

UNIVERSITÉ DE M'SILA
FACULTÉ DES MATHÉMATIQUES ET DE L'INFORMATIQUE
DEPARTEMENT DES MATHÉMATIQUES

Mémoire

Présenté pour l'obtention du diplôme de **Master**

Domaine: Mathématiques et Informatiques

Filière: Mathématiques

Option: Mathématiques Appliquées et discrètes

Par

ZANAT NASSIRA

THÈME

Restauration des images de documents anciens

Soutenu le : ../06/ 2016

Devant le jury composé de :

- 1).....
- 2).....
- 3).....

Dirigé par:

*Mr.*Ben Hamidouch.N

Année: **2015/2016**

Remerciements

Je tiens à remercier, en premier lieu, **Mon Dieu** qui m'a donné la force de rédiger ce modeste travail.

Je tiens à remercier les membres du jury, qui ont accepté d'évaluer mon travail de mémoire.

Je tiens à remercier Ben Hamidouch.N directeur de mon mémoire, pour sa disponibilité et ses conseils judicieux tout au long de ce travail.

Je tiens à exprimer tout mes respects à mes parents, mes frères , ma soeur et mes frères qui m'ont toujours encouragé.

Je remercie tous les professeurs du département de Mathématiques, sans oublier aussi mes collègues et amies, ainsi tous ceux qui ont participé de loin ou de près à l'élaboration de ce mémoire.

Dédicaces

Je dédie ce modeste travail :

A mes parents ma mère et mon père.

-A ma soeur

-A mes frères.

-A toute la famille.

-A toute mes amies.

Je tiens à remercier l'ensemble de tous les étudiants et étudiantes de ma promotion,

En fin je dédie ce mémoire à mes collègues et tous ceux qui me sont chers.

Table des figures

Figure (1.1.1): Pixels et niveaux de gris

Figure (1.4.1): Exemple de dégradations chimiques

Figure (1.4.2):Exemples de dégradations biologiques

Figure (1.4.3):Exemple de dégradatiobs d'origine humaine

Figure(1.5.2):Dégradations du fond sur une image

Figure(1.5.2):Exemple de dégradations des formes

Figure (2.1 .a):Suppression des taches en transparence par recalage[Sha-00]

Figure (2.1 .b):Suppression des taches en transparence par recalage[DUB- 01]

Figure (2.2):(a)Image d'origine-(b)Image de recto séparée du verso[WOL- 06]

Figure(2.3 .a) Confuse : (e) «e» ou «c», (f) «e» ou «o»

Figure(2.4) Représentation directionnelle des vecteurs $\vec{\eta}$, $\vec{\xi}$

Table des matières

Introduction	1
1 Dégradation des images de documents	2
1.1 Les images numériques	2
1.1.1 Les images binaires (noir ou blanc)	2
1.1.2 Les images en niveaux de gris	3
1.1.3 Les images couleurs	3
1.2 Les documents anciens	4
1.3 Qu'est ce qu'une dégradation?	4
1.4 Dégradations physiques du document	4
1.4.1 Dégradations chimiques	4
1.4.2 Dégradations biologiques	7
1.4.3 Dégradations physiques d'origine humaine	8
1.5 Proposition d'une classification non-causale des dégradations	9
1.5.1 Dégradations uniformes	10
1.5.2 Dégradations non-uniformes	10
2 Restauration des images de documents	13
2.1 Restauration par recalage	13
2.2 Restauration par analyse des formes	15
2.3 Restauration par prototypage	15
2.4 Restauration des formes par diffusion	16
2.4.1 Flou:Rehaussement	23

3 Application sur documents anciens	27
3.1 Débruitage linéaire(restauration)	30
3.2 Débruitage non linéaire(restauration)	32
Conclusion générale	35
Annexe	36
Bibliographie	40

Introduction

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) aux supports souples (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 XVème siècle a marqué le début d'une nouvelle ère de production et de diffusion massive du savoir. La numérisation est la technologie qui permet de convertir des documents physiques en images électroniques. La numérisation des documents anciens est un enjeu important pour les services d'archives, les bibliothèques, les historiens,...

Dans ce memoire nous allons présenter le principe de restauration d'image appliquée aux documents anciens avec quelques applications. Le traitement d'image est une étape importante de la numérisation des documents anciens. Les traitements présentés ici permettent de restaurer ou nettoyer l'image, d'extraire les différentes structures du document (illustrations, éléments graphiques et textuels,...). Les méthodes de restauration des images de documents anciens proposés dans la littérature ont commencées récemment. Leydier et al [LEY- 04] ont réalisé un système de segmentation adaptatif basée sur une sérialisation de l'algorithme des k-moyennes appliquée séquentiellement. [WAN-03] ; ont développé une méthode de séparation du texte à partir d'un fond bruité avec effet de transparence.

Ce mémoire est construit en trois chapitre:

Le premier chapitre: On 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 deuxième chapitre: On traite des méthodes existantes de restauration des images dégradées de documents.

Le troisième chapitre: On présente quelques applications des techniques qu'on a développé dans le chapitre deux.

On termine notre mémoire par une conclusion générale.

Chapitre 1

Dégradation des images de documents

1.1 Les images numériques

Une image est un signal bidimensionnel. Une image analogique est par exemple celle formée sur la rétine de l'œil ou l'images obtenue par la photographie argentique classique.

Une image **numérique** est un signal numérique composé d'unités élémentaires (appelées pixels) qui représentent chacun une portion de l'image.

Une image numérique est définie par :

le nombre de pixels qui la composent en largeur et en hauteur (qui peut varier presque à l'infini).

l'étendue des teintes de gris ou des couleurs que peut prendre chaque pixel (on parle de dynamique de l'image).

I est l'image numérique : $I [i, j] = n$ est la valeur du niveau de gris. Lorsque $n \in [N_{\min}, N_{\max}]$, N_{\max} , N_{\min} est le nombre de niveaux de gris .

1.1.1 Les images binaires (noir ou blanc)

Images les plus simples, un pixel peut prendre uniquement les valeurs noir ou blanc. C'est typiquement le type d'image que l'on utilise pour scanner du texte quand celui ci est composé d'une seule couleur.

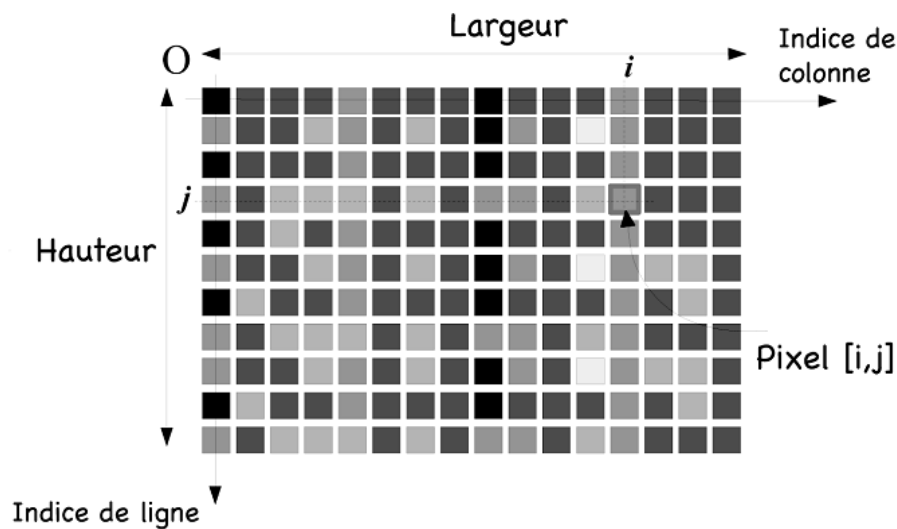


Fig.1.1.1: Pixels et niveaux de gris

1.1.2 Les images en niveaux de gris

En général, les images en niveaux de gris renferment 256 teintes de gris. Par convention la valeur zéro représente le noir (intensité lumineuse nulle) et la valeur 255 le blanc (intensité lumineuse maximale). Le nombre 256 est lié à la quantification de l'image. En effet chaque entier représentant un niveau de gris est codé sur 8 bits. Il est donc compris entre 0 et $2^8 - 1 = 255$. C'est la quantification la plus courante. On peut coder une image en niveaux de gris sur 16 bits ($0 \leq n \leq 2^{16} - 1$) ou sur 2 bits : dans ce dernier cas le «niveau de gris» vaut 0 ou 1 : il s'agit alors d'une image binaire (Noir et Blanc).

1.1.3 Les images couleurs

L'espace couleur est basé sur la synthèse additive des couleurs, c'est à dire que le mélange de trois composantes (par exemple (R, V, B)) donne une couleur.

On garde l'information couleur, ou intensité lumineuse et chromaticité. Un pixel est codé par trois valeurs numériques. La signification de ces valeurs dépend du type de codage choisi. Le plus utilisé pour le maniement des images numériques est l'espace couleur Rouge,

Vert,Bleu (R,V,B), (RGB en anglais). La restitution des couleurs sur écran utilise cette représentation (synthèse additive).

1.2 Les documents anciens

Les documents anciens forment une bonne partie de notre patrimoine culturel et scientifique. Ces manuscrits encourent une dégradation progressive et sont menacés d'un réel danger de disparition. Le traitement automatique de ces documents en vue de leur restauration offre un avantage certain, mais il est confronté à de nombreuses difficultés dues au mauvais état de conservation de ces manuscrits et à la complexité de leur contenu. En effet, les documents anciens possèdent de nombreuses particularités qui invalident les techniques classiques de filtrage. Les détériorations les plus communes que l'on trouve dans les images de documents anciens sont : l'effet de transparence, la présence de taches d'humidité absorbé par le papier, les variations de la couleur du papier, la présence de plis et de déchirures, et les déformations dues à la courbure naturelle des pages.

1.3 Qu'est ce qu'une dégradation?

Les dégradations ont plusieurs origines et l'accumulation des défauts rend difficile leurs séparations. Les dégradations ont des effets plus destructeurs sur destructeurs sur les images de documents que sur toutes autres images dites "naturelles".

1.4 Dégradations physiques du document

1.4.1 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 1.4.1).

Les processus chimiques ont des chances de s'intensifier avec l'accroissement de la température, l'humidité, la luminosité et la pollution atmosphérique.



Fig 1.4.1 :Exemples des dégradations chimiques

La température et l'humidité

Influence de la température : Les matériaux organiques qui constituent le papier possèdent des caractéristiques différentes de dilatation en fonction de la température.

Des températures trop basses peuvent :

Fragiliser les matériaux (gèle) et entraîner des fissurations et des craquellements.

Des températures trop élevées peuvent :

Accélérer la dégradation des papiers acides par catalyse des phénomènes d'hydrolyse acide, d'oxydation et les effets photochimiques. Le papier devient cassant et perd sa résistance mécanique.

Influence de l'humidité relative: Une humidité relative trop élevée entraîne :

Une prolifération de moisissures. En effet, le taux d'humidité relative et la température jouent un rôle prépondérant dans le cycle vital des champignons et des bactéries présents dans les documents d'archive.

Une humidité relative trop basse entraîne :

Une déshydratation des structures et rend ces matériaux fragiles ces matériaux ont besoin d'un taux d'humidité avoisinant les 50% pour garder souplesse et élasticité. Les colles pour la reliure se dessèchent et deviennent cassantes.

La lumière La lumière fait partie des éléments qui contribuent à la destruction du papier.

Une source lumineuse qu'elle soit naturelle ou artificielle, émet des radiations visibles et des radiations invisibles, infrarouges et ultraviolettes.

Source naturelle : Le soleil est la principale source de lumière naturelle.

Sources artificielles : Actuellement, il existe de nombreux types de sources artificielles que l'on peut néanmoins distinguer par deux caractéristiques : les lampes à incandescence (qui ont des effets thermiques) et les tubes fluorescents (qui ont des effets photochimiques).

La pollution atmosphérique La pollution atmosphérique est un autre facteur important de dégradation.

Elle existe sous :

Forme gazeuse : dioxyde de soufre ou ozone provenant des voitures et de l'industrie, formaldéhyde se dégageant de certains matériaux (bois, textiles, papiers) utilisés lors pour le stockage.

Forme solide : (suie, particules poussiéreuses).

Forme liquide : (graisse ou de sueur) qui déposée à la surface d'un document lors de sa manipulation.

1.4.2 Dégradations biologiques

Les dégradations biologiques des documents (Fig 1.4.2) sont causées par des organismes «les parasites ». Les dommages d'origine biologique diffèrent des précédents quant à la leurs effets. Ils peuvent être causés par différents agents:

Les micro-organismes Micro-organismes (bactéries et champignons que l'on trouve le plus fréquemment dans les archives) provoquent des taches dont la couleur varie avec l'organisme pathogène. Ces taches envahissent divers types de matériaux et provoquent des altérations irréversible.

Au cours de sa croissance, le mycélium s'insinue entre les fibres et attaque le papier à cause d' un véritable processus de digestion. Cette action peut-être lente : il faut parfois plusieurs mois pour détecter les dégâts.Certains champignons ne perforent pas le papier mais sécrètent un pigment qui colore les fibres. Ces piqûres créent dans le papier des taches de couleurs.

Les insectes Les insectes (les insectes provoquent également une destruction moins étendue et rapide de nombreux documents) 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 Les rongeurs (phénomène rare) nuisent en particulier à la partie extérieure des ouvrages. Ils grignotent les couvertures, les reliures et les bords des livres.Ils causent des dommages importants, Ils attaquent le papier mécaniquement sans en ingérer une quantité appréciable .

Les dommages des rongeurs sont facilement détectables grâce aux marques de leurs dents et à la présence de leurs excréments sur les papiers.



Fig 1.4.2:Exemples de dégradations biologiques

1.4.3 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 1.4.3). 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.



Fig 1.4.3: Exemples de dégradations d'origine humaine

Dégradation numérique

Les dégradations numériques sont les conséquences des choix technologiques délibérés comme le paramétrage du matériel. 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.

1.5 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.

1.5.1 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).

1.5.2 Dégradations non-uniformes

Par dégradations non-uniformes on entend toutes altérations qui varient suivant la position dans l'image. 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 des fond

Les dégradations du fond (Fig 1.5.2.a) 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. 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.



Fig 1.5.2.a: Dégradations du fond sur une image

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 (Fig 1.5.2.b) 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.

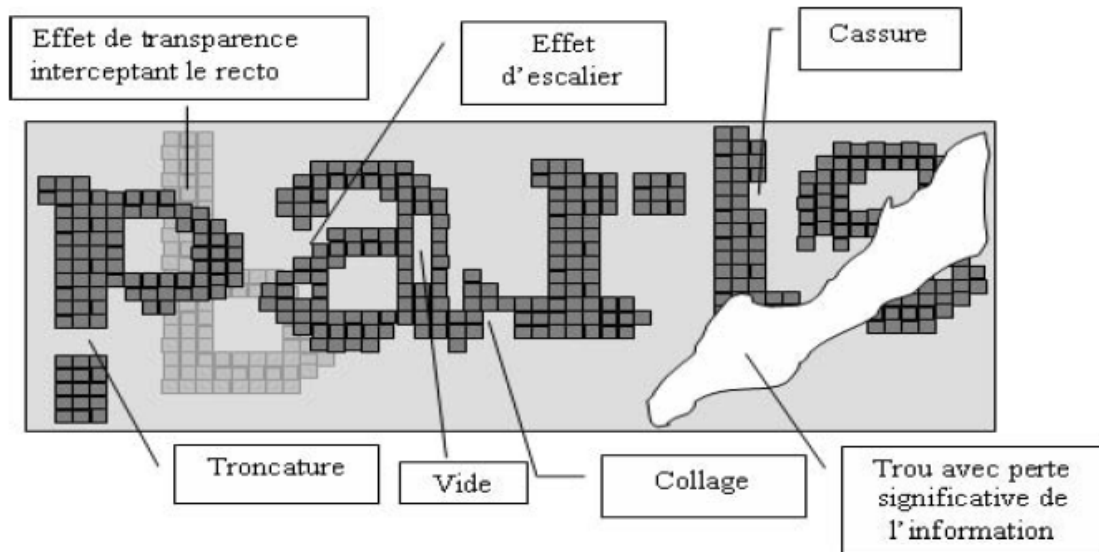


Fig 1.5.2.b:Exemple de dégradations des formes

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 de dégradations et de leurs causes, nous avons proposé une classification qui regroupe les dégradations qui peuvent être traitées avec les mêmes méthodes.

Chapitre 2

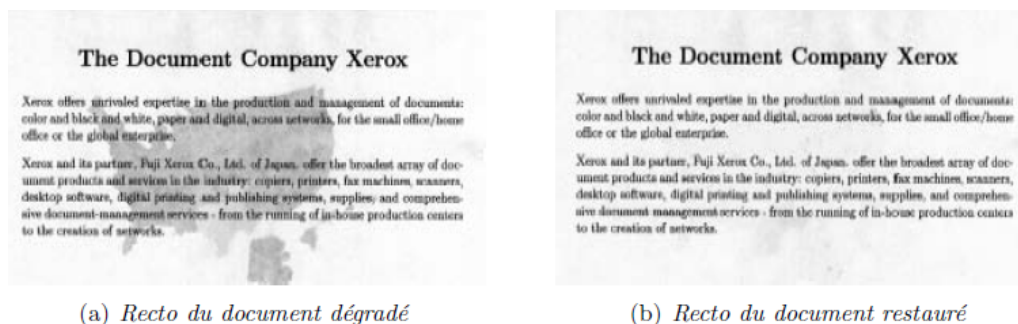
Restauration des images de documents

Les systèmes qui traitent des documents dégradés disposent souvent d'un module de restauration qui permet de reproduire une représentation la plus proche possible, de la qualité de l'image originale avant sa dégradation. Les défauts attachés aux documents sont de différents types, et peuvent être classés en deux groupes : dégradations du fond, et dégradation de formes. Nous présentons dans cette section quelques méthodes de restauration.

2.1 Restauration par recalage

Le principe des méthodes utilisant le recalage pour la séparation entre le recto/ verso est de 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 (2.1.a) le seul résultat publié par l'auteur ; il faut remarquer que les deux parties à séparer sont de natures différentes (texte et image).

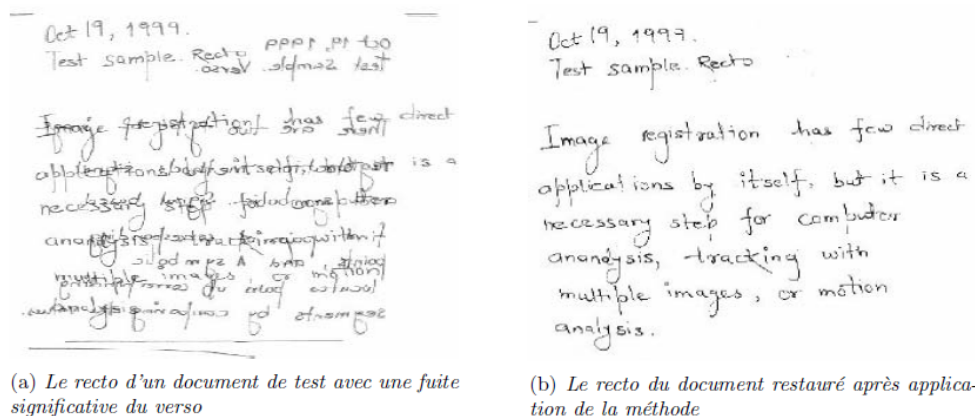


(a) Recto du document dégradé

(b) Recto du document restauré

Fig 2.1.a: 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 (2.1.b) montre un exemple d'application de cette méthode.



(a) Le recto d'un document de test avec une fuite significative du verso

(b) Le recto du document restauré après application de la méthode

Fig 2.1.b: Suppression des taches en transparence par recalage [Dub-01]

2.2 Restauration par analyse des formes

La méthode de Wolf [WOL- 06], comme représentant de ce type de méthodes est basée sur une régularisation de la segmentation par les champs de Markov pour la séparation entre le recto et le verso. 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. Ils tendent à régulariser la segmentation pour préserver la continuité des traits du recto qui croisent ceux du verso (Fig 2.2).

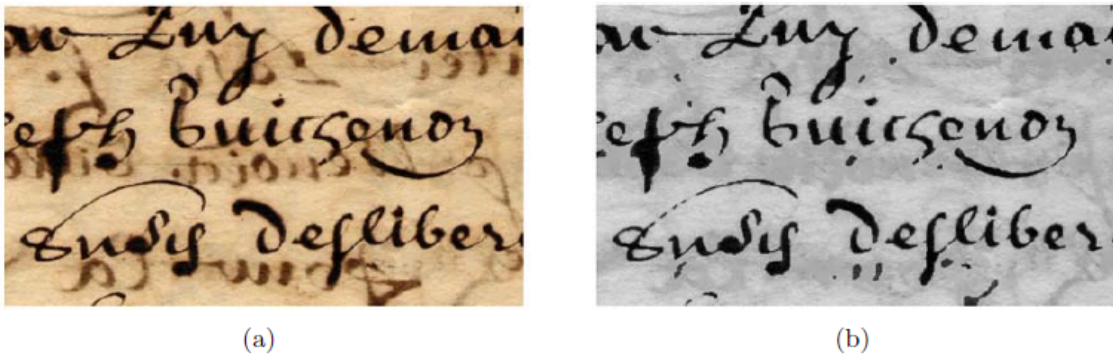


Fig 2.2: (a) Image d'origine, (b) Image finale de recto séparée du verso [Wol-06]

2.3 Restauration par prototypage

Cette approche est appliquée généralement sur les caractères d'un document imprimé [DRI-07]. Elle consiste à collectionner toutes les formes similaires d'une page, ensuite on construit le modèle de la forme idéale d'une lettre à partir de la superposition de toutes les formes dégradées de la même lettre. La figure (2.3.a) montre différentes occurrences dégradées du caractère «e».

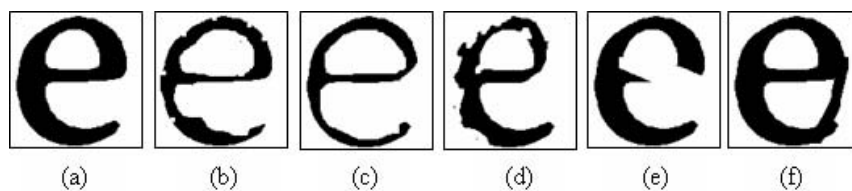


Fig 2.3.a: Cas de dégradations de la lettre «e» : (a) caractère idéal ; (b),(c) et (d) caractères dégradés avec structure du «e» conservée, (e) et (f) caractères dégradés avec structure confuse : (e) «e» ou «c», (f) «e» ou «o»

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 moyennner 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 2.3.b*)

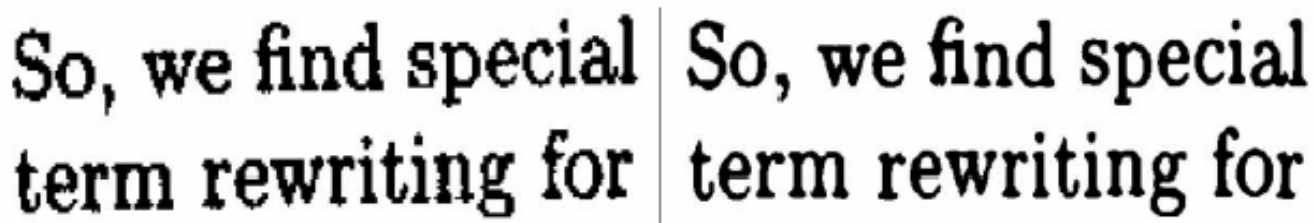


Fig 2.3.b:Restauration par prototypage des formes dégradées de lettres

2.4 Restauration des formes par diffusion

Pour montrer la différence entre la restauration et le débruitage prenons comme exemple le cas d'une image I_0 affectée par un bruit b additif en plus d'un ou optique H .

Nous pouvons distinguer le débruitage d'images qui consiste à retrouver l'image I_0 sans bruit à partir du modèle :

$$I = I_0 + b \quad (2.4.1)$$

et la restauration d'images qui consiste à récupérer I_0 à partir du modèle de formation du ou tout en prenant en compte le bruit :

$$I = H * I_0 + b \quad (2.4.2)$$

La restauration permet donc de retrouver l'image source en essayant de prendre en compte les différents phénomènes qui entrent en jeu dans le processus de formation d'images et qui peuvent altérer la qualité de celle-ci

Dans certains cas, le problème peut être mal posé : le problème n'admet pas de solution, en admet plusieurs ou il n'existe pas de solution stable. Dans ce cas, il devient nécessaire d'introduire des contraintes (a priori) permettant de régulariser la solution.

Dans ce qui suit, nous allons montrer quelques méthodes existantes permettant d'améliorer la qualité d'images bruitées, puis nous nous intéresserons aux méthodes de restauration d'images dégradées par les conditions météorologiques.

Le débruitage d'image est une tâche qui a suscité beaucoup d'intérêt dans le domaine du traitement d'image.

Le bruit peut apparaître sous différentes formes : aléatoire, centré, additif, multiplicatif ou convolutif.

Filtres classiques

En général, les filtres d'images peuvent être soit linéaires (moyenne, convolution) ou non linéaires (médiane). Lorsque le bruit est un signal à hautes fréquences, un filtrage passe bas permet de le réduire mais ajoute du ou et atténue les contours. En particulier, ils atténuent le bruit dans les zones uniformes, en revanche, au niveau des zones texturées ou proche des contours, les effets sont visibles. Plus la taille du filtre est grande plus la dégradation de l'image est importante.

Toutefois, pour des grandes quantités de bruit cet algorithme atteint ses limites. Dans la suite, nous allons montrer des méthodes de débruitage qui sont plus robustes.

Diffusion et filtrage

Ce type de filtrage est inspiré de la physique (propagation des ondes et equation de la chaleur). Il s'agit de propager l'information contenue dans les pixels vers les zones voisines suivant l'équation de diffusion de la chaleur (équations aux dérivées partielles : EDP). La diffusion, en physique, a pour but d'équilibrer les différences de concentrations (niveaux de gris, dans le cas d'images) sans aucun changement de la masse.

La loi de Fick décrit la propriété d'équilibre :

$$j = -D\nabla I \quad (2.4.3)$$

avec j le flux, ∇I le gradient de I (l'image bruitée dans notre cas) et D le tenseur de diffusion.

Lorsque $D = g$ est un scalaire, appelé la diffusivité, il s'agit de diffusion isotropique. Dans ce cas, j et ∇I sont parallèles. Sinon, il s'agit de diffusion anisotropique.

L'équation de continuité décrivant la conservation de la masse est la suivante :

$$I_t = -\operatorname{div} j \quad (2.4.4)$$

avec t le temps.

à partir des équations (2.4.3) et (2.4.4), nous déduisons l'équation de diffusion :

$$I_t = \operatorname{div}(D\nabla I) \quad (2.4.5)$$

Nous pouvons réécrire cette équation sous la forme :

$$I_t = D\Delta(I) \quad (2.4.6)$$

Nous distinguons fondamentalement deux types de diffusion :

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.

Diffusion isotrope il s'agit d'une diffusion isotropique globale telle que la diffusivité est constante (2.4.7). Ce type de diffusion est optimal dans les zones uniformes mais oute les contours et les textures et les rend difficilement identifiable.

$$\begin{cases} I_t = \Delta I = \operatorname{div}(\nabla I) \\ I_0 = I_{\text{bruitée}} \end{cases} \quad (2.4.7)$$

Koenderink [Koe-84] est le premier à remarquer que l'évolution de cette équation est équivalente à un lissage gaussien de variance $\sigma = \sqrt{2t}$ (2.4.8).

$$I = I_0 * G_{\sigma=\sqrt{2t}} \quad (2.4.8)$$

avec:

$I(x; y; t)$: l'image restaurée (intensité lumineuse).

$I_0(x; y; t)$: l'image bruitée.

$G(x; y; t)$: un opérateur de lissage.

t : paramètre de contrôle de l'importance du lissage.

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.

Diffusion anisotrope 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 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 (2.4.9).

$$\begin{cases} I_t = \operatorname{div}(c \|\nabla I\|) \times \nabla I \\ I_0 = I_{\text{bruitée}}. \end{cases} \quad (2.4.9)$$

Avec :

$c(\cdot)$: une fonction de diffusion positive décroissante.

div : l'opérateur de divergence.

∇ : l'opérateur gradient.

En développant la divergence de (2.4.9) celle ci devient :

$$\begin{cases} I_t = c(\|\nabla I\|)\Delta I + \nabla(c(\|\nabla I\|))\nabla I \\ I_0 = I_{bruitée} \end{cases} \quad (2.4.10)$$

Si C est égale à «1», on retrouve une diffusion isotrope par l'équation de la chaleur.

Il s'agit cette fois d'un lissage conditionnel, dont le comportement est fonction de la norme du gradient de l'image, grâce à la fonction $c(\cdot)$. Le principe est de diffuser fortement dans les zones à faibles gradients (zones homogènes), et faiblement dans les zones à forts gradients (contours). Un seuil k permet de distinguer les zones à faible gradient de celles à fort gradient.

Perona - Malik proposent les fonctions suivantes :

$$\begin{aligned} c(\|\nabla I\|) &= \frac{1}{1+(\frac{\|\nabla I\|}{k})^2} \\ c(\|\nabla I\|) &= \exp(-(\frac{\|\nabla I\|}{k})^2) \end{aligned} \quad \text{ou}$$

Le choix de seuil k n'est d'ailleurs pas évident ; il peut être fixé arbitrairement ou en estimant le bruit. Pour le déterminer, Perona-Malik proposent de se baser sur un histogramme cumulé du gradient. 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 (2.4.9) est identique à l'équation de la chaleur (2.4.7) qui réalise une forte diffusion dans les zones éloignées des contours. Inversement, les parties de l'image où $\|\nabla I\| \geq k$ sont considérés comme étant des variations d'intensité 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.

Afin de mieux comprendre le comportement diffusif de ce modèle, nous allons le réécrire sous la forme des dérivées secondes dans la directionnelles dans la direction du gradient $\vec{\eta} = \left(\frac{I_x}{\|\nabla I\|}, \frac{I_y}{\|\nabla I\|}\right)^T$ et la direction du orthogonale $\vec{\xi} = \left(-\frac{I_y}{\|\nabla I\|}, \frac{I_x}{\|\nabla I\|}\right)^T$

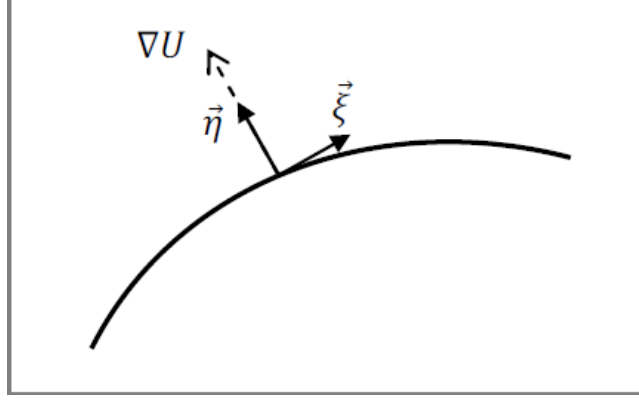


Fig 2.4.1: Représentation directionnelle des vecteurs $\vec{\eta}$, $\vec{\xi}$

Pour formuler l'équation de diffusion selon les directions par $\vec{\eta}$ et $\vec{\xi}$, posons $\nabla I = (I_x, I_y)$, le vecteur gradient de I et $\|\nabla I\| = \sqrt{I_x^2 + I_y^2}$ sa norme.

Dans le repère orthonormé (\vec{i}, \vec{j}) , le développement de (2.4.10) conduit à :

$$\begin{aligned} I_t &= c(\|\nabla I\|)\Delta I + \nabla(c(\|\nabla I\|))\nabla I \\ &= c(\|\nabla I\|)(I_{xx} + I_{yy}) + \frac{\partial}{\partial x}c(\|\nabla I\|)I_x + \frac{\partial}{\partial y}c(\|\nabla I\|)I_y \end{aligned} \quad (2.4.11)$$

Le deuxième terme s'écrit :

$$\begin{aligned} \frac{\partial}{\partial x}c(\|\nabla I\|)I_x + \frac{\partial}{\partial y}c(\|\nabla I\|)I_y &= c'(\|\nabla I\|)\frac{\partial}{\partial x}(\|\nabla I\|)I_x + c'(\|\nabla I\|)\frac{\partial}{\partial y}(\|\nabla I\|)I_y \\ &= \frac{c'(\|\nabla I\|)}{\|\nabla I\|}(I_{xx}I_x + I_{yx}I_y)I_x + \frac{c'(\|\nabla I\|)}{\|\nabla I\|}(I_{xy}I_x + I_{yy}I_y)I_y \\ &= \frac{c'(\|\nabla I\|)}{\|\nabla I\|}(I_{xx}I_x^2 + 2I_{xy}I_x + I_{yy}I_y^2) \end{aligned} \quad (2.4.12)$$

Dans le repère $(\vec{\xi}, \vec{\eta})$ la dérivée première directionnelle selon $\vec{\eta}$ s'exprime par :

$$I_\eta = \nabla I \cdot \vec{\eta} = \|\nabla I\| \quad (2.4.13)$$

la dérivée seconde s'écrit:

$$\begin{aligned}
 I_{\eta\eta} &= \nabla I_\eta \cdot \vec{\eta} \\
 &= \frac{1}{\|\nabla I\|} \left[\frac{\partial}{\partial x} (\|\nabla I\|) I_x + \frac{\partial}{\partial y} (\|\nabla I\|) I_y \right] \\
 &= \frac{(I_{xx} I_x + I_{yx} I_x) I_x + (I_{xy} I_x + I_{yy} I_y) I_y}{\|\nabla I\|^2} \\
 &= \frac{I_{xx} I_x^2 + 2I_{xy} I_x I_y + I_{yy} I_y^2}{\|\nabla I\|^2}
 \end{aligned} \tag{2.4.14}$$

le laplacien étant invariant au changement de base, le premier terme (2.4.11) peut s'écrire:

$$\begin{aligned}
 c(\|\nabla I\|) \Delta I &= c(\|\nabla I\|) (I_{xx} + I_{yy}) \\
 &= c(\|\nabla I\|) (I_{\xi\xi} + I_{\eta\eta})
 \end{aligned} \tag{2.4.15}$$

En combinant (2.4.12)(2.4.14)(2.4.15) nous obtenons:

$$\begin{aligned}
 I_t &= \operatorname{div}(c \|\nabla I\| \nabla I) \\
 &= c(\|\nabla I\|) I_{\xi\xi} + c(\|\nabla I\|) I_{\eta\eta} + c'(\|\nabla I\|) \|\nabla I\| I_{\eta\eta}
 \end{aligned} \tag{2.4.16}$$

réécrivons(2.4.10) sous une autre forme dans le repère orthonomé

$$\begin{cases} I_t = c_\xi I_{\xi\xi} + c_\eta I_{\eta\eta} \\ c_\xi = c(\|\nabla I\|) \\ c_\eta = (c(\|\nabla I\|) \times \|\nabla I\|)' \\ c_\eta = c(\|\nabla I\|) + c'(\|\nabla I\|) \|\nabla I\| \end{cases} \tag{2.4.17}$$

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.

Cette démarche introduite par Perona et Malik a fait l'objet de plusieurs études critiques [Wei-97, 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 (2.4.18).

$$\begin{cases} I_t = \operatorname{div}(C(\|G_\sigma \otimes \nabla I\|) \times \nabla I) \\ I_0 = I_{bruitée} \end{cases} \quad (2.4.18)$$

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 $G_\sigma \otimes \nabla I$ 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.

2.4.1 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.

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

Dans l'équation (2.4.19), 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.

$$\begin{cases} I_t = +F(I, \nabla I, \dots) \\ I_0 = I_{bruitée} \end{cases}, \quad \begin{cases} I_t = -F(I, \nabla I, \dots) \\ I_0 = I_{flou} \end{cases} \quad (2.4.19)$$

Par exemple, l'équation inverse de la chaleur renforce toutes les variations, celles issues des contours comme celles dues aux bruits de l'image (2.4.20).

$$\begin{cases} I_t = -\Delta I = -\operatorname{div}(\nabla I) \\ I_0 = I_{flou} \end{cases} \quad (2.4.20)$$

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».

Renforcement des contours : Filtre de choc

Les filtres de choc sont des équations basées sur les EDP morphologiques que nous allons brièvement introduire ; ces filtres 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 :

$$I_t = c \|\nabla I\| \quad (2.4.21)$$

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

$$\begin{aligned} \text{Dilatation : } I_t &= \|\nabla I\| \\ \text{Erosion : } I_t &= -\|\nabla I\| \end{aligned} \quad (2.4.22)$$

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 .

$$\begin{aligned} \text{Dilatation : } I^{n+1}(x, y) &= I^n(x, y) + dt \|\nabla I^n(x, y)\| \\ \text{Erosion : } I^{n+1}(x, y) &= I^n(x, y) - dt \|\nabla I^n(x, y)\| \end{aligned} \quad (2.4.23)$$

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.

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 (2.4.1). Aucun extremum local erroné n'est créé. Ce schéma préserve donc la totale variation.

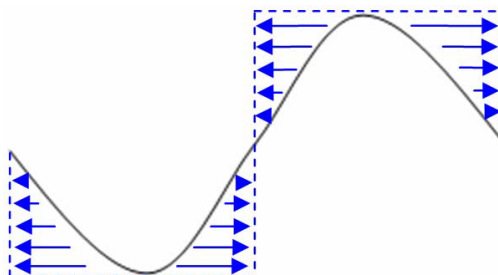


Fig 2.4.2.a: Principe de rehaussement par filtre de choc dans le cas 1-D

L'équation proposée par Rudin (2.4.24.a) 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.

$$\text{FiltredechocdeRudin} : \begin{cases} I_t = -\text{sign}(\Delta I) \|\nabla I\| \\ I_0 = I_{\text{flou}} \end{cases} \quad (2.4.24)$$

La figure (2.4.1.b) 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 de manque de netteté des contours.

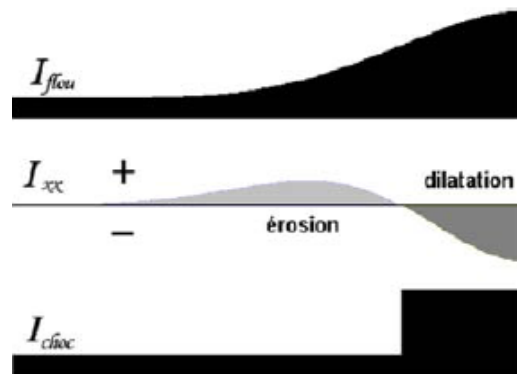


Fig 2.4.2.b: Principe du filtre de choc de Rudin dans le cas 1-D

Chapitre 3

Application sur documents anciens

Dans ce chapitre nous allons présente quelques applications des techniques qu'on a développé dans le chapitre deux.

Nous avons pris des documents dégradés d'un ancien livre; et nous avons essayer de les traiter par différentes technique de restauration.

La procedure de restauration est comme suit: nous commençons tout d'abord par introduire le document sous forme d'une image JPG

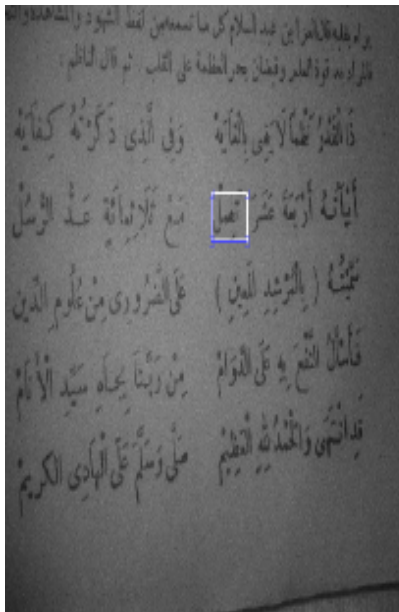
يراه يقليه قال العزرا بن عبد السلام كل ما تسمعه من لفظ الشهود والمجاهدة والت
فلراد به قوة العلم وفضلان بحر العظمة على القلب . ثم قال الناظم :

ذَا الْقَدْرُ نَظْمًا لَا يَفِي بِالْفَائِدَةِ وَفِي الَّذِي ذَكَرْتُهُ كِفَايَةٌ
أَيَاتُهُ أَرْبَعَةٌ عَشَرَ تَصِلُ مَعَ ثَلَاثِينَ عَدُّ الرُّسُلِ
سَمِيَّتُهُ (بِالْمُرْتَدِّ الْمَعِينِ) عَلَى الضَّرُورِيِّ مِنْ عُلُومِ الدِّينِ
فَأَسْأَلُ النَّفْعَ بِهِ عَلَى الدَّوَامِ مِنْ رَبِّنَا بِجَاهِ سَيِّدِ الْأَنَامِ
قَدِ انْتَهَى وَالْحَمْدُ لِلَّهِ الْعَظِيمِ صَلَّى وَسَلَّمَ عَلَى الْهَادِي الْكَرِيمِ

لَمْ تَسْمَعْ إِذَا نَادَى
كَسْمَاعِي لَا مَقْدَاحَ
وَيَسْبِيحُ الْمُتَّقِينَ وَالرُّسُلَ

pour le traiter on doit le convertir en niveau de gris sous Matlab avec le programme suivant:

```
i1=imread('N.jpg');
i2=imread('W.jpg');
%pour convertie une image de colours vers une image de niveau de gris
j1=rgb2gray(i1);
figure(1);
imshow(j1);impixelregion;
title('Image en nivea de gris')
mettre le document couleur avec le niveau de gris
```



129	114	120	125	122
130	133	140	124	131
125	134	137	136	129
132	126	118	127	118

3.1 Débruitage linéaire(restauration)

On applique le filtre moyenne (linéaire) sur l'image bruitée pour réduction de bruit

Le code correspondant à ce filtre est le suivant :

```
% Filtre moyenne de taille 7 x 7
```

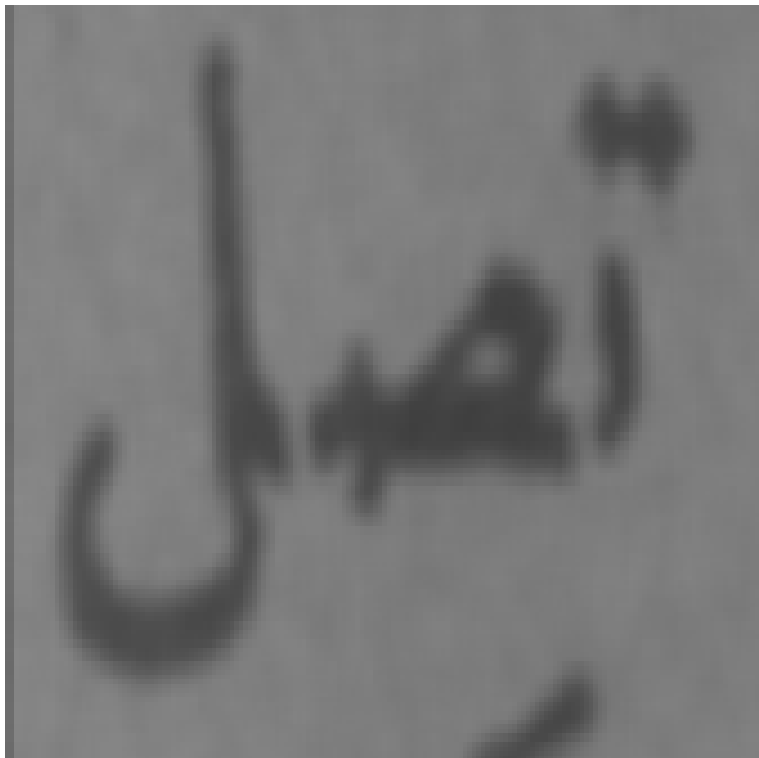
```
h = fspecial('average',[7 7]);
```

```
k=imfilter(j1,h);
```

```
figure(3),
```

```
imshow(k);
```

```
title('image restaurée')
```



filtre moyenne

On applique le filtre gaussien (linéaire) sur l'image bruitée pour réduction de bruit

```
% Filtre gaussien de taille 15 x 15
```

```
% et d'ecart type 1
```

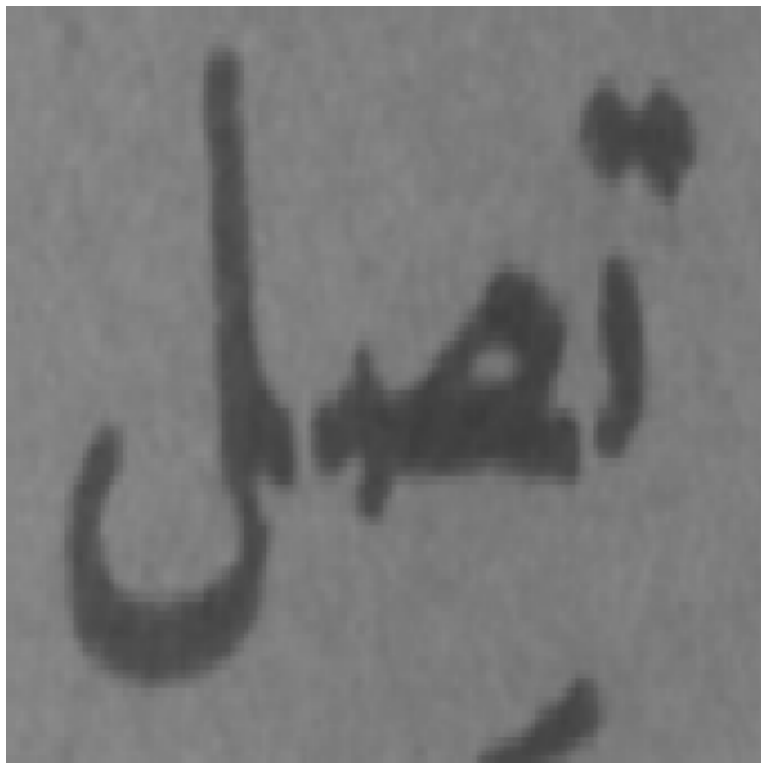
```
h = fspecial('gaussian',[15 15],1);
```

```
n=imfilter(j1,h);
```

```
figure(4);
```

```
imshow(n);
```

```
title('image restaurée')
```



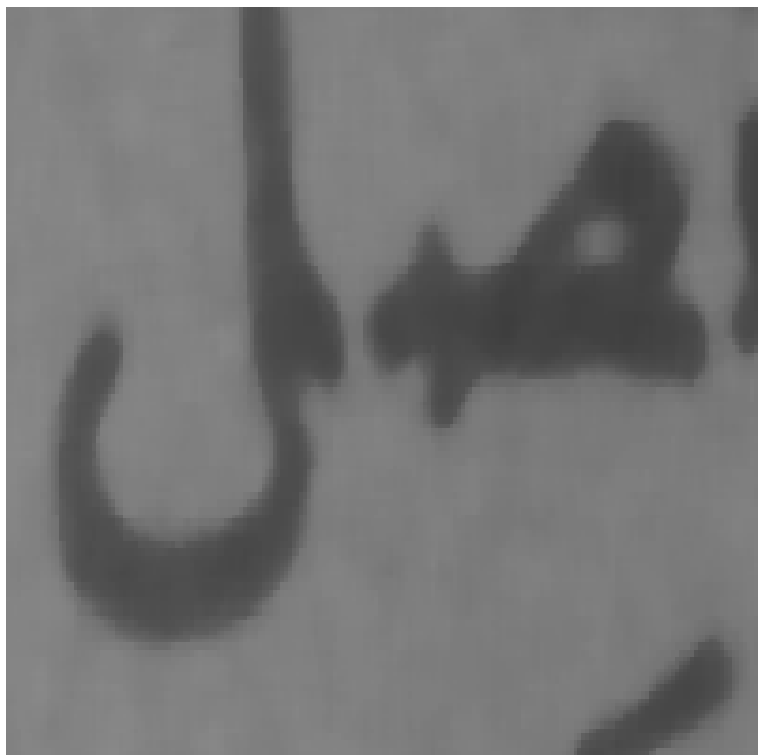
filtre gaussien

3.2 Débruitage non linéaire(restauration)

On applique le filtre médiane (non linéaire) sur l'image bruitée pour réduction de bruit

Le code correspondant à ce filtre est le suivant :

```
% Filtre médiane de taille 7 x 7  
w2=medfilt2(j1,[7 7]);  
figure(5),imshow(w2);title('image restaurée')
```

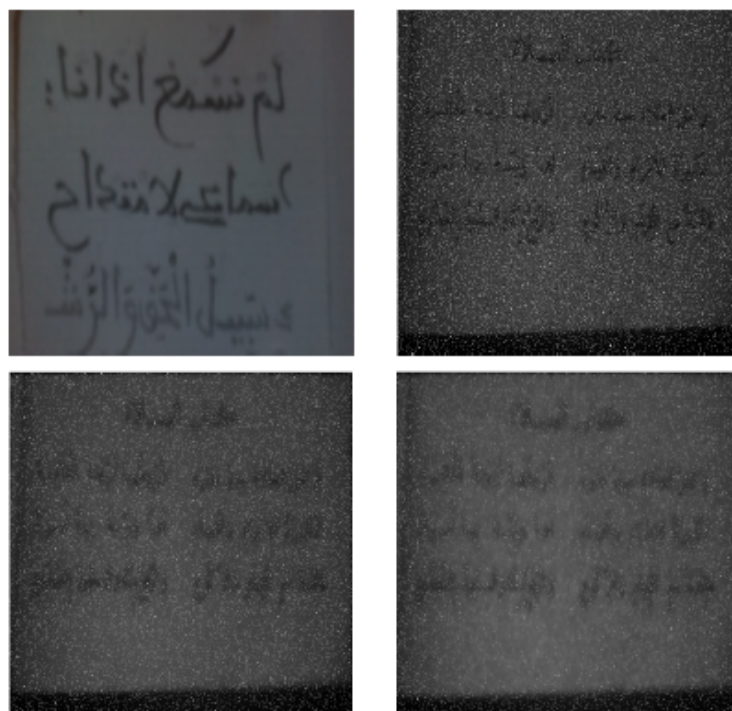


filtre médiane

remarque

On considère que le résultat par le filtre médiane est le plus performant.

Nous allons appliquer le filtre (perona et malik) l'image bruitée pour réduction de bruit



filtre(perona-malik)

Remarque

perona et malik ont des lacunes telles que son inefficacité dans les zones où le bruit présente de grosses discontinuités

Avantages et inconvénients du modèle de diffusion au traitement des images de documents

- Avantages(perona et malik)

Réduit les bruits additifs ou ponctuels

Préserve les contours et structures.

Modèle générique adapté à tous les types de documents.

- Inconvénients(perona et malik)

Ne lisse pas les contours.

Ne renforce pas les traits.

Paramètre k très critique.

Ne renforce pas les traits.

Conclusion générale

Cette étude a présenté une méthode de restauration des images de documents anciens. Les images de documents anciens possèdent de nombreuses dégradations. Ces dégradations freinent actuellement le développement des projets de numérisation et limitent considérablement l'accès de ces images au grand public. Notre travail est une contribution à la restauration des images de documents anciens et présente les différentes étapes nécessaires au développement de notre algorithme. Ce travail porte sur des techniques à base d'équations aux dérivés partielles pour le traitement d'images. Notre méthode, est fondée sur une équation de diffusion non linéaire pour la restauration des images de documents anciens. Et nous a permis de moitiriser les techniques de restaurant d'image appliquée aux documents anciens, ce ci nous ouvent de nouvelles perspectives petits de troits des problèmes plus complexes.

Annexe

Code matlab (Perona et Malik)

```

function diff_im = anisodiff2D(im, num_iter, delta_t, kappa, option)
%ANISODIFF2D Conventional anisotropic diffusion
% DIFF_IM = ANISODIFF2D(IM, NUM_ITER, DELTA_T, KAPPA, OPTION) per-
foms
% conventional anisotropic diffusion (Perona & Malik) upon a gray scale
% image. A 2D network structure of 8 neighboring nodes is considered for
% diffusion conduction.
% IM - gray scale image (MxN).
% NUM_ITER - number of iterations.
% DELTA_T - integration constant (0 <= delta_t <= 1/7).
% Usually, due to numerical stability this
% parameter is set to its maximum value.
% KAPPA - gradient modulus threshold that controls the conduction.
% OPTION - conduction coefficient functions proposed by Perona & Malik:
% 1 -  $c(x,y,t) = \exp(-(\text{nabla}I/\text{kappa})^2)$ ,
% privileges high-contrast edges over low-contrast ones.
% 2 -  $c(x,y,t) = 1./(1 + (\text{nabla}I/\text{kappa})^2)$ ,
% privileges wide regions over smaller ones.
% OUTPUT DESCRIPTION:
% DIFF_IM - (diffused) image with the largest scale-space parameter.
% Convert input image to double.
im = double(im);
% PDE (partial differential equation) initial condition.
diff_im = im;
% Center pixel distances.
dx = 1;dy = 1;dd = sqrt(2);

```

```

% 2D convolution masks - finite differences.
hN = [0 1 0; 0 -1 0; 0 0 0];
hS = [0 0 0; 0 -1 0; 0 1 0];
hE = [0 0 0; 0 -1 1; 0 0 0];
hW = [0 0 0; 1 -1 0; 0 0 0];
hNE = [0 0 1; 0 -1 0; 0 0 0];
hSE = [0 0 0; 0 -1 0; 0 0 1];
hSW = [0 0 0; 0 -1 0; 1 0 0];
hNW = [1 0 0; 0 -1 0; 0 0 0];
% Anisotropic diffusion.
for t = 1:num_iter
% Finite differences. [imfilter(...,'conv') can be replaced by conv2(...,'same')]
nablaN = imfilter(diff_im,hN,'conv');
nablaS = imfilter(diff_im,hS,'conv');
nablaW = imfilter(diff_im,hW,'conv');
nablaE = imfilter(diff_im,hE,'conv');
nablaNE = imfilter(diff_im,hNE,'conv');
nablaSE = imfilter(diff_im,hSE,'conv');
nablaSW = imfilter(diff_im,hSW,'conv');
nablaNW = imfilter(diff_im,hNW,'conv');
% Diffusion function.
if option == 1
cN = exp(-(nablaN/kappa).^2);
cS = exp(-(nablaS/kappa).^2);
cW = exp(-(nablaW/kappa).^2);
cE = exp(-(nablaE/kappa).^2);
cNE = exp(-(nablaNE/kappa).^2);
cSE = exp(-(nablaSE/kappa).^2);
cSW = exp(-(nablaSW/kappa).^2);
cNW = exp(-(nablaNW/kappa).^2);

```

```

elseif option == 2
cN = 1./(1 + (nablaN/kappa).^2);
cS = 1./(1 + (nablaS/kappa).^2);
cW = 1./(1 + (nablaW/kappa).^2);
cE = 1./(1 + (nablaE/kappa).^2);
cNE = 1./(1 + (nablaNE/kappa).^2);
cSE = 1./(1 + (nablaSE/kappa).^2);
cSW = 1./(1 + (nablaSW/kappa).^2);
cNW = 1./(1 + (nablaNW/kappa).^2);
end
% Discrete PDE solution.
diff_im = diff_im + ...
delta_t*(...
(1/(dy^2))*cN.*nablaN + (1/(dy^2))*cS.*nablaS + ...
(1/(dx^2))*cW.*nablaW + (1/(dx^2))*cE.*nablaE + ...
(1/(dd^2))*cNE.*nablaNE + (1/(dd^2))*cSE.*nablaSE + ...
(1/(dd^2))*cSW.*nablaSW + (1/(dd^2))*cNW.*nablaNW );
% Iteration warning.
fprintf('\rIteration %d\n',t);
end

```

```
clear all;clc;clf
I=imread('H.jpg');
figure(1)
imshow(I)
orig=I;
I1=rgb2gray(I);
I=im2double(I1);
imgb=imnoise(I1,'salt & pepper',0.09);
imshow(I)
figure(2)
s =imgb;
num_iter = 70;
delta_t = 1/7;
kappa = 15;
option = 2;
ad = anisodiff2D(s,num_iter,delta_t,kappa,option);
figure, imshow(s,[]),title('image restaurée')
```

Bibliographie

- [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.
- [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.
- [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.
- [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.
- [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.
- [Koe-84] Koenderink J., *The Structure of Images*, *Biol. Cybern*, vol.50, pp. 363-370, 1984.
- [LEY-04] Y Leydier., F. LeBourgeois, H. Emptoz. Serialized k-means for adaptative color image segmentation : Application to document images and others, *Document*

-
- Analysis Systems VI, 6th international workshop, pp. 252-263, Florence, ITALY, 2004.
- [Osh-88b] S. Osher, L. Rudin. Featured-oriented image enhancement using shock filters, SIAM J. on Numerical Analysis (27), pp. 919-940, 1988.
- [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. Dans Bart M. Ter Haar Romeny Editor, Geometry-driven diffusion in computer vision. Dordrecht : Kluwer Academic, 1994, pp. 229-254, 1994.
- [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.
- [Tri-03] E. Trinh. De la numérisation à la consultation des documents anciens, Thèse de doctorat dans le cadre d'un projet européen Débora et d'un projet industriel, Book Restorer, INSA de Lyon, France, 2003.
- [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.
- [Wol-06] C. Wolf. Document Ink bleed-through removal with two hidden Markov random fields and a single observation field, Technical Report RRLIRIS- 2006-019. Lyon : LIRIS, INSA Lyon, 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.
- [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.