, la texture du papier n'est pas vraiment une information importante dans une image de documents et seuls les traits doivent être reconstruits. Cependant, la texture du support papier peut devenir un obstacle à la reconstruction si celle-ci possède de fortes variations et perturbe le calcul des dérivées. Pour les documents dont le support papier est fortement texturé, nous recommandons d'appliquer une forte diffusion anisotropique sur les bords de la zone à reconstruire avant de réaliser l'inpainting.

L'inpainting fonctionne correctement sur des images contenant des variations d'intensités et des contours clairement définis. C'est donc une méthode particulièrement bien adaptée aux images de documents dont les traits forment des contours bien marqués qui peuvent guider la diffusion.

Cependant, cette approche présente de sérieuses limites dont les principales sont les suivantes :

L'inpainting repose sur la continuité des isophotes dans la zone à reconstruire. Cette continuité n'existe pas toujours pour les caractères dont les traits forment des structures complexes. Il faut donc que la zone à reconstruire ne se situe ni sur les extrémités des lettres ni sur des jonctions entre les traits (Fig. 8.19).



Figure 8.18 : Limite de l'inpainting par manque de continuité dans les isophotes

La largeur de la zone à reconstruire doit être inférieure à la largeur des traits des caractères. Avec le modèle de diffusion guidée par la courbure des lignes de niveaux (CDD), on gagne un tout petit peu sur la longueur admissible en fonction de la courbure locale des extrémités. Mais, si le masque est trop large par rapport à l'épaisseur des traits, alors il n'y aura pas de réparation (Fig.8.19). En fonction de la résolution, l'inpainting est donc limitée à la réparation des rayures ou des craquelures de largeur inférieure à l'épaisseur des traits des caractères.



Figure 8.19 : Limites de l'inpainting pour un masque plus large que l'épaisseur des traits

L'inpainting de Beltramio et celui de Shan détruisent les parties irrégulières des formes. On dit qu'une forme est régulière si en chaque point de son contour on peut

entièrement placer deux boules parallèles de rayon r respectivement à l'intérieur et à l'extérieur de la forme. Une forme continue peut présenter des parties régulières et des parties irrégulières (Fig. 8.20). Les parties irrégulières d'une forme sont situées près des frontières de forte courbure ou situées à une distance inférieure à r d'une autre partie du même objet [BRE-03].

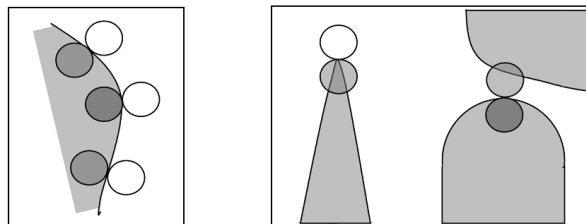


Figure 8.20 : De gauche à droite respectivement : Forme régulière que l'inpainting pourra reconstruire et Formes irrégulières que les modèles actuels ne pourront pas reconstruire

Pour le modèle de Beltramio, c'est la diffusion anisotrope qui permet de ne pas mélanger les courbes de niveaux dans la zone à repeindre. Pour le modèle de Shan, c'est la diffusion basée sur la courbure qui arrondit les angles droits et gomme les formes irrégulières comme les angles droits. Des travaux sont actuellement en cours par [KOR-06] pour reconstruire des parties irrégulières par un système de vote afin de garantir une bonne direction de diffusion. Ces directions sont calculées à partir de l'axe principal des tenseurs lissés T . Nous trouvons ces travaux très prometteurs pour la restauration des structures complexes des caractères (Fig.8.21).

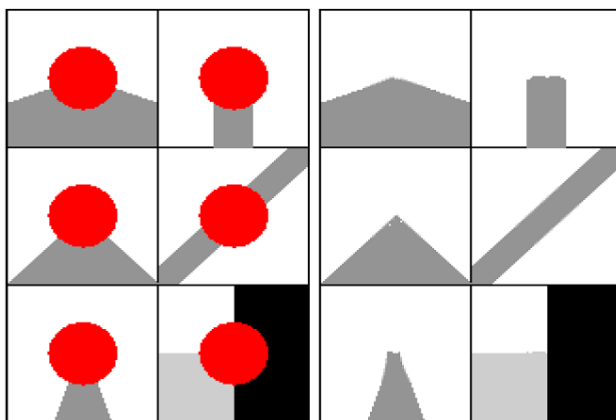


Figure 8.21 : Travaux en cours sur la reconstruction de structures complexes et des parties irrégulières des formes [KOR-06]

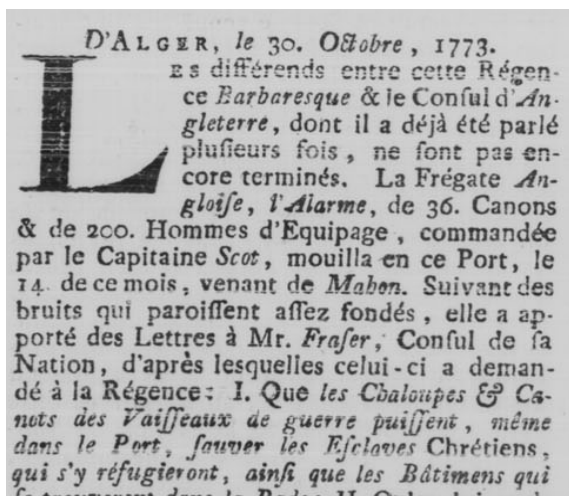
L'inpainting nécessite la définition d'un masque. Dans le cas de la restauration des images de documents, il faut trouver un traitement qui permette de localiser et segmenter correctement la zone à reconstruire. Si les trous ou les taches peuvent être assez facilement segmentés, les dégradations comme les rayures ou l'effacement partiel de l'encre, sont, en revanche, très difficiles à localiser. Si on surdimensionne le masque comme autour de la lettre «A» dans la figure 8.22, alors la reconstruction devient hasardeuse et les singularités du contour sont perdues.



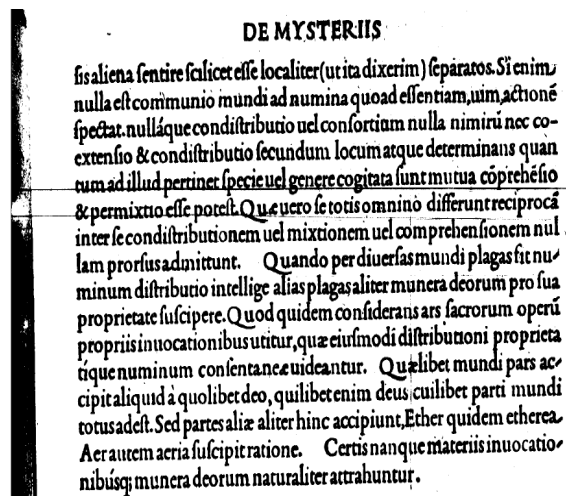
Figure 8.22 : De gauche à droite respectivement, l'image initiale, l'image dégradée, localisation des défauts à reconstruire en rouge, l'image reconstruite après 100 itérations.

8.6 Conclusion

La reconstruction par inpainting pourrait être utilisée dans de nombreuses applications comme la réparation des rayures de microfilms (Fig.8.23). Mais, nous avons été confronté aux problèmes de la détection automatique de ces rayures afin de créer les masques qui délimitent les zones à reconstruire. La segmentation des rayures dans les gazettes (Fig.8.23(a)) nécessite des méthodes morphologiques comme le chapeau haut de forme ou la transformée de Hough pour détecter des lignes dans les images binaires de microfilms (Fig.8.23(b)).



(a)



(b)

Figure 8.23 : (a) Extrait des gazettes de Leydes, (b) Extrait d'une image de microfilm de la BNF.

Chapitre 9

Proposition d'un Nouveau Modèle de Diffusion Adapté aux Images de Documents

« La clef de toutes les sciences est sans contredit le point d'interrogation ; nous devons la plupart des grandes découvertes au comment ? »

Honoré de Balzac

Résumé

Dans ce chapitre, nous présentons un modèle de diffusion que nous avons élaboré pour les images de documents et la démarche méthodologique qui a guidé sa conception. Ce modèle préserve les singularités de l'image et renforce la continuité des traits.

Sommaire

9.1	Les limites des modèles existants	134
9.2	Proposition d'un nouveau modèle de diffusion	135
9.3	Notre modèle versus les autres modèles de diffusion	139
9.4	Expérimentations et paramétrages	140
9.5	Conclusion	146

9.1 Les limites des modèles existants

Tous les modèles de diffusion respectent les fortes variations et donc les contours, mais ils modifient les détails de l'image comme les zones de fortes courbures ou les angles droits qui sont des singularités. Nous définissons les singularités de l'image par les points centrés autour des fortes variations des orientations des normales aux isophotes; les singularités d'une lettre se situent donc aux extrémités des traits, autour de fortes courbures et aux coins en angles droits ou aigus (Fig.9.1.(a)). Les images de documents présentent donc de nombreuses singularités. La perte des singularités peut transformer la lettre «t» en la lettre «l» (Fig.9.1.(b)). Les singularités contiennent donc des informations indispensables pour la lisibilité des caractères par une machine ou par l'Homme.

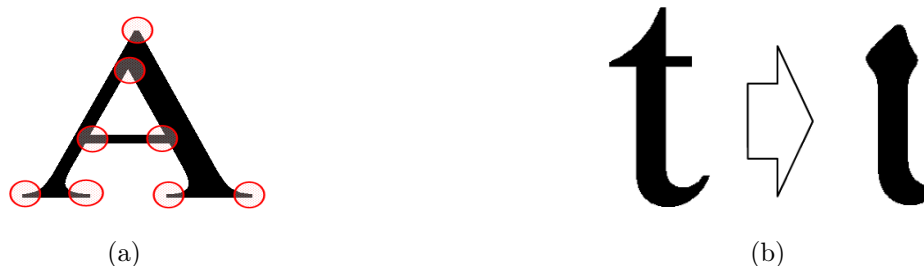
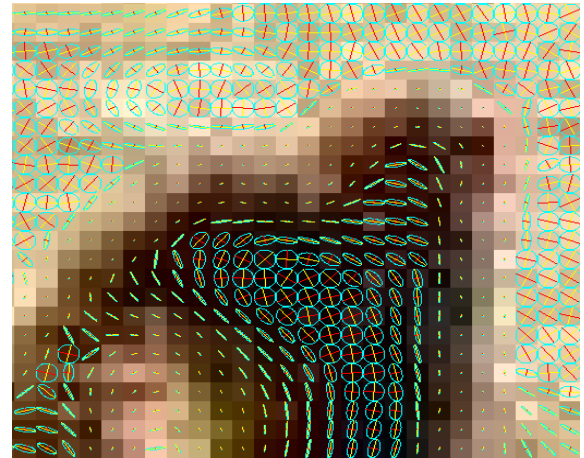
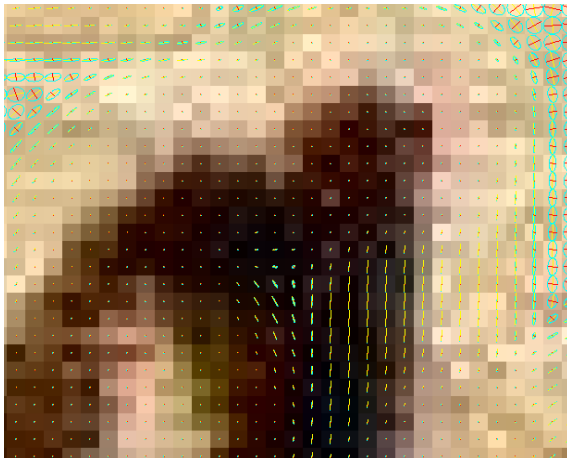
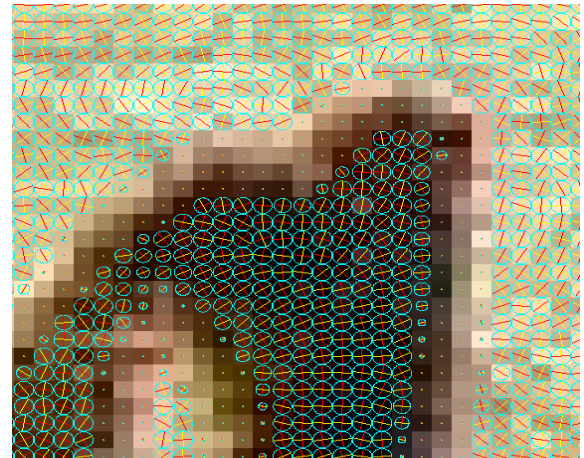


Figure 9.1 : (a) Position des singularités sur le contour d'une lettre, (b) Perte des singularités d'une image modifie les détails utiles à sa reconnaissance.

En visualisant le comportement de certaines matrices de diffusivité (Fig.9.2), nous avons réalisé qu'elles ne respectaient pas la géométrie des singularités. Les matrices de diffusion des modèles de Beltrami, Tschumperlé et Perona-Malik deviennent nulles à proximité de la moindre variation des intensités sans aucune distinction entre un contour ou un point singulier. En outre, les matrices de diffusivité de Perona-Malik sont isotropes et ne permettent pas de renforcer la continuité des traits. Le modèle de Beltrami est bien anisotrope et il peut renforcer les traits, mais ses matrices de diffusivité sont trop rapidement nulles à l'approche d'une variation. Pour le modèle de Tschumperlé, ses matrices de diffusivité ont du mal à respecter les angles droits et tendent à prendre des trajectoires arrondies autour d'une singularité. Enfin, le modèle de Weickert affiche des matrices jamais nulles et totalement anisotropes qui permettent de renforcer la cohérence des traits au détriment des singularités. Le tableau 9.1 résume les inconvénients et les avantages des modèles de diffusion sur les images de documents.

Modèle de Diffusion	Lisse les contours	Renforce la cohérence des traits	Préserve les singularités
Perona-Malik-Catte	non	non	oui
Weickert	oui	oui	non
Tschumperlé	oui	non	Modérément
Beltrami	oui	Modérément	oui

Tableau 9.1 : Récapitulatif des propriétés des modèles de diffusion sur les images de documents.

(a) Matrices de diffusivité D de Weickert(b) Matrices de diffusivité D de Tschumperlé(c) Matrices de diffusivité D de Beltrami(d) Matrices de diffusivité D de Perona-Malik**Figure 9.2** : Détails des matrices de diffusivité autour d'une singularité et d'un contour.

9.2 Proposition d'un nouveau modèle de diffusion

9.2.1 Vers un modèle de diffusion qui préserve les singularités

Nous voulons développer une matrice de diffusivité D qui préserve les singularités de l'image tout en conservant la propriété d'une diffusion qui renforce la cohérence des traits et qui lisse les contours. D'après notre analyse, le filtre idéal pour traiter les documents serait basé sur un modèle de diffusion qui reprendrait les propriétés du modèle de Perona-Malik ou de Beltrami et celui de Weickert. Nous aurions alors un filtre qui, à la fois, lisse les contours, renforce la cohérence des traits et préserve les singularités. Comme le modèle de Tschumperlé, basé sur la trace de la Hessienne est incompatible avec le modèle de Weickert basé sur la divergence du gradient, nous nous sommes donc concentré sur la combinaison entre le modèle scalaire de Perona-Malik et le modèle matriciel de Weickert. Nous savons qu'il est contradictoire de simultanément renforcer la continuité des traits coupés et de préserver les singularités et donc les extrémités des traits. Mais si les discontinuités des traits ne sont pas franches ou nettes, il est théoriquement possible de faire les deux opérations

simultanément. Nous avons étudié les différences entre le modèle de Perona-Malik et celui de Weickert. Comme nous l'avons remarqué précédemment, la formulation de Perona-Malik est un cas particulier du modèle de Weickert (Eq.7.17). Les différences se situent à deux niveaux :

- Le modèle de Weickert construit une matrice de diffusivité orientée dans la direction principale Θ_+ des tenseurs. C'est ce qui permet de renforcer la cohérence et la continuité des traits. Inversement, le modèle de Perona-Malik utilise une matrice de diffusivité sans direction particulière et donc $D = C(\|\nabla I\|) \times \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.
- Le modèle de Weickert utilise des coefficients de diffusivités $\mu_{+/-}$ toujours fortement anisotropes qui ne permettent pas de suivre la forme des singularités. Même si les matrices de Weickert deviennent presque nulles à l'approche d'une singularité, la forme anisotrope des matrices de diffusivité finissent par dégrader les singularités de l'image. Dans le modèle de Perona-Malik, les singularités sont respectées car les coefficients de diffusivités deviennent totalement nuls à l'approche d'une importante variation sans aucune distinction entre un contour ou une singularité. Le coefficient de diffusivité de Perona-Malik est une fonction décroissante de $\|\nabla I\|$ qui tend vers zéro aussi bien pour les contours que pour les singularités de l'image.

Pour tirer l'avantage des deux modèles, il suffit de conserver l'orientation de la diffusion dans la direction principale Θ_+ des tenseurs pour renforcer les traits et de modifier les coefficients de diffusivités $\mu_{+/-}$ de façon à préserver les singularités de l'image. Il faut donc trouver des coefficients de diffusivités $\mu_{+/-}$ qui s'annulent seulement à l'approche des singularités et non des contours. Il faut donc différencier les zones proches d'un contour où les valeurs de λ_+ sont grandes et λ_- faibles, des zones de singularités où les deux valeurs propres λ_+ et λ_- affichent des valeurs élevées simultanément. Pour cela, nous proposons de dissocier les valeurs λ_+ et λ_- dans les coefficients de diffusivités $\mu_{+/-}$. Nous avons retenu les coefficients de diffusivité de Perona-Malik $f_1(u) = \exp\left(-\left(\frac{u}{K}\right)^2\right)$ ou $f_2(u) = \frac{1}{1+\left(\frac{u}{K}\right)^2}$ pour qu'ils s'annulent rapidement lorsque l'argument u possède des valeurs supérieures au seuil K . Après de nombreuses tentatives, nous avons élaboré un modèle stable, qui satisfait nos attentes, à savoir un renforcement et un lissage des contours des traits et une préservation des singularités.

$$f_1 \begin{cases} \mu_- = \exp\left(-\frac{\lambda_-}{K_-}\right) \\ \mu_+ = \exp\left(-\frac{\lambda_+}{K_+}\right) \end{cases} \text{ ou } f_2 \begin{cases} \mu_- = \frac{1}{1+\left(\frac{\lambda_-}{K_-}\right)^2} \\ \mu_+ = \frac{1}{1+\left(\frac{\lambda_+}{K_+}\right)^2} \end{cases}$$

Il utilise deux seuils K_+ et K_- pour définir respectivement la diffusion à réaliser le long de l'axe principal Θ_+ et l'axe secondaire Θ_- . Sous sa forme complète, notre modèle de diffusion s'écrit donc comme suit :

$$\begin{cases} I_t = \operatorname{div}(D^* \times \nabla I) \\ I_0 = I_{\text{bruitée}} \end{cases} \text{ avec } \begin{cases} D^* = \exp\left(-\frac{\lambda_-}{K_-}\right) \times \Theta_+ \Theta_+^t + \exp\left(-\frac{\lambda_+}{K_+}\right) \times \Theta_- \Theta_-^t \\ \text{ou} \\ D^* = \frac{1}{1+\left(\frac{\lambda_-}{K_-}\right)} \times \Theta_+ \Theta_+^t + \frac{1}{1+\left(\frac{\lambda_+}{K_+}\right)} \times \Theta_- \Theta_-^t \\ \Theta_{+/-} , \quad \lambda_{+/-} \text{ vecteurs et valeurs propres de } T_\sigma^p \end{cases}$$

9.2.2 Fonctionnement de notre modèle de diffusion

Comme cette EDP est conçue pour lisser les contours, réduire le bruit des zones uniformes et préserver les singularités, nous allons expliquer le fonctionnement de cette EDP pour chacune des trois situations suivantes :

- À l'approche d'un contour $\lambda_+ \gg \lambda_-$, $\lambda_+ > K_+$ et $\lambda_- < K_-$, par conséquent, le coefficient de diffusivité μ_+ dans l'axe θ_- est proche de 0 et μ_- celui dans la direction θ_+ est proche de 1 pour une diffusion maximale. À l'approche des contours, on a des matrices de diffusivité D^* proches de celles de Weickert qui renforcent la cohérence des traits et accentuent le lissage le long des contours (Fig.9.3).

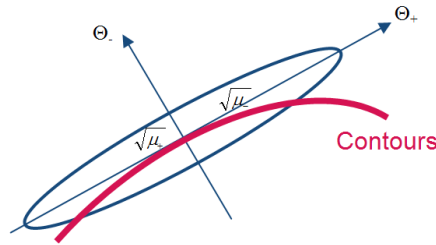


Figure 9.3 : Matrice de diffusion D^* le long d'un contour

- Dans une zone uniforme de l'image contenant de faibles variations des intensités, les tenseurs n'affichent pas une direction privilégiée et donc $\lambda_+ < K_+$ et $\lambda_- < K_-$. La matrice de diffusivité D^* devient totalement isotrope car les deux coefficients de diffusivité sont proches de 1. Nous avons une diffusion équivalente à celle de l'équation de la chaleur qui nivelle totalement les faibles variations (Fig.9.4).

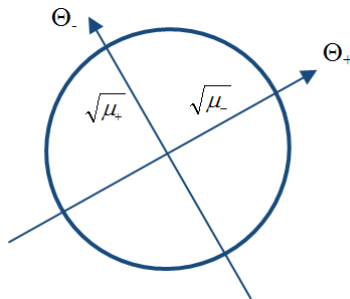


Figure 9.4 : Matrice de diffusion D^* sur les zones de faibles variations

- Dans une zone de forte singularité les valeurs (Fig.9.5), λ_+ et λ_- sont très élevées et sont supérieures aux deux seuils $\lambda_+ > K_+$ et $\lambda_- > K_-$. Les coefficients de diffusivité μ_+ et μ_- sont proches de zéro. La matrice de diffusion D^* devient une matrice presque nulle stoppant ainsi la diffusion sur les angles droits, les coins et les points de fortes courbures. C'est l'effet du modèle de Perona qui apparaît dans ce cas.

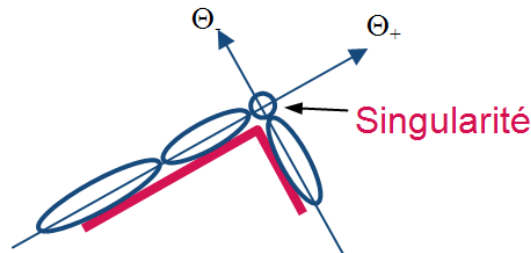
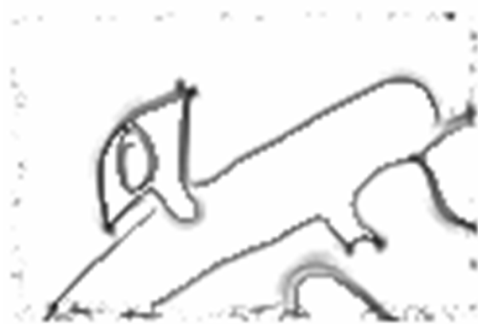


Figure 9.5 : Matrice de diffusion D^* à la rencontre d'une singularité.

Notre modèle permet, grâce aux deux paramètres K_+ et K_- , de contrôler la diffusion selon la présence d'un contour ou d'une singularité; ces paramètres K_+ et K_- permettent de définir le seuil de sensibilité nécessaire pour la détection d'une variation de l'image autour respectivement d'un contour ou d'une singularité. La figure 9.6 montre les valeurs des coefficients de diffusivité μ_+ et μ_- simultanément égales à 1 (niveaux de gris élevés à blanc) dans les zones uniformes correspondant à un fort lissage isotrope. Le long des contours, les valeurs des coefficients de diffusivité sont différentes avec $\mu_+ \approx 0$ (gris foncé à noir) et $\mu_+ \approx 1$ (gris clair à blanc) indiquant ainsi une diffusion fortement anisotrope dans la direction des contours afin de lisser et renforcer la continuité des traits. Enfin, les coefficients de diffusivité sont simultanément proches de zéro avec $\mu_+ \approx \mu_- \approx 0$ (gris foncé à noir) autour des singularités stoppant ainsi la diffusion pour les préserver.



(a) Coefficient de diffusivité $\mu_+ = \exp\left(-\frac{\lambda_+}{K_+}\right)$ le long de θ_- . Les valeurs basses (gris foncé à noir) indiquent les contours et les singularités. Les matrices de diffusivités D^* deviennent totalement anisotrope le long des contours car μ_+ est presque nul et μ_- est proche de 1.



(b) Coefficient de diffusivité $\mu_- = \exp\left(-\frac{\lambda_-}{K_-}\right)$ le long de θ_+ . Les valeurs basses (gris foncé à noir) indiquent la position des singularités. Les matrices de diffusivités D^* deviennent totalement nulles sur les singularités car les coefficients μ_+ et μ_- sont simultanément nuls.

Figure 9.6 : Affichage des valeurs des coefficients de diffusivité sur une image filtrée.

La figure 9.7 montre que nous obtenons bien une diffusion nulle autour de la singularité et que les matrices D^* suivent les angles droits. Nos matrices de diffusion D^* sont totalement isotropes sur les zones à faibles variations et deviennent rapidement anisotropes à l'approche d'un contour seulement. On remarquera que les matrices de diffusivité D^* suivent bien les contours en angle droit jusqu'à l'intersection où elles deviennent nulles pour préserver les coins.

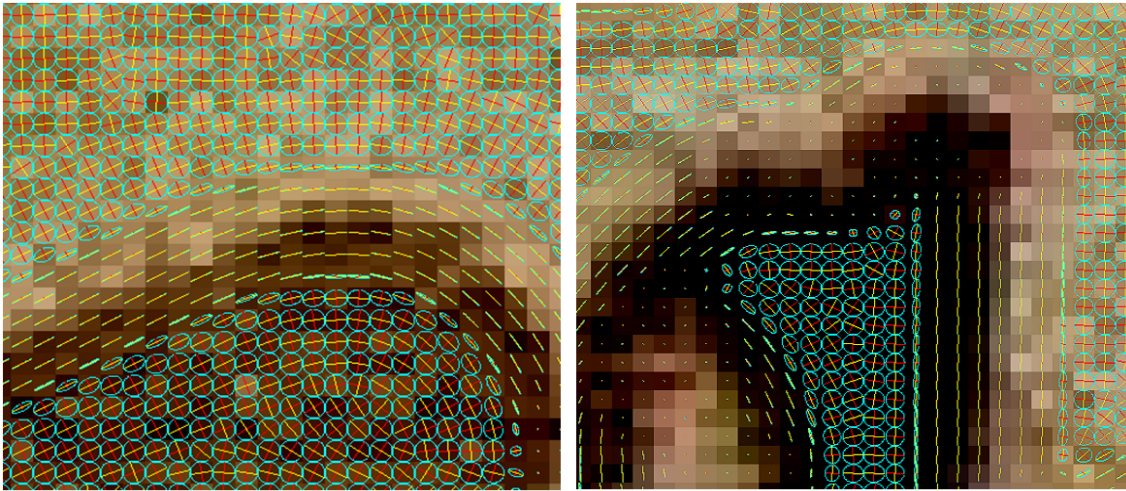


Figure 9.7 : Matrices D^* autour d'un contour et d'une singularité avec f_1

9.3 Notre modèle versus les autres modèles de diffusion

Les différences entre notre modèle et ceux de Perona-Malik, de Weickert ou de Tschumperlé se situent dans l'expression des coefficients de diffusivité $\mu_{+/-}$ qui lient les valeurs λ_+ et λ_- . En effet, les autres modèles de diffusion ont des coefficients $\mu_{+/-}(\lambda_+, \lambda_-)$ qui dépendent de la somme $\lambda_+ + \lambda_-$ ou de la différence $\lambda_+ - \lambda_-$ ou même du rapport des valeurs propres λ_+ et λ_- . L'explication de ces liaisons est simple :

- La somme $\lambda_+ + \lambda_-$ représente la magnitude au carré du gradient car $\Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ est l'écriture tenseur $T = \begin{pmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{pmatrix}$ dans la base de vecteurs propres (θ_-, θ_+) et comme la trace est invariante quelque soit le changement de base alors $\lambda_+ + \lambda_- = I_x^2 + I_y^2 = \|\nabla I\|^2$.
- La différence des valeurs propres $\lambda_+ - \lambda_-$ mesure la directivité ou l'anisotropie locale de l'image. Si l'image est localement anisotrope, alors $\lambda_+ \gg \lambda_-$ et la différence atteint de grandes valeurs. Inversement, si l'image est isotrope c'est-à-dire sans directivité précise ou bien avec deux directions de même magnitude (par exemple un coin), alors $\lambda_+ \approx \lambda_-$ et la différence atteint zéro. La différence $\lambda_+ - \lambda_-$ ne permet pas de différencier une singularité, c'est-à-dire plusieurs directions privilégiées avec de grandes valeurs de λ_+ et λ_- , d'une zone «plate» isotrope sans réelles variations et donc avec des petites va-

leurs de λ_+ et λ_- . La formule de Beltrami, qui reprend le rapport des valeurs propres λ_+ et λ_- , fait en quelque sorte une autre mesure de l'anisotropie locale de l'image.

Dans notre modèle, nos coefficients séparent clairement les valeurs propres λ_+ et λ_- . Par conséquent, les coefficients de diffusivité sont des fonctions d'une seule valeur propre μ_+ (λ_+) et μ_- (λ_-). Contrairement aux autres modèles de diffusion qui effectuent une combinaison des valeurs propres, notre proposition permet de distinguer le cas où $\lambda_+ > K_+$ et $\lambda_- < K_-$ le long d'un contour, du cas d'une singularité où $\lambda_+ > K_+$ et $\lambda_- > K_-$. Ainsi, en séparant le rôle des valeurs propres, nous pouvons distinguer le modèle de diffusion pour une singularité par rapport au modèle de diffusion pour un contour ou une zone homogène.

Enfin, il est difficile de comparer notre modèle avec celui de Beltrami (Tab.7.1) qui ne permet pas d'établir directement une relation entre les coefficients de diffusivité μ_+ et μ_- à cause de la pondération $\frac{1}{\sqrt{(1+\lambda_+)(1+\lambda_-)}}$ devant le terme de la divergence.

Notre modèle est aussi différent du filtre directionnel proposé par Terebes [TER-02] pour restaurer les images de documents imprimés et en particulier les gravures en taille douce de la renaissance. Les auteurs proposent un filtre de lissage des traits à partir d'une variante de Perona-Malik en calculant le gradient dans la direction de θ qui a des liens évidents de parenté avec le calcul de l'axe principal θ_+ des tenseurs lissés T_σ^ρ de Weickert.

$$\begin{cases} I_t &= \operatorname{div} (C (\|G_\sigma \otimes \nabla I\|) \times \nabla I^\theta) \\ I_0 &= I_{\text{bruitée}} \end{cases} \quad \text{avec} \quad \nabla I^\theta = \begin{pmatrix} I_x \cos \theta \\ I_y \sin \theta \end{pmatrix} \quad (9.1)$$

Ce filtre ne correspond qu'à un cas particulier de notre modèle avec les coefficients $\mu_+ = \mu_- = \exp - \left(\frac{\|G_\sigma \otimes \nabla I\|}{K} \right)^2 = \exp - \left(\frac{\lambda_+ + \lambda_-}{K^2} \right)$. Les auteurs de [TER-02], travaillant aussi dans le domaine du document, avaient les mêmes objectifs que nous, à savoir maintenir la séparabilité entre les traits parallèles et garantir les structures. Notre modèle de diffusion, plus complet, préserve aussi très bien la séparation des traits et les structures, mais en plus il renforce la continuité des traits. En effet, le modèle de diffusion de [TER-02], basé sur un modèle de diffusion scalaire de Perona-Malik et non sur un modèle de diffusion matricielle de Weickert, ne peut pas renforcer la continuité des traits. Les résultats présentés dans la section suivante (Fig.9.13 et Fig.9.14) montrent les différences entre notre modèle de diffusion et celui du filtre directionnel de [TER-02].

9.4 Expérimentations et paramétrages

K_+ est le seuil de préservation de la variation principale (contours), plus il est grand, moins les contours d'objets peu contrastés seront préservés. Cela s'explique par le fait que λ_+ est grand à l'approche d'un contour quelque soit λ_- . Ainsi, sur la figure 9.8, la terminaison du «a» disparaît car la variation des contours dans cette zone est inférieure au seuil K_+ .

K_- est le seuil de préservation des singularités, plus il est grand, moins on préserve les singularités de l'image. En effet, λ_- est grand lorsque l'on s'approche d'une singularité car cette valeur indique une structure plus complexe qu'une simple ligne ou une zone uniforme.

Si $K_- \gg K_+$, alors on a un effet de diffusion extrême dans la direction des traits qui renforce leurs continuités. C'est l'effet de «renforcement de la cohérence» de Weickert.



Figure 9.8 : Effet de K_+ sur la préservation des contours, de gauche à droite respectivement, l'image originale : l'image traitée avec $\sigma = 0.5$, $\rho = 1.5$, $K_+ = 60$, $K_- = 600$ et l'image traitée avec $\sigma = 0.5$, $\rho = 1.5$, $K_+ = 600$, $K_- = 60$.

Si $K_+ \gg K_-$, alors on a un effet de préservation extrême des singularités de l'image et une disparition des contours de faibles variations. Mais, si la valeur de K_+ est trop élevée, alors les contours ne seront plus préservés.

La figure 9.9 montre le comportement de notre modèle de diffusion suivant les valeurs de K_+ et de K_- avec le choix de la fonction f_1 .

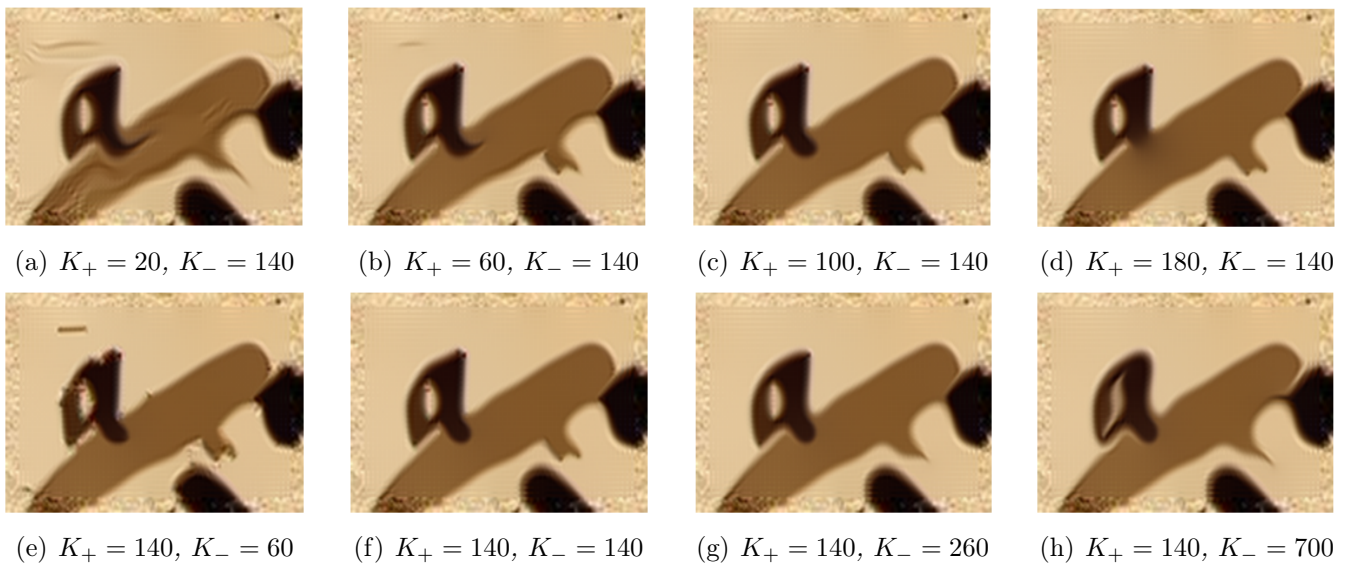


Figure 9.9 : Effets des paramètres K_+ et K_- sur la diffusion avec $\sigma = 0.5$, $\rho = 1.5$ avec f_1

En comparaison avec les autres modèles de diffusion (Fig. 9.10), nous remarquons que notre modèle maintient les singularités et restaure beaucoup mieux les formes des caractères.

9.4.1 Résultats sur les images de documents

La figure 9.11 montre que l'application de notre filtre avec un seuil $K_+ = K_- = 128$ contribue à la préservation simultanée des singularités et des contours. Sur l'image restaurée, les traits extrêmement fins en prolongement des terminaisons des lettres sont effacés. Cela s'explique par des valeurs des seuils K_+ et K_- trop élevées, choisies pour ne pas faire apparaître les traits de l'arrière plan qui servent de guide au copiste pour aligner ses lettres.

En baissant le seuil de détection des contours K_+ à 60, on renforce les traits fins des terminaisons (Fig.9.12.a). Avec un seuil K_+ trop bas, les guides d'alignements apparaissent

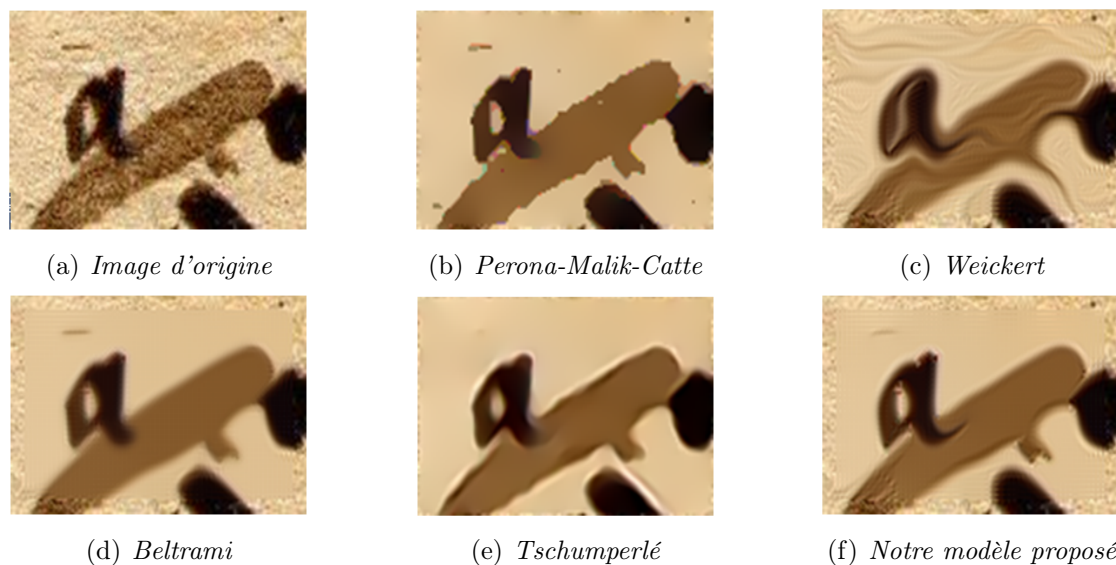


Figure 9.10 : Comparaison de différents modèles de diffusion

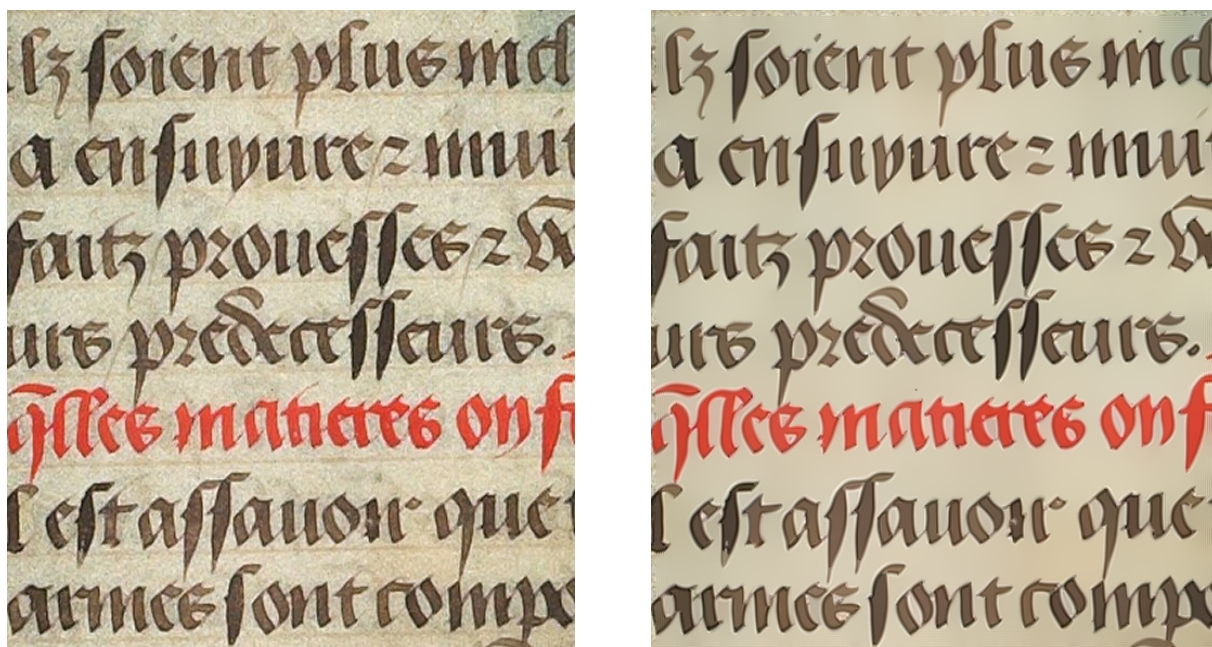


Figure 9.11 : L'image originale et l'image traitée avec f_1 pour $\sigma = 0.5$, $\rho = 1.5$, $K_+ = K_- = 128$

et les traits fins de faibles contrastes vont s'étendre au-delà des limites d'origine. L'effet de renforcement des lignes va se poursuivre en prolongeant ces traits fins car leurs extrémités ne représentent pas une véritable singularité, contrairement aux frontières des caractères (Fig.9.12.b). Comparé au filtre de Weickert, notre filtre conserve les structures et la topologie des caractères dont les contours possèdent des singularités qui bloquent le processus de diffusion et assure l'extension des traits dans leurs prolongements. En revanche, les traits fins qui ne sont pas terminés par une singularité subissent des déformations dans la direction initiale du trait. Sans la préservation des structures par le blocage de la diffusion autour des

singularités, nous aurions une diffusion équivalente à celle de Weickert et qui déforme les lignes jusqu'à les idéaliser sous la forme de traits parallèles.

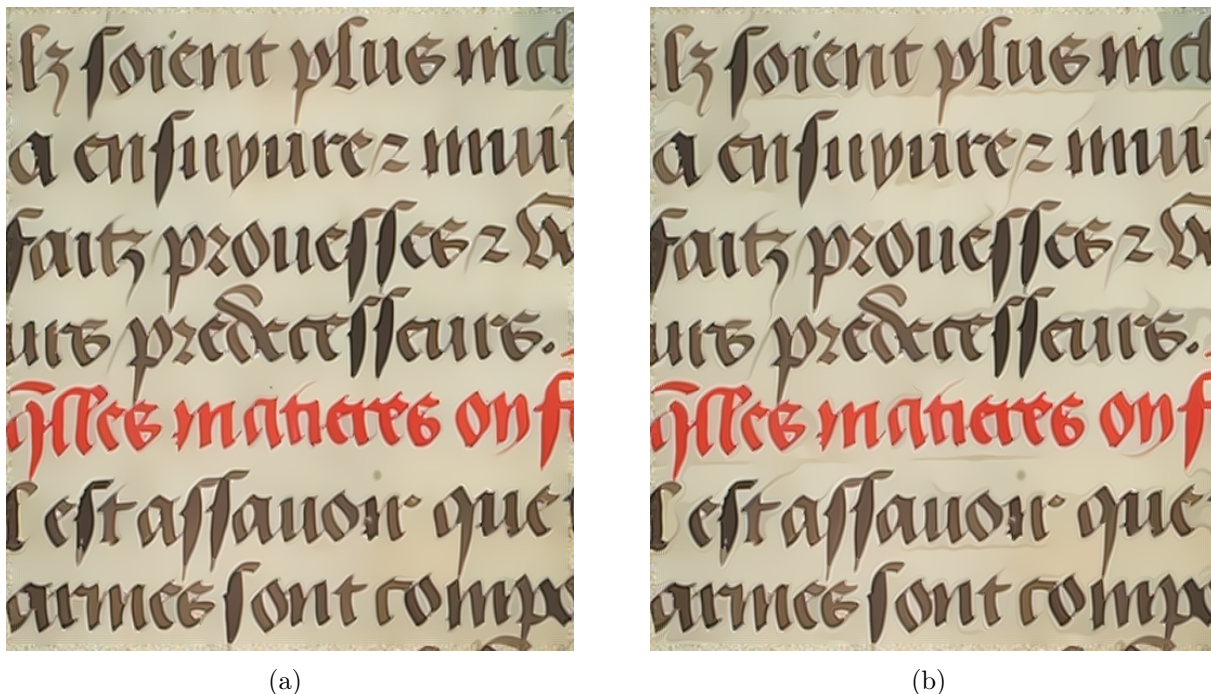


Figure 9.12 : (a) $\sigma = 0.5$, $\rho = 1.5$, $K_+ = 60$, $K_- = 128$ et f_1 . L'image montre une préservation correcte des terminaisons fines des traits, (b) $\sigma = 0.5$, $\rho = 1.5$, $K_+ = 20$, $K_- = 128$ et f_1 . La réglure apparaît et les terminaisons se prolongent dans leurs continuités par un effet de fibre comparable au modèle de Weickert.

Sur des images des illustrations en taille douce des documents imprimés (Fig.9.13 et Fig.9.14), les tests montrent bien que notre filtre et le filtre directionnel proposé par [TER-02] sont les seuls capables de préserver la continuité des traits, de lisser les contours et de maintenir les traits parallèles. En revanche, seul notre filtre apporte, en plus, une préservation des singularités et donc des détails des traits quand ils ne sont pas droits (voir la fresque au centre, les chapiteaux des colonnes, des intersections des traits dans la figure 9.13).

Sur des images très dégradées (Fig.9.15) où la couleur du texte manuscrit n'est pas très différente de celle du support papier, notre modèle et celui de Beltrami conservent mieux la topologie des caractères par rapport aux autres modèles.

D'autres tests réalisés sur d'autres images dégradées de documents (Fig.9.16, Fig.9.17, Fig.9.18, Fig.9.20 et Fig.9.21) prouvent les performances de notre filtre et son aptitude à mieux préserver les singularités et les contours tout en lissant les zones bruitées. Un meilleur rendu est assuré après l'application de notre filtre.

Le modèle de diffusion proposé permet de corriger les effets d'escaliers dus à la diminution de la résolution. Les résultats apportés, sur la figure 9.22, suite à l'application de notre modèle apporte une nette amélioration par rapport aux résultats obtenus respectivement par le modèle de Perona-Malik-Catté et celui de Tschumperlé. La figure 9.23 montre aussi l'apport de notre filtre pour la correction des dégradations numériques comme les artefacts de la compression JPEG.

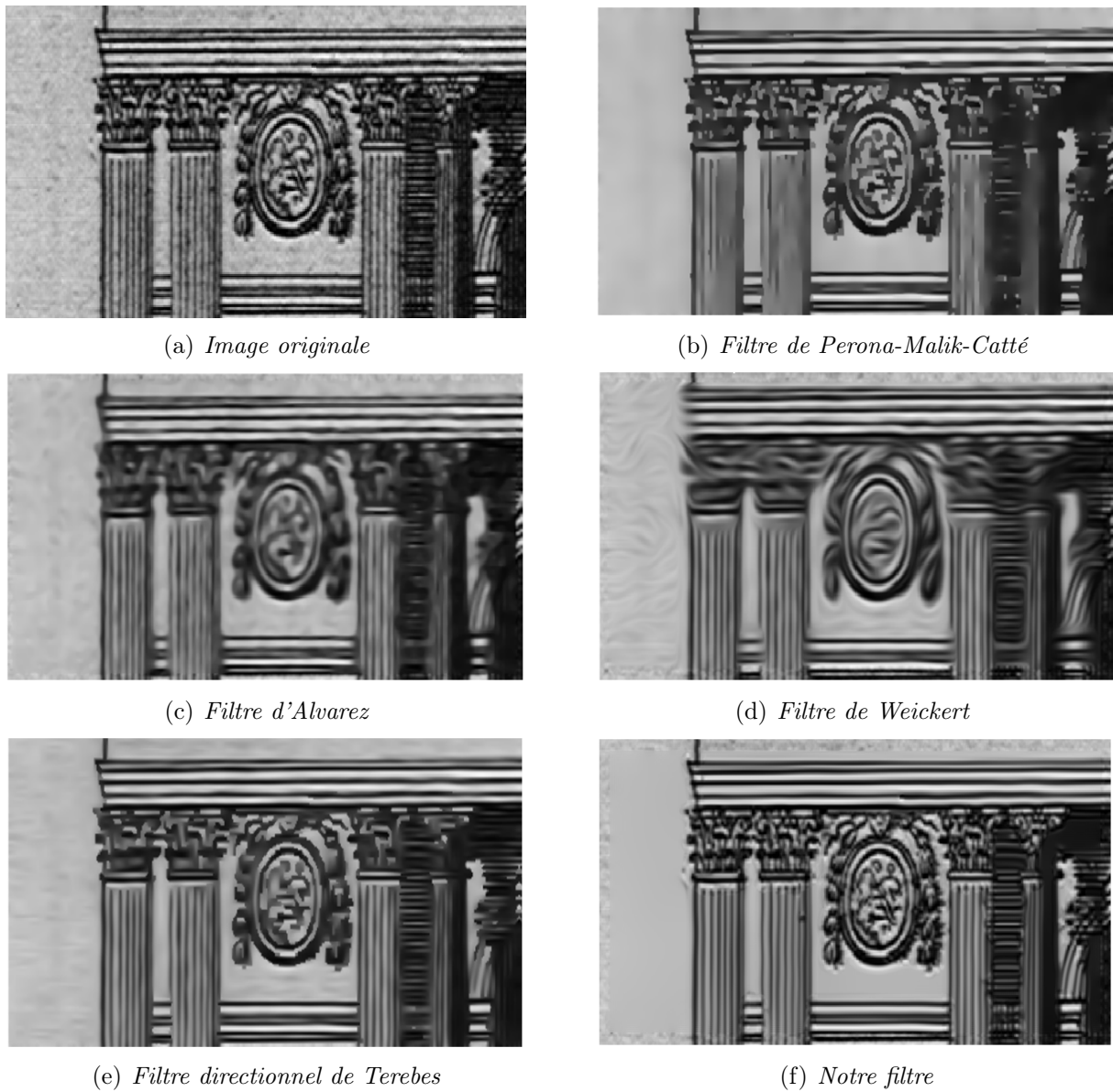


Figure 9.13 : Comparaison des différents modèles de diffusion. Pour notre filtre, les paramètres choisis sont $\sigma = 0.5$, $\rho = 3$, $K_+ = K_- = 20$

9.4.2 Application aux autres images

Notre filtre, créé au départ pour restaurer les formes dégradées des caractères et pour garder leurs topologies tout en renforçant les lignes, est assez générique pour s'appliquer aux images naturelles de différentes origines. Le rendu visuel des images traitées est bon car les singularités dans les images naturelles sont aussi très importantes (Fig. 9.24). Le renforcement des traits permet de corriger les artefacts JPEG (Fig.9.24.a), de conserver les structures fines (Fig.9.24.b) et de renforcer les structures des lignes tout en effectuant une forte diffusion dans les zones non-texturées et bruitées (Fig.9.24.c).

Sur l'image de l'empreinte digitale (Fig.9.25), notre filtre renforce bien les traits et se comporte presque de la même manière que le filtre de renforcement des cohérences des structures

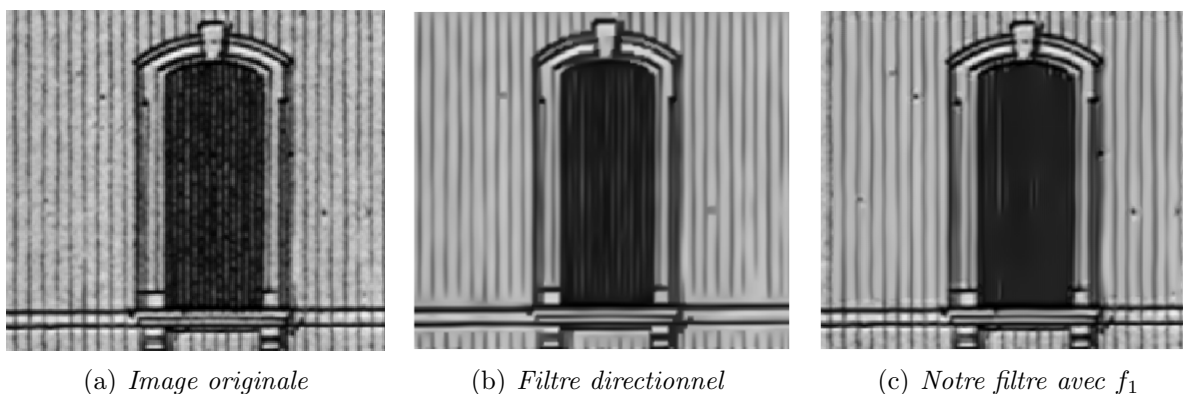


Figure 9.14 : Comparaison entre le filtre directionnel proposé dans [TER-02] et notre filtre de diffusion

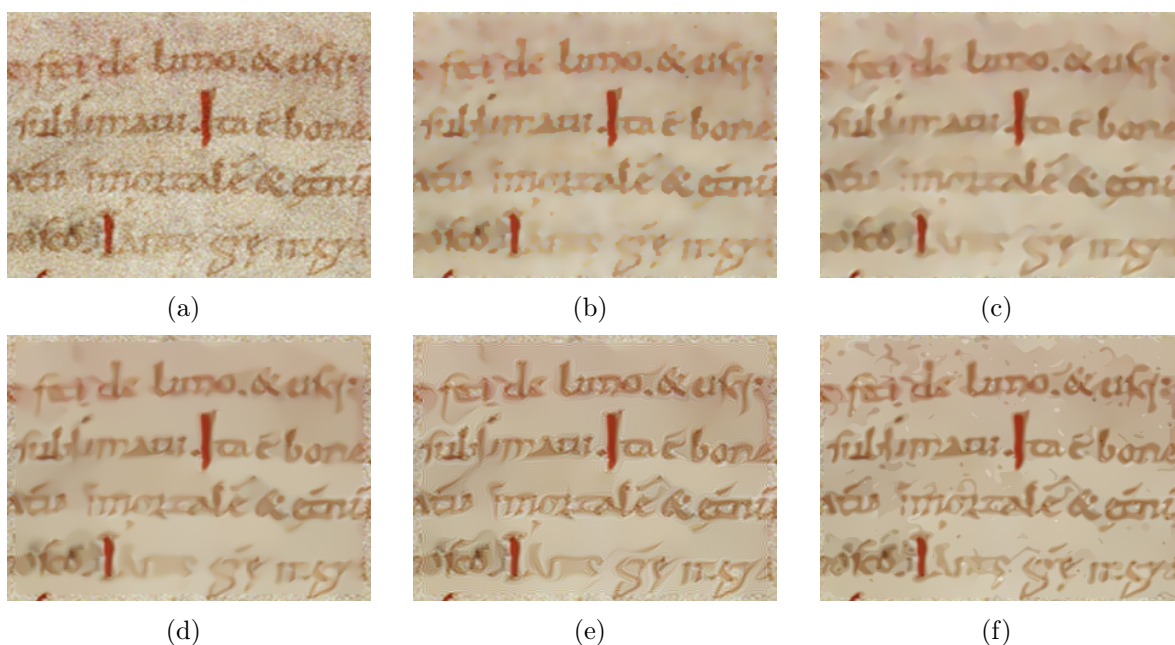


Figure 9.15 : Comparaison des différents modèles de restauration sur une image fortement dégradée de l'IRHT : (a) Image originale (IRHT), (b) Perona-Malik-Catté $k = 10$, $\sigma = 0.5$, (c) D. Tschumperlé, (d) Beltrami, (e) Notre filtre pour $\sigma = 0.5$, $\rho = 3$, $K_+ = 10$, $K_- = 20$ et f_2 , (f) Notre filtre pour $\sigma = 0.5$, $\rho = 1$, $K_+ = K_- = 20$ et f_2

de Weickert. La seule différence réside dans le maintien des véritables discontinuités des traits car elles sont terminées par des singularités qui bloquent la diffusion et la prolongation des traits.

Certes notre filtre de diffusion préserve les singularités et renforce la cohérence des traits, mais il ne convient pas aux images où le bruit possède une magnitude supérieure à celle des contours des objets. On ne peut pas, à la fois, chercher à préserver les singularités et enlever un bruit impulsionnel de forte magnitude créant des singularités dans l'image. Par conséquent, le filtre proposé est adapté aux images raisonnablement bruitées et contenant des traits à renforcer et des contours à lisser.



Figure 9.16 : De gauche à droite respectivement : l'image originale (source Châtillon-sur-chalaronne) et l'image traitée par notre filtre pour $\sigma = 0.5$, $\rho = 1.5$, $K_+ = 40$, $K_- = 140$

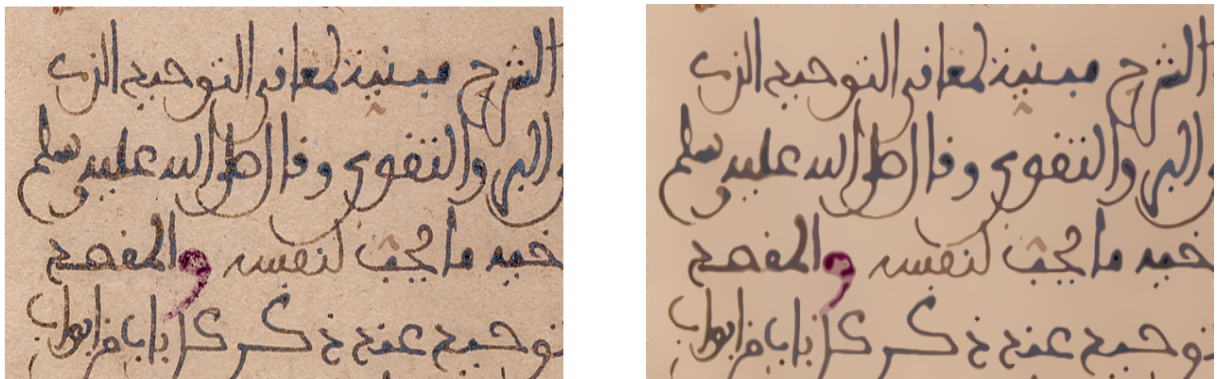


Figure 9.17 : de gauche à droite respectivement : l'image originale (source Tombouctou) et l'image traitée par notre filtre pour $\sigma = 0.5$, $\rho = 1.5$, $K_+ = 40$, $K_- = 140$

9.5 Conclusion

Nous avons proposé un modèle de diffusion qui cherche impérativement à préserver les singularités des images pour conserver la lisibilité des textes. Ce modèle de diffusion est une synthèse de plusieurs modèles existants et il cherche à tirer les avantages de chacun d'entre eux. Ainsi, notre modèle permet de filtrer les contours de l'image et de renforcer la continuité des traits tout en préservant les singularités et donc les structures complexes. Il est donc particulièrement adapté au traitement des images de documents. Mais, il est aussi suffisamment générique pour traiter avec succès des images naturelles. En effet, presque tous les modèles de diffusion actuels préservent essentiellement les contours et ne cherchent pas vraiment à adapter la diffusion pour les singularités afin de les conserver. Or les singularités jouent un rôle tout aussi important que celui des contours pour la lisibilité des textes comme pour la qualité visuelle d'une image quelconque. Le domaine d'application de ce filtre n'est donc pas limité au traitement des documents.



Figure 9.18 : de gauche à droite respectivement : l'image originale (source British Libraries) et image traitée avec $\sigma = 0.5$, $\rho = 0.5$, $K_+ = 20$, $K_- = 80$.



Figure 9.19 : de gauche à droite respectivement : l'image originale et l'image traitée par notre filtre pour $\sigma = 0.5$, $\rho = 0.5$, $K_+ = 1$, $K_- = 4$.



Figure 9.20 : De gauche à droite respectivement : l'image originale (source IRHT) et l'image traitée par notre filtre pour $\sigma = \rho = 0.5$, $K_+ = K_- = 30$

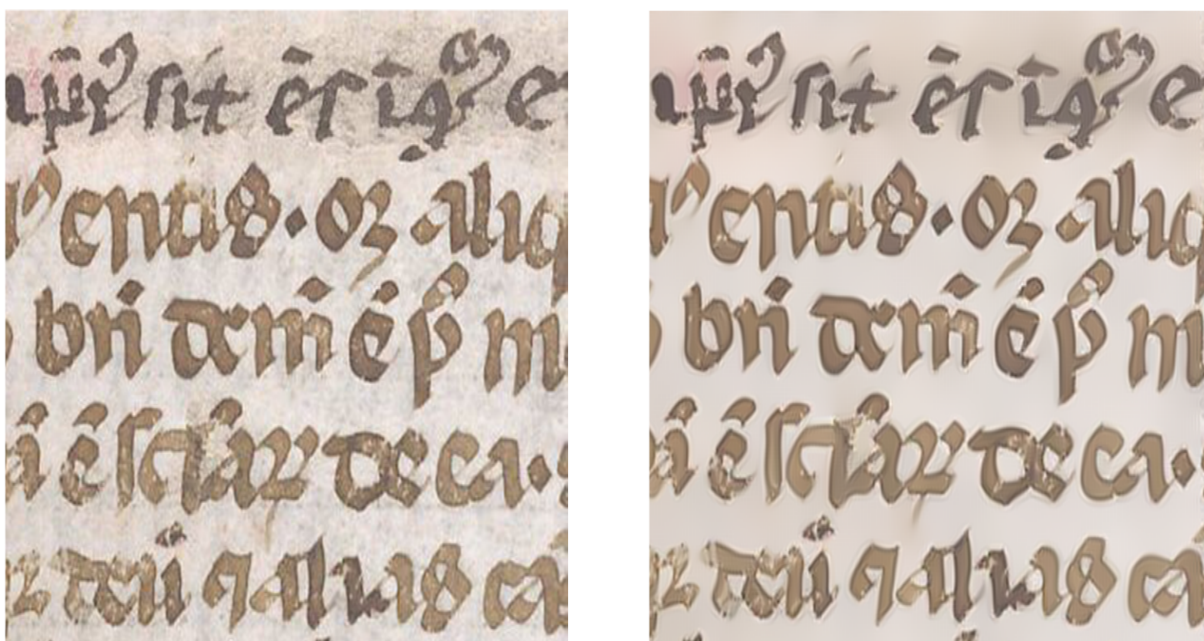


Figure 9.21 : De gauche à droite respectivement : l'image originale (source British Libraries) et l'image traitée avec $\sigma = \rho = 0.5$, $K_+ = K_- = 60$



Figure 9.22 : (a) image originale en très basse résolution, (b) image traitée par le filtre de Perona-Malik-Catté, (c) image traitée par le filtre de Tschumperlé, (d) image traitée par notre filtre pour $K_+ = 100$, $K_- = 500$



Figure 9.23 : De gauche à droite respectivement : l'image originale fortement compressée par JPEG et l'image restaurée avec $K_+ = 20$, $K_- = 300$



Figure 9.24 : Test de notre filtre sur des images naturelles : de gauche à droite respectivement : l'image originale et sa version traitée. Les valeurs utilisées de haut en bas respectivement ; $\sigma = 0.5$, $\rho = 1.5$, $K_+ = K_- = 20$; $\sigma = 0.5$, $\rho = 1.5$, $K_+ = K_- = 40$; $\sigma = 1$, $\rho = 2$, $K_+ = K_- = 180$



Figure 9.25 : De gauche à droite respectivement : l'image originale et sa version traitée avec $\sigma = 0.5$, $\rho = 4$, $K_+ = 20$, $K_- = 140$



Figure 9.26 : De haut en bas, l'image originale et l'image traitée avec $\sigma = 0.5$, $\rho = 1.5$, $K_+ = 2$, $K_- = 8$.

Quatrième partie

Dégradations Non-Uniformes : Restauration par une Approche Unificatrice

Chapitre 10

Segmentation/Diffusion : Vers un Cadre Commun

« Le hasard est le plus grand romancier du monde ;
pour être fécond, il n'y a qu'à l'étudier. »

Honoré de Balzac

Résumé

Dans cette dernière partie, nous allons démontrer qu'il existe un cadre commun entre la segmentation par Mean-Shift et les méthodes de diffusion. Le cadre proposé permet de démontrer qu'une approche de segmentation colorimétrique peut très bien se combiner avec les méthodes spatiales de diffusion et produire de résultats meilleurs que ceux que l'on obtiendrait en appliquant chacune des méthodes séparément.

Sommaire

10.1 Introduction	156
10.2 Étude comparative entre le Mean-Shift global et la diffusion	156
10.3 Proposition d'une combinaison coopérative	158
10.4 Combinaison coopérative versus combinaison séquentielle	159
10.5 Conclusion	162

10.1 Introduction

Dans ce chapitre, nous abordons la mise en place d'un cadre formel commun dédié au couplage de l'approche locale de diffusion avec une approche globale de segmentation, celle basée sur le Mean-Shift global ; l'objectif est de démontrer que ces deux approches peuvent s'écrire sous une même formulation et que cette combinaison coopérative donne des résultats meilleurs que ceux que l'on obtiendrait en appliquant les méthodes séparément. Du point de vu de la restauration des images de documents, cette coopération permet aussi de restaurer simultanément les dégradations du support papier et les déformations des traits des caractères.

Nous fournissons une première proposition de couplage basée sur le modèle de diffusion scalaire de Perona-Malik dont la formulation simple est très proche de celle du Mean-Shift. Le cadre commun, une fois validé pour la diffusion scalaire, peut être aisément étendu pour la diffusion tensorielle. Pour parvenir à définir une approche unificatrice regroupant segmentation et diffusion, nous proposons de commencer par réécrire ces deux processus sous une forme discrète.

10.2 Étude comparative entre le Mean-Shift global et la diffusion

Le Mean-Shift global se base uniquement sur la distribution des couleurs de l'image entière, alors que la diffusion est une méthode locale qui utilise les informations spatiales entre les pixels. En effet, la diffusion anisotrope permet de lisser le bruit tout en préservant la position et le contraste des contours. En plus des propriétés de réduction du bruit, la diffusion scalaire de Perona converge aussi vers une image segmentée. Généralement, elle peut être utilisée comme une étape de pré-segmentation ou de segmentation même. Toutefois, lorsque les régions de l'image ne sont pas constantes en intensité et qu'elles sont assez fortement bruitées, le modèle de diffusion de Perona-Malik tend à sur-segmenter l'image.

Comme la méthode du Mean-Shift et celle de la diffusion de Perona-Malik aboutissent toutes les deux à une image segmentée à partir d'informations différentes, il nous est apparu évident que nous devons chercher à les combiner ensemble. Il existe une complémentarité entre ces deux approches dont nous pouvons tirer profit pour améliorer le résultat final de la segmentation. Nous allons commencer par réécrire ces deux processus sous une forme discrète. Dans un premier temps, nous rappelons la formulation discrète de la diffusion scalaire de Perona-Malik et celle Mean-Shift global. Puis, après avoir mis en évidence les points communs et les différences, nous proposons une combinaison simple sous une forme discrète des deux approches.

10.2.1 Formulation discrète de la diffusion scalaire de Perona-Malik

Dans cette section, nous allons présenter la version discrète de l'équation de Perona et Malik. Rappelons que cette équation mène une diffusion anisotrope scalaire. Elle représente une variation de l'équation de diffusion de la chaleur, pour laquelle le flux de diffusion est pondéré par une fonction de diffusion, permettant de limiter des échanges d'intensité dans les zones de fort gradient. L'équation de Perona-Malik n'est autre que le calcul pour chaque

pixel de la moyenne de son voisinage pondérée par la norme du gradient. La version discrète de cette équation s'écrit donc, pour tout pixel p de l'image I et à une itération n donnée, sous la forme :

$$\vec{I}^{(n+1)}(p) = \vec{I}^{(n)}(p) + \overrightarrow{VDiff}^{(n)}(p) \quad (10.1)$$

Nous notons $\overrightarrow{VDiff}^{(n)}(p)$ le coefficient de vélocité de la diffusion au point p et à l'itération n . Ce coefficient mesure la différence entre les valeurs des intensités de l'image pour deux itérations successives. Il est donné par :

$$\overrightarrow{VDiff}^{(n)}(p) = \sum_{h \in N_s(p)} \left(C^* \left(\frac{\overrightarrow{\nabla I_n^n}(h)}{K} \right) \times \overrightarrow{\nabla I_n^n}(h) \right) \quad (10.2)$$

La fonction C^* est une fonction normalisée qui dépend de la norme du gradient. Elle représente la fonction clé du processus de diffusion de Perona-Malik. C'est elle qui stoppe la diffusion dans les zones de fort gradient. Cette fonction vérifie les propriétés suivantes : $C^*(0) = 1$, $C^*(x) \geq 0$ et $C^*(x) \xrightarrow{x \rightarrow \infty} 0$.

Elle peut être formulée comme suit :

$$C^*(x) = \frac{1}{|N_s|} \exp(-\|x\|^2) \quad (10.3)$$

K est une constante arbitrairement définie. Elle permet de distinguer les zones à faible gradient de celles à fort gradient. Autrement dit, les points dont la norme du gradient est inférieure à K sont assimilés à du bruit alors que les autres sont considérés comme étant des contours.

$$\overrightarrow{\nabla I_n^n}(h) = \vec{I}^{(n)}(p) - \vec{I}^{(n)}(h) \quad \forall h \in N_s(p) \quad (10.4)$$

$\overrightarrow{\nabla I_n^n}$ est une mesure du gradient de l'image donnée par la différence entre l'intensité d'un point courant p et l'intensité de son voisin proche h . Ce voisin est un élément du voisinage spatial $N_s(p)$ qui regroupe les quatre voisins connectés et entourant le pixel courant p . Son cardinal est $|N_s|$ et il est défini par :

$$N_s(p) = \{h/d_4(p, h) \leq 1\} \quad (10.5)$$

10.2.2 Formulation discrète simplifiée du Mean-Shift global

Le Mean-Shift global est une procédure itérative de recherche de modes ou maxima locaux. Son objectif est de faire converger les points d'une image, ayant une distance colorimétrique en dessous d'un certain seuil arbitraire, vers un même mode. Pour cela, chaque point de l'image est remplacé, au fil des itérations, par la moyenne de ses écarts avec les points du nuage. Autrement dit, à chaque itération, chaque point est décalé vers le centre de gravité pondéré de l'ensemble des points du nuage. Nous pouvons réécrire la procédure du Mean-Shift sous une formulation similaire à celle définie pour la diffusion scalaire. Nous proposons ainsi de l'écrire sous la forme suivante :

$$\vec{I}^{(n+1)}(p) = \vec{I}^{(n)}(p) + \overrightarrow{VMS}^{(n)}(p) \quad (10.6)$$

Nous notons $\overrightarrow{VMS}^{(n)}$ le vecteur Mean-Shift que nous considérons comme une vitesse au point p et à l'itération n . Il est donné par :

$$\overrightarrow{VMS}^{(n)}(p) = \left(\sum_{k \in N_c(p)} \left(G_k^* \left(\frac{\overrightarrow{\nabla I_o^n}(k)}{\sigma_R} \right) \times \overrightarrow{I}^{(n)}(k) \right) \right) - \overrightarrow{I}^{(n)}(p) \quad (10.7)$$

La fonction G_k^* est une fonction normalisée qui correspond au noyau Gaussien, choisi comme noyau par défaut le long de cette étude. Elle est donnée par :

$$G_k^*(w_k) = \frac{\exp(-\|w_k\|^2)}{\sum_{k \in N_c(p)} \exp(-\|w_k\|^2)} \quad (10.8)$$

Cette fonction permet de pondérer les intensités des voisins d'un point courant. L'objectif est d'éliminer tous les points ayant une distance colorimétrique supérieure à une valeur donnée. On note d_c la distance colorimétrique sur les trois composantes couleurs rouge, verte et bleue de chaque pixel de l'image. Contrairement au voisinage spatial pour la diffusion, le Mean-Shift utilise un voisinage colorimétrique N_c défini par :

$$N_c(p) = \{h/d_c(I(h), I(p)) \leq \sigma_R\}. \quad (10.9)$$

Nous notons $\overrightarrow{\nabla I_o^n}$ le gradient de densité colorimétrique de Fukugana donné par :

$$\overrightarrow{\nabla I_o^n}(k) = \overrightarrow{I}^{(n)}(p) - \overrightarrow{I}^{(0)}(k) \quad \forall k \in N_c(p) \quad (10.10)$$

10.3 Proposition d'une combinaison coopérative

10.3.1 Principe de notre proposition

Les formulations données dans la section précédente font clairement apparaître le lien qui existe entre le Mean-Shift et l'approche scalaire de diffusion. Elles nous amènent à réfléchir sur la mise en place d'une combinaison possible entre ces deux approches. L'approche du Mean-Shift utilise un voisinage colorimétrique qui se base sur la distribution des couleurs de l'image entière. Quant à l'approche de diffusion, elle utilise un voisinage spatial qui décrit l'information locale de l'image. Combiner l'approche de diffusion avec l'approche de segmentation colorimétrique par MeanShift dans un même cadre, nous mène à combiner l'information locale et globale et donc à mixer un voisinage spatial et un voisinage colorimétrique.

Deux stratégies de coopération sont ainsi envisageables : soit une combinaison séquentielle soit une combinaison coopérative. Si une combinaison séquentielle se résume à la simple juxtaposition de boîtes noires, la combinaison coopérative consiste plutôt à l'encapsulation d'une méthode dans une autre. Cette dernière solution fait appel, à chaque itération, aux deux approches à la fois. Elle a tendance à fusionner les résultats respectifs de la diffusion et du Mean-Shift. Ces deux algorithmes s'exécutent dans ce cas en parallèle et leur fonctionnement est indépendant l'un de l'autre. Cette combinaison est rendue possible par le fait que le Mean-Shift et la diffusion utilisent le même schéma itératif de résolution par la méthode de descente de gradient.

Étant données les formulations des équations (10.1) et (10.6), nous proposons d'écrire notre approche unificatrice sous la forme suivante :

$$\vec{I}^{(n+1)}(p) = \vec{I}^{(n)}(p) + \alpha \overrightarrow{VDiff}^{(n)}(p) + \beta \overrightarrow{VMS}^{(n)}(p) \quad (10.11)$$

Nous notons que les valeurs de vitesse $\overrightarrow{VDiff}^{(n)}$ et $\overrightarrow{VMS}^{(n)}$ sont respectivement pondérées par les valeurs α et β , avec $\alpha, \beta \in [0, 1]$. Pour $\alpha = 0$, nous avons la procédure classique de segmentation par un Mean-Shift et pour $\beta = 0$, nous avons plutôt une procédure de diffusion. En plus des valeurs de α et β , deux autres paramètres sont également spécifiques à ce cadre, ce sont les valeurs arbitraires des seuils à définir : K pour la diffusion scalaire de Perona-Malik et σ_R pour la segmentation par Mean-Shift. En total, nous avons essentiellement quatre paramètres à définir et qui sont nécessaires à l'exécution de l'équation (10.11).

Nous pouvons étendre cette combinaison à d'autres modèles de diffusion notamment les modèles tensoriels plus performants pour lisser les contours et renforcer la continuité des traits.

10.3.2 Implémentation et discrétisation

Dans cette section, nous allons présenter l'algorithme détaillé de la combinaison Mean-Shift/diffusion. L'algorithme 1 décrit la combinaison coopérative couplant la méthode du Mean-Shift (Algorithme 2) et la méthode de la diffusion scalaire de Perona-Malik (Algorithme 3). Nous utilisons le vecteur de vitesse de la diffusion et le vecteur de Mean-Shift pour mettre à jour chaque pixel de l'image dans un schéma classique de résolution par la méthode de descente de gradient. La solution est atteinte lorsqu'il n'y a plus de changements dans l'image, c'est à dire lorsque les normes des vitesses de la diffusion et des vecteurs de Mean-Shift sont toutes inférieures à une valeur epsilon donnée.

10.4 Combinaison coopérative versus combinaison séquentielle

La combinaison séquentielle entre la méthode du Mean-Shift et la méthode de diffusion anisotrope consiste à démarrer le processus par l'une de ces deux méthodes et le résultat obtenu sera utilisé comme initialisation pour l'autre méthode. Le but final est d'incorporer le résultat de segmentation (respectivement diffusion) dans le processus de diffusion (respectivement segmentation) dans le cas d'une combinaison procédant par segmentation puis diffusion (respectivement diffusion puis segmentation). L'algorithme détaillé d'une combinaison séquentielle de la diffusion scalaire suivie par le Mean-Shift est donné ci-dessous (Algorithme 4).

Notre proposition ne correspond pas à une application séquentielle de chacune des méthodes prises séparément. En effet, la figure 10.1.a montre le cas d'application d'une diffusion scalaire suivie d'un Mean-Shift et la figure 10.1.b montre le cas d'application d'un Mean-Shift suivi d'une diffusion scalaire. Quel que soit le cas de figure (Fig.10.1), nous notons que la combinaison séquentielle des deux méthodes ne donne pas de bons résultats car elles sont appliquées séparément de façon indépendante. Une segmentation par Mean-Shift augmente le bruit colorimétrique dans l'image segmentée qui ne peut plus être correctement diffusée sans

```

[Initialisation]
 $I_0 \leftarrow I, Norme\_MS \leftarrow 1, Norme\_Diff \leftarrow 1$ 
[Boucle générale]
Tant que (( $Norme\_MS > 0.01$ ) ET ( $Norme\_Diff > 0.01$ )) faire
    [Commander la diffusion] [Soit  $K$  le paramètre défini par l'utilisateur]
     $\overrightarrow{VDiff} \leftarrow Algo\_Diffusion\_Scalaire(K)$ 
    [Commander le Mean-Shift] [Soit  $\sigma_R$  le paramètre défini par l'utilisateur]
     $\overrightarrow{VMS} \leftarrow Algo\_Mean\_Shift(\sigma_R)$ 
    [Mettre à jour les canaux]
     $Norme\_MS \leftarrow 0, Norme\_Diff \leftarrow 0$ 
    Pour chaque pixel  $p$  de l'image faire
        Pour chaque composante couleur  $c$  faire
             $I_0(p, c) \leftarrow I_0(p, c) + \alpha * VDiff(p, c) + \beta * VMS(p, c)$ 
             $Norme\_MS \leftarrow Norme\_MS + VMS(p, c)$ 
             $Norme\_Diff \leftarrow Norme\_Diff + VDiff(p, c)$ 
        Fin Pour
    Fin Pour
    [Soit  $N$  le nombre total des pixels de l'image]
     $Norme\_MS \leftarrow Norme\_MS/N, Norme\_Diff \leftarrow Norme\_Diff/N$ 
Fin Tant que

```

Algorithme 1: Algorithme de restauration par combinaison segmentation/diffusion

```

Pour chaque pixel  $p$  de l'image faire
    [Pour les 3 canaux RVB de l'image]
    Pour chaque composante couleur  $c$  faire
        [Pour les 4 voisins de chaque point  $p$  de l'image]
        Pour chaque voisin  $v(p)$  faire
             $Ecart \leftarrow I(v(p)) - I(p)$ 
             $VDiff(p, c) \leftarrow VDiff(p, c) + \exp(-\frac{Ecart^2}{K^2}) * Ecart$ 
        Fin Pour
         $VDiff(p, c) / = 4$ 
    Fin Pour
Fin Pour

```

Algorithme 2: Algorithme du calcul des vitesses $Vdiff$ au point p pour le canal couleur c du modèle de diffusion scalaire de Perona-Malik ($Algo_Diffusion_Scalaire$)

```

Pour chaque pixel  $p$  de l'image faire
  |  $Norme \leftarrow 0$ 
  | [Pour les 3 canaux RVB de l'image]
  | Pour chaque composante couleur  $c$  faire
  | |  $Norme \leftarrow Norme + |VMS(p, c)|$ 
  | Fin Pour
  | [Test pour éviter de reshifter un point qui est stationnaire ]
  | Si ( $Norme > 0.0001$ ) Alors
  | |  $Somme \leftarrow 0$ 
  | | Pour chaque pixel  $p1$  de l'image faire
  | | |  $Ecart \leftarrow 0$ 
  | | | Pour chaque composante couleur  $c$  faire
  | | | | [Il faut shifter sur l'image originale]
  | | | |  $Ecart \leftarrow Ecart + (I(p, c) - I_0(p1, c))^2$ 
  | | | | Fin Pour
  | | | | Si ( $Ecart < \sigma_R^2$ ) Alors
  | | | | |  $G \leftarrow \exp(-\frac{Ecart}{\sigma_R^2})$ 
  | | | | | Pour chaque composante couleur  $c$  faire
  | | | | | |  $VMS(p, c) \leftarrow VMS(p, c) + G * I_0(p1, c)$ 
  | | | | | | Fin Pour
  | | | | |  $Somme \leftarrow Somme + G$ 
  | | | | Fin Si
  | | | Fin Pour
  | | | Si ( $Somme > 0$ ) Alors
  | | | | Pour chaque composante couleur  $c$  faire
  | | | | |  $VMS(p, c) \leftarrow VMS(p, c) / Somme$ 
  | | | | |  $VMS(p, c) \leftarrow VMS(p, c) - I(p, c)$ 
  | | | | | Fin Pour
  | | | | Fin Si
  | | | Fin Si
  | | Fin Si
  | Fin Pour

```

Algorithme 3: Algorithme du calcul du vecteur de Mean-Shift global VMS au point p pour le canal couleur c (*Algo_Mean_Shift*)

la rendre totalement flou. Inversement, l'application d'une diffusion suivie d'une segmentation par Mean-Shift conduit à segmenter une image diffusée qui a perdu toute la richesse de ses informations colorimétriques, informations nécessaires au bon fonctionnement de la segmentation.

En revanche, une combinaison coopérative (Fig.10.2) permet à chaque itération d'échanger les informations colorimétriques et spatiales pour converger vers une meilleure solution.

Elle ne se fonde plus seulement sur les propriétés propres de chaque méthode, mais sur la synergie des deux méthodes simultanément. La figure 10.2.b montre la carte des vecteurs contradictoires du Mean-Shift et de la diffusion scalaire. Sur l'image de départ, correspondant à l'itération 0, nous notons la présence d'un grand nombre de vecteurs contradictoires. Ses vecteurs se corrigent par réaction interactive entre le Mean-Shift et la diffusion qui réussit à réduire ses vecteurs au fil des itérations. À l'itération 100, les principaux vecteurs se focalisent essentiellement sur les contours des régions de l'image. En outre, l'illustration des normes respectives de la diffusion (Fig.10.2.c) et du Mean-Shift (Fig.10.2.d) met en valeur l'interaction entre ces deux méthodes. Si on a tendance à perdre les contours de l'image au bout de 20 itérations avec la diffusion, la méthode du Mean-Shift les conserve. Aux itérations de 60 à 100, les normes des vecteurs de Mean-Shift et des vitesses de la diffusion sont élevées pour les mêmes points de l'image soulignant ainsi la cohérence entre les deux approches.

10.5 Conclusion

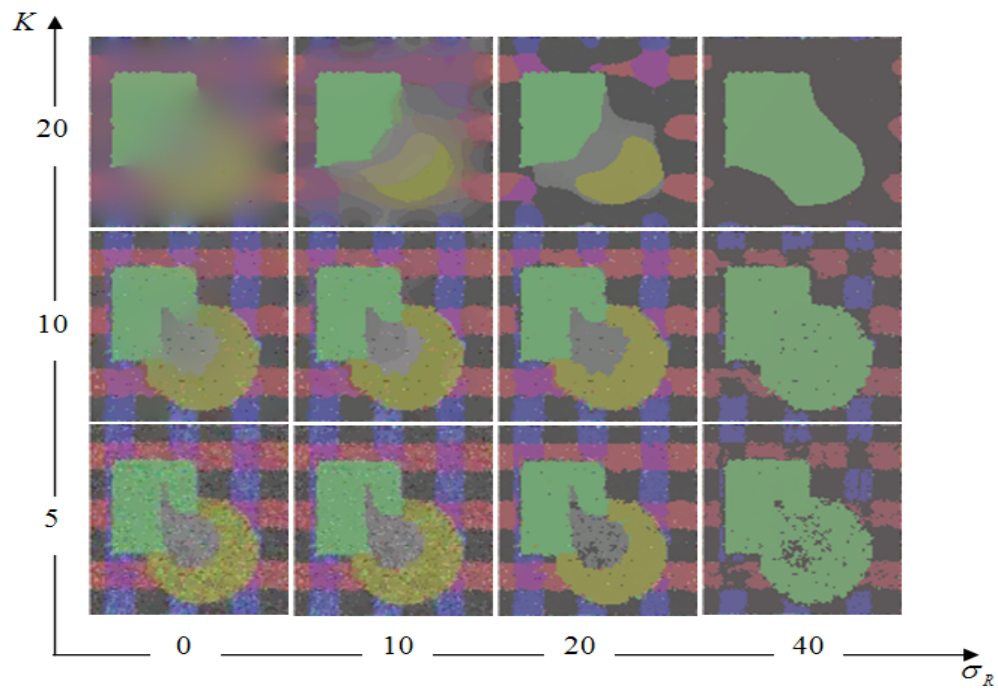
Nous nous sommes intéressé dans ce chapitre à prouver l'existence d'un cadre commun entre le processus de segmentation par Mean-Shift et le processus de diffusion scalaire de Perona-Malik. Nous avons ainsi proposé une méthode qui combine d'une part les avantages d'une approche globale de segmentation et d'autre part les avantages d'une approche locale et adaptative de diffusion. Il en découle donc un modèle qui élimine progressivement le bruit sur une image tout en classant les pixels de celle-ci. Étant donné que ces deux processus reposent sur le principe de descente de gradient, leur combinaison est possible. Mais cette combinaison n'est pas forcément stable et peut diverger ou osciller dans un état instable sans jamais converger s'il y a une forte opposition entre les deux méthodes. Dans le chapitre suivant, nous allons démontrer que le modèle de combinaison est stable et qu'il converge toujours vers une solution finale. Nous étudierons aussi les interactions entre les paramètres de la combinaison (α, β) et les paramètres (K, σ_R) de chacune des méthodes. Enfin nous présenterons les résultats expérimentaux obtenus sur des images de documents ainsi que sur des images naturelles.

```

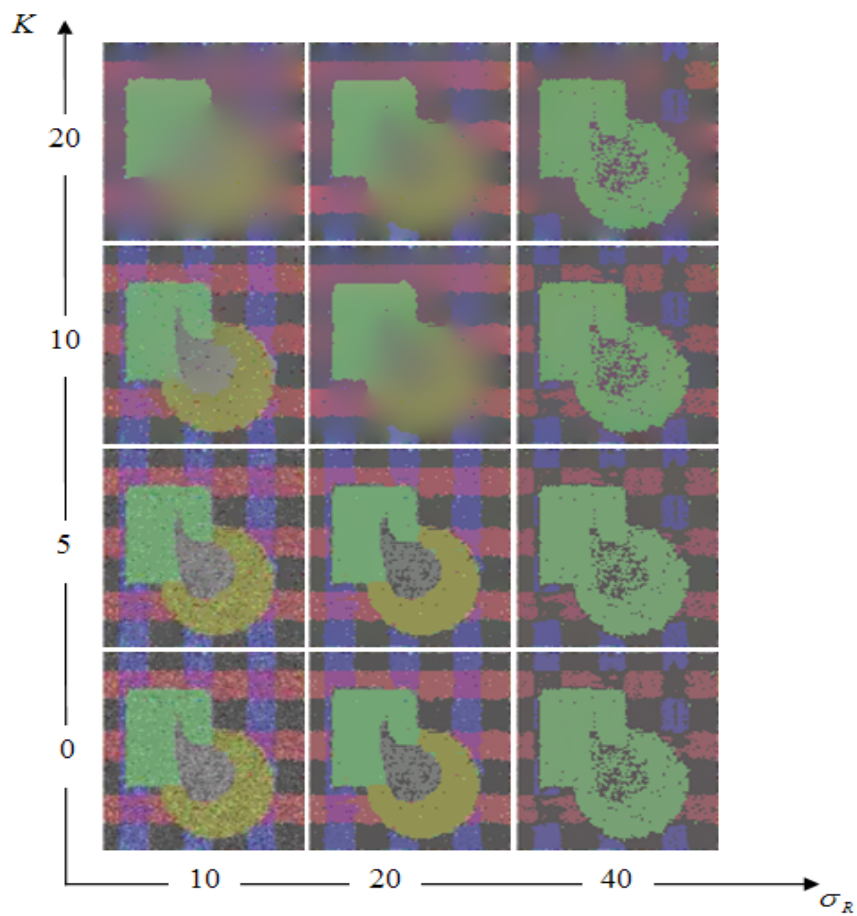
[Initialisation]
 $I_0 \leftarrow I$ 
 $Norme\_MS \leftarrow 1$ 
[Boucle générale]
[Commander la diffusion]
[Soit  $K$  et  $NIter$  les paramètres définis par l'utilisateur]
Tant que ( $NIter > 0$ ) faire
     $\overrightarrow{VDiff} \leftarrow Algo\_Diffusion\_Scalaire(K)$ 
    [Mettre à jour les canaux]
     $\overrightarrow{I_0} \leftarrow \overrightarrow{I_0} + \alpha * \overrightarrow{VDiff}$ 
     $NIter \leftarrow NIter - 1$ 
Fin Tant que
[Commander le Mean-Shift]
[Soit  $\sigma_R$  le paramètre défini par l'utilisateur]
Tant que ( $Norme\_MS > 0.01$ ) faire
     $\overrightarrow{VMS} \leftarrow Algo\_Mean\_Shift(\sigma_R)$ 
    [Mettre à jour les canaux]
     $Norme\_MS \leftarrow 0$ 
    Pour chaque pixel  $p$  de l'image faire
        Pour chaque composante couleur  $c$  faire
             $I_0(p, c) \leftarrow I_0(p, c) + \beta * VMS(p, c)$ 
             $Norme\_MS \leftarrow Norme\_MS + VMS(p, c)$ 
        Fin Pour
    Fin Pour
    [Soit  $N$  le nombre total des pixels de l'image]
     $Norme\_MS \leftarrow Norme\_MS / N$ 
Fin Tant que

```

Algorithme 4: Algorithme d'une combinaison séquentielle de la diffusion scalaire suivie par le Mean-Shift

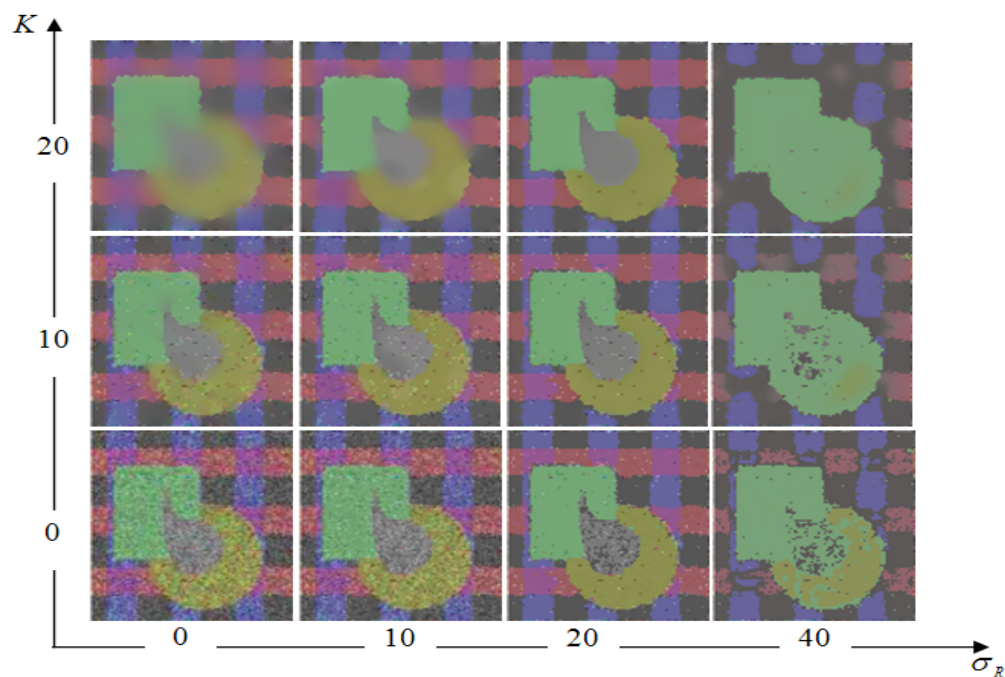


(a) Cas d'application d'une diffusion scalaire suivie d'un Mean-Shift

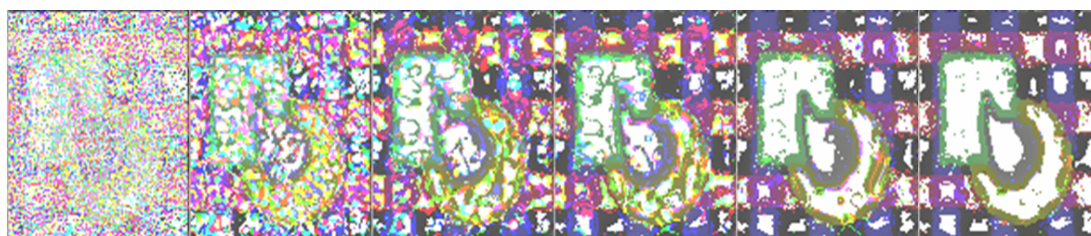


(b) Cas d'application d'un Mean-Shift suivi d'une diffusion scalaire

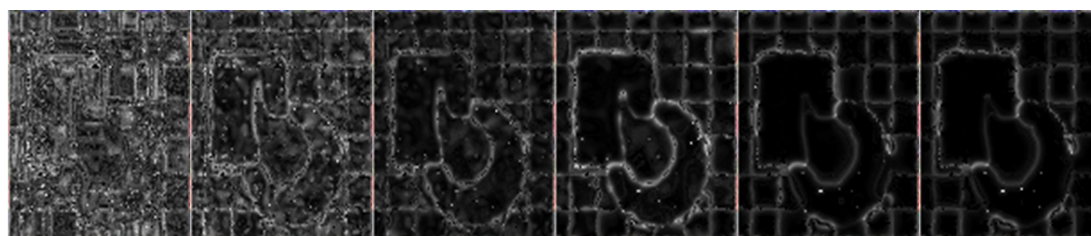
Figure 10.1 : Application de la combinaison séquentielle entre la diffusion scalaire ($\alpha = 0.2$) et le Mean-Shift ($\beta = 1.0$) sur une image de test



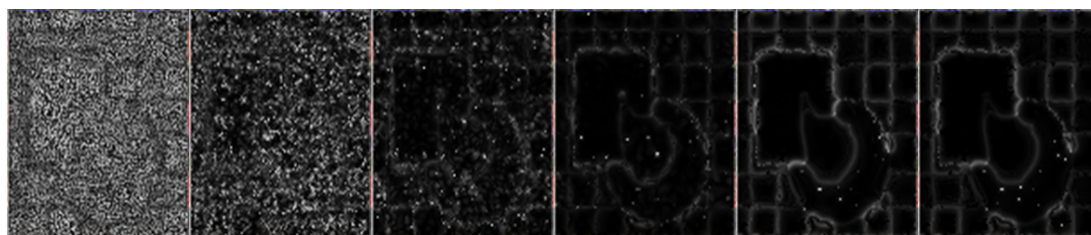
(a) Résultat de la combinaison diffusion et MeanShift pour différentes valeurs de K et σ_R



(b) Affichage des vecteurs contradictoires Mean-Shift/diffusion scalaire



(c) Evolution de la norme des vecteurs de Mean-Shift



(d) Evolution de la norme des vitesses de la diffusion scalaire

Figure 10.2 : Cas d'une combinaison coopérative $(\alpha, \beta) = (0.2, 0.5)$: (a) pour différentes valeurs de (K, σ_R) et illustration avec $(K, \sigma_R) = (20, 20)$ pour (b), (c) et (d) avec respectivement de gauche à droite 0, 5, 10, 20, 60, 100 itérations

Chapitre 11

Segmentation/Diffusion : Évaluation des Performances

« Savoir où l'on veut aller, c'est très bien ; mais il faut encore montrer qu'on y va. »

Émile Zola

Résumé

Ce chapitre présente une évaluation des performances de l'approche unificatrice décrite dans le chapitre précédent. Nous montrerons, à travers des exemples de tests, que cette proposition peut nous guider vers des solutions satisfaisantes pour la restauration des dégradations non-uniformes des images de documents ; les dégradations du fond et des formes. Nous prouverons également que cette approche est suffisamment générique pour restaurer et /ou la segmenter des images naturelles.

Sommaire

11.1 Introduction	168
11.2 Étude du choix des paramètres	168
11.3 Étude de la convergence	170
11.4 Mise en œuvre pratique	171
11.5 Conclusion	179

11.1 Introduction

Nous avons défini dans le chapitre précédent (Chapitre 10) une approche unificatrice entre la diffusion d'une part et la segmentation par Mean-Shift d'autre part. Cette approche est introduite comme suit :

$\vec{I}^{(n+1)}$	L'image à l'itération n
\parallel	
$\vec{I}^{(n)}$	L'image à l'itération $n + 1$
+	
$\alpha \overrightarrow{VDiff}_K^{(n)}$	Vélocité pondérée de la diffusion
+	
$\beta \overrightarrow{VMS}_{\sigma_R}^{(n)}$	Vélocité pondérée de la segmentation

Dans ce chapitre, nous proposons de tester cette approche unificatrice sur les images couleurs et de mettre en évidence l'influence des différents paramètres utilisés. Nous testerons l'influence des couples (σ_R, K) et (α, β) dans le cas de l'utilisation du modèle scalaire de diffusion de Perona-Malik. Nous rappelons que α et β sont les paramètres de pondération respectifs de la diffusion et du Mean-Shift, K est une constante prédéfinie permettant de distinguer les zones à faible gradient de celles à fort gradient dans le cas de la diffusion scalaire et σ_R est une constante qui définit le voisinage colorimétrique associé à chaque point de l'image lors d'une segmentation par Mean-Shift.

Nous verrons ensuite que l'approche de diffusion scalaire de Perona-Malik peut être remplacée par d'autres approches de diffusion notamment des approches tensorielles qui peuvent mener à de meilleurs résultats.

11.2 Étude du choix des paramètres

11.2.1 Cas de (σ_R, K)

Les paramètres σ_R et K sont respectivement reliés à la diffusion scalaire de Perona-Malik et à la segmentation par Mean-Shift. La définition de ces paramètres ne peut être faite que de façon expérimentale pour une image donnée. Néanmoins, nous avons une idée de leur ordre de valeurs.

Le paramètre σ_R doit être choisi selon la précision des couleurs sur l'image segmentée. En effet, pour une valeur de σ_R très faible, les éléments caractéristiques de l'image s'affichent à un niveau de détails fin. Par contre, une valeur élevée de σ_R risque d'uniformiser des zones entières de couleur et de perdre les fines nuances de couleur. Généralement, une valeur de σ_R de l'ordre de 20 dans l'espace RVB a tendance à conserver les détails de l'image et à enlever les traces du bruit colorimétrique (Fig.11.2.1).

La constante K permet de distinguer les zones homogènes à faible gradient des zones de contours correspondant à des valeurs de gradients très élevées. La figure 11.2.1 regroupe un ensemble d'images obtenues après application de notre proposition avec différentes valeurs

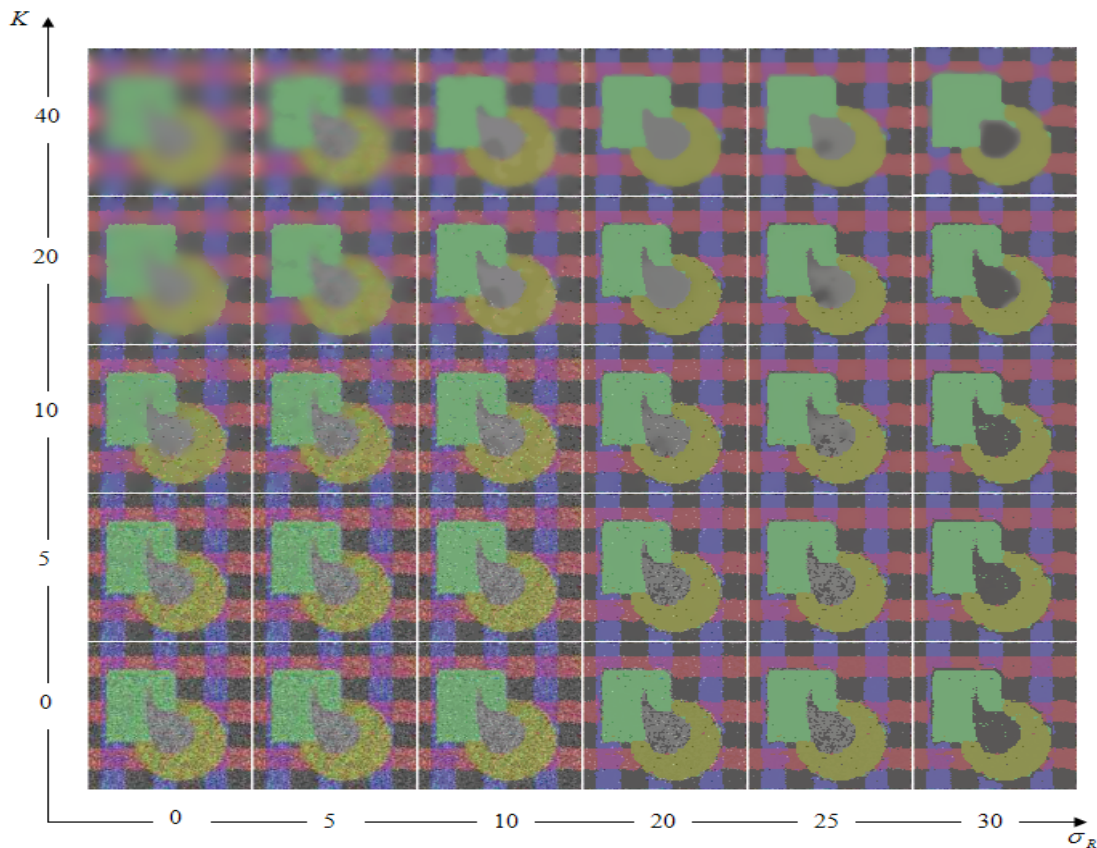


Figure 11.1 : Étude de l'influence des paramètres (σ_R, K) pour $(\alpha = \beta = 0.2)$.

de K (suivant les lignes) et de σ_R (suivant les colonnes) pour $\alpha = \beta$. Le cas $(\sigma_R = 0, K = 0)$ correspond à l'image de départ non traitée sans diffusion ni segmentation. La ligne $K=0$ (respectivement la colonne $\sigma_R=0$) montre les images obtenues suite à l'application d'une segmentation par Mean-Shift (respectivement d'une diffusion scalaire de Perona-Malik).

Pour (σ_R, K) égal à $(20, 25)$, cette combinaison diffusion/segmentation réussit à restaurer l'image tout en conservant ses principaux détails. En projetant suivant la ligne et la colonne, nous pouvons comparer ce résultat avec celui obtenu par une application directe du Mean-Shift ou par une application directe de la diffusion scalaire de Perona-Malik. La combinaison des deux approches aboutit à un résultat nettement meilleur à celui d'une application directe de chaque approche séparément.

11.2.2 Cas de (α, β)

Les paramètres α et β sont les coefficients pondérateurs de la diffusion et de la segmentation par Mean-Shift respectivement. Le paramètre α dépend de la nature de l'image ; pour une image bruitée, il faut donner un valeur α plus élevée par rapport à celle de β . Il est certain que, pour mettre en valeur l'apport de la diffusion, il ne faut pas choisir une valeur de β inférieure à celle de α . La différence entre ces deux coefficients ne doit pas être très élevée si on veut préserver une véritable combinaison. La figure 11.2 montre quelques résultats

obtenus avec l'approche unificatrice pour différentes valeurs de α et β .

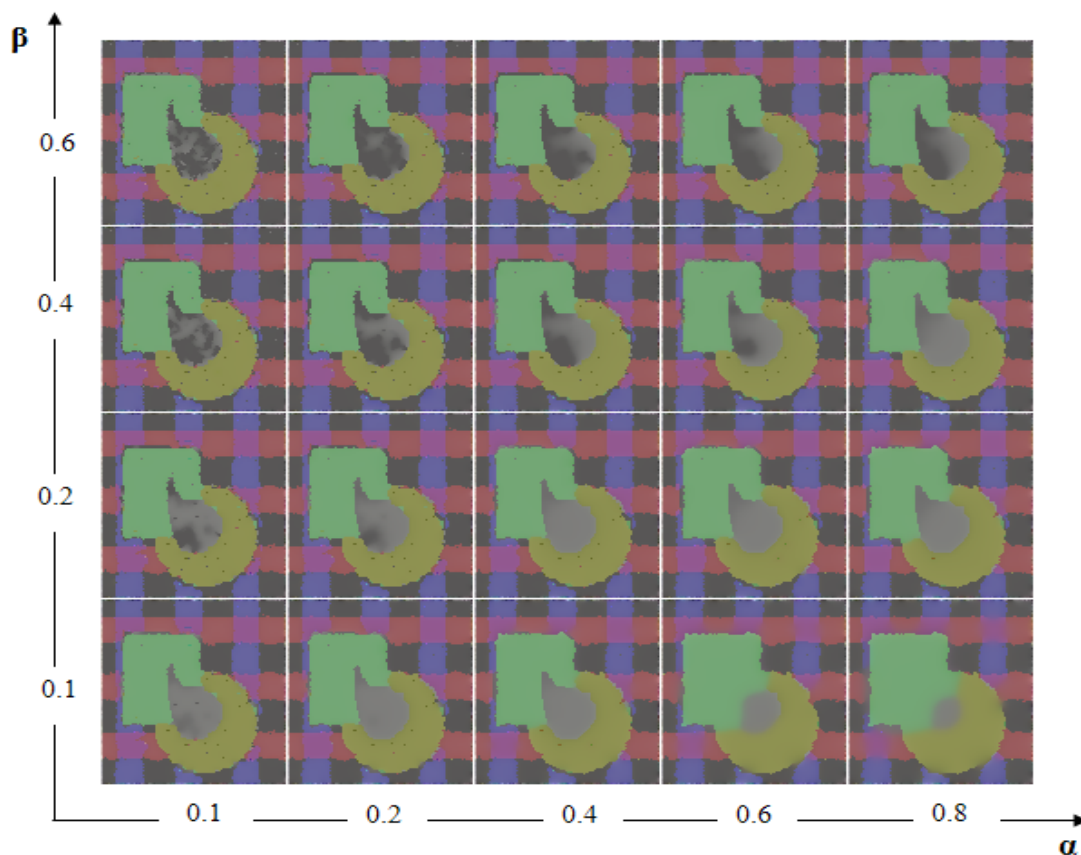


Figure 11.2 : Étude de l'influence des paramètres (α, β) pour $(\sigma_R = 25, K = 20)$.

11.3 Étude de la convergence

Après avoir vérifié que la combinaison d'une diffusion et d'une segmentation colorimétrique améliore les performances des traitements, nous devons étudier si ce nouveau modèle converge toujours vers une solution correspondant à une image à la fois diffusée et segmentée. C'est la raison pour laquelle nous devons vérifier que les normes des vitesses de la diffusion et des vecteurs de Mean-Shift forment des courbes strictement décroissantes qui convergent simultanément vers zéro. Si les deux méthodes ne se combinent pas, alors nous aurions des courbes non-décroissantes représentant une opposition entre les deux méthodes. Pour un point donné, nous pouvons avoir une contradiction entre une classification colorimétrique à partir de la distribution globale des couleurs de l'image entière et le processus de diffusion à partir des informations locales. L'étude du comportement analytique des vitesses de la diffusion et des vecteurs de Mean-Shift durant le processus de traitement démontre bien qu'il y a globalement une décroissance stricte des deux courbes à la même vitesse vers zéro. La figure 11.3 affiche les courbes de l'évolution des vitesses du Mean-Shift et de la diffusion pour différents cas : suivant les valeurs de K (Fig.11.3.a), les valeurs de σ_R (Fig.11.3.b), les valeurs de α (Fig.11.3.c) et les valeurs de β (Fig.11.3.d). Quel que soit le paramètre étudié,

toutes les courbes sont strictement décroissantes et tendent vers 0. Tout cela tend à montrer que notre modèle est stable et converge toujours vers une unique solution. Cela démontre que la combinaison d'une diffusion et d'une segmentation colorimétrique est viable et que ces deux approches se complètent et ne rentrent pas en conflit.

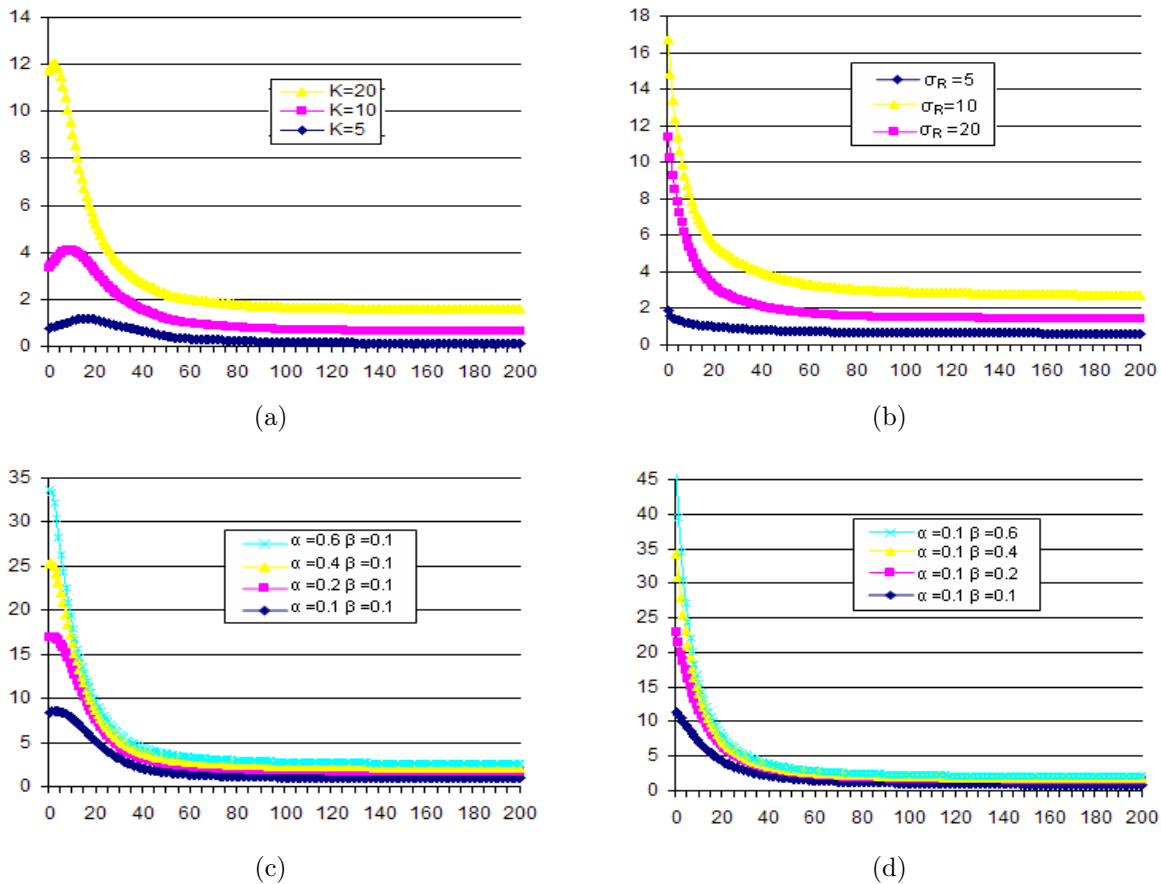


Figure 11.3 : (a) Évolution de la norme moyenne des vitesses de la diffusion pour $\sigma_R = 25$ et pour différentes valeurs de K , (b) Évolution de la norme moyenne des vecteurs du Mean-Shift pour $K = 20$ et pour différentes valeurs de σ_R , (c) Évolution de la norme des vitesses de la diffusion pour $\beta = 0.1$ et pour différentes valeurs de α , (d) Évolution de la norme des vecteurs de Mean-Shift pour $\alpha = 0.1$ et pour différentes valeurs de β .

11.4 Mise en œuvre pratique

Afin d'évaluer les performances, nous proposons une série de tests menée sur des images dégradées de documents. Pour prouver la généralité de notre proposition, nous allons aussi tester ses performances sur des images naturelles.

11.4.1 Application sur des images de documents

Tout au long de ce chapitre, nous nous sommes focalisés sur l'étude du cadre commun proposé entre le Mean-Shift et la diffusion scalaire de Perona-Malik. Le choix de cette ap-

proche de diffusion n'est pas le seul et d'autres modèles de diffusion comme les modèles tensoriels peuvent être également utilisées. La figure 11.4 montre un cas d'application de notre proposition avec trois approches différentes de diffusion, une scalaire de Perona-Malik (Fig.11.4.g), les deux autres tensorielles, celle de Tschumperlé (Fig.11.4.k) et celle que nous avons élaboré Chapitre 9.

Nous observons qu'une diffusion tensorielle assure une meilleure restauration des formes. La diffusion scalaire génère sur l'image résultat des effets d'escalier sur les contours des lettres. Les figures 11.4(h) et 11.4(n) montrent les cartes de différence colorimétrique entre les pixels des images respectivement traitées par les approches de Perona-Malik et de Tschumperlé. Le résultat de l'application de la diffusion toute seule sur l'image de test tend à reconstruire une image constante par morceaux ; la figure 11.4(e) (respectivement 11.4(m)) correspond au résultat d'une diffusion scalaire de Perona-Malik (respectivement d'une diffusion tensorielle de Tschumperlé). Une nette amélioration du résultat de la segmentation est obtenue avec l'approche unificatrice. En outre, une segmentation directe par un Mean-Shift (Fig.11.4(c)) sur l'image de test donnée sur la figure 11.4(a) génère sur l'image résultat (Fig.11.4(c)) des effets de blocs qui fragmentent une région plate. Comme nous l'espérions, la restauration obtenue avec notre cadre commun apparaît visuellement meilleure que celle obtenue précédemment en utilisant la procédure de segmentation ou de diffusion toute seule.

D'autres tests (Fig.11.6, Fig.11.7 et Fig.??) menés avec l'approche de diffusion tensorielle de Weickert et notre propre modèle de diffusion montrent l'efficacité de cette approche unificatrice sur des dégradations non-uniformes.

Comme nous avons pu le voir à travers les exemples, la combinaison entre les approches de diffusion et de segmentation par Mean-Shift permet de résoudre le problème des images sur-segmentées, tel est le cas des images dégradées suite à l'oxydation de l'encre¹ (Fig.11.6) et des images bruitées soit par la texture non-uniforme du papier (Fig.11.4) soit par des taches sur le support (Fig.11.7). Le résultat obtenu par l'approche unificatrice permet de restaurer les formes des caractères tout en réduisant les défauts du support papier. La segmentation associée à la diffusion permet de séparer finement les différentes couleurs du texte de la couleur du papier, tout en conservant la continuité des traits.

Enfin l'approche unificatrice s'applique aussi à notre modèle de diffusion (Chapitre 9). Sur une image de la British Library (Fig.11.4.1), nous pouvons clairement distinguer l'application de notre modèle de diffusion seul (Fig.11.4.1) et le résultat d'une combinaison avec le Meanshift (Fig.11.4.1). Enfin le cadre commun permet de simultanément de séparer le recto et le verso tout en restaurant les formes (Fig.11.4.1). Le texte du recto s'obtient en regroupant les means couleurs correspondant aux couleurs du recto (Fig.11.4.1). les traits du recto apparaissent plus continus et plus lissés qu'avec les méthodes de séparations décrites dans les chapitres 5 et 6.

¹L'oxydation correspond au vieillissement des encres qui a pour conséquence leurs effacements progressif ou une modification partielle de leurs couleurs.

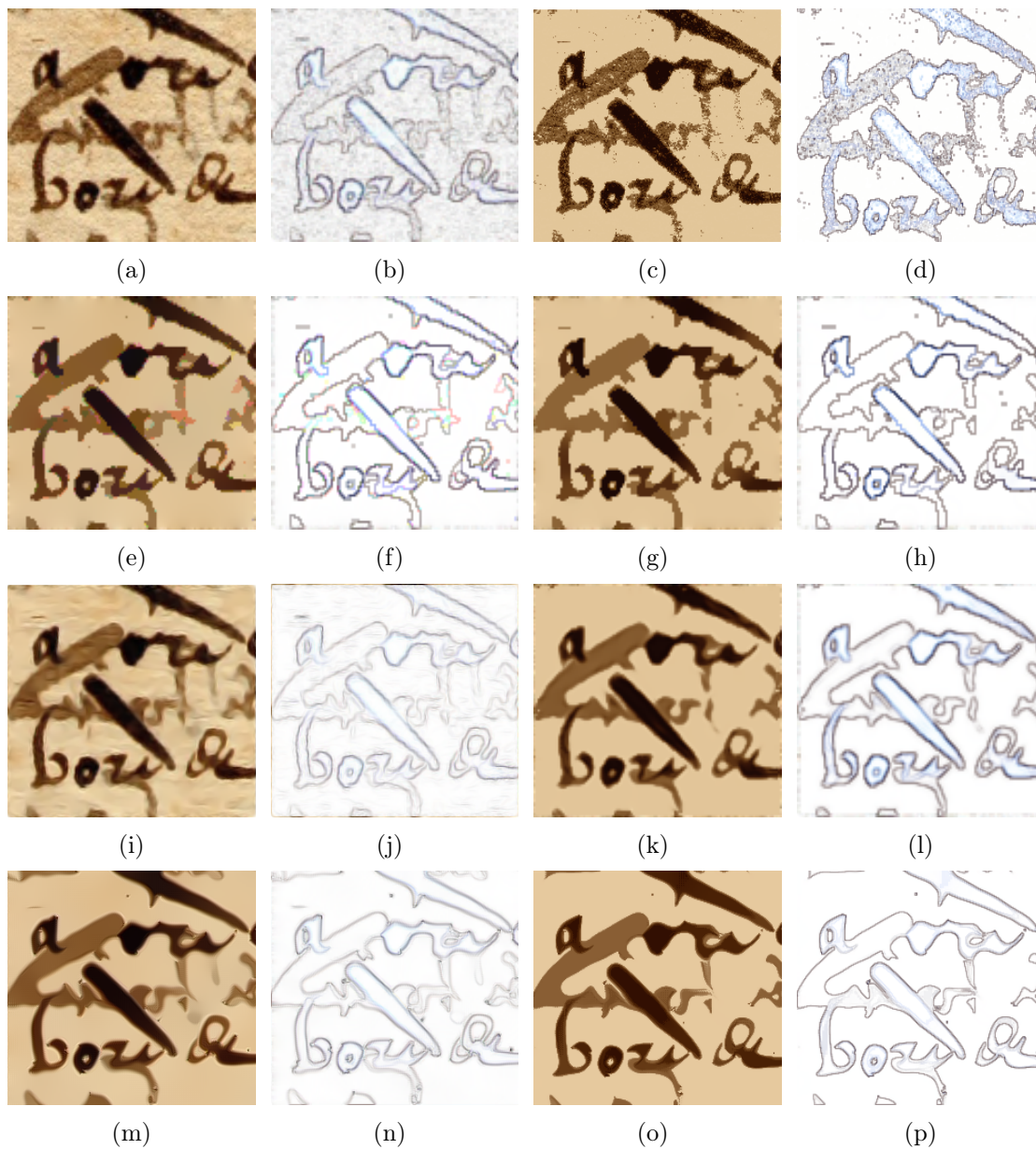


Figure 11.4 : (a) Image dégradée de document, (c) Segmentation globale par Mean-Shift, (e) Diffusion scalaire de Perona-Malik, (g) Approche unificatrice utilisant la diffusion scalaire de Perona-Malik, (i) Diffusion tensorielle de Tschumperlé (k) Approche unificatrice avec diffusion tensorielle de Tschumperlé, (m) Résultat du modèle de diffusion décrit dans (Chapitre 9), (o) Approche unificatrice utilisant notre modèle de diffusion, (b), (d), (f), (h), (j), (l), (n) et (p) sont les contours des images (a), (c), (e), (g), (i), (k), (m) et (o) respectivement

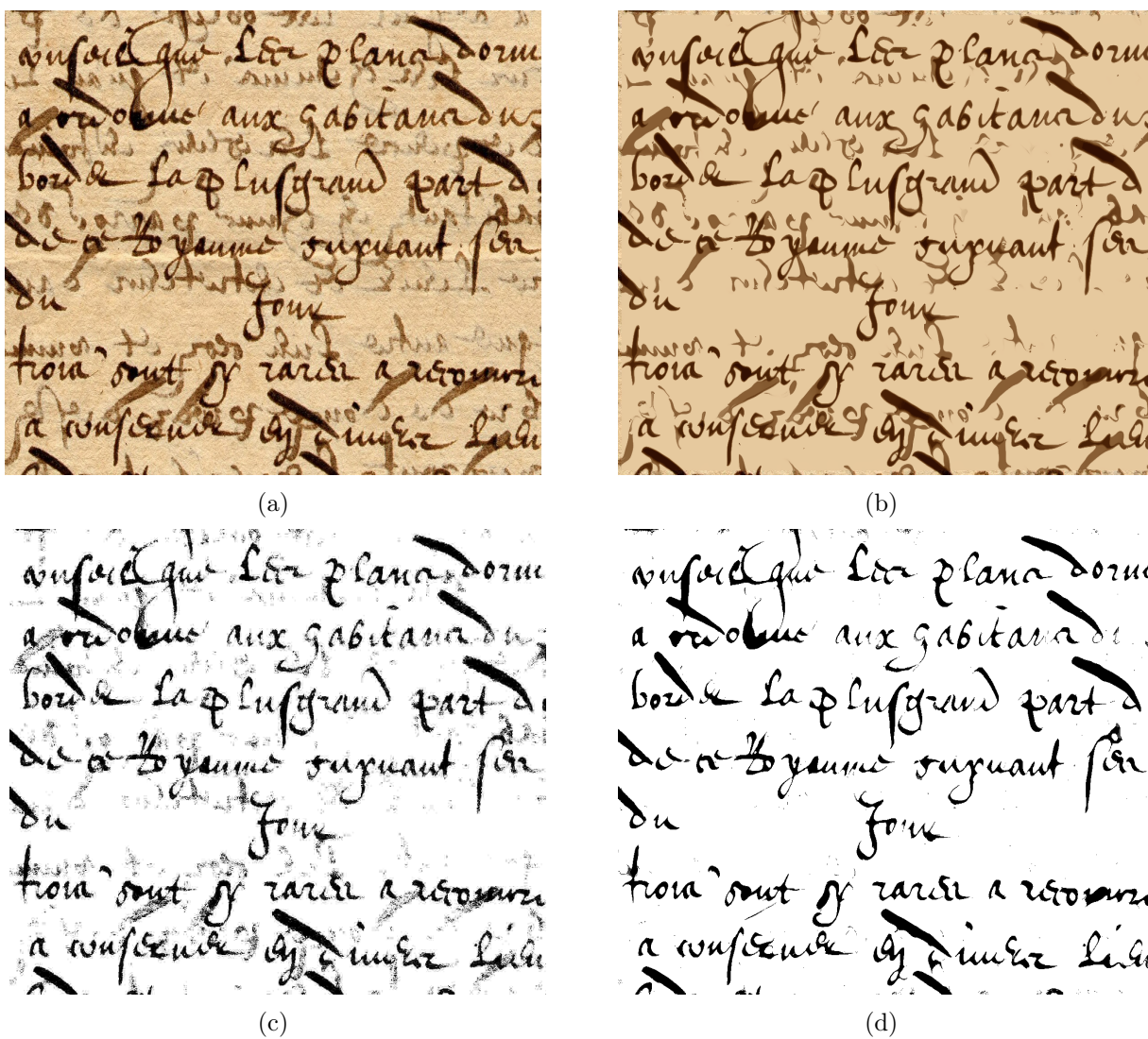


Figure 11.5 : (a) Image dégradée, (b) Résultat du cadre commun avec le modèle de diffusion décrit dans (Chapitre 9), (c) Résultat d'une binarisation de (a), (d) Résultat d'une binarisation de (b) avec le même seuil que (c)



Figure 11.6 : (a) Image de manuscrit de Tombouctou, (d) Diffusion de Weickert seule, (g) segmentation par Mean-Shift seule, (j) Approche unificatrice et (c), (f), (i) et (l) sont les contours des images (b), (e), (h) et (k).

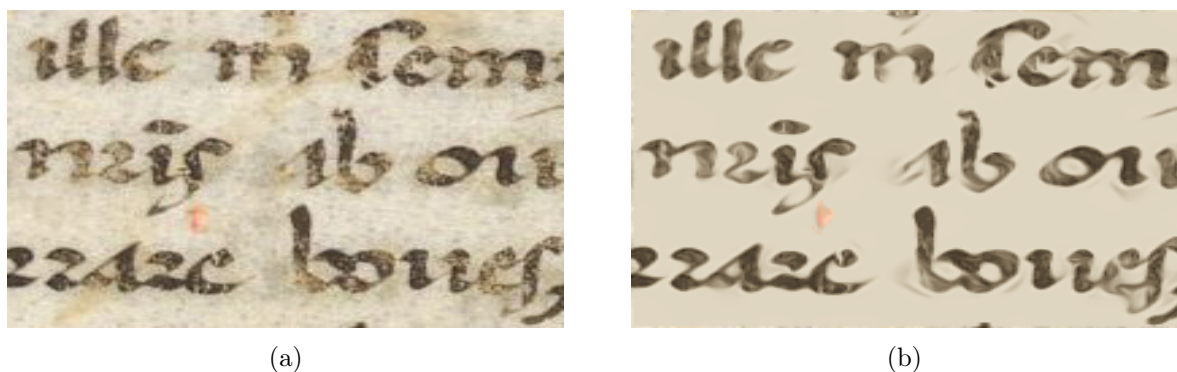


Figure 11.7 : (a) Image dégradée, (b) Approche unificatrice avec une diffusion de Weickert.

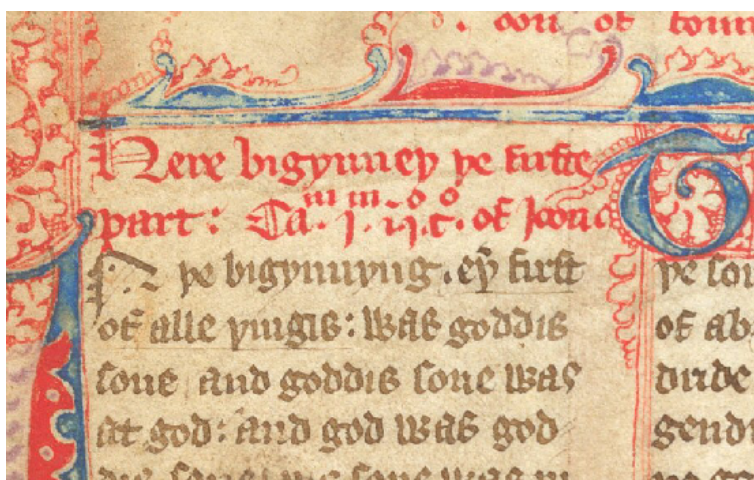


Figure 11.8 : Image originale (Arundel254f19 Source British Libraries).

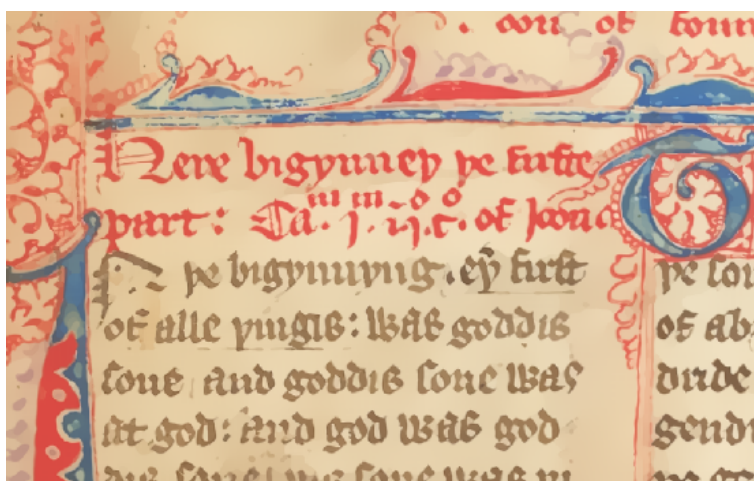


Figure 11.9 : Application du modèle de diffusion décrit en (Chapitre 9) seul avec $\sigma = 0.5$, $\rho = 0$, $K_+ = K_- = 1$.

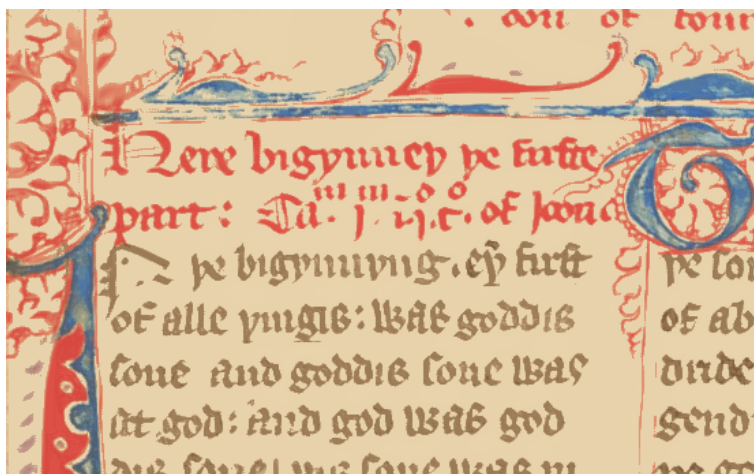


Figure 11.10 : combinaison avec le MeanShift et notre modèle de diffusion avec les mêmes paramètres.

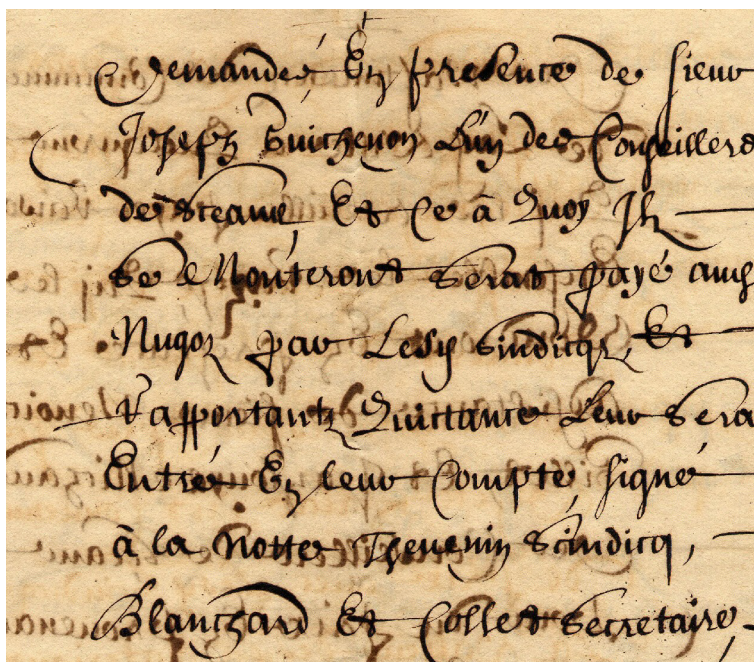


Figure 11.11 : Image originale des archives de châtilion sur Chalaronne.

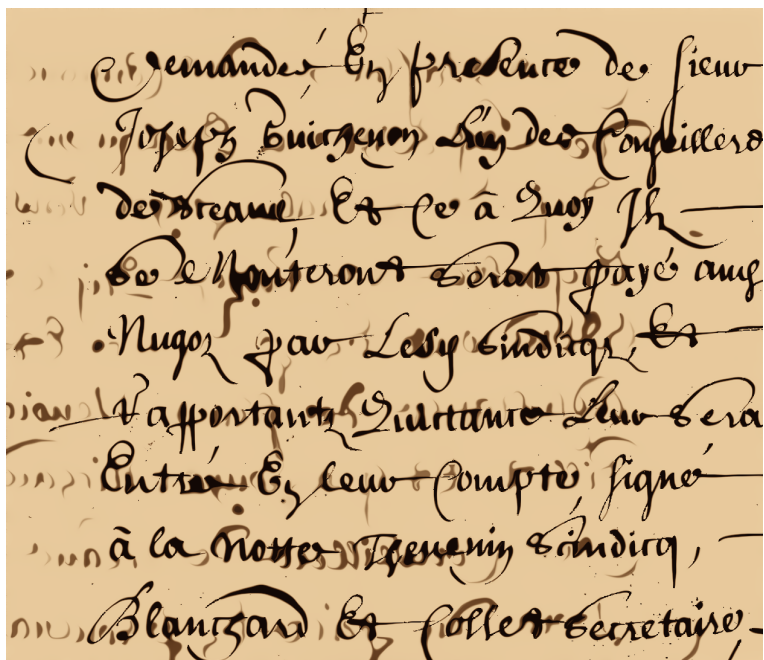


Figure 11.12 : Image traitée par l'approche unificatrice $K_+ = 60$ $K_- = 120$

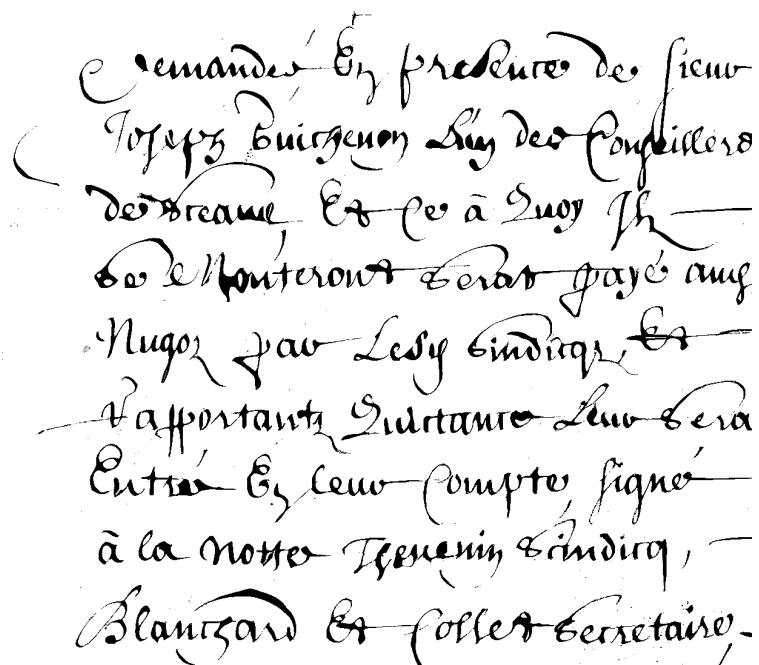


Figure 11.13 : Image du recto

11.4.2 Application sur des images naturelles

Nous présentons des résultats obtenus par l'approche unificatrice sur des images naturelles. Dans chacun des exemples, nous avons présenté l'image de départ et sa version traitée et leurs contours. Les figures 11.14 et 11.15 montrent une bonne stabilité vis à vis du bruit et une excellente conservation des formes et des contours. L'avantage de l'approche unificatrice devient plus évident sur des images très bruitées. Nous apportons ainsi sur une image naturelle (Fig.11.15(c)) un bruit Gaussien et nous observons le comportement de notre méthode pour des différents taux de bruit. Nous comparons ce comportement par rapport à celui obtenu avec un Mean-Shift seul (Fig.11.16). Pour différents taux du bruit Gaussien (15%, 25%, 40% et 50%), notre proposition est capable de correctement segmenter les images bruitées avec un minimum d'artefacts.

11.5 Conclusion

Ce chapitre a présenté une évaluation des performances de la combinaison entre le Mean-Shift et les modèles de diffusion scalaires ou tensoriels. Si le Mean-Shift est une approche globale permettant de segmenter l'image par classification globale des couleurs, la diffusion est capable de filtrer sélectivement le bruit tout en préservant les contours et les formes. Les différents résultats expérimentaux présentés prouvent que notre proposition est une approche générique capable de résoudre à la fois le problème des dégradations du fond et des formes. Elle permet de réaliser une segmentation des images de documents couleurs tout en renforçant la continuité des traits et préservant la forme des contours des lettres. Cette proposition est également performante pour la segmentation des images naturelles couleurs. Une étude comparative, plus large, avec d'autres méthodes de segmentation devrait être envisagée.



Figure 11.14 : (a) image naturelle, (b) image traitée par l'approche unificatrice, (c) contours de (a), (d) contours de (b).

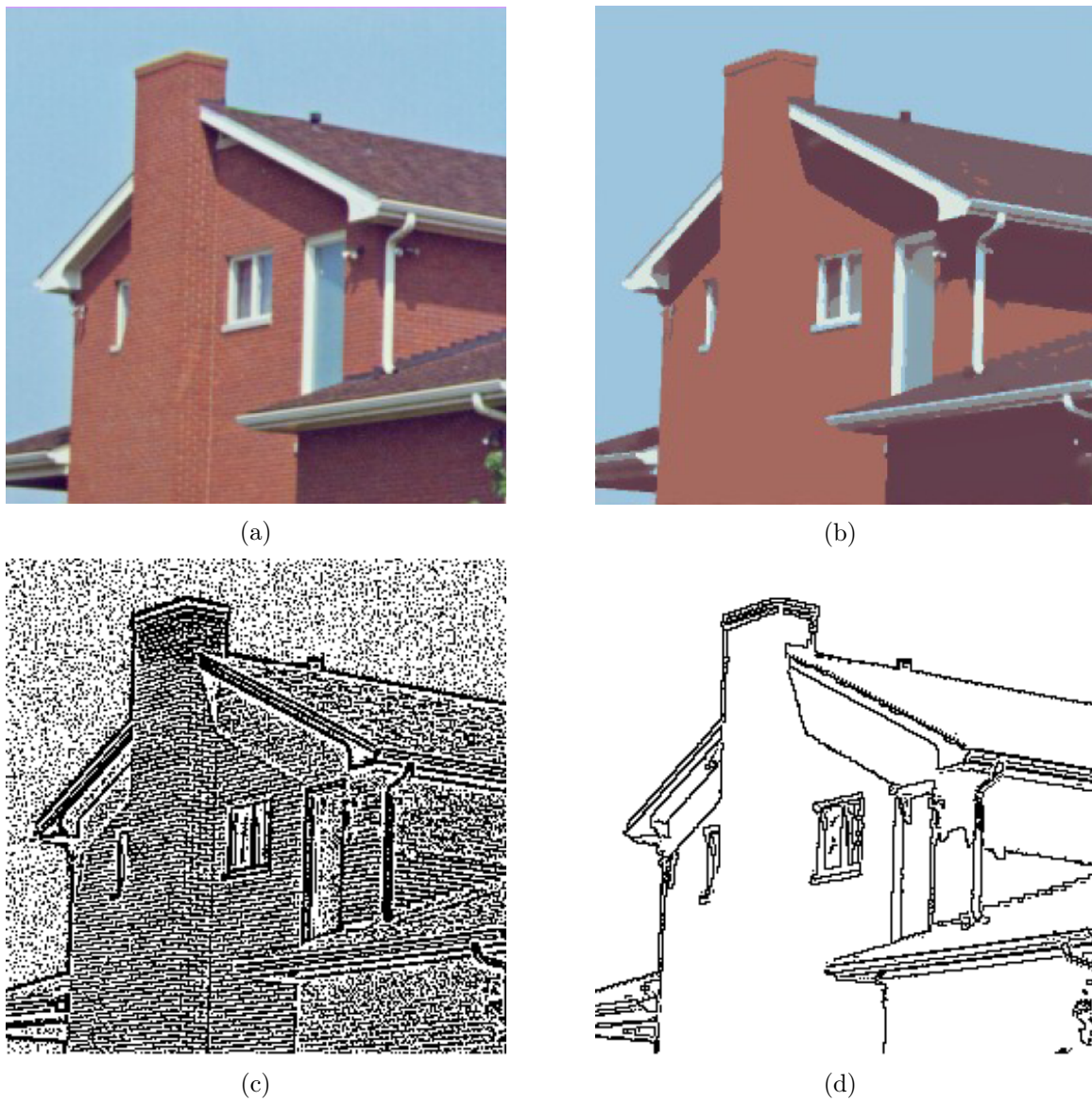


Figure 11.15 : (a) Image naturelle, (b) image traitée par l'approche unificatrice, (c) affichage des contours de (a), (d) affichage des contours de (b).

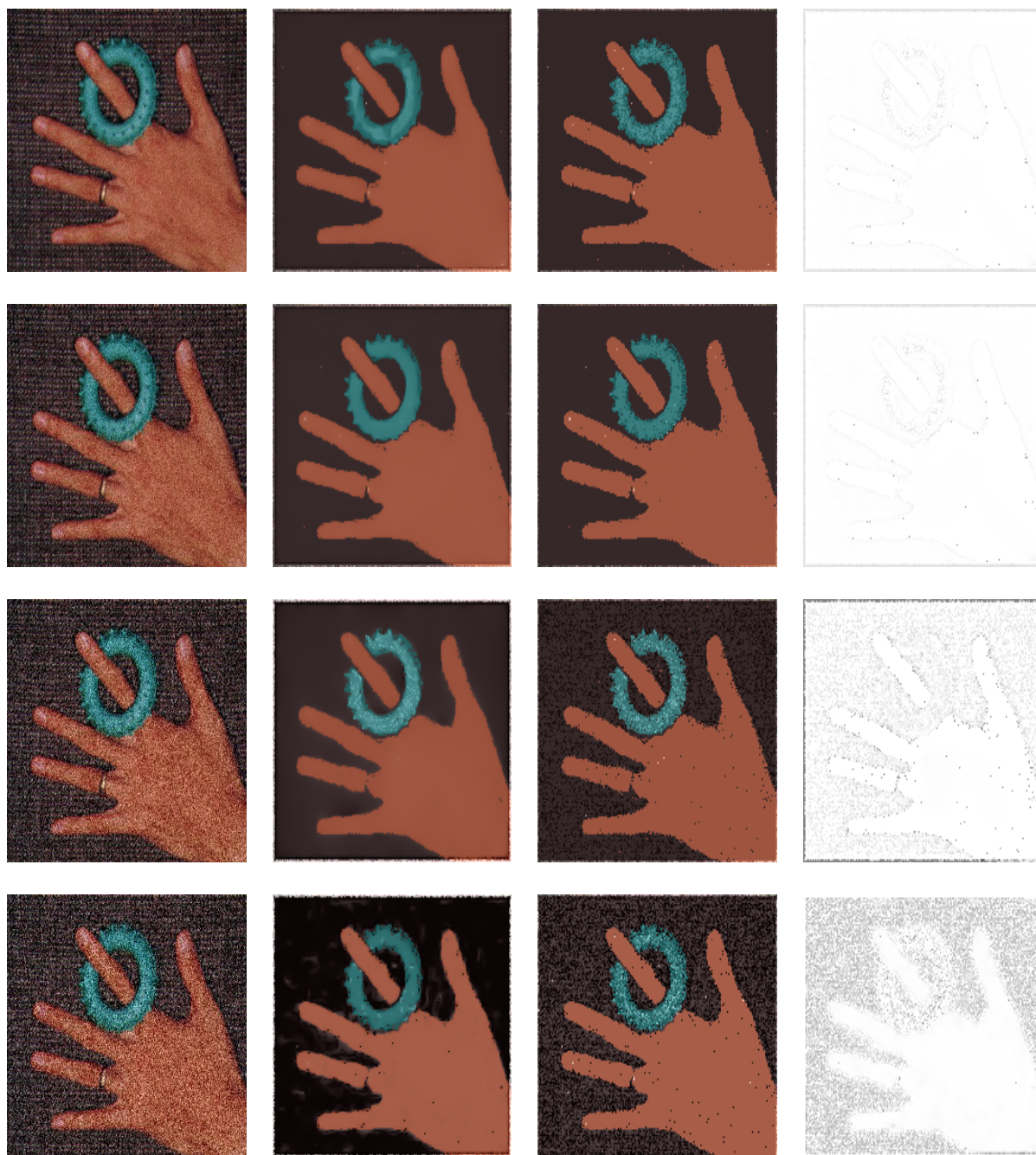


Figure 11.16 : Robustesse de l'approche unificatrice pour différents taux de bruit Gaussien (15%, 25%, 40% et 50%), et de gauche à droite, l'image bruitée, l'image restaurée par notre proposition, l'application directe du Mean-Shift global et l'image de différence entre les deux images précédemment citées.

Chapitre 12

Conclusion Générale

« C'est là en effet un des grands et merveilleux caractères des beaux livres que pour l'auteur ils pourraient s'appeler conclusions et pour le lecteur incitations. »

Marcel Proust

12.1 Résumé de nos contributions

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

En résumé, nos principales contributions sont les suivantes :

Proposition d'une typologie des dégradations orientée traitement d'image (Chapitre 2) et l'étude des travaux antérieurs sur la restauration des images de documents (Chapitre 3)

Nous avons proposé une typologie non-causale des dégradations qui permet de les regrouper dans des classes associées à des traitements adaptés. Après avoir resitué notre étude par rapport aux travaux existants, nous avons choisi de traiter les dégradations du support papier comme les taches ou le passage du verso sur le recto par des méthodes de séparation des couches d'information et les dégradations des traits par les approches de diffusion.

Proposition de deux approches de segmentation par analyse de la couleur pour la restauration des dégradations du fond : une première approche globale (Chapitre 5) et une deuxième approche duale (Chapitre 6)

Pour séparer les couches d'information entre le texte et le fond, nous avons testé une première approche basée sur la classification récursive des couleurs à partir de l'algorithme K -moyennes combinée à une décorrélation de l'information par l'Analyse en Composantes Principales. Cette approche est limitée aux documents qui utilisent une seule couleur d'encre. La deuxième proposition, que nous avons appelé le Mean-Shift local-global, est une amélioration de la méthode de classification des couleurs par l'algorithme du Mean-Shift et qui permet de réduire sa complexité de calcul en introduisant des informations locales. Nous avons aussi démontré qu'une classification entièrement basée sur la recherche des modes de la fonction de densité de la distribution des couleurs de l'image entière donne des résultats satisfaisants pour différencier les couleurs du texte original du recto des autres couleurs présentes sur le support papier comme les traits du verso ou les taches.

Étude (Chapitres 7 et 8) et adaptation (Chapitre 9) des approches de diffusion pour la restauration des dégradations des formes

Nous avons étudié les différentes méthodes de diffusion sur des images dégradées de documents. Nous avons ainsi démontré que certaines approches de diffusion, notamment les approches tensorielles de diffusion anisotrope, sont suffisamment efficaces pour restaurer les formes dégradées des lettres. Elle permet d'améliorer les performances des logiciels OCR et d'augmenter la lisibilité des textes. Nous avons évalué aussi la reconstruction virtuelle des parties manquantes de l'image par les méthodes de diffusion comme l'inpainting. Enfin, nous avons développé notre propre modèle de diffusion adapté au traitement des images de documents qui préserve les singularités et la topologie des formes. Ce modèle repose sur une dissociation du comportement de la diffusion suivant le contenu de l'image ce qui permet de lisser les contours, renforcer la continuité des traits et de préserver les singularités de l'image. Nous avons démontré que la préservation des singularités est très importante pour conserver les détails nécessaires à la lisibilité des textes et pour garder la structure des caractères. Notre modèle de diffusion peut s'appliquer sur une grande variété d'images de documents ainsi qu'aux images naturelles.

Proposition d'une méthode unificatrice entre la segmentation colorimétrique par Mean-Shift et la diffusion (Chapitre 10) et évaluation de ses performances (Chapitre 11)

Nous avons proposé de combiner les méthodes de diffusion avec une méthode de segmentation colorimétrique comme le Mean-Shift pour tirer les avantages des deux méthodes. Ce cadre commun est une solution possible au paradigme local/global des méthodes de restauration. Les deux méthodes sont complémentaires et reposent sur des informations indépendantes qui peuvent se combiner pour offrir de meilleurs résultats. Nous avons abouti ainsi à une approche générique capable de restaurer simultanément les dégradations non-uniformes du fond et des formes.

12.2 Perspectives de recherche

Les travaux réalisés ouvrent diverses perspectives de recherche qui sont à la fois méthodologiques et applicatives.

Perspectives en terme d'applications

Nous avons décrit que quelques applications de la diffusion dans le domaine de la restauration. Pour conserver la concision du mémoire, les dégradations numériques n'ont pas été développées dans notre étude. Nous nous sommes, en effet, plus concentrés sur la restauration des dégradations physiques des documents. Ainsi, tous les tests sur l'augmentation de la résolution des images par diffusion (image magnification en anglais) et la correction des artéfacts de la compression JPEG par les méthodes à base de diffusion ont été retirés de ce mémoire. En amont du domaine de la restauration, la diffusion ouvre aussi la voie sur de nombreuses perspectives dans le domaine de l'amélioration des images pour rendre les textes plus lisibles ou embellir les images. De même l'usage de la diffusion comme prétraitement pour améliorer les performances des OCR est une piste qui doit être approfondie. Nous envisageons aussi poursuivre nos travaux sur les méthodes de reconstruction des parties manquantes de l'image par inpainting et plus particulièrement sur la reconstruction des parties irrégulières d'une forme et la segmentation des zones à reconstruire.

La diffusion peut être aussi utilisée pour augmenter les performances des méthodes de compression des images de documents. Nous voyons deux pistes de recherche à investir. La première consiste à appliquer une diffusion de façon à augmenter le signal par rapport au bruit et lisser les contours des traits avant d'appliquer une méthode quelconque de compression. La diffusion peut être utilisée dans ce cas pour différencier l'information utile qui doit être préservée du reste de l'information qui doit être éliminée. Les méthodes de restauration des dégradations du fond comme la séparation entre le recto et le verso ou l'élimination des taches qui sont décrites dans cette thèse peuvent aussi contribuer à éliminer une information inutile et augmenter ainsi le taux de compression. La seconde piste consiste à utiliser la géométrie simple et redondante des formes des caractères pour ne conserver que les lignes aléatoires ou régulières des pixels de l'image originale et reconstruire l'image entière par diffusion.

La diffusion peut aussi contribuer à extraire des caractéristiques robustes et des zones d'intérêts pertinentes pour améliorer les mesures de similarité d'images de texte. Ces travaux sont directement en relation avec ceux qui sont actuellement conduits au laboratoire sur la similarité d'images de mots (Wordspotting) ou d'images de caractères (Word-retrieval sur les manuscrits ou Transcription Assistée par Ordinateur des documents imprimés).

Perspectives méthodologiques

Du point de vue méthodologique, nous voulons approfondir nos travaux sur le paradigme local/global (Chapitre 4) et sur la complémentarité entre les informations colorimétriques et spatiales dans les mécanismes de segmentation ou de diffusion. La proposition d'un cadre commun diffusion/segmentation par Mean-Shift (Chapitre 10) est une première étape. Nous souhaitons aller plus loin et développer d'autres modèles combinés diffusion/segmentation dans un espace spatio-colorimétrique qui mélange à la fois les informations spatiales et colorimétriques.

EDP : Définitions et notations

Cette annexe introduit les différentes notations utilisées tout au long de ce chapitre et rappelle quelques connaissances de base. La profusion de notations différentes d'un article à l'autre, nous oblige à définir, une fois pour toute, les notations qui seront utilisées dans ce mémoire.

A.1 Définition des EDP

Soit Ω l'ensemble de tous les points (x, y) possibles de l'image, on définit une image I comme une fonction, à l'instant t par

$$\begin{aligned}
 I & : \quad \Omega \times \mathbb{R} & \longrightarrow & \quad \mathbb{R} \\
 & ((x, y), t) & \longrightarrow & \quad I(t, x, y)
 \end{aligned}
 \tag{A.1}$$

Une Équation aux Dérivées Partielles (EDP) est donnée par la relation F entre la dérivée $\frac{\partial I}{\partial t}$ et les dérivées successives de l'image I suivant x et y jusqu'à l'ordre n qui représente le degré de l'EDP.

$$\left\{ \begin{array}{l}
 I(0, x, y) = I_0(x, y) \\
 \frac{\partial I}{\partial t}(t, x, y) = F\left(I, \frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}, \frac{\partial^2 I}{\partial x^2}, \frac{\partial^2 I}{\partial y^2}, \frac{\partial^2 I}{\partial x \partial y}, \dots, \frac{\partial^n I}{\partial x^p \partial y^{n-p}}\right)
 \end{array} \right.
 \tag{A.2}$$

Les conditions initiales, à l'instant $t = 0$, sont données par l'image dégradée $I(0, x, y) = I_0(x, y)$. La définition de ces conditions initiales, à partir de laquelle un système évolue, permet de converger vers une solution unique. La démonstration de l'existence de relations entre la formulation des EDP et l'espace multi-échelle en traitement des images est donnée dans [DER-96].

A.2 Application aux images discrètes

Les EDP étant définies dans le domaine continu, l'application aux images numériques nécessite la discrétisation des équations. En écrivant la dérivée $\frac{\partial I}{\partial t}$ sous la forme d'un taux

A.2. APPLICATION AUX IMAGES DISCRÈTES

d'accroissement $\frac{\partial I}{\partial t} \approx \frac{I(t+dt, x, y) - I(t, x, y)}{dt}$, et en appliquant le schéma de résolution d'Euler, une EDP définie dans le domaine continu peut s'écrire dans le domaine discret sous la forme d'une suite d'images $I^n(x, y)$ construite par récurrence à partir de l'image à traiter.

$$\begin{aligned} \begin{cases} I(0, x, y) &= I_0(x, y) \\ \frac{\partial I}{\partial t}(t, x, y) &= F(I) \end{cases} &\Leftrightarrow \begin{cases} I(0, x, y) &= I_0(x, y) \\ I(t+dt, x, y) &= I(t, x, y) + dt \times F(I) \end{cases} \\ &\Leftrightarrow \begin{cases} I^0(x, y) &= I_0(x, y) \\ I^{n+1}(x, y) &= I^n(x, y) + dt \times F(I^n) \end{cases} \end{aligned}$$

L'incrément dt ne doit pas être trop petit pour que le nombre d'itérations nécessaire ne soit pas trop grand. Mais il ne doit absolument pas être trop grand si on ne veut pas osciller autour de la solution et rendre le système numériquement instable. Ce schéma de résolution par récurrence impose de convertir les valeurs d'intensités de chaque pixel en réel alors qu'elles sont généralement codées sur 256 valeurs discrètes.

La discrétisation des dérivées de l'image [DER-90, LUC-97, SCH-00] doit être minutieusement choisie pour éviter les dérives numériques. Les dérivées sont obtenues par convolution en prenant soin d'éviter les masques 3x3 comme le Sobel, le Prewitt qui font intervenir un lissage dans la direction opposée à celle de la dérivée. Comme les dérivées doivent être de préférence centrées, nous devons utiliser les dérivées simples basées sur les différences premières :

$$\frac{\partial I}{\partial x} \approx I \otimes \frac{1}{2} [-1 \ 0 \ 1] = \frac{1}{2} (I(t, x+1, y) - I(t, x-1, y)) \quad (\text{A.3})$$

$$\frac{\partial I}{\partial y} \approx I \otimes \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \frac{1}{2} (I(t, x, y+1) - I(t, x, y-1)) \quad (\text{A.4})$$

Les dérivées secondes sont calculées numériquement de préférence avec les convolutions suivantes :

$$\begin{aligned} \frac{\partial^2 I}{\partial x^2} &\approx I \otimes \frac{1}{4} [1 \ -2 \ 1] \\ &\approx \frac{1}{4} (I(t, x+1, y) - 2 \times I(t, x, y) + I(t, x-1, y)) \end{aligned} \quad (\text{A.5})$$

$$\begin{aligned} \frac{\partial^2 I}{\partial y^2} &\approx I \otimes \frac{1}{4} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \\ &\approx \frac{1}{4} (I(t, x, y+1) - 2 \times I(t, x, y) + I(t, x, y-1)) \end{aligned} \quad (\text{A.6})$$

A.3. DÉRIVÉES SUCCESSIVES D'UNE IMAGE ET LEURS SIGNIFICATIONS

$$\begin{aligned}
 \frac{\partial^2 I}{\partial x \partial y} &\approx I \otimes \frac{1}{4} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} \\
 &\approx \frac{1}{4} (I(t, x-1, y-1) + I(t, x+1, y+1)) \\
 &\quad - \frac{1}{4} (I(t, x-1, y+1) + I(t, x+1, y-1))
 \end{aligned} \tag{A.7}$$

A.3 Dérivées successives d'une image et leurs significations

Pour simplifier les notations tout au long de ce mémoire, nous n'écrirons pas les variables des fonctionnelles et remplacerons systématiquement $I(t, x, y)$ par I . De même, nous noterons les dérivées partielles (premières et secondes) de l'image I par les notations spécifiées dans le système suivant :

$$\begin{cases} \frac{\partial I}{\partial t}(t, x, y) = I_t \\ \frac{\partial I}{\partial x}(t, x, y) = I_x \\ \frac{\partial I}{\partial y}(t, x, y) = I_y \end{cases} \quad \begin{cases} \frac{\partial^2 I}{\partial x^2}(t, x, y) = I_{xx} \\ \frac{\partial^2 I}{\partial y^2}(t, x, y) = I_{yy} \\ \frac{\partial^2 I}{\partial x \partial y}(t, x, y) = I_{xy} \end{cases} \tag{A.8}$$

Les isophotes représentent les courbes d'intensité constante de l'image où, pour (x, y) dans Ω , $I(t, x, y) = c$. Le long de chaque isophote, aussi appelée ligne de niveaux, on peut définir le gradient de l'image $\nabla I = \begin{pmatrix} I_x \\ I_y \end{pmatrix}$ qui est parallèle à la direction du changement maximal d'intensité et perpendiculaire à la courbe des isophotes et donc à la variation minimale des intensités de l'image (Fig.A.1).

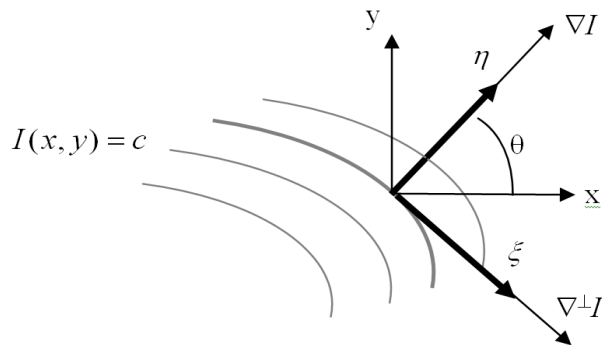


Figure A.1 : Représentation des isophotes, du vecteur gradient et des tangentes aux isophotes.

La magnitude $\|\nabla I\|$ indique le niveau des variations de l'image et θ représente l'orientation du vecteur gradient dans le repère d'origine (Eq. A.9).

A.3. DÉRIVÉES SUCCESSIVES D'UNE IMAGE ET LEURS SIGNIFICATIONS

$$\begin{cases} \|\nabla I\| = \sqrt{I_x^2 + I_y^2} \\ \theta = \arctan\left(\frac{I_x}{I_y}\right) \end{cases} \begin{cases} I_x = \|\nabla I\| \cos \theta \\ I_y = \|\nabla I\| \sin \theta \end{cases} \quad (\text{A.9})$$

La transformation inverse qui permet de retrouver les dérivées de l'image à partir de la magnitude et de l'orientation du gradient est évidente et est donnée par le système A.9. La magnitude du gradient est un indicateur important pour les EDP car elle permet de différencier les zones homogènes de l'image où $\|\nabla I\| < \epsilon$ des zones de fortes variations d'intensités où $\|\nabla I\| \geq \epsilon$, en particulier autour des contours des objets. Toutefois, la magnitude du gradient ne permet cependant pas de différencier les variations d'intensités issues du bruit de l'image de celles des contours des traits.

On notera η le vecteur gradient normalisé (Eq. A.10) et ξ le vecteur tangent aux isophotes (Eq. A.11) avec $\xi \perp \eta$.

$$\eta = \frac{\nabla I}{\|\nabla I\|} = \frac{1}{\|\nabla I\|} \begin{pmatrix} I_x \\ I_y \end{pmatrix} \quad (\text{A.10})$$

$$\xi = \frac{\nabla^\perp I}{\|\nabla I\|} = \frac{1}{\|\nabla I\|} \begin{pmatrix} -I_y \\ I_x \end{pmatrix} \quad (\text{A.11})$$

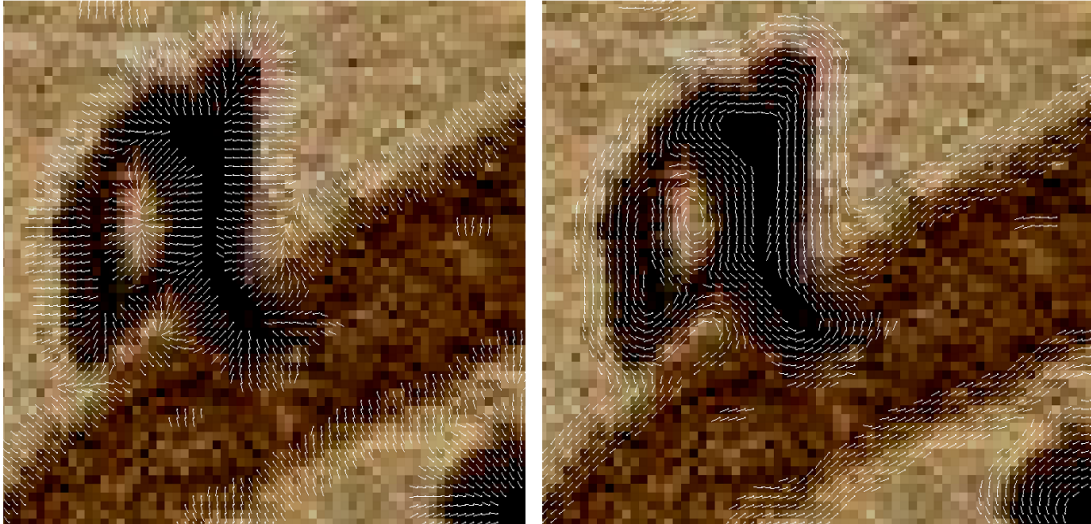


Figure A.2 : Pour $\|\nabla I\| > \epsilon$, affichage des images, de gauche à droite respectivement, du vecteur gradient η et de la tangente aux isophotes ξ .

Les vecteurs (η, ξ) forment un nouveau repère en chaque point (x, y) de Ω (aussi appelé (v, w) dans le repère de Gauge) dans lequel on peut définir les dérivées de l'image dans les directions du gradient η et dans la direction des tangentes aux isophotes ξ . Le nouveau repère (η, ξ) n'est qu'une rotation d'angle θ du repère cartésien original :

A.3. DÉRIVÉES SUCCESSIVES D'UNE IMAGE ET LEURS SIGNIFICATIONS

$$\begin{aligned} I_\eta &= \cos \theta \times I_x + \sin \theta \times I_y \\ I_\xi &= \cos \theta \times I_y - \sin \theta \times I_x \end{aligned} \quad (\text{A.12})$$

Ainsi, pour exprimer les relations entre les coordonnées dans l'ancien repère spatial et les coordonnées dans le nouveau repère (η, ξ) , il suffit d'appliquer une rotation d'angle θ .

Les dérivées premières I_η et I_ξ représentent respectivement les variations d'intensités de l'image le long du vecteur gradient η et dans la direction des tangentes aux isophotes ξ . Intuitivement, il est facile de concevoir que $I_\xi = 0$ car la variation des intensités de l'image suivant la tangente aux isophotes est nulle. En effet, une isophote ou bien une ligne de niveau d'une image est toujours tangente à la direction où les variations d'intensités de l'image est nulle. De même, la dérivée I_η est égale à $\|\nabla I\|$ car la dérivée dans la direction du gradient indique la variation maximale des intensités.

On définit la matrice Hessienne, la matrice symétrique des dérivées secondes :

$$H = \begin{pmatrix} I_{xx} & I_{xy} \\ I_{yx} & I_{yy} \end{pmatrix} \quad (\text{A.13})$$

D'après la formule de Taylor Lagrange, la variation $I(x+h, y+k) - I(x, y)$ à l'ordre 2 s'écrit :

$$\begin{aligned} I(x+h, y+k) - I(x, y) &= (h, k) \nabla I + \frac{1}{2} (h, k) H \begin{pmatrix} h \\ k \end{pmatrix} + O(h^2 + k^2) \\ &= hI_x + kI_y + \frac{1}{2} (h^2 I_{xx} + 2hk I_{xy} + k^2 I_{yy}) + O(h^2 + k^2) \end{aligned}$$

Les dérivées premières et les dérivées secondes apportent la majeure partie de l'information dans l'estimation des variations de l'image. Les dérivées d'ordre supérieur représentent une information moindre. De plus, elles sont numériquement plus instables et sensibles au bruit. D'ailleurs, le Laplacien de l'image I notée aussi $\nabla I = I_{xx} + I_{yy} = \text{trace}(H)$ joue un rôle très important dans les EDP pour l'estimation du bruit dans une image. Le passage par zéro du Laplacien permet aussi de détecter précisément les contours des objets, c'est-à-dire les maxima et minima des gradients de l'image.

La dérivée seconde I_{uu} dans une direction u quelconque est donnée par le changement de base d'une forme quadratique

$$I_{uu} = \frac{\partial^2 I}{\partial u^2} = u^\perp H u = \begin{pmatrix} u_1 & u_2 \end{pmatrix} \begin{pmatrix} I_{xx} & I_{xy} \\ I_{yx} & I_{yy} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = u_1^2 I_{xx} + 2u_1 u_2 I_{xy} + u_2^2 I_{yy}$$

Ainsi, on définit les dérivées de l'image dans le repère (η, ξ) par $I_{\xi\xi} = \xi^t H \xi$, $I_{\eta\eta} = \eta^t H \eta$ et $I_{\eta\xi} = \eta^t H \xi$ avec H la matrice Hessienne de I . On peut retrouver les dérivées secondes dans le nouveau repère (η, ξ) en fonction des dérivées secondes et premières de l'image dans le repère d'origine par le calcul direct en coordonnées cartésiennes ou en coordonnées polaires :

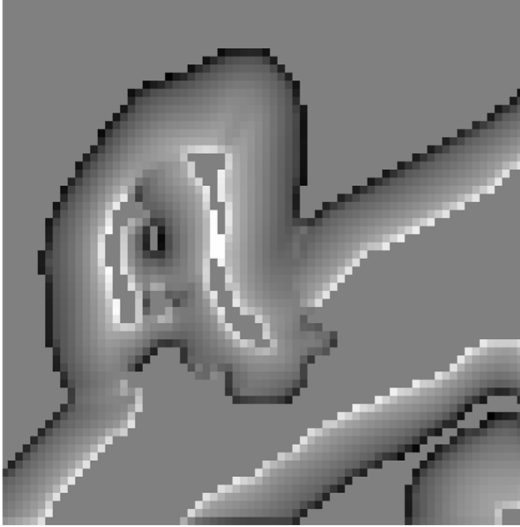
$$\begin{aligned} I_{\xi\xi} &= \frac{1}{\|\nabla I\|^2} \begin{pmatrix} -I_y & I_x \end{pmatrix} \begin{pmatrix} I_{xx} & I_{xy} \\ I_{yx} & I_{yy} \end{pmatrix} \begin{pmatrix} -I_y \\ I_x \end{pmatrix} = \frac{I_x^2 I_{yy} - 2I_x I_y I_{xy} + I_y^2 I_{xx}}{I_x^2 + I_y^2} \\ &= I_{yy} \cos^2 \theta - 2 \cos \theta \sin \theta I_{xy} + I_{xx} \sin^2 \theta \end{aligned}$$

A.3. DÉRIVÉES SUCCESSIVES D'UNE IMAGE ET LEURS SIGNIFICATIONS

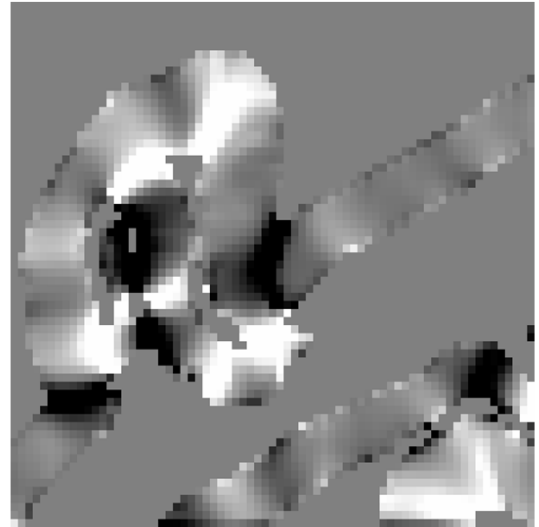
$$I_{\eta\eta} = \frac{1}{\|\nabla I\|^2} \begin{pmatrix} I_x & I_y \end{pmatrix} \begin{pmatrix} I_{xx} & I_{xy} \\ I_{yx} & I_{yy} \end{pmatrix} \begin{pmatrix} I_x \\ I_y \end{pmatrix} = \frac{I_x^2 I_{xx} + 2I_x I_y I_{xy} + I_y^2 I_{yy}}{I_x^2 + I_y^2} \\ = I_{xx} \cos^2 \theta + 2 \cos \theta \sin \theta I_{xy} + I_{yy} \sin^2 \theta$$

$$I_{\eta\xi} = \frac{1}{\|\nabla I\|^2} \begin{pmatrix} I_x & I_y \end{pmatrix} \begin{pmatrix} I_{xx} & I_{xy} \\ I_{yx} & I_{yy} \end{pmatrix} \begin{pmatrix} -I_y \\ I_x \end{pmatrix} = \frac{(I_x^2 - I_y^2) I_{xy} + I_x I_y (I_{yy} - I_{xx})}{I_x^2 + I_y^2}$$

La dérivée seconde $I_{\eta\eta}$ de I dans la direction du gradient η est utilisée pour localiser précisément le contour (lorsque $I_{\eta\eta} = 0$). La dérivée seconde $I_{\xi\xi}$ de I dans la direction des tangentes aux isophotes ξ mesure la courbure des courbes d'iso-intensités (Fig.A.3). On obtient $I_{\xi\xi} = 0$ pour les isophotes ayant la forme d'une ligne et $I_{\xi\xi}$ est fortement positive ou négative pour les isophotes respectivement convexes ou concaves.



(a) $I_{\eta\eta} \approx 0$ localise les contours, $I_{\eta\eta}$ est affichée avec des valeurs entre 0 et 255 décalée de +128 pour afficher les valeurs négatives entre 0 et 127 et les valeurs positives entre 129 et 255. La position exacte des contours est donc affichée avec une valeur de 128 exactement.



(b) Le signe de $I_{\xi\xi}$ mesure la courbure des isophotes. $I_{\xi\xi}$ est affichée avec des valeurs entre 0 et 255 décalée de +128. Les valeurs négatives de $I_{\xi\xi}$ (concavités) sont comprises entre 0 et 127 et des valeurs positives de $I_{\xi\xi}$ (convexités) entre 129 et 255.

Figure A.3 : Affichage de $I_{\eta\eta}$ et $I_{\xi\xi}$ sur l'image de document de la figure A.2

??

On note généralement κ la courbure des isophotes donnée par :

$$\kappa = \frac{I_{\xi\xi}}{I_\eta} = \frac{I_x^2 I_{yy} - 2I_x I_y I_{xy} + I_y^2 I_{xx}}{(I_x^2 + I_y^2)^{3/2}} \quad (\text{A.14})$$

La dérivée $I_{\eta\xi}$ joue le même rôle que $I_{\xi\xi}$ pour les lignes orthogonales aux isophotes (aussi appelées lignes de flux). $I_{\eta\xi}$ mesure la courbure des lignes orthogonales aux lignes de niveaux.

A.4. LE TENSEUR DE DÉRIVÉES PREMIÈRES ET SON INTERPRÉTATION GÉOMÉTRIQUE

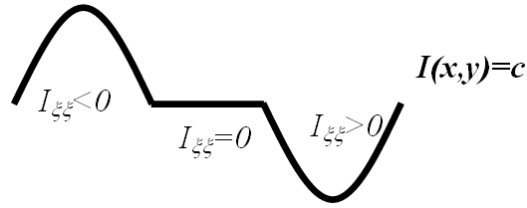


Figure A.4 : *Signe de la courbure et convexité de l'isophote*

On notera le vecteur nabla par $\nabla = \left(\frac{\partial}{\partial x} \right)$ et la divergence d'un vecteur $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ par $div(u) = \nabla \cdot \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial y}$.

Cette divergence joue un rôle très important dans la plupart des EDP dans les sections suivantes. On remarquera que la divergence du gradient représente une autre écriture du Laplacien (Eq.A.15).

$$div(\nabla I) = \nabla \cdot \begin{pmatrix} I_x \\ I_y \end{pmatrix} = \frac{\partial I_x}{\partial x} + \frac{\partial I_y}{\partial y} = I_{xx} + I_{yy} = \Delta I \quad (\text{A.15})$$

A.4 Le tenseur de dérivées premières et son interprétation géométrique

Le tenseur de dérivées premières est une notion importante dans la géométrie différentielle car elle représente la structure locale de l'image. On retrouve par exemple le tenseur des dérivées premières dans le détecteur de points d'intérêts de Harris [BRE-03]. On retrouve en plus les tenseurs dans de nombreuses EDP. On définit le tenseur T des dérivées premières par le produit des gradients (Eq.A.16).

$$T = \nabla I \times \nabla I^t = \begin{pmatrix} I_x \\ I_y \end{pmatrix} (I_x \ I_y) = \begin{pmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{pmatrix} \quad (\text{A.16})$$

Le tenseur en géométrie différentielle joue le même rôle que la matrice de covariance en statistiques. Il construit une matrice symétrique à partir de données corrélées qui pourront être analysées plus facilement par analyse factorielle. Cette analyse factorielle permettra de trouver, après une rotation, un nouveau repère dans lequel les données seront au maximum décorréelées.

En géométrie différentielle, le tenseur T est une matrice 2x2 définie semi-positive (i.e. ${}^t u T u \geq 0, \forall u$) et symétrique ($T = T^t$) donc diagonalisable. On peut donc trouver une matrice de passage P et une matrice diagonale Λ vérifiant

$$T = P \Lambda P^{-1} = P \Lambda P^t \quad (\text{A.17})$$

P est formé de deux vecteurs propres orthonormaux $P = (u^\rho u^{\rho\perp}) = \begin{pmatrix} u_1 & -u_2 \\ u_2 & u_1 \end{pmatrix}$ vérifiant $\|u\| = \|u^\perp\| = 1$ qui représente une matrice de rotation d'angle θ donc $P =$

A.4. LE TENSEUR DE DÉRIVÉES PREMIÈRES ET SON INTERPRÉTATION GÉOMÉTRIQUE

$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$. La matrice diagonale Λ est composée des deux valeurs propres avec $\Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$. Le tenseur T s'écrit alors :

$$\begin{aligned} T &= P\Lambda^t P = \begin{pmatrix} u_1 & -u_2 \\ u_2 & u_1 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} u_1 & u_2 \\ -u_2 & u_1 \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 u_1^2 + \lambda_2 u_2^2 & \lambda_1 u_1 u_2 - \lambda_2 u_1 u_2 \\ \lambda_1 u_1 u_2 - \lambda_2 u_1 u_2 & \lambda_1 u_2^2 + \lambda_2 u_1^2 \end{pmatrix} \end{aligned} \quad (\text{A.18})$$

Généralement, on utilise plutôt l'écriture vectorielle du changement de base que l'écriture matricielle (Eq. A.19).

$$\begin{aligned} D &= \begin{pmatrix} \lambda_1 u_1^2 + \lambda_2 u_2^2 & \lambda_1 u_1 u_2 - \lambda_2 u_1 u_2 \\ \lambda_1 u_1 u_2 - \lambda_2 u_1 u_2 & \lambda_1 u_2^2 + \lambda_2 u_1^2 \end{pmatrix} = \lambda_1 \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \begin{pmatrix} u_1 & u_2 \end{pmatrix} + \\ &\quad \lambda_2 \begin{pmatrix} -u_2 \\ u_1 \end{pmatrix} \begin{pmatrix} -u_2 & u_1 \end{pmatrix} = \lambda_1 u \times u^t + \lambda_2 u^\perp \times u^{\perp t} \end{aligned} \quad (\text{A.19})$$

Comme le tenseur est une matrice définie semi-positive, les valeurs propres λ_1 et λ_2 sont positives ou nulles. Dans cette base de vecteurs propres, le tenseur T représente alors généralement une ellipse dont l'axe principale u forme un angle θ avec l'axe des abscisses. Les valeurs propres donne les longueurs des axes de cette ellipse (Fig.A.5).

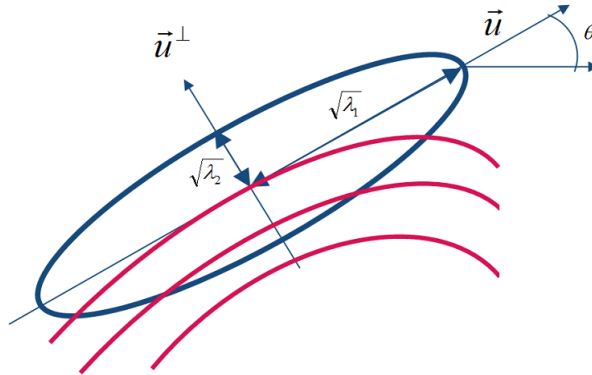


Figure A.5 : Interprétation géométrique du tenseur des dérivées premières

En notant le tenseur sous la forme d'une matrice 2×2 par $T = \begin{pmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{pmatrix}$, le calcul numérique des valeurs et des vecteurs propres impose une notation différente de façon à lever l'ambiguïté de la plus grande valeur propre. On notera les valeurs propres λ_+ λ_- (Eq.A.20) à la place de λ_1 λ_2 et les vecteurs propres associés que l'on notera dorénavant Θ_+ Θ_- à la place de u u^\perp (Eq.A.21).

$$\lambda_{+/-} = \frac{J_{11} + J_{22} \pm \sqrt{(J_{11} - J_{22})^2 + 4J_{12}^2}}{2} \quad (\text{A.20})$$

$$\Theta_{+/-} // \left(J_{22} - J_{11} \pm \sqrt{(J_{11} - J_{22})^2 + 4J_{12}^2} \right) \quad (\text{A.21})$$

Avec cette nouvelle notation, la plus grande valeur propre sera toujours λ_+ et le vecteur propre associé Θ_+ . Ainsi, le tenseur T s'écrira dorénavant

$$T = \lambda_- \times \Theta_- \Theta_-^t + \lambda_+ \times \Theta_+ \Theta_+^t \quad (\text{A.22})$$

Les vecteurs propres $\Theta_{+/-}$ donnent les orientations locales des variations de l'image et les valeurs propres $\lambda_{+/-}$ correspondantes (Eq.A.22) définissent l'énergie associée, c'est-à-dire l'amplitude de ces variations. Selon les valeurs de $\lambda_{+/-}$, divers cas peuvent se présenter :

- $\lambda_+ \approx \lambda_-$ avec des valeurs faibles, alors il n'y a pas de direction privilégiée, la zone de l'image est homogène, l'ellipse est donc un cercle isotrope ;
- $\lambda_+ \gg \lambda_-$ avec $\lambda_- \gg 1$ et λ_+ avec des valeurs faibles, alors la totalité de l'information est portée par l'axe principal du vecteur propre Θ_+ associé à λ_+ et on est en présence d'un fort contour
- $\lambda_+ \gg \lambda_- \gg 1$, alors l'information est dispersée sur les deux axes des deux vecteurs propres à directions orthogonales. Nous sommes alors à l'angle d'un coin dans l'image.

Enfin, nous définissons G_ρ (Eq.A.23) la fonction Gaussienne correspondant à une loi normale centrée d'écart-type σ (Fig.A.6).

$$G_\sigma(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (\text{A.23})$$

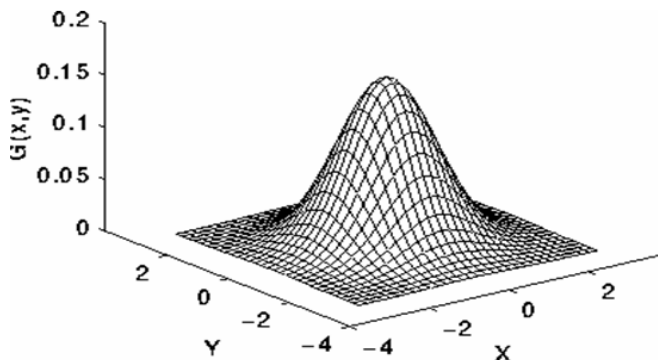


Figure A.6 : Représentation du noyau Gaussien 2D

A.5 L'application des EDP aux images couleurs

Si l'application d'un traitement par EDP sur des images en niveaux de gris est directe, son application sur des images couleurs ou plus généralement sur des signaux multicanaux ne l'est pas. En effet, appliquer directement les EDP des images en niveaux de gris pour chaque canal séparément a pour inconvénient de créer une décorrélation entre les différents canaux. Plusieurs indicateurs sont localisés indépendamment sur chacune des sources engendrant, le plus souvent, l'apparition de certains artefacts liés au rehaussement de faux contours propres

A.5. L'APPLICATION DES EDP AUX IMAGES COULEURS

à chaque canal. Si on considère une image couleur RVB classique à trois canaux, une image I se définit de la façon suivante :

$$\begin{aligned} I : \quad \Omega &\longrightarrow \mathbb{R}^3 \\ (x, y) &\longrightarrow (R(x, y), V(x, y), B(x, y)) \end{aligned} \quad (\text{A.24})$$

Par conséquent, le gradient couleur ∇I peut se définir par la matrice Jacobienne J de l'image I :

$$J = \begin{pmatrix} \frac{\partial R}{\partial x} & \frac{\partial R}{\partial y} \\ \frac{\partial V}{\partial x} & \frac{\partial V}{\partial y} \\ \frac{\partial B}{\partial x} & \frac{\partial B}{\partial y} \end{pmatrix} \quad (\text{A.25})$$

Une solution à ce problème consiste à appliquer un traitement par EDP en utilisant des indicateurs communs à toutes les composantes de l'image. Ces indicateurs sont calculés d'une manière conjointe en faisant appel au tenseur des dérivées premières définie comme suit [ZEN-86] :

$$T = {}^t J \times J = \begin{pmatrix} \frac{\partial R}{\partial x} & \frac{\partial V}{\partial x} & \frac{\partial B}{\partial x} \\ \frac{\partial R}{\partial y} & \frac{\partial V}{\partial y} & \frac{\partial B}{\partial y} \end{pmatrix} \begin{pmatrix} \frac{\partial R}{\partial x} & \frac{\partial R}{\partial y} \\ \frac{\partial V}{\partial x} & \frac{\partial V}{\partial y} \\ \frac{\partial B}{\partial x} & \frac{\partial B}{\partial y} \end{pmatrix} \quad (\text{A.26})$$

$$T = \begin{pmatrix} \frac{\partial R^2}{\partial x} + \frac{\partial V^2}{\partial x} + \frac{\partial B^2}{\partial x} & \frac{\partial R}{\partial x} \frac{\partial R}{\partial y} + \frac{\partial V}{\partial x} \frac{\partial V}{\partial y} + \frac{\partial B}{\partial x} \frac{\partial B}{\partial y} \\ \frac{\partial R}{\partial y} \frac{\partial R}{\partial x} + \frac{\partial V}{\partial y} \frac{\partial V}{\partial x} + \frac{\partial B}{\partial y} \frac{\partial B}{\partial x} & \frac{\partial R^2}{\partial y} + \frac{\partial V^2}{\partial y} + \frac{\partial B^2}{\partial y} \end{pmatrix} = \begin{pmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{pmatrix} \quad (\text{A.27})$$

Ce tenseur T se diagonalise dans une base orthonormale (Eq.A.18) et peut s'écrire sous la forme de l'équation A.22.

Bibliographie

- [ALL-03] B. ALLIER. Contribution à la numérisation des collections : apports des contours actifs, Thèse de doctorat, Institut National des Sciences Appliquées de Lyon, 2003.
- [ALV-93] L. Alvarez, F. Guichard, P. L. Lions, J. M. Morel. Axioms and Fundamental Equations of Images Processing, *Archive for Rational Mechanics and Analysis*, 123 (3), pp 199-257, 1993.
- [ALV-94] L. Alvarez, L. Mazorra. Signal and image restoration using shock filters and anisotropic diffusion. *SIAM Journal on Numerical Analysis*, vol. 31, n°2, pp.590-605, 1994.
- [AND-99] J. André, M.A. Chabin. Numériser les documents anciens : et après ?, 1999, *Les documents anciens, numéro spécial de Document Numérique*, vol. 3, n°1-2, pp. 7-11.
- [AU-02] G. Aubert, P. Kornprobst. *Mathematical Problems in Image Processing : Partial Differential Equations and the Calculus of Variations*, volume 147 of *Applied Mathematical Sciences*. Springer-Verlag, 2002.
- [BAI-00] H. Baird. State of the Art of Document Image Degradation Modeling. In *Proceedings of 4th IAPR International Workshop on Document Analysis Systems*, vol. 1, Rio de Janeiro (Brazil), 2000.
- [BAI-92] H . S. Baird. Document Image Defect Models, dans *Structured Document Image Analysis -Proc. of the IAPR 1990 Workshop on SSPR* : Springer, pp. 546-556, 1992.
- [BAI-93] H. Baird. Document Image defect Models and Their Uses, *Proc. of the Internatuonal Conférence on Document Analysis and recognition*, pp. 62-67, Tsukuba, Japan, 1993.
- [BAR-02] B. Nicolas, J. Guillemain , G. Péoc'h, P.Ract. La renaissance du livre ancien : bilan du projet DEBORA et perspectives d'avenir, Diplôme de conservateur de bibliothèque : Mémoire de recherche, Lyon, 2002.
- [BAL-00] C. Ballester, M. Bertalmio, V. Caselles, J. Verdera. Filling-in by joint interpolation of vector fields and grey levels. *University of Minnesota IMA TR*, 2000.

- [BEL-06] A. Belaïd, H. Cecotti. Reconnaissance de caractères : évaluation des performances, Dans Les documents écrits : de la numérisation à l'indexation par le contenu, pp. 311-361, 2006.
- [BEN-05] B. Bringier. Élaboration de modèles d'évaluation de la qualité du rendu couleur, Thèse de doctorat, Université de Poitier, 2005.
- [BER-00] M. Bertalmio, G. Sapiro, V. Caselles, C. Ballester. Image inpainting, Proceedings of the SIGGRAPH, pp. 417-424, ACM Press, Addison Wesley Longman, 2000.
- [BON-01] N.BONNET, J.CUTRONA. Improvement of unsupervised multi-component image segmentation through fuzzy relaxation, IASTED International Conference on Visualization, Imaging and Image Processing, 2001.
- [BOT-98] L. Bottou, P. Haffner, P.G. Howard, P. Simard, Y. Bengio, Y. LeCun, High quality document image compression with DjVu, journal of electronics imaging, 7(3) :410-428, 1998.
- [BOU-06] F. Le Bourgeois, La reconnaissance des structures, chapitre 3, dans le livre Les documents écrits de la numérisation à l'indexation par le contenu, traité IC2, Hermes Lavoisier, pp. 87-178, 2006.
- [BOU-07] F. Le Bourgeois, H. Emptoz. DEBORA : Digital AccEss to BOoks of the RenAissance, International Journal on Document Analysis and Recognition (IJ DAR), Special issue on the analysis of historical documents, Springer, Vo. 9, Nř 2-4, pp. 193-221, 2007.
- [BRE-03] S. Bres, JM. Jolion, F. Le Bourgeois. Traitement et analyse des images numériques, Hermès, ISBN 2-7462-0741-9, 2003.
- [BRO-92] R. Brockett, P. Maragos. Evolution equations for continuous-scale morphology, ICASSP, Vol. 3, pp. 125-128, 1992.
- [CAL-04] J. R. Caldas Pinto¹, Pedro Pina², Lourenço Bandeira¹, Luís Pimentell¹, Mário Ramalho. Underline Removal on Old Documents , ICIAR 2004, pp. 226-233, 2004.
- [CAO-03] H. Cao, X. Ding, C. Liu. Rectifying the Bound Document Image Captured by the Camera : A Model Based Approach, ICDAR , pp. 71-75, Edinburgh, Scotland, 2003.
- [CAT-92] F. Catté, J. M. Morel, P. L. Lions, T. Coll, Image Selective smoothing and edge detection by nonlinear diffusion, SIAM J. Numer. Anal., 29 :182-193, 1992.
- [CHA-94] P. Charbonnier, G. Aubert, M. Blanc-Féraud, M. Barlaud. Two deterministic half-quadratic regularization algorithms for computed imaging. Dans Proceedings of the International Conference on Image Processing, volume II, pp. 168-172, 1994.
- [CHA-97] P. Charbonnier, L. Blanc-Féraud, G. Aubert, M. Barlaud. Deterministic edge preserving regularization in computed imaging, IEEE Transactions on Image Processing, 6(2) :298.311, 1997.

- [CHA-00] T. Chan and J. Shen, Non-texture inpaintings by curvature-driven diffusions; technical Report 00-35, Department of Mathematics, UCLA, Los Angeles, 2000.
- [CHA-02] T.Chan, J.Shen, Mathematical models for local nontexture inpaintings, SIAM J.Appl. Math. 62 (3) pp. 1019-1043, 2002.
- [COC-95] J.P. Cocquerez, S. Philipp. Analyse d'images : filtrage et segmentation, Paris, Masson, 1995.
- [COU-00] O. Coulon, S.R. Arridge. Dual echo MR image processing using multispectral probabilistic diffusion coupled with shock filters, MIUA '2000, British Conference on Medical Image Understanding and Analysis, London, United-Kingdom, 2000.
- [CHE-95] Y. Cheng. Mean shift, mode seeking, and clustering. Pattern Analysis and Machine Intelligence, IEEE Transactions, vol.17, pp. 790-799, 1995.
- [CHO-97] S.I.Cho, H.Saito, S. Ozawa. A divide-and-Conquer Strategy In recovering Shape Of Book From Shading, Dans Proceedings of the 9th International Conference on Image Analysis and Processing (Volume II), volume 1311 de Lecture Notes in Computer Science, pp. 262-269, Florence, Italie, 1997.
- [COH-95] B. Cohen, I. Dinstein, M. Eyal. Computerized Classification of Color Textured Perthite Images, IEEE Proceedings of ICPR, 1996.
- [COM-99] D. Comaniciu, P. Meer. Mean Shift Analysis and Applications, Dans Proceeding Computer Vision, pp. 1197-1203, 1999.
- [COM-02] D. Comaniciu, P. Meer. Mean shift : A robust approach toward feature space analysis, IEEE Transactions on PAMI, vol. 24, no. 5, pp. 603-619, 2002.
- [COM-97] D. Comaniciu, P. Meer. Robust Analysis of Feature Spaces : Color Image Segmentation, IEEE Conf. Computer Vision and Pattern Recognition (CVPR'97), San Juan, Puerto Rico, pp. 750-755, 1997.
- [COU-04] F. Courteille, A. Crouzil, J.-D.Durou, P. Gurdjos. Towards shape from Shading under realistic photographic conditions, Dans Proceedings of the 17th International Conference on Pattern recognition (Volume II), pp. 277-280, Cambridge, Royaume-Uni, 2004.
- [DER-90] R. Deriche. Fast algorithms for low-level vision, IEEE Transactions on Pattern Analysis and Machine Intelligence, 1(12) :78.88, 1990.
- [DER-95] R. Deriche, O. Faugeras. Les EDP en traitement des images et vision par ordinateur, Technical report 2697, INRIA, 1995.
- [DER-96] R. Deriche, O. Faugeras. Les équations aux dérivées partielles en traitement des images et vision par ordinateur, Traitement du Signal, Vol. 13, pp. 551-577, 1996.
- [DRI-07] F. Drira, F. Le Bourgeois, H. Emptoz. OCR Accuracy Improvement Through a PDE-based Approach, Dans The 9th IEEE International Conference on Document Analysis and Recognition (ICDAR), Brazil, vol. 2, ISBN 0-7695-2822-8, pp. 1068-1072, 2007.

- [DUB-01] E. Dubois, A. Pathak. Reduction of bleed-through in scanned manuscripts documents, Dans Proceeding IS&T Image processing, Image Quality, Image capture Systems, pp. 177-180, Montreal, Canada, 2001.
- [DUD-98] O. Duda, P. E. Hart, D. G. Stork. Pattern classification and scene analysis, Wiley et sons, 1998.
- [EMP-02] E. Hubert, Bouché Richard, Lebourgeois Frank et al. DEBORA, Lyon, Projet européen n° LB5608A, 2000.
- [EMP-03] H. Emptoz, F. Le Bourgeois, V. Eglin, Y Leydier. La reconnaissance dans les images numérisées, OCR et transcription, reconnaissance des structures fonctionnelles et des méta-données, dans l'ouvrage La numérisation des textes et des images : Techniques et réalisations, éditions Presse de Lille 3, ISBN 2-84467-050-4, pp. 105-129, 2003.
- [EYN-07] L. Eynard, F. Le Bourgeois, H. Emptoz. Segmentation Of Broken Characters Using Pattern Matching, Dans Ninth International Conference on Pattern Recognition and Information Processing, PRIP'2007, ISBN 978-985-6744-29, pp. 101-107, Minsk, Belarus, 2007.
- [FED-03] C. Feddern, J. Weickert, B. Burgeth. Level-set methods for tensor valued images. Dans Proceeding Second IEEE Workshop on Geometric and Level Set Methods in Computer Vision, pp. 65-72, Nice, France, 2003.
- [FUK-75] K. Fukunaga, L.D. Hostetler. The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition, IEEE Trans. Information Theory, vol. 21, pp. 32-40, 1975.
- [FUK-90] K. Fukunaga. Introduction to Statistical Pattern Recognition, Boston, MA : Academic Press, 1990.
- [GAR-06] U. Garain, T. Paquet, L. Heutte. On foreground-background separation in low quality document images, International Journal on Document Analysis and Recognition, vol. 8, no.1, pp. 47-63, 2006.
- [GAS-03] F. Gasparini, R. Schettini, P.Gallina. Tunable cast remover for digital photographs. In Proceedings of SPIE, volume 5008 of Color Imaging VIII : Processing, Hardcopy, and Applications, pp. 92 :100, 2003.
- [GRE-90] P.J. Green. Bayesian reconstruction from emission tomography data using a modified em algorithm. IEEE Trans. Med. Imaging, MI-9(1) :84.93, 1990.
- [GUI-95] F. Guichard, J. Morel. Introduction to partial differential equations in image processing, Tutorial Notes, IEEE Int. Conf. Image Proc., Washington, 1995.
- [HIR-01] T. Hiroshi, H. Hiroya, N. Koichi. Image processing device for providing correction of a curvature distortion of a bound portion of a spread book, Brevet US6330050, Ricoh Company, LTD, 2001.
- [HOB-97] J. D. Hobby, T.K.Ho. Enhancing Degraded Document Images via Bitmap Clustering and Averaging , Dans Proceeding of the International

- Conference on Document Analysis and Recognition, 4th ICDAR, Germany, pp. 394-400, 1997.
- [HSI-89] J. Y. Hsiao , A. A. Sawchuk. Supervised Textured Image Segmentation Using Feature Smoothing and Probabilistic Relaxation Techniques, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.11 no.12, pp. 1279-1292, 1989.
- [HYV-99A] A. Hyvärinen, Survey on independent component analysis, Neural Computing Surveys, 2 :94-128, 1999, <http://www.icsi.berkeley.edu/jagota/NCS>.
- [HYV-99B] A. Hyvarinen, E. Oja. A fast fixed-point algorithm for independent component analysis, Neural Computation, 9(7) :1483-1492, 1997. Reprinted in Unsupervised Learning, G. Hinton and T. J. Sejnowski, eds., MIT Press, 1999.
- [JAC-96] P. Jackway, M. Deriche. Scale-space properties of the multiscale morphological dilation-erosion, IEEE Transactions on Pattern Analysis and Machine Intelligence 18(1) : 38-51, 1996.
- [KAN-93] T. Kanungo, R.M.haralick, I. Phillips. Global and Local Document Degradation Models, Dans Proceedings of the 2nd International Conference on Document Analysis and Recognition, pp. 730-734, 1993.
- [KAN-94] T. Kanungo, R. M. Haralick, H. S. Baird, W. Stuetzle, D. Madigan. Document Degradation Models : Parameter Estimation and Model Validation. Dans Proceedings of IAPR Workshop on Machine Vision Applications, Kawasaki (Japan), pp. 552-557, 1994.
- [KAN-00] T. Kanungo, R. M. Haralick, H. S. Baird, W. Stuezle, D. Madigan. A Statistical, Nonparametric Methodology for Document Degradation Model Validation. IEEE Trans. on Pattern Analysis and Machine Intelligence, 22(11) :1209-1223, 2000.
- [KAS-99] M. Kashimura, T. Nakajima, N. Onda, H. Saito, S. Ozawa. Practical introduction of image processing technology to digital archiving of rare books , Dans Proceedings of the international conference on signal processing application technology, Toronto, Canada, 1999.
- [KIM-92] N. N. Kimia, A. Tannenbaum, S. W. Zucker. On the evolution of curves via a function of curvature, Journal of mathematical analysis and applications, 163(2), 1992.
- [KIM-00] R. Kimmel, R. Malladi, N. Sochen. Images as embedded maps and minimal surfaces : movies, color, texture, and volumetric medical images. International Journal of Computer Vision, 39(2) :111.129, 2000.
- [KOE-84] Koenderink J., The Structure of Images, Biol. Cybern, pp. 363-370, 1984.
- [KOR-96] P. Kornprobst, R. Deriche, G. Aubert. Image restoration via PDE's. In First Annual Symposium on Enabling Technologies for Law Enforcement and Security - SPIE Conference 2942 : Investigative Image Processing., Boston, Massachusetts, USA, 1996.

- [KOR-97A] P. Kornprobst, R. Deriche, G. Aubert. Nonlinear operators in image restoration. In Proceedings of the International Conference on Computer Vision and Pattern Recognition, pp. 325-331, Puerto Rico, 1997.
- [KOR-97B] P. Kornprobst, R. Deriche, G. Aubert. Image coupling, restoration and enhancement via PDE's, Dans Proceeding Int. Conf. on Image Processing, pp. 458-461, Santa-Barbara (USA), 1997.
- [KOR-98] P. Kornprobst. Contributions à la Restauration d'Images et à l'Analyse de Séquences : Approches Variationnelles et Solutions de Viscosité. Thèse de doctorat, Université de Nice-Sophia Antipolis, 1998.
- [KOR-06] P. Kornprobst, G. Aubert. Explicit Reconstruction for Image inpainting, Rapport INRIA N° 5905, 2006.
- [KUM-03] S. Kumar, A. C. Loui, M. Hebert. An Observation-Constrained Generative Approach for Probabilistic Classification of Image Regions. Image and Vision Computing, vol. 21, no. 1, pp 87-97, 2003.
- [LEY-04] Y Leydier., F. LeBourgeois, H. Emptoz. Serialized k-means for adaptive color image segmentation : Application to document images and others, Document Analysis Systems VI, 6th international workshop, pp. 252-263, Florence, ITALY, 2004.
- [LI-96] Y. Li, D. Lopresti, G. Nagy, A. Tomkins. Validation of Image Defect Models for Optical Character Recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 18, Issue 2, pp. 99 - 108, 1996.
- [LU-02] H. Lu et al. An Objective Distortion Measure for Binary Document Images Based on Human Visual Perception, in proc. of 16th International Conference on Pattern Recognition (ICPR'02), Vol. 4, pp. 402-406, 2002.
- [LUC-97] L. Lucido, R. Deriche, L. Alvarez, V. Rigaud. Sur quelques schémas numériques de résolution d'équations aux dérivées partielles pour le traitement d'images. Rapport de Recherche 3192, INRIA, 1997.
- [NAG-92] N.G.NAGY. What does a machine need to read a document ? Proc. of the 1st Annual Symposium on Document Analysis and Information Retrieval, University of Nevada, Las Vegas, Nevada, pp. 1-10, 1992.
- [NIE-94] W.Niessen , B.T. H.Romery, L.Florack, A.Salden, M.Vergever. Nonlinear Diffusion of Scalar Images Using Well-Posed Differential Operators, IEEE, Editor, IEEE Conf. On Comp. Vision and Pattern Recognition, pp. 92-120, Seattle, 1994.
- [OSH-88A] S. Osher, J. A. Sethian. Fronts propagating with curvature-dependent speed : algorithms based on Hamilton-Jacobi formulations. J. Comput. Phys., 79(1) :12-49, 1988.
- [OSH-88B] S. Osher, L. Rudin. Featured-oriented image enhancement using shock filters, SIAM J. on Numerical Analysis (27), pp. 919-940, 1988.
- [PAR-97] R.J.Parker. Algorithms for image processing and computer vision, New York : Wiley, xi, 1997.

- [PARZ-61] E. Parzen. On estimation of a probability density functions and modes, *Ann Math. Stat.* (33), pp. 1065-1076, 1961.
- [PER-90] P. Perona, J.Malik. Scale-Space and Edge Detection Using Anisotropic Diffusion, *IEEE Trans. Pattern. Analysis and Machine Intelligence*, pp. 629-639, 1990.
- [PER-94] P.Perona, T. Shiota, J.Malik. Anisotropic Diffusion, Kluwer Academic In Bart M. Ter Haar Romeny Editor, pp. 229-254, 1994.
- [SCH-94] M. Schmitt, J.Mattioli. *Morphologie mathématique* , Masson, 1994.
- [SCH-00] H. Scharr, J. Weickert. An anisotropic diffusion algorithm with optimized rotation invariance. In G. Sommer, N. Krüger, C. Perwass (Eds.), *Mustererkennung 2000*, Springer, Berlin, pp. 460-467, 2000.
- [SOC-96] N. Sochen, R. Kimmel, R. Malladi. From high energy physics to low level vision. Technical Report 39243, LBNL report, UC Berkeley, 1996.
- [SOC-98] N. Sochen, R. Kimmel, R. Malladi. A geometrical framework for low level vision. *IEEE Transaction on Image Processing*, Special Issue on PDE based Image Processing, 7(3) :310.318, 1998.
- [SER-04] N. Serrano, A. E. Savakis, J. Luo. Improved scene classification using efficient low-level features and semantic cues. *Pattern Recognition*, vol. 37, no. 9, pp. 1773-1784, 2004.
- [SET-85] J. A. Sethian. Curvature and the evolution of fronts. *Comm. in Math. Phys.*, 1985.
- [SET-87] J. Sethian. Levels sets and fast marching methods, Thèse de doctorat, Stanford Univ., Palo Alto, Calif., 1987.
- [SHA-00] G. Sharma. Cancellation of show-through in duplex scanning, Appears in *Proceedings IEEE International Conference on Image processing*, vol. 2, pp.609-612, 2000.
- [SHI-96] Z. Shi, V. Govindaraju. character image enhancement by selective region-growing *Pattern Recognition Letters*, no. 17, pp. 523-527, 1996.
- [SMI-04] E. Smigiel, A. belaid, H. Hamza, Self-organizing Maps and Ancient Documents, *Document Analysis Systems VI*, 6th international workshop, pp.125-134, Florence, ITALY, 2004.
- [STE-95] *Mémoire du monde : Principes directeurs pour la sauvegarde du patrimoine documentaire /document élaboré pour l'UNESCO au nom de l'IFLA par Stephen Foster, Jan Lyall, Duncan Marshall et Roslyn Russel.* - Paris : UNESCO, 1995.
- [SMI-04] E. Smigiel, A. belaid, H. Hamza. Self-organizing Maps and Ancient Documents, *Document Analysis Systems VI*, 6th international workshop, pp.125-134, Florence, ITALY, 2004.
- [STA-00] J.-L. Starck, P. Querre, D. Donoho Karhunen-Loeve Multispectral and Multiscale Image Restoration, *SPIE conference on Signal and Image Processing : Wavelet Applications in Signal and Image Processing VIII*, San Diego, pp. 721-732, 2000.

- [TAK-01] H. Takahashi, H. Hiroya, N. Koichi, Image processing device for providing correction of a curvature distortion of a bound portion of a spread book. Brevet US6330050, Ricoh Company, LTD, 2001.
- [TER-02] R. Terebes, O. Laviolle, P. Baylou, M. Borda. Mixed anisotropic diffusion, Proceedings of IEEE International Conference on Pattern Recognition, Quebec, Canada, vol. 3, pp. 760-764, 2002.
- [TIK-77] A. Tikhonov, V. Arsenin. Solution of Ill-Posed Problems, Winston and Wiley, 1977.
- [TOM-98] C. Tomasi, R. Manduchi. Bilateral Filtering for Gray and Color Images, Proceedings of the IEEE International Conference on Computer Vision, Bombay, India, 1998.
- [TON-04] A. Tonazzini, L. Bedini, E. Salerno. Independent Component Analysis for Document Restoration, International Journal on Document Analysis and Recognition (IJ DAR), pp. 17-27, 2004.
- [TOR-03] A. B. Torralba. Contextual Priming for Object Detection. International Journal of Computer Vision, vol. 53, no. 2, pp. 169 :191, 2003.
- [TRE-04] A. Trémeau, C. Fernandez-Maloigne, P. Bonton. Image numérique couleur : de l'acquisition au traitement, 2004.
- [TRI-03] E. Trinh. De la numérisation à la consultation des documents anciens, Thèse de doctorat dans le cadre d'un projet européen Dépora et d'un projet industriel, Book Restorer, INSA de Lyon, France, 2003.
- [TSC-01] D. Tschumperlé, R. Deriche. Constrained and Unconstrained PDEs for Vector Image Restoration. Dans Proceedings of Scandinavian Conference on Image Analysis (SCIA'2001), pp. 153-160, 2001.
- [TSC-02A] D. Tschumperlé, R. Deriche. Diffusion PDE's on Vector-valued Images : Local Approach and Geometric Viewpoint. Dans IEEE Signal Processing Magazine, Vol 19. No 5. pp. 16-25, 2002.
- [TSC-02B] D. Tschumperlé. PDE's based regularization of multivalued images and applications. Thèse de doctorat, Université de Nice-Sophia Antipolis/France, 2002.
- [TSC-05] D. Tschumperlé, R. Deriche. Vector-Valued image regularisation with PDE's : A common framework for different applications. IEEE transactions on Pattern Analysis and Machine Intelligence, vol.27, No 4, 2005.
- [TSI-03] Y. Tsin, V. Ramesh, T. Kanade. Statistical Calibration of the CCD Imaging Process. In IEEE International Conference on Computer Vision (ICCV), volume 1, pp. 480-487, Vancouver, Canada, 2001.
- [Tso-04] Y.C.Tsoi, M.S. Brown. Geometric and Shading correction of imaged print materials : a unified approach using boundary, Dans proceedings of the IEEE conference on Computer vision and Pattern recognition, Washington, D.C, Etats-Unis, 2004.

- [WAD-95] T. Wada, H. Ukida, T. Matsuyama. Shape from shading with interreflections under proximal light source : 3D shape reconstruction of unfolded book surface from a scanner image . Dans Proceedings of the 5th IEEE International Conference on Computer Vision, pp. 66-71, Boston, Massachusetts, Etats-unis, 1995.
- [WAD-97] T. Wada, H. Ukida, T. Matsuyama. Shape from Shading with Interreflections Under a Proximal Light Source : Distortion-Free Copying of an Unfolded Book, *International Journal of Computer Vision*, 24(2) : 125-135, 1997.
- [WAN-03] Q. Wang, T. Xia, C. L. Tan, L. Li. Directional Wavelet Approach to Remove Document Image Interference, *ICDAR 2003*, pp. 736-740, Edinburgh, 2003.
- [WEI-94] J. Weickert. Scale-space properties of nonlinear diffusion filtering with a diffusion tensor. Report No. 110. Laboratory of Technomathematics, University of Kaiserslautern, Germany, 1994.
- [WEI-96] J. Weickert. Anisotropic Diffusion in Image Processing, Laboratory of Technomathematics, Thèse de doctorat, University of Kaiserslautern, Germany, 1996.
- [WEI-97A] J. Weickert. a review of nonlinear diffusion filtering, dans *Scale-Space Theory in Computer Vision (Scale -Space)*, volume 1252 de *Lecture Notes in Computer Science*, Utrecht, the Netherlands, 1997.
- [WEI-97B] J. Weickert. Coherence-enhancing diffusion of colour images. 7th National Symposium on Pattern Recognition and Image Analysis, 1997.
- [WEI-98] J. Weickert. *Anisotropic Diffusion in Image Processing* , Stuttgart, teubner-Verlag, 1998.
- [WEI-01] J. Weickert, C. Schnörr. A theoretical framework for convex regularizers in PDE-based computation of image motion. *The International Journal of Computer Vision*, 45(3) :245.264, 2001.
- [WEI-03] J. Weickert. Coherence enhancing shock filters. Dans *Pattern Recognition, Lecture Notes in Computer Science*, Vol. 2781, Springer, Proceedings of DAGM 2003, pp. 1-8, Germany, 2003.
- [WHI-96] A. Whichello, H. Yan. Linking Broken Character Borders with Variable Sized Masks to Improve Recognition, *Pattern Recognition*, vol. 29, no. 8, pp. 1429-1435, 1996.
- [WOL-06] C. Wolf. Document Ink bleed-through removal with two hidden Markov random fields and a single observation field, *Technical Report RR-LIRIS-2006-019*, 2006
- [WU-02] C. Wu, G. Agam. Document Image De-warping for Text/Graphics Recognition, *Proc. of Joint IAPR and SPR 2002 Windsor*, Ontario, Canada, pp. 348-357, 2002.
- [YAM-04] A. Yamashita, A. Kawarago, T. Kanako, K.T.Miura. Shape Reconstruction and Image restoration for Non-Flat Surfaces of Documents with

- a Stereo Vision System, Dans Proceedings of the 17th International Conference on Pattern recognition (Volume I), pp. 482-485, Cambridge, Royaume-Uni, 2004.
- [YU-01] D.Yu, H.Yan, Reconstruction of Broken Handwritten digits based on structural morphological features, Pattern recognition, no. 34, pp. 235-254, 2001.
- [ZEN-86] S. Di Zenzo. Note : A note on the gradient of a multi-image, Computer Vision, Graphics, and Image Pro-cessing, vol. 33, no. 1, pp. 116-125, 1986.
- [ZHA-01] Z. Zhang, C.L. Tan, Recovery of Distorted Document Images from Bound Volumes. International Conference on Document Analysis and Recognition, ICDAR'01, Seattle, WA, pp. 429-433, 2001.
- [ZHA-03] Z. Zhang, C. L. Tan. Correcting Document Image Warping Based on Regression of Curved Text Lines, ICDAR 2003, pp. 589-593, Edinburgh, 2003.
- [ZHE-01] Q. Zheng, T. Kanungo, Morphological Degradation Models and Their Use in Document Image Restoration , Rapport LAMP-TR-065/CS-TR-4218/CAR-TR-962, University of Maryland, College Park, USA, 2001.