

Université Mohamed Khider – Biskra

Faculté des Sciences et de la technologie

Département : Génie électrique

Ref :.....



جامعة محمد خيضر بسكرة

كلية العلوم و التكنولوجيا

قسم: الهندسة الكهربائية

المرجع:.....

Thèse présentée en vue de l'obtention  
Du diplôme de  
**Doctorat en sciences en : Génie électrique**

**Spécialité (Option) : Electronique**

**Débruitage et simplification d'images**

Présentée par :

**Nacira DIFFELLAH**

Soutenue publiquement le 10 janvier 2021

**Devant le jury composé de :**

|                            |                           |               |                            |
|----------------------------|---------------------------|---------------|----------------------------|
| Dr. Zitouni Othmane        | Maitre de Conférences 'A' | Président     | Université de Biskra       |
| Dr. Baarir Zine-Eddine     | Professeur                | Rapporteur    | Université de Biskra       |
| Dr. Taleb-Ahmed Abdelmalik | Professeur                | Co-Rapporteur | Université de Valenciennes |
| Dr. Ouafi Abdelkarim       | Professeur                | Examineur     | Université de Biskra       |
| Dr. Ferdi Youcef           | Professeur                | Examineur     | ENSB Constantine           |
| Dr. Doghmane Noureddine    | Professeur                | Examineur     | Université de Annaba       |

---

# Débruitage et simplification d'images

---

|   |          |
|---|----------|
| Page de garde                                     | 1        |
| Sommaire  | I        |
| Dédicace  | V        |
| Remerciements                                     | VI       |
| Résumé  | VIII     |
| Abstract  | IX       |
| Résumé en arabe                                   | X        |
| Liste des figures                                 | XI       |
| Liste des tableaux                                | XV       |
| Liste des algorithmes                             | XVI      |
| Nomenclature                                      | XVII     |
| <b>I Introduction</b>                             | <b>1</b> |
| I.1 Introduction . . . . .                        | 1        |
| I.2 Dégradation de l'image . . . . .              | 1        |
| I.3 Types de bruit . . . . .                      | 2        |
| I.4 Filtrage . . . . .                            | 7        |
| I.5 Mesures de qualité utilisées . . . . .        | 8        |
| I.5.1 Critère subjectif (aspect visuel) . . . . . | 9        |

|            |  |           |
|------------|--|-----------|
| I.5.2      | Critères objectifs . . . . .   | 9         |
| I.6        | Contribution et organisation de la thèse . . . . .   | 12        |
| <b>II</b>  | <b>Etat de l'art en restauration d'images</b>  | <b>14</b> |
| II.1       | Introduction . . . . .   | 14        |
| II.2       | Définitions utiles . . . . .   | 14        |
| II.2.1     | Région homogène-Région hétérogène . . . . .  | 14        |
| II.2.2     | Coefficient de variation . . . . .   | 15        |
| II.2.3     | Equation aux Dérivées Partielles ( <i>EDP</i> ) . . . . .  | 16        |
| II.2.4     | Variation totale . . . . .   | 18        |
| II.3       | Modèle mathématique . . . . .  | 18        |
| II.4       | Filtre moyenneur . . . . .   | 20        |
| II.5       | Filtre median . . . . .  | 21        |
| II.6       | Filtrage par les méthodes adaptatives : Filtres basés sur la minimisation<br>de l'erreur quadratique moyenne . . . . . | 22        |
| II.6.1     | Filtre de Lee . . . . .  | 22        |
| II.6.2     | Filtre de Frost . . . . .  | 23        |
| II.6.3     | Filtre de Kuan . . . . .   | 25        |
| II.6.4     | Filtres de Lee et Frost modifiés . . . . .   | 26        |
| II.7       | Filtrage par les méthodes de diffusion . . . . .   | 28        |
| II.7.1     | Diffusion isotrope . . . . .   | 28        |
| II.7.2     | Diffusion anisotrope . . . . .   | 30        |
| II.7.3     | Régularisation du gradient avec une Gaussienne . . . . .   | 32        |
| II.7.4     | Diffusion tensorielle . . . . .  | 34        |
| II.7.5     | Filtrage d'image par méthode de Yu et Acton . . . . .  | 34        |
| II.8       | Filtrage par approche variationnelle . . . . .   | 37        |
| II.8.1     | Méthode des moindres carrés . . . . .  | 37        |
| II.8.2     | Le modèle de Tykhonov . . . . .  | 39        |
| II.8.3     | Modèle de Rudin, Osher et Fatemi (ROF) . . . . .   | 40        |
| II.9       | Conclusion . . . . .   | 43        |
| <b>III</b> | <b>Méthodes proximales pour la restauration d'images</b>   | <b>45</b> |
| III.1      | Introduction . . . . .   | 45        |
| III.2      | Rappels de topologie dans $\mathbb{R}^n$ . . . . .   | 45        |
| III.2.1    | Boule, ouvert, fermé . . . . .   | 45        |
| III.2.2    | Minorant/Majorant, Infimum/Supremum, Minimum/Maximum . . . . .   | 46        |
| III.3      | Quelques rappels d'analyse convexe . . . . .   | 47        |
| III.3.1    | Domaine, fonction propre, fonction indicatrice, épigraphe, lignes de<br>niveau . . . . .                               | 47        |
| III.3.2    | Fonction semi-continue inférieurement, fonction coercive . . . . .   | 50        |

|  |           |
|--|-----------|
| III.3.3 Ensemble convexe, fonction convexe, fonction concave . . . . .                                 | 50        |
| III.3.4 Différentiabilité . . . . .  | 52        |
| III.3.5 Fonction affine, Enveloppe supérieure (fonction $\Gamma^\circ(H)$ ) . . . . .                  | 55        |
| III.4 Algorithmes d'optimisation . . . . .   | 56        |
| III.4.1 Conditions d'optimalité . . . . .  | 56        |
| III.4.2 Algorithmes avec calcul de gradient . . . . .  | 57        |
| III.4.2.1 Algorithme d'optimisation convexe 'algorithmes à direc-<br>tions de descente' . . . . .      | 57        |
| III.4.2.2 Algorithme à descente de gardient . . . . .  | 58        |
| III.4.2.3 Méthode de gradient conjugué . . . . .   | 61        |
| III.4.3 Algorithmes sans calcul du gradient . . . . .  | 63        |
| III.4.3.1 Rappel sur les multiplicateurs de Lagrange . . . . .   | 63        |
| III.4.3.2 Conditions d'optimalité du premier ordre : conditions de<br>Karush-Kuhn-Tucker KKT . . . . . | 64        |
| III.5 Algorithme proximal . . . . .  | 65        |
| III.5.1 Algorithme du point proximal . . . . .   | 66        |
| III.5.2 Application du point proximal à l'optimisation convexe . . . . .                               | 68        |
| III.6 Application du proximal au débruitage d'images . . . . .   | 73        |
| III.6.1 Problématique . . . . .  | 73        |
| III.6.2 Algorithme Forward-Backward (FB) appliqué au débruitage d'images<br>(bruit additif) . . . . .  | 74        |
| III.6.2.1 Algorithme explicite-implicite (FB : Forward-Backward) . . . . .                             | 74        |
| III.6.2.2 Résolution du problème . . . . .   | 76        |
| III.6.3 Algorithme Douglas-Rachford (DR) appliqué au débruitage d'images<br>bruit additif . . . . .    | 81        |
| III.6.3.1 Algorithme de Douglas-Rachford (DR) . . . . .  | 81        |
| III.6.3.2 Résolution du problème . . . . .   | 82        |
| III.7 Conclusion . . . . .   | 92        |
| <b>IV Méthode proposée et résultats</b> . . . . .  | <b>93</b> |
| IV.1 Introduction . . . . .  | 93        |
| IV.2 Méthodes de comparaison . . . . .   | 94        |
| IV.2.1 Filtrage d'image par la méthode de Yu et Acton . . . . .  | 94        |
| IV.2.2 Filtrage d'image par la méthode de ROF multiplicative . . . . .                                 | 94        |
| IV.2.3 Diffusion non linéaire complexe améliorée : diffusion complexe adap-<br>tative . . . . .        | 96        |
| IV.2.4 Modèle variationnel convexe . . . . .   | 97        |
| IV.2.5 Modèle variationnel convexe multiplicatif . . . . .   | 98        |
| IV.3 Méthode proposée . . . . .  | 98        |

|   |   |            |
|---|---|------------|
| IV.3.1                                      | Rappel de la loi Gamma . . . . .  | 99         |
| IV.3.2                                      | Bruit multiplicatif de la loi Gamma . . . . .   | 100        |
| IV.3.3                                      | Estimateur du Maximum A Posteriori (MAP) . . . . .  | 101        |
| IV.3.4                                      | Algorithme proposé . . . . .  | 103        |
| IV.3.4.1                                    | Première formulation de débruitage par une regulariza-<br>tion variationnelle lorsque le bruit multiplicatif suit la loi<br>Gamma . . . . . | 103        |
| IV.3.4.2                                    | Deuxième formulation de débruitage par une regulariza-<br>tion variationnelle lorsque le bruit multiplicatif suit la loi<br>Gamma . . . . . | 104        |
| IV.3.4.3                                    | L'algorithme de Douglas-Rachford . . . . .  | 105        |
| IV.3.4.4                                    | La fonction Edata . . . . .   | 105        |
| IV.3.4.5                                    | Existence de ProxEdata . . . . .  | 107        |
| IV.3.4.6                                    | Recherche du ProxEreg . . . . .   | 107        |
| IV.4  | Résultats de simulation et interprétations . . . . .  | 110        |
| IV.5  | Conclusion . . . . .  | 129        |
| <b>Conclusion et perspectives</b>           |   | <b>130</b> |
| A   | Euler explicite-Euler implicite . . . . .   | 133        |
| <b>Annexes</b>                              |   | <b>133</b> |
| B   | Visualisation de l'enveloppe de Moreau Yosida et le proximal de $f(x) =  x $<br>pour $\lambda = 1$ . . . . .                                | 136        |
| C   | Filtrage d'image par RLO . . . . .  | 141        |
| D   | Visualisation des Tracés de la fonction $y =  x $ , sa dérivée et sa tangente . . . . .   | 145        |
| Bibliographie . . . . .                     |   | 148        |
| Couverture avec les trois résumés . . . . . |   | 155        |

---

## DÉDICACE

Je dédie cette thèse à

la mémoire de mon père,  
ma famille qui m'a soutenue tout au long de ce parcours.

*Nacira DIFFELLAH*

---

## REMERCIEMENTS

Mes remerciements s'adressent en premier lieu, au grand **Dieu** pour la volonté, la santé, le courage et la patience qui m'avaient donné pour mener ce travail à terme.

Il est de devoir d'adresser mes remerciements les plus sincères à Monsieur **Abdelmalik TALEB-AHMED**, professeur à Université Polytechnique Hauts de France Valenciennes Lille France, pour m'avoir accueilli au sein de son laboratoire de recherche 'IEMN DOAE UMR CNRS 8520', pour son humanité et surtout pour tous ses conseils et ses directives si précieux, pour sa disponibilité et sa patience, qu'il trouvera mes expressions de reconnaissance, et du grand respect.

J'adresse aussi, mes sincères remerciements les plus profonds à Monsieur **Zine-Eddine-BAARIR**, professeur à l'Université Mohamed Khider Biskra mon directeur de thèse, qu'il trouvera mes expressions de reconnaissance, de profonde gratitude et du grand respect.

J'exprime ma profonde reconnaissance à Monsieur **Foued DERRAZ**, Docteur à l'Université Abou Bekr Belkaid Tlemcen pour son aide, ses conseils et ses directives.

J'adresse aussi, mes sincères remerciements les plus profonds à Monsieur **Othmane ZITOUNI**, Maître de conférence à l'Université de Biskra, de m'avoir fait l'honneur d'accepter de présider le jury de cette thèse.

J'exprime ma gratitude la plus profonde et mon meilleur respect à tous les membres qui vont composer le jury : **Noureddine DOGHMANE** Professeur à l'université de Annaba, **Youcef FERDI** Professeur à l'École Nationale Supérieure de Biotechnologie Constantine et **Abdelkarim OUAFI** Professeur à l'université de Biskra.

Je tiens particulièrement à remercier Monsieur **Tewfik BEKKOUCHE**, Docteur enseignant chercheur à l'université de Bordj Bou Arréridj pour son aide si précieuse, ses conseils et sa disponibilité.

J'ai eu la grande chance de pouvoir terminer les dernières retouches de la rédaction de ma thèse au Laboratoire de monsieur **Lahcene ZIET**, Professeur à l'université Ferhat Abbas de Sétif. Il règne dans ce laboratoire une ambiance extraordinaire rendant les conditions de travail idéales, qu'il trouve ici mes remerciements les plus sincères.

Ma gratitude va également à Monsieur **Salah MOKHENACHE**, Docteur enseignant chercheur à l'université Ferhat Abbas de Sétif pour son aide et ses encouragements.

Je souhaite également exprimer ma gratitude à tous les enseignants qui ont participé à ma formation, à ceux qui ont exercé (et, qui exercent) leur travail avec amour et conscience.

Mes remerciements s'adressent à tous mes collègues enseignants, du département d'électronique de l'université BBA, en particulier **Fouzia HAMADACHE**, ma compagne du parcours, pour leurs encouragements et leurs soutien.

Je remercie chaleureusement et profondément ma **FAMILLE** chacun son nom.

Enfin à tous ceux qui m'ont soutenu de près ou de loin et à tous ceux qui m'ont incité même involontairement à faire mieux, veuillez trouver ici le témoignage de ma profonde gratitude.

*Nacira DIFFELLAH*

[nacirapush@gmail.com](mailto:nacirapush@gmail.com)

[nacira.diffellah@univ-bba.dz](mailto:nacira.diffellah@univ-bba.dz)

L'axe primordial de notre thèse est le filtrage des images dégradées par un bruit multiplicatif de type Gamma. L'application de l'estimateur du Maximum A Posteriori MAP nous a permis de créer un modèle basé sur un critère énergétique, cette énergie est la somme de deux termes le premier est le terme d'attache aux données lié au bruit et le second est le terme de régularisation qui n'est que la variation totale, le minimiseur de l'énergie correspond à l'image recherchée. La méthode proposée pour résoudre ce problème est liée à l'optimisation convexe fondée sur les algorithmes proximaux, nous avons d'abord recherché et montré l'existence et l'unicité du proximal du terme d'attache aux données, puis nous avons appliqué l'algorithme de Douglas-Rachford pour trouver la solution et garantir sa convergence. Afin de comparer notre approche, nous avons présenté cinq méthodes récemment introduites qui utilisent les Equations aux Dérivées Partielles EDP et les approches variationnelles. Cette étude comparative nous a conduit à tirer la conclusion que la méthode proposée donne des résultats satisfaisants de filtrage d'images par rapport aux autres méthodes. Des critères quantitatifs ont été utilisés pour mesurer les performances tels que : le rapport signal sur bruit (PSNR), l'indice de similarité structurelle (SSIM) et le rapport visuel signal sur bruit (VSNR) et la qualité visuelle. L'évaluation de l'approche proposée montre qu'elle fournit des résultats meilleurs que ceux obtenus par les autres méthodes testées. La méthode proposée est programmée sous UNLocBoX qui est une boîte à outils d'optimisation convexe MATLAB basée sur les méthodes de fractionnement proximal.

**Mots clés :** bruit multiplicatif Gamma, filtrage, variation totale, algorithme proximal, mesures de qualité,...

This thesis considers the restoration problem of images which are corrupted by gamma multiplicative noise. Our problem is assumed as a regularization problem in total variation (TV) framework with the data fitting term which is deduced by maximizing the a-posteriori probability density (MAP estimation). We need to evaluate the proximal operator of a data fitting term then we adapt numerically the Douglas - Rachford (DR) splitting method to solve the problem.

To validate the effectiveness of the proposed method, we compare it with other variational models. The measures of image quality such as : PSNR (Peak Signal-to-Noise Ratio), VSNR (Visual Signal-to-Noise Ratio) and SSIM (Structural SIMilarity index) explain that the proposed model has good performance. The proposed method is implemented under UNLocBoX which is, a MATLAB convex optimization toolbox based on proximal splitting methods.

**Key words :** Gamma multiplicative noise, filtering, total variation, proximal algorithm, quality measure,...

## المخلص

تمركز بحثنا في الدكتوراه حول ترشيح الصور المشوشة بواسطة ضوضاء مضاعفة من نوع جاما، باستعمال زيادة كثافة الاحتمال الخلفي MAP، حيث قمنا باستخراج نموذج تم تقسيمه إلى طرفين، الأول له علاقة مباشرة بالضوضاء، أما الثاني فهو يمثل التباعد أو الاختلاف الكلي VT.

ولحلّ المشكل المطروح، قمنا بإيجاد الكمية الأقرب للطرف الاول، و برهنّا أنّ حلّها موجود و وحيد، بعد ذلك استعملنا نظرية دوغلاس و راش فورد لإيجاد الحل الذي يضمن التقارب، ولإثبات نجاعة الطريقة المقترحة، قمنا بمقارنة نتائجنا مع النتائج المتحصّل عليها سابقا، فوجدنا أنّ طريقتنا جدّ فعّالة لنزع الضوضاء من الصور، سواء بالعين المجردة أو استعمال مقاييس الجودة مثل PSNR , SSIM و VSNR .

## الكلمات المفتاحية:

ضوضاء جاما المضاعفة، الترشيح، التباين الكلي، الخوارزمية القريبة، مقياس الجودة...

## LISTE DES FIGURES

|      |   |    |
|------|---|----|
| I.1  | Compression de l'image 'Cameraman.tif' . . . . .  | 2  |
| I.2  | Densité de probabilité de la distribution gaussienne . . . . .  | 3  |
| I.3  | Dégradation de l'image '106025.jpg' par le bruit gaussien . . . . .   | 3  |
| I.4  | Floutage de l'image '106025.jpg' . . . . .  | 4  |
| I.5  | Densité de probabilité de la distribution de poisson . . . . .  | 5  |
| I.6  | Dégradation de l'image '106025.jpg' par le bruit poivre et sel . . . . .  | 5  |
| I.7  | Densité de probabilité de la distribution Gamma . . . . .   | 6  |
| I.8  | Dégradation de l'image '106025.jpg' par le bruit multiplicatif de type Gamma . . . . .  | 7  |
| I.9  | Restauration de l'image '106025.jpg' de gauche à droite. Image : originale, bruiée par poivre et sel, restaurée par filtre moyenneur , restaurée par filtre médian . . . . .  | 8  |
|      |   |    |
| II.1 | Exemple de la variation totale d'une fonction . . . . .   | 18 |
| II.2 | Restauration par le filtre moyenneur de l'image Clock.tif' : de gauche à droite : originale, corrompue par un bruit multiplicatif, restaurée pour les fenêtres : $W = 3$ , $W = 5$ , $W = 7$ , $W = 9$ , $W = 11$ , $W = 13$ , $W = 15$ . . . . . | 21 |
| II.3 | Restauration par le filtre Median de l'image Clock.tif' : de gauche à droite : originale, corrompue par un bruit multiplicatif, restaurée pour les fenêtres : $W = 3$ , $W = 5$ , $W = 7$ , $W = 9$ , $W = 11$ , $W = 13$ , $W = 15$ . . . . .    | 22 |
| II.4 | Restauration par le filtre de Lee de l'image Clock.tif' : de gauche à droite : originale, corrompue par un bruit multiplicatif, restaurée pour les fenêtres : $W = 3$ , $W = 5$ , $W = 7$ , $W = 9$ , $W = 11$ , $W = 13$ , $W = 15$ . . . . .    | 23 |
| II.5 | Restauration par le filtre de Frost de l'image Clock.tif' : de gauche à droite : originale, corrompue par un bruit multiplicatif, restaurée pour les fenêtres : $W = 3$ , $W = 5$ , $W = 7$ , $W = 9$ , $W = 11$ , $W = 13$ , $W = 15$ . . . . .  | 25 |

|        |  |    |
|--------|--|----|
| II.6   | Restauration par le filtre de Kuan de l'image Clock.tiff' : de gauche à droite : originale, corrompue par un bruit multiplicatif, restaurée pour les fenêtres : $W = 3, W = 5, W = 7, W = 9, W = 11, W = 13, W = 15$ . . . . . | 26 |
| II.7   | Classification des pixels selon la valeur du coefficient de variation local, pour les filtres de Lee, Frost et Kuan modifiés. . . . .  | 27 |
| II.8   | Restauration par équation de la chaleur de l'image Clock.tiff' : de gauche à droite : originale, dégradée par un bruit additif gaussien, restaurée pour : $t = 2, t = 20, t = 100$ . . . . .                                   | 29 |
| II.9   | Restauration par Pérona-Malik de l'image Clock.tiff' : de gauche à droite : originale, dégradée par un bruit additif gaussien, restaurée pour : $t = 10, t = 20, t = 30, t = 50, t = 100, t = 200, t = 400$ . . . . .          | 32 |
| II.10  | Restauration par Catté de l'image Clock.tiff' : de gauche à droite : originale, dégradée par un bruit additif gaussien, restaurée pour : $t = 2, t = 4, t = 10, t = 15, t = 20, t = 50, t = 100$ . . . . .                     | 33 |
| II.11  | Restauration par EED de l'image Clock.tiff' : de gauche à droite : originale, dégradée par un bruit additif gaussien, restaurée pour : $t = 5, t = 10, t = 20, t = 30, t = 60, t = 80, t = 100$ . . . . .                      | 35 |
| II.12  | Restauration par SRAD de l'image Clock.tiff' : de gauche à droite : originale avec la zone homogène, bruitée par speckle, restaurée pour : $t = 5, t = 10, t = 20, t = 100$ . . . . .  | 37 |
| II.13  | Restauration de l'image Clock.tiff' par Tykhonov : de gauche à droite : originale, dégradée par un bruit additif gaussien, restaurée pour : $\lambda = 0.3, \lambda = 1, \lambda = 5$ . . . . .                                | 41 |
| II.14  | Restauration par TV de l'image Clock.tiff' : de gauche à droite : originale, dégradée par un bruit additif gaussien, restaurée pour : $\rho = 10, \rho = 30, \rho = 50$ . . . . .  | 43 |
| III.1  | Fonction valeur absolue et son domaine . . . . .   | 48 |
| III.2  | Fonction indicatrice et son domaine . . . . .  | 49 |
| III.3  | Epigraphe des fonctions valeur absolue et indicatrice . . . . .  | 49 |
| III.4  | Illustration du domaine et l'ensemble des lignes de niveau $lev_{\leq \delta} f$ de la fonction $f : H \rightarrow ]-\infty, +\infty]$ . . . . .   | 50 |
| III.5  | Exemple de convexité . . . . .   | 51 |
| III.6  | Exemple d'une fonction convexe . . . . .   | 51 |
| III.7  | Exemple de sous-différentiel d'une fonction . . . . .  | 52 |
| III.8  | La fonction $f(x) =  x $ et son sous-différentiel $f(x) = sgn(x)$ . . . . .  | 53 |
| III.9  | Exemple de sous-différentiel et hyperplans d'une fonction $f(x) \geq f(x_0) + \langle v, x - x_0 \rangle$ . . . . .  | 54 |
| III.10 | Organigramme de la méthode de la descente de gradient à pas fixe . . . . .   | 59 |

|        |  |    |
|--------|--|----|
| III.11 | Organigramme de la méthode de la descente de gradient à pas optimal . . .  | 60 |
| III.12 | Organigramme du gradient conjugué . . . . .  | 61 |
| III.13 | Approximation de Moreau-Yosida $f_\gamma(x)$ de $f(x) =  x $ pour $\gamma = 0.25, 0.50, 1, 2, 4$<br>. . . . .  | 69 |
| III.14 | Proximal de $f(x) =  x $ pour $\gamma = 0, 1, 3, 3, 4, 5$ . . . . .  | 70 |
| III.15 | Représentation de $f(x) =  x $ , son enveloppe de Moreau Yosida et son<br>opérateur proximal pour $\gamma = 1.5$ . . . . .   | 71 |
| III.16 | Visualisation de l'enveloppe de Moreau Yosida et le proximal de $f(x) =  x $<br>pour $\gamma = 1$ . . . . .  | 72 |
| III.17 | Débruitage de l'image Cameraman de dimension $256 \times 256$ par l'algorithme<br>Forward-Backward lorsque l'image est corrompue par un bruit gaussien de<br>variance $\sigma = 10, \sigma = 20, \sigma = 30$ . . . . .                  | 77 |
| III.18 | Débruitage de l'image Clock de dimension $256 \times 256$ par l'algorithme<br>Forward-Backward lorsque l'image est corrompue par un bruit gaussien<br>de variance $\sigma = 10, \sigma = 20, \sigma = 30$ . . . . .                      | 78 |
| III.19 | Débruitage de l'image parrot de dimension $256 \times 256$ par l'algorithme<br>Forward-Backward lorsque l'image est corrompue par un bruit gaussien<br>de variance $\sigma = 10, \sigma = 20, \sigma = 30$ . . . . .                     | 79 |
| III.20 | Débruitage de l'image Satellite de dimension $512 \times 512$ par l'algorithme<br>Forward-Backward lorsque l'image est corrompue par un bruit gaussien de<br>variance $\sigma = 10, \sigma = 20, \sigma = 30$ . . . . .                  | 80 |
| III.21 | Débruitage de l'image Cameraman de dimension $256 \times 256$ par l'algorithme<br>Douglas-Rachford avec contraintes lorsque l'image est corrompue par un<br>bruit gaussien de variance $\sigma = 10, \sigma = 20, \sigma = 30$ . . . . . | 84 |
| III.22 | Débruitage de l'image Clock de dimension $256 \times 256$ par l'algorithme<br>Douglas-Rachford avec contraintes lorsque l'image est corrompue par un<br>bruit gaussien de variance $\sigma = 10, \sigma = 20, \sigma = 30$ . . . . .     | 85 |
| III.23 | Débruitage de l'image parrot de dimension $256 \times 256$ par l'algorithme<br>Douglas-Rachford avec contraintes lorsque l'image est corrompue par un<br>bruit gaussien de variance $\sigma = 10, \sigma = 20, \sigma = 30$ . . . . .    | 86 |
| III.24 | Débruitage de l'image Satellite de dimension $512 \times 512$ par l'algorithme<br>Douglas-Rachford avec contraintes lorsque l'image est corrompue par un<br>bruit gaussien de variance $\sigma = 10, \sigma = 20, \sigma = 30$ . . . . . | 87 |
| III.25 | Débruitage de l'image Cameraman de dimension $256 \times 256$ par l'algorithme<br>Douglas-Rachford sans contraintes lorsque l'image est corrompue par un<br>bruit gaussien de variance $\sigma = 10, \sigma = 20, \sigma = 30$ . . . . . | 88 |
| III.26 | Débruitage de l'image Clock de dimension $256 \times 256$ par l'algorithme<br>Douglas-Rachford sans contraintes lorsque l'image est corrompue par un<br>bruit gaussien de variance $\sigma = 10, \sigma = 20, \sigma = 30$ . . . . .     | 89 |

|        |  |     |
|--------|--|-----|
| III.27 | Débruitage de l'image parrot de dimension $256 \times 256$ par l'algorithme Douglas-Rachford sans contraintes lorsque l'image est corrompue par un bruit gaussien de variance $\sigma = 10, \sigma = 20, \sigma = 30$ . . . . .    | 90  |
| III.28 | Débruitage de l'image Satellite de dimension $512 \times 512$ par l'algorithme Douglas-Rachford sans contraintes lorsque l'image est corrompue par un bruit gaussien de variance $\sigma = 10, \sigma = 20, \sigma = 30$ . . . . . | 91  |
| IV.1   | Coefficient de diffusion $D$ en fonction de $\Delta u$ de selon les valeurs de $k$ . . . . .   | 96  |
| IV.2   | Tracés de la fonction $\log x + \frac{c}{x}$ pour $c = 1$ , sa dérivée et sa tangente . . . . .  | 106 |
| IV.3   | Tracés de la fonction $y =  x $ , sa dérivée et sa tangente . . . . .  | 108 |
| IV.4   | Tracés de la fonction $y =  x $ , sa dérivée et sa tangente en $x_0 = 0$ . . . . .   | 109 |
| IV.5   | Collection des images utilisées. . . . .   | 111 |
| IV.6   | Résultats de restauration de l'image 'Girl' pour $L = 1$ . . . . .   | 112 |
| IV.7   | Résultats de restauration de l'image 'Girl' pour $L = 4$ . . . . .   | 113 |
| IV.8   | Résultats de restauration de l'image 'Girl' pour $L = 10$ . . . . .  | 114 |
| IV.9   | Résultats de restauration de l'image 'Cap' pour $L = 1$ . . . . .  | 115 |
| IV.10  | Résultats de restauration de l'image 'Cap' pour $L = 4$ . . . . .  | 116 |
| IV.11  | Résultats de restauration de l'image 'Cap' pour $L = 10$ . . . . .   | 117 |
| IV.12  | Résultats de restauration de l'image 'Subway' pour $L = 1$ . . . . .   | 118 |
| IV.13  | Résultats de restauration de l'image 'Subway' pour $L = 4$ . . . . .   | 119 |
| IV.14  | Résultats de restauration de l'image 'Subway' pour $L = 10$ . . . . .  | 120 |
| IV.15  | Résultats de restauration de l'image 'Parrot' pour $L = 1$ . . . . .   | 121 |
| IV.16  | Résultats de restauration de l'image 'Parrot' pour $L = 4$ . . . . .   | 122 |
| IV.17  | Résultats de restauration de l'image 'Parrot' pour $L = 10$ . . . . .  | 123 |
| IV.18  | Résultats de restauration de l'image 'Cameraman' pour $L = 1$ . . . . .  | 124 |
| IV.19  | Résultats de restauration de l'image 'Cameraman' pour $L = 4$ . . . . .  | 125 |
| IV.20  | Résultats de restauration de l'image 'Cameraman' pour $L = 10$ . . . . .   | 126 |
| A.1    | Subdivision de l'intervalle de temps . . . . .   | 133 |
| D.1    | Tracés de la fonction $y =  x $ , sa dérivée et sa tangente lorsque $x$ négatif . . . . .  | 145 |
| D.2    | Tracés de la fonction $y =  x $ , sa dérivée et sa tangente au point $x = 0$ . . . . .   | 146 |
| D.3    | Tracés de la fonction $y =  x $ , sa dérivée et sa tangente lorsque $x$ positif . . . . .  | 147 |

---

## LISTE DES TABLEAUX

|   |     |
|---|-----|
| IV.1 Les deux formulations de la distribution Gamma . . . . .                     | 99  |
| IV.2 Avantages et inconvénients de l'utilisation de la variation totale . . . . . | 103 |
| IV.3 Dimension des images test . . . . .  | 111 |
| IV.4 Mesures de qualité des différentes images . . . . .                          | 128 |

---

## LISTE DES ALGORITHMES

|       |  |     |
|-------|--|-----|
| II.1  | Diffusion isotrope : Equation de la chaleur                    | 29  |
| II.2  | Diffusion anisotropique de Pérona-Malik                        | 31  |
| II.3  | Diffusion de Yu et Acton SRAD                                  | 36  |
| III.1 | Directions de descente   | 57  |
| III.2 | Descente de gradient   | 58  |
| III.3 | Descente de gradient à pas fixe                                | 59  |
| III.4 | Descente de gradient à pas optimal                             | 60  |
| III.5 | Gradient conjugué selon Fletcher et Reeves ou Polak et Ribière | 62  |
| III.6 | Point proximal   | 67  |
| III.7 | Forward-Backward (Explicite-Implicite)                         | 76  |
| III.8 | Douglas Rachford   | 82  |
| IV.1  | Méthode de ROF multiplicative                                  | 95  |
| IV.2  | Squelette de l'algorithme proposé : Formulation1               | 104 |
| IV.3  | Squelette de l'algorithme proposé : Formulation2               | 104 |
| IV.4  | Algorithme détaillé de Douglas-Rachford                        | 105 |
| IV.5  | Algorithme de la méthode proposée                              | 110 |
| C.1   | Méthode variationnelle   | 141 |

## Notations générales

|                                |  |
|--------------------------------|--|
| $x$                            | vecteur $x$ de coordonnées spatiales dans une image            |
| $ x $                          | valeur absolue du réel $x$                                     |
| $\ v\ $                        | norme du vecteur $v$   |
| $\nabla$                       | opérateur gradient   |
| $div$                          | opérateur divergence   |
| $\Delta$                       | opérateur laplacien  |
| $*$                            | opérateur de convolution                                       |
| $T$                            | opérateur de transposition                                     |
| $\times$                       | produit matriciel  |
| $\cdot$                        | produit scalaire   |
| $u_x, u_y$                     | dérivée de $u$ respectivement selon les variables $x$ et $y$   |
| $u = (u_i)_{1 \leq i \leq N}$  | vecteur associé à une image composée de $N$ pixels.            |
| $u$                            | image originale (considérée comme image de référence).         |
| $\hat{u}$                      | image restaurée/reconstruite.                                  |
| $Arg \min f$                   | ensemble des minimiseurs de la fonction $f$ .                  |
| $arg \min f$                   | minimiseur de la fonction $f$ en cas d'unicité.                |
| $I_d$ ou $I$                   | opérateur linéaire identité                                    |
| $\ \cdot\ _p$                  | norme $\ell_p$   |
| $\langle \cdot, \cdot \rangle$ | produit scalaire   |
| $\ x\ $                        | norme euclidienne $\ x\  = \langle x, x \rangle^{\frac{1}{2}}$ |

## Notations spécifiques au proximal

|                         |   |
|-------------------------|---|
| $H, G$                  | espaces de Hilbert réels séparables..   |
| $L : H \rightarrow G$   | opérateur linéaire borné.   |
| $L^* : G \rightarrow H$ | adjoint de l'opérateur linéaire borné $L : H \rightarrow G$ .   |
| $\Gamma_0(H)$           | ensemble des fonctions de $H$ dans $]-\infty, +\infty]$ , convexes, semi-continues inférieurement et propres (non égales à $+\infty$ ). |
| $\mathbb{R}^X$          | espace euclidien de dimension $X$ .   |
| $\text{dom } f$         | domaine de la fonction $f : H \rightarrow ]-\infty, +\infty]$ .   |
| $\text{int } S$         | intérieur de l'ensemble $S \subset H$ .   |
| $\text{ri } S$          | intérieur relatif de l'ensemble $S \subset H$ .   |
| $\text{sri } S$         | intérieur relatif fort de l'ensemble $S \subset H$ .  |
| $C$                     | ensemble convexe fermé non vide de $H$ .  |
| $P_C$                   | projecteur sur l'ensemble convexe fermé non vide $C \subset H$  |
| $C$                     | ensemble convexe  |
| $\text{epi } f$         | épigraphe de la fonction $f$ .  |
| $\bar{\mathbb{R}}$      | $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$   |
| $\text{prox}_f x$       | Opérateur proximal de $f(x)$  |

## Abbreviation

|        |   |
|--------|---|
| $dB$   | décibel   |
| $DR$   | Douglas-Rachford  |
| $EDP$  | Equations aux Dérivées Partielles                             |
| $FB$   | Forward-Backward  |
| $MAP$  | Maximum A posteriori  |
| $MSE$  | Mean Square Error 'Erreur quadratique moyenne'                |
| $PSNR$ | Peak Signal to Noise Ratio 'Rapport signal sur bruit'         |
| $ROF$  | Rudin Osher Fatemi  |
| $SSIM$ | Structural SIMilarity 'indice de similarité structurelle'     |
| $VSNR$ | Visual Signal to Noise Rate 'Rapport signal sur bruit visuel' |
| $VT$   | Variation Totale  |

Une heure de lecture est le souverain  
remède contre les dégoûts de la vie

---

Montesquieu

## I.1 Introduction

En traitement d'images, il est souvent indispensable d'enlever le bruit de l'image vu que ce bruit est une anomalie qui dégrade sa qualité, tout en rendant difficile son interprétation visuelle. Le filtrage a pour but d'éliminer l'effet de ces perturbations en essayant de ne pas toucher aux informations essentielles de l'image (contours, dynamique, textures etc.) En général, l'image observée est contaminée par un bruit qui est soit additif ou multiplicatif. Les traitements de restauration sont généralement nécessaires pour améliorer la qualité des images observées.

## I.2 Dégradation de l'image

Nous pouvons définir le bruit comme étant une dégradation dans l'image, provoquée par une perturbation externe. Les raisons de la perte de l'information peuvent être dues à différents facteurs :

1. Les problèmes de transmission : c'est un cas fréquent en astronomie ou en aérospatial. Le signal/image est difficile à capter et donc le récepteur terrestre n'en reçoit qu'une partie.
2. Les problèmes optiques : cette limite de l'acquisition vient directement du système lui-même, que ce soit un appareil photo numérique (CCD), un télescope ou un outil

médical tel qu'un IRM.

3. Problèmes de compression : le récepteur ne reçoit pas le signal direct mais une version comprimée pour une meilleure transmission du signal. Un bon exemple pour les utilisateurs d'Internet est le téléchargement des images en JPEG [Salmon, 2010]. Un exemple de compression d'image est montré par la figure I.1



FIGURE I.1 – Compression de l'image 'Cameraman.tif'

4. Problèmes d'échantillonnage : le bruit d'échantillonnage est causé par la résolution spatiale avec laquelle le champ de vision observé par la caméra est échantillonné. Ce bruit a les effets suivants :
  - affecte la netteté des variations rapides (contours) d'illuminance dans l'image,
  - amène des effets de recouvrement spectral (aliasing) lorsque le critère d'échantillonnage de Nyquist n'est pas respecté.

### I.3 Types de bruit

Les sources de bruit sont difficiles à modéliser. Par conséquent, pour supprimer un bruit, il est important de connaître comment ce bruit se mélange aux données utiles (bruit additif, multiplicatif, etc.). Il est également pertinent de connaître la loi de distribution que suit le modèle de bruit (distribution uniforme, gaussienne, de Rayleigh, etc.)

Les images peuvent être entachées de dégradations de nature différente suivant les conditions d'acquisition. Les images peuvent être dégradées par différentes perturbations telles que :

1. Le bruit additif gaussien : dans de nombreux cas nous considérons que l'image non dégradée  $u(x, y)$  est affectée d'un bruit additif  $\eta(x, y)$  :

$$f(x, y) = u(x, y) + \eta(x, y) \quad (\text{I.1})$$

Le bruit suit la loi normale, sa densité de probabilité est donnée par :

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (\text{I.2})$$

où  $\mu$  est l'espérance mathématique et  $\sigma$  est l'écart type.

Pour  $P(x)$ , on note que la gamme des valeurs de  $x$  où  $-\infty < x < +\infty$ , la fonction de densité de probabilité est différente de zéro, dans les directions positives et négatives. Mais, si on considère une image comme une carte d'intensité, alors les valeurs doivent être non négatives.

La représentation graphique de la distribution gaussienne est montrée par la figure I.2 :

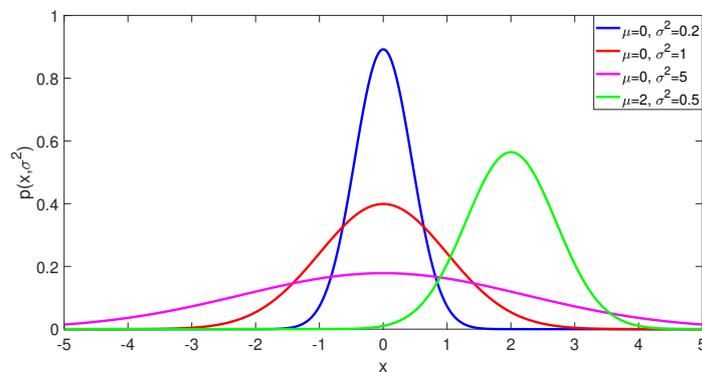


FIGURE I.2 – Densité de probabilité de la distribution gaussienne

La figure I.3 montre un exemple de dégradation de l'image '106025.jpg'<sup>1</sup> par un bruit qui suit la loi gaussienne :



FIGURE I.3 – Dégradation de l'image '106025.jpg' par le bruit gaussien

1. <https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/BSDS300/html/dataset/images/gray/train-126-150.html>

2. Le flou (modélisé par un opérateur de convolution) : un mouvement lors de l'acquisition d'une photographie pourra conduire à une image floue. Le flou dans les images est un phénomène essentiellement convolutif, le modèle simplifié de la formation de l'image  $u$  s'écrit :

$$f = K * u \quad (\text{I.3})$$

où  $K$  est le noyau de convolution.

Un exemple de floutage de l'image '106025.jpg' est illustré dans la figure I.4



FIGURE I.4 – Floutage de l'image '106025.jpg'

3. Le bruit impulsionnel dit poivre et sel : le bruit poivre et sel qui apparaît dans une image numérique est dû soit à des erreurs de transmission de données, soit au dysfonctionnement ou à la présence de particules fines sur les éléments du capteur de la caméra ou à des emplacements mémoire défectueux dans le matériel.

Le bruit poivre et sel<sup>2</sup> suit une distribution de Poisson donnée par :

$$p(k) = \frac{e^{-\lambda} \lambda^k}{k!} \quad (\text{I.4})$$

où  $p(k)$  est la probabilité d'avoir  $k$  pixels affectés par le bruit dans une fenêtre d'une certaine dimension  $\lambda$  et  $k$  est le nombre moyen de pixels affecté dans une fenêtre de la même taille qui est aussi la variance de la distribution de Poisson.

2. [https://fr.wikipedia.org/wiki/Bruit\\_poivre\\_et\\_sel](https://fr.wikipedia.org/wiki/Bruit_poivre_et_sel)

La figure I.5 illustre un exemple de la distribution de poisson.

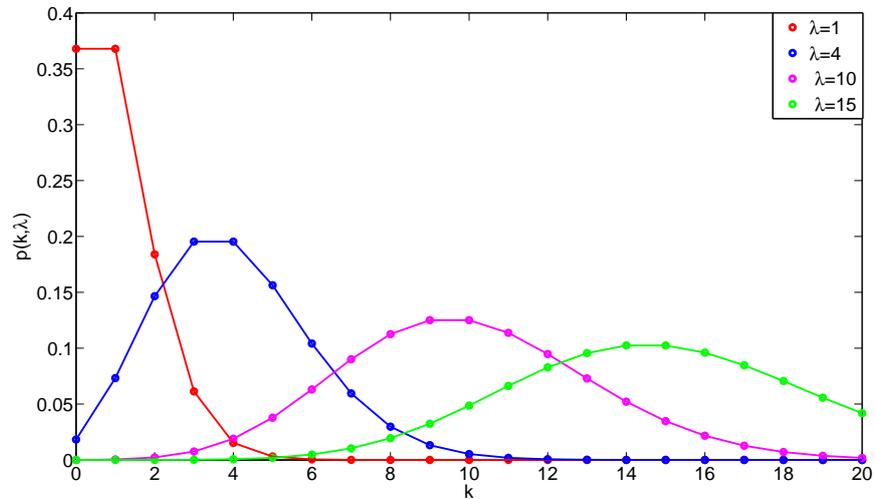


FIGURE I.5 – Densité de probabilité de la distribution de poisson

Le bruit poivre et sel dans une image se traduit dans le cas d'une photo numérique par l'apparition de pixels noirs et blancs comme le montre la figure I.6

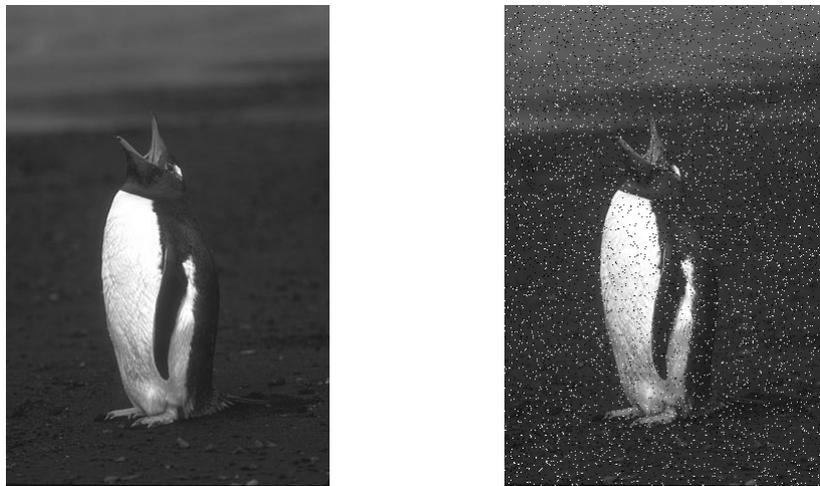


FIGURE I.6 – Dégradation de l'image '106025jpg' par le bruit poivre et sel

- Le bruit multiplicatif : le bruit multiplicatif est un phénomène commun dans tous les systèmes d'imagerie cohérents tels que le laser, l'acoustique, les ultrasons, et les images SAR (Synthetic Aperture Radar).

En général, le bruit multiplicatif est modélisé selon l'expression suivante :

$$f(x, y) = u(x, y) \times \eta(x, y) \quad (\text{I.5})$$

où le bruit est une variable aléatoire de moyenne égale à 1.

Le bruit multiplicatif<sup>3</sup> est beaucoup plus difficile, la principale caractéristique de ce bruit est que les pixels d'une zone homogène seront d'autant plus bruités que leur niveau de gris est élevé<sup>4</sup>.

La densité de probabilité d'une distribution Gamma est donnée par la formule suivante :

$$p(x) = \frac{\lambda^a}{\Gamma(a)} x^{a-1} e^{-\lambda x}, \quad x \in \mathbb{R}^+ \quad (\text{I.6})$$

Cette mesure est identifiée par la notation à  $\Gamma(a, \lambda)$

$\Gamma(a)$  est le facteur de normalisation apparaissant dans la formule I.6 et la valeur en  $a > 0$  de la fonction  $\Gamma$  (Gamma) d'Euler :

$$\Gamma(a) = \int_0^{+\infty} t^{a-1} e^{-t} dt \quad (\text{I.7})$$

ou bien

$$p(x, k, \theta) = \frac{x^{k-1} e^{-\frac{x}{\theta}}}{\theta^k \Gamma(k)} \quad (\text{I.8})$$

La figure I.7 illustre la distribution de la loi Gamma :

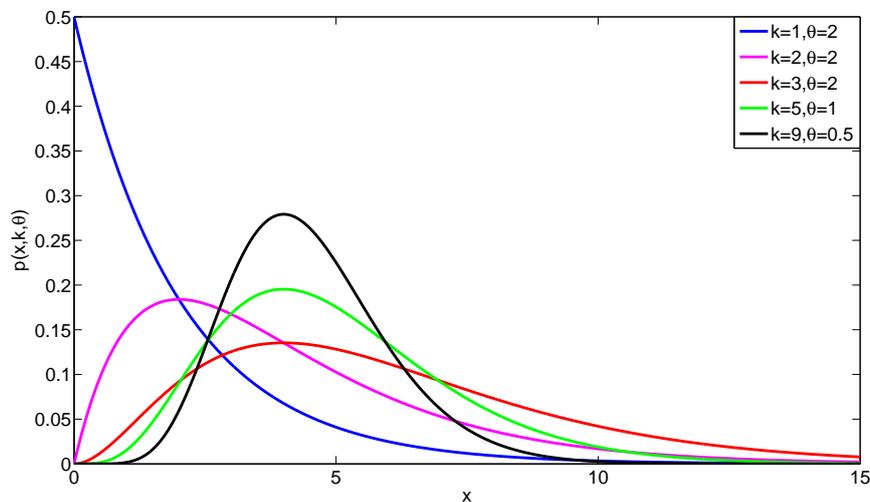


FIGURE I.7 – Densité de probabilité de la distribution Gamma

Un exemple de dégradation d'image par un bruit multiplicatif de type Gamma est illustré par la figure I.8 :

3. Des termes différents sont employés dans des domaines différents : granularité est le terme employé en optique, tavelure en astronomie, chatoiement en radar. Le terme anglo-saxon consacré est speckle. [Bloch et al., 2005]

4. <http://www.tsi.enst.fr/pages/enseignement/ressources/mti/bruit>



FIGURE I.8 – Dégradation de l'image '106025jpg' par le bruit multiplicatif de type Gamma

## I.4 Filtrage

Le filtrage d'image a pour but d'améliorer la qualité d'une image numérique. Le terme qualité peut prendre plusieurs sens suivant la raison qui nécessite d'effectuer cette amélioration. Par exemple, on parle de qualité visuelle lorsqu'on cherche à avoir un meilleur aspect de l'image ou de qualité du signal lorsqu'on cherche à atténuer les distorsions liées aux techniques de numérisation.

Le filtrage manipule uniquement les données de l'image numérisée. On ne fait aucune supposition sur ce que représente l'image. Le résultat obtenu suite à un filtrage dépend donc énormément de la qualité de l'image d'origine. Si le signal est très détérioré, le filtrage seul sera très peu efficace. Dans ce genre de situation le filtrage est juste une première étape dans un traitement plus complexe qui fera intervenir des informations sémantiques sur l'image (modèle, ...). Chaque filtre cherche à atténuer un type de défaut bien précis. Il n'y a pas de filtre universel capable de corriger tous les défauts. Il faut choisir les bons filtres suivant les défauts que nous désirions corriger.

Dans la littérature du traitement d'image, plusieurs méthodes de débruitage d'image ont été proposées et développées. Chaque approche ayant ses propres hypothèses, ses avantages, et ses limites. Toutefois, beaucoup de ces méthodes enlèvent en plus du bruit, des détails de l'image. De ce fait, des algorithmes de débruitage continuent à être développés pour remédier à ce problème.

La figure I.9 illustre les résultats de réduction de bruit d'une image dégradée par un bruit de type poivre et sel avec un filtre linéaire moyonneur ( $3 \times 3$ ) puis un filtre non linéaire médian ( $3 \times 3$ ).

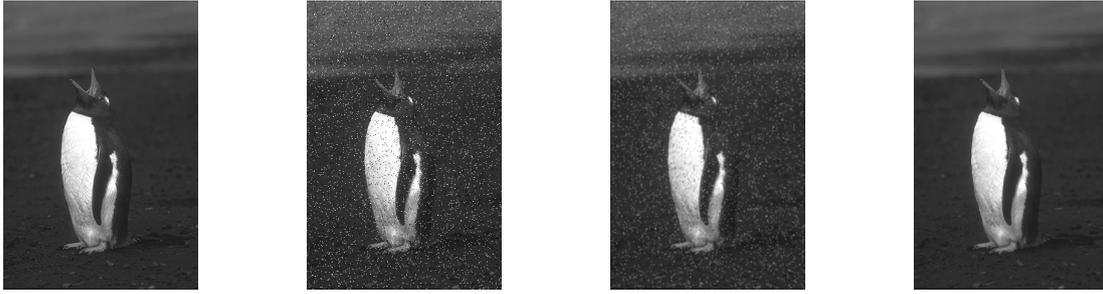


FIGURE I.9 – Restauration de l'image '106025jpg' de gauche à droite. Image : originale, bruitée par poivre et sel, restaurée par filtre moyennneur , restaurée par filtre médian

Le bruit 'poivre et sel' est faiblement réduit dans la troisième image et est visiblement réduit dans la quatrième image donc tout dépend de la méthode de restauration choisie.

Pour pouvoir juger le succès d'un filtre, une analyse de ses performances est nécessaire, pour se faire, nous allons limiter notre étude sur les étapes suivantes :

1. Utiliser une image originale :  $u$ .
2. Dégrader l'image originale :  $f$ .
3. Restaurer l'image dégradée en appliquant les techniques de filtrage existant :  $\hat{u}$ .
4. Quantifier le résultat, évaluation objective de la qualité d'images restaurées en utilisant par exemple : l'erreur quadratique moyenne  $MSE$ , le rapport signal sur bruit  $PSNR$ , l'indice de similarité structurelle  $SSIM$ , le rapport signal sur bruit visuel  $VSNR$ , etc (Voir section I.5). A ces critères quantifiables, on peut ajouter le critère psychovisuel qui est non quantifiable.

## I.5 Mesures de qualité utilisées

Les critères d'évaluation des performances des filtres ont fait l'objet de plusieurs travaux. La pertinence des méthodes de débruitage dépend de deux critères, subjectif et objectif permettant d'estimer l'efficacité de ces méthodes. Le critère subjectif représente l'aspect visuel [Djenabou, 2018].

Ces critères recherchés ont pour objectifs de mesurer le degré d'amélioration de l'image qui peut être selon l'application : la qualité visuelle de l'image, l'élimination ou la réduction du bruit, la préservation des détails et la préservation ou l'amélioration de la qualité du contraste.

Les critères retenus sont les suivants :

### I.5.1 Critère subjectif (aspect visuel)

L'oeil humain est un outil essentiel pour apprécier la qualité d'une image. Il va permettre à l'utilisateur d'identifier le contenu des images, la netteté de celles-ci, la présence d'artefacts et la qualité des contours. Il est donc capital que les méthodes de débruitage prennent en compte le système optique humain. Cependant, cette évaluation ne peut être que subjective puisqu'il n'existe aucune mesure correcte pouvant traduire fidèlement la perception de l'oeil humain [Ghazel, 2004].

### I.5.2 Critères objectifs

La performance d'une méthode de restauration peut être calculée à l'aide d'un des indicateurs suivants :

#### 1. Erreur quadratique moyenne (*Mean Square Error MSE*)

L'image restaurée  $\hat{u}$  est toujours comparée à l'originale  $u$  pour déterminer le rapport de différence. Ce critère est le plus utilisé. Il est basé sur la mesure de l'erreur quadratique moyenne (MSE) calculée entre les pixels originaux et débruités :

$$MSE(u, \hat{u}) = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N (u(i, j) - \hat{u}(i, j))^2. \quad (\text{I.9})$$

avec  $u$  l'image originale,  $\hat{u}$  l'image débruitée,  $M$  le nombre de lignes de l'image,  $N$  le nombre de colonnes de l'image et  $(i, j)$  le positionnement des pixels.

La différence  $u(i, j) - \hat{u}(i, j)$  est considérée comme un bruit.

La valeur la plus faible est à retenir, pas d'échelle de valeur. Pour corriger cet inconvénient (pour pouvoir effectuer des comparaisons), on emploie plutôt une version corrigée du rapport signal sur bruit, notée PSNR.

#### 2. Rapport signal sur bruit (*Peak Signal to Noise Ratio PSNR*)

Le PSNR est l'une des métriques les plus connues et les plus utilisées, c'est une mesure de distorsion utilisée en image numérique. Elle est basée sur l'erreur quadratique moyenne et est donnée par :

$$PSNR(u, \hat{u}) = 10 \log \frac{L_d^2}{MSE} \quad (\text{I.10})$$

où  $L_d$  est la dynamique du signal (la valeur maximum possible pour un pixel), dans le cas standard d'une image codée sur 8-bits,  $L_d = 255$ .

Sa valeur étant indéfinie lorsque les deux images comparées sont identiques et une valeur de PSNR infini correspond à une image non dégradée et cette valeur décroît en fonction de la dégradation.

Le problème avec  $MSE$  et  $PSNR$  est qu'une légère modification d'une image, parfois même imperceptible par l'oeil humain, peut avoir un effet majeur sur ces métriques. Par conséquent, d'autres métriques ont été proposés. Une des plus utilisée est  $SSIM$

### 3. L'indice de similarité structurelle ( *Structural SIMilarity SSIM* )

L'idée de  $SSIM$  [Wang et al., 2004] est de mesurer la similarité de structure entre les deux images, plutôt qu'une différence pixel à pixel comme le fait par exemple le  $PSNR$ . L'hypothèse sous-jacente est que l'oeil humain est plus sensible aux changements dans la structure de l'image.

$SSIM$  est une mesure de similarité  $S(u, \hat{u})$  entre deux images numériques  $u$  et  $\hat{u}$ , elle est basée sur trois mesures de comparaison : la luminance  $l(u, \hat{u})$ , le contraste  $c(u, \hat{u})$  et la structure  $s(u, \hat{u})$ .

La métrique  $SSIM$  est calculée sur plusieurs fenêtres d'une image<sup>5</sup>, la mesure entre deux fenêtres  $u$  et  $\hat{u}$  de taille  $M \times N$  est :

$$SSIM(u, \hat{u}) = \frac{(2\mu_u\mu_{\hat{u}} + C_1)(2\sigma_{u\hat{u}} + C_2)}{(\mu_u^2 + \mu_{\hat{u}}^2 + C_1)(\sigma_u^2 + \sigma_{\hat{u}}^2 + C_2)} \quad (\text{I.11})$$

La luminance est estimée par la mesure de l'intensité moyenne :

$$\mu_u = \frac{1}{N} \sum_{i=1}^N u_i \quad (\text{I.12})$$

$$\mu_{\hat{u}} = \frac{1}{N} \sum_{i=1}^N \hat{u}_i \quad (\text{I.13})$$

Le contraste est mesuré par :

$$\sigma_u = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (u_i - \mu_u)^2} \quad (\text{I.14})$$

$$\sigma_{\hat{u}} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (\hat{u}_i - \mu_{\hat{u}})^2} \quad (\text{I.15})$$

$$\sigma_{u\hat{u}} = \frac{1}{N-1} \sum_{i=1}^N (u_i - \mu_u)(\hat{u}_i - \mu_{\hat{u}}) \quad (\text{I.16})$$

$\mu_u$  est la moyenne de  $u$ ,  $\mu_{\hat{u}}$  la moyenne de  $\hat{u}$ ,  $\sigma_u^2$  la variance de  $u$ ,  $\sigma_{\hat{u}}^2$  la variance de  $\hat{u}$ ,  $\sigma_{u\hat{u}}$  la covariance de  $u$  et  $\hat{u}$ .

5. [https://fr.wikipedia.org/wiki/Structural\\_Similarity](https://fr.wikipedia.org/wiki/Structural_Similarity)

$C_1$  et  $C_2$  sont donnés respectivement par les équations I.17 et I.18 :

$$C_1 = (K_1 L)^2 \quad (\text{I.17})$$

$$C_2 = (K_2 L_d)^2 \quad (\text{I.18})$$

$K_1$  et  $K_2$  sont des constantes de très faible valeur :  $K_1 \ll 1$  et  $K_2 \ll 1$

Les deux constantes  $K_1$  et  $K_2$  sont fixées par les auteurs [Wang et al., 2004] à 0.01 et 0.03, respectivement.  $L_d$  est la dynamique des niveaux de gris.

La fonction *SSIM* doit satisfaire les conditions suivantes :

- $SSIM(u, \hat{u}) = SSIM(\hat{u}, u)$
- $SSIM(u, \hat{u}) \leq 1$
- $SSIM(u, \hat{u}) = 1$  si  $\hat{u} = u$

Le résultat doit tendre vers 1 pour une qualité optimale.

Les constantes  $C_1$  et  $C_2$  évitent l'instabilité des fonctions de comparaison lorsque  $\mu_u^2 + \mu_{\hat{u}}^2$  et  $\sigma_u^2 + \sigma_{\hat{u}}^2$  sont très proches de zéro.

#### 4. Rapport signal sur bruit visuel (*Visual Signal to Noise Rate VSNR*)

Dans [Chandler and Hemami, 2007], les auteurs ont proposé une mesure de qualité basée sur l'analyse des coefficients d'ondelettes. La première étape consiste à calculer un seuil de détection dans le domaine transformée en utilisant quelques caractéristiques du Système Visuel Humain (*SVH*). L'image est considérée comme visuellement identique si le niveau de dégradation est inférieur à ce seuil.

Un index de qualité est défini comme suit :

$$VSNR = 10 \log \frac{C^2(u)}{VD^2} \quad (\text{I.19})$$

où  $C(u)$  est le contraste moyen de l'image originale et  $VD$  est la distortion perceptuelle.

**Remarque I.5.1** *Ces mesures objectives ne représentent pas forcément bien la qualité perçue.*

**Remarque I.5.2** *Quoique les travaux présentés en littérature semblent satisfaisants, le problème du choix du filtre et de l'analyse de ses performances reste posé puisqu'il est difficile de trouver une méthode universelle applicable à tout type d'images. Ce challenge lancera d'autres sujets de recherche.*

## I.6 Contribution et organisation de la thèse

La thèse est organisée en quatre chapitres :

Dans le présent chapitre, nous avons introduit des définitions relatives à la dégradation d'image, les différents types de dégradation, la notion de filtrage, les performances d'une image ainsi que la méthode d'évaluation de la qualité d'image qui peut être basée sur un critère subjectif (aspect visuel) ou sur des critères objectifs tels que : MSE, PSNR, SSIM et VSNR.

Dans le chapitre nommé état de l'art en restauration d'images, nous commençons par les définitions utiles à la compréhension des méthodes de filtrage utilisées telles que : la notion du coefficient de variation, les équations aux dérivées partielles, la variation totale. Ensuite, nous décrivons théoriquement et pratiquement quelques méthodes de filtrage citées dans la littérature comme le filtre moyenneur et le filtre median, nous passons après aux filtres adaptatifs basés sur la minimisation de l'erreur quadratique moyenne, nous citons : filtre de Lee, filtre de Frost, filtre de Kuan, filtres de Lee et Frost modifiés, ces filtres sont appliqués aux images dégradées par un bruit multiplicatif. Puis quelques méthodes de filtrage de diffusion sont élaborées : diffusion isotrope, diffusion anisotrope, régularisation du gradient avec une gaussienne, diffusion tensorielle lorsque l'image est corrompues par un bruit additif gaussien et filtrage d'image par méthode de Yu et Acton pour le cas d'une image dégradée par un bruit multiplicatif. Le chapitre se termine par le filtrage par approche variationnelle : nous décrivons la méthode des moindres carrés suivie par le modèle de Tykhonov et nous cloturons par le modèle de Rudin Osher Fatemi ROF.

Le chapitre suivant décrit les méthodes proximales pour la restauration d'images : nous abordons les algorithmes d'optimisation tels que : les algorithmes avec calcul de gradient et ceux sans calcul de gradient. Ensuite, nous entamons deux applications différentes appliquées au débruitage d'images : la première est l'application de l'algorithme Forward-Backward (FB) et la seconde est l'utilisation de l'algorithme Douglas-Rachford (DR), lorsque le bruit est gaussien additif.

Le quatrième chapitre est consacré au vif de notre contribution dans cette thèse, qui consiste à créer un modèle énergétique convexe de débruitage des images dégradées par un bruit multiplicatif de type Gamma en se basant sur l'estimateur du Maximum A Posteriori MAP. Cette énergie, est composée de deux termes, le premier est le terme d'attache aux données lié au bruit et le second est le terme de régularisation qui n'est que la variation totale. Pour trouver la solution de la problématique évoquée, nous avons proposé la minimisation de cette énergie en appliquant les algorithmes proximaux. En effet, afin de montrer l'efficacité de la méthode proposée, une étude comparative est faite en intégrant

un ensemble de méthodes de filtrage existant dans la littérature , ces méthodes utilisent les Equations aux Dérivées Partielles EDP et d'autres parmi elles utilisent les approches variationnelles. Ces comparaisons sont appuyées sur une évaluation objective utilisant les métriques du PSNR, SSIM et VSNR.

Enfin, la thèse se termine par une conclusion générale qui récapitule les travaux réalisés et propose quelques visions pour les travaux futurs.

Notre contribution réside dans la recherche et la démonstration de l'existence et l'unicité du proximal du terme d'attache aux données, puis l'intégration du résultat trouvé dans l'algorithme de Douglas-Rachford pour trouver la solution et garantir sa convergence.

Notre contribution a été concrétisée par la publication [[Diffellah et al., 2019](#)].

Science sans conscience n'est que  
ruine de l'âme.

---

François Rabelais

## II.1 Introduction

La résolution de quelques problèmes de base en traitement d'images par des méthodes mathématiques : équations aux dérivées partielles, en abrégé *EDP*, dont certains sont exprimés sous forme variationnelle, est un domaine de recherche très actif. Les problèmes issus de la restauration des images sont des problèmes largement traités dans la littérature scientifique. Ces méthodes permettent un développement dans le domaine continu.

La restauration par filtrage d'images 'débruitage' est l'un des objectifs majeurs du traitement d'images [Aubert and Kornprobst, 2006]. Le problème de restauration par filtrage (denoising) consiste à retrouver l'image originale non bruitée  $u$  à partir de l'image dégradée observée  $f$ .

L'objectif de ce chapitre est de présenter les principaux algorithmes utilisés en filtrage du speckle.

Nous allons tout au long de ce chapitre essayer de présenter les bases théoriques et l'implémentation numérique de quelques méthodes de filtrage citées dans la littérature.

## II.2 Définitions utiles

### II.2.1 Région homogène-Région hétérogène

**Définition II.1 (Région homogène)** Une région de l'image est dite homogène [Maher, 1997] si elle correspond à une zone de terrain qui est uniforme avec un coefficient de rétrodiffusion constant. Dans cette région de l'image, l'intensité doit être constante<sup>1</sup>.

**Définition II.2 (Région hétérogène)** Une région de l'image est dite hétérogène [Maher, 1997] si elle correspond à une zone de terrain qui contient plusieurs objets. Autrement dit, c'est une région qui couvre plusieurs sous-régions homogènes séparées par des contours (frontières entre les régions)<sup>2</sup>.

### II.2.2 Coefficient de variation

**Définition II.3 (Coefficient de variation)** Le coefficient de variation d'une variable aléatoire [Maher, 1997]<sup>3</sup> est défini comme étant le rapport entre l'écart type  $\sigma$  et la moyenne  $\mu$  de cette variable :

$$C_v = \frac{\sigma}{\mu} \quad (\text{II.1})$$

Nous différencions deux coefficients de variations [Tauber, 2005] :

- le coefficient de variation global (*CV* global), noté  $\gamma_g$
- le coefficient de variation local (*CV* local), noté  $\gamma_\ell$

**Définition II.4 (Coefficient de variation global)** Soit  $\Omega_h$  une région homogène de l'image observée ; le coefficient de variation global (*CV* global) est défini comme :

$$\gamma_g = \left( \frac{f_{\Omega_h}}{\overline{f_{\Omega_h}^2}} \right)^2 \quad (\text{II.2})$$

où  $f_{\Omega_h}$  est l'ensemble des intensités des pixels de  $\Omega_h$  et  $\overline{f_{\Omega_h}}$  est la moyenne.

Nous pouvons démontrer facilement que le coefficient de variation global peut être aussi égal à :

$$\gamma_g = \left( \frac{\text{var}(\eta)}{\overline{\eta}^2} \right)^2 \quad (\text{II.3})$$

$\overline{\eta}$  est la moyenne de  $\eta$  calculée sur  $\Omega_h$ .

L'équation II.2 montre que le *CV* global ne dépend ni de l'image réelle ni de l'image

---

1. Surfaces homogènes : surfaces dont la fonction de luminance est constante.  
 2. Surfaces hétérogènes : surfaces dont la fonction de luminance rencontre de fortes variations (zones texturées, arêtes, contours, cibles ponctuelles.)  
 3. Le coefficient de variation est une mesure du degré d'homogénéité des surfaces.

observée, mais uniquement du bruit multiplicatif. C'est dans ce sens qu'il en est une caractéristique globale.

**Propriété II.1** *Le coefficient de variation global  $\gamma_g$  est une caractérisation globale du speckle présent dans l'image.*

**Définition II.5 (Coefficient de variation local)** *Le coefficient de variation local (CV local) est défini comme une estimation locale du CV global :*

$$\gamma_\ell = \frac{1}{|\eta_c|} \sum_{p \in \eta_c} \left( \frac{f_p - \bar{f}_c}{\bar{f}_c} \right)^2 \quad (\text{II.4})$$

où  $\eta_c$  est le voisinage du pixel central  $c$ .  $\bar{f}_c$  est la valeur moyenne de l'intensité des pixels de  $\eta_c$ . (Comme exemple nous pouvons considérer le cas de 4-connexité et celui de 8-connexité.)

**Propriété II.2** *Dans une fenêtre contenant des pixels d'une région homogène, la variance de l'image réelle s'annule, on retrouve  $\gamma_\ell \approx \gamma_g$ .*

**Propriété II.3** *Dans une fenêtre contenant des pixels de contours ou des textures. Plus le contour ou la zone texturée est important, plus la variance de l'image réelle est élevée donc le CV local est élevé, on retrouve  $\gamma_\ell \gg \gamma_g$ .*

**Conséquence II.1** *Le coefficient de variation local est un détecteur de contours dans les images contenant du speckle.*

**Conséquence II.2** *Le coefficient de variation est une mesure statistique qui permet non seulement de caractériser le bruit dans les images contenant du speckle, mais aussi de différencier les zones homogènes des contours avec précision.*

### II.2.3 Equation aux Dérivées Partielles (EDP)

**Définition II.6** *Une Equation aux Dérivées Partielles (EDP) est une équation mathématique contenant en plus de la variable dépendante  $u$  et les variables indépendantes  $(x, y, \dots)$  une ou plusieurs dérivées partielles. Cette équation est ainsi de la forme :*

$$F \left( x, y, \dots, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \dots, \frac{\partial^2 u}{\partial x^2}, \frac{\partial^2 u}{\partial x \partial y}, \frac{\partial^2 u}{\partial y^2}, \dots \right) = 0 \quad (\text{II.5})$$

où  $F$  est une fonction de plusieurs variables.

La résolution d'EDP en traitement d'images s'effectue généralement par des méthodes numériques ou variationnelles.

- Les méthodes numériques : ces méthodes changent une EDP en un système d'équations aux différences qui peuvent être résolues sur le plan informatique par des tech-

niques itératives, ces techniques sont souvent les seules qui donnent des résultats satisfaisants.

- Les méthodes variationnelles reformulent l'équation comme un problème de minimisation, le minimum de l'expression est alors solution de l'*EDP*.

Pour obtenir un problème discret nous remplaçons les dérivées partielles par les différences finies :

- Discretisation spatiale

$$\frac{\partial u(x, y; t)}{\partial t} \xrightarrow{\text{Discretisation}} \frac{u_{i,j}^{n+1} - u_{i,j}^n}{\delta t} \quad (\text{II.6})$$

- Différence finie progressive

$$\frac{\partial u(x, y; t)}{\partial x} \xrightarrow{\text{Discretisation}} \frac{u_{i+1,j}^n - u_{i,j}^n}{\delta x} \quad (\text{II.7})$$

$$\frac{\partial u(x, y; t)}{\partial y} \xrightarrow{\text{Discretisation}} \frac{u_{i,j+1}^n - u_{i,j}^n}{\delta y} \quad (\text{II.8})$$

- Différence finie rétrograde

$$\frac{\partial u(x, y; t)}{\partial x} \xrightarrow{\text{Discretisation}} \frac{u_{i,j}^n - u_{i-1,j}^n}{\delta x} \quad (\text{II.9})$$

$$\frac{\partial u(x, y; t)}{\partial y} \xrightarrow{\text{Discretisation}} \frac{u_{i,j}^n - u_{i,j-1}^n}{\delta y} \quad (\text{II.10})$$

- Différence finie centrée

$$\frac{\partial u(x, y; t)}{\partial x} \xrightarrow{\text{Discretisation}} \frac{u_{i+1,j}^n - u_{i-1,j}^n}{2\delta x} \quad (\text{II.11})$$

$$\frac{\partial u(x, y; t)}{\partial y} \xrightarrow{\text{Discretisation}} \frac{u_{i,j+1}^n - u_{i,j-1}^n}{2\delta y} \quad (\text{II.12})$$

- Différence finie centrée d'ordre deux

$$\frac{\partial u(x, y; t)}{\partial x} \xrightarrow{\text{Discretisation}} \frac{u_{i+1,j}^n - 2u_{i,j}^n + u_{i-1,j}^n}{\delta x^2} \quad (\text{II.13})$$

$$\frac{\partial u(x, y; t)}{\partial y} \xrightarrow{\text{Discretisation}} \frac{u_{i,j+1}^n - 2u_{i,j}^n + u_{i,j-1}^n}{\delta y^2} \quad (\text{II.14})$$

Le choix du schéma numérique doit prendre en compte le type d'évolution de l'équation pour avoir des chances que le système soit stable.

La technique des *EDPs* est appliquée dans la section II.7.

## II.2.4 Variation totale

La variation totale discrète se calcule par une différence finie (semblable à une dérivée d'un signal continu) :

$$VT(f) = \sum_i |y_{i+1} - y_i| = \sum_i |f(x_{i+1}) - f(x_i)| \quad (\text{II.15})$$

La variation totale sur l'intervalle de définition est la mesure de la trajectoire de la courbe de la fonction  $y = f(x)$ . La figure II.1 illustre un exemple concret de la variation totale sur une dimension. Lorsque le point vert se déplace sur le graphe de la fonction représentée

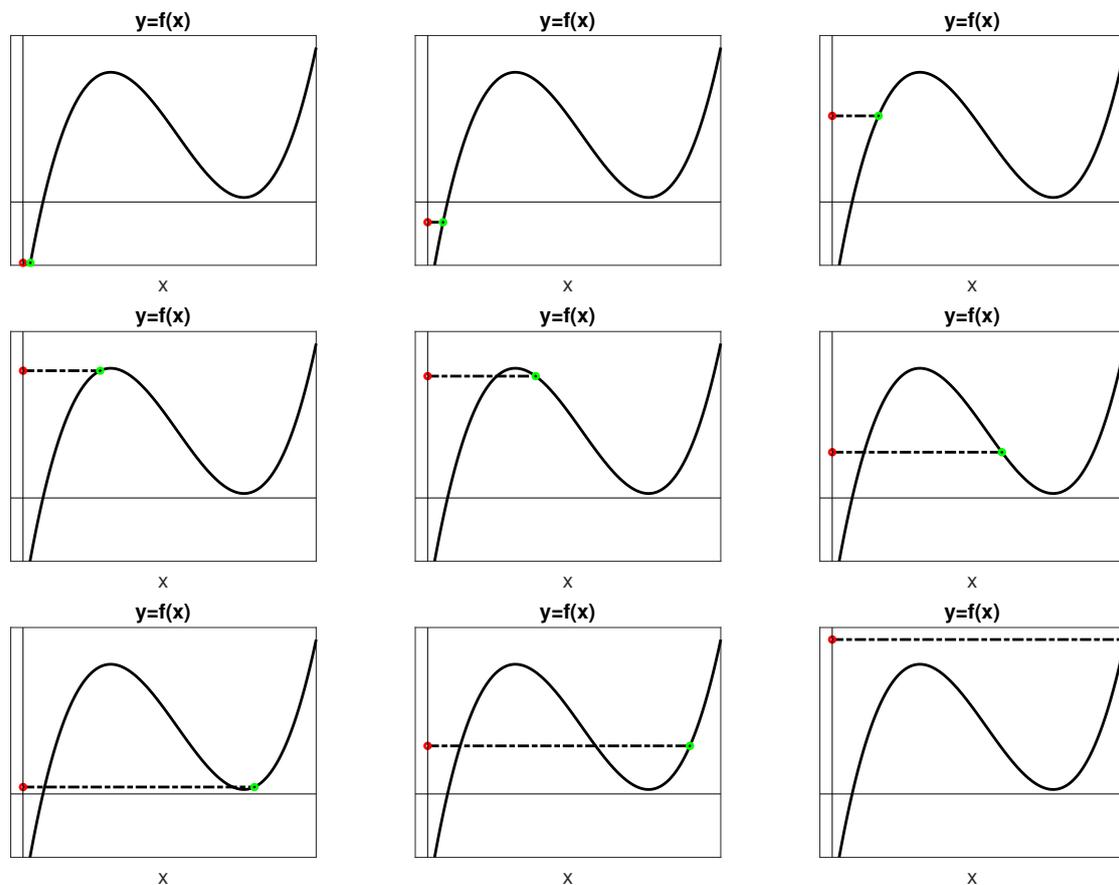


FIGURE II.1 – Exemple de la variation totale d'une fonction

par la figure II.1, la longueur du chemin parcouru par la projection sur l'axe  $y$  de ce déplacement (montré par un point rouge), représente la variation totale de la fonction.

## II.3 Modèle mathématique

Le modèle d'une image corrompue par un bruit multiplicatif est donné par :

$$f = u \times \eta \quad (\text{II.16})$$

$(x, y)$  représente les coordonnées spatiales de l'image (pixels), et  $f$  est l'intensité de l'image observée au point  $(x, y)$ .  $u$  correspond à l'information que l'on veut mesurer (intensité non bruitée).  $\eta$  est un processus aléatoire qui correspond au bruit, ou à l'élément perturbateur du signal. La composante de bruit  $\eta$ , est donc ici multipliée au signal au lieu d'être additionnée comme c'est généralement le cas. Le processus de filtrage consiste à estimer l'image  $u$  pour chaque pixel de l'image connaissant l'intensité observée  $f$ . On note par  $\hat{u}$  la valeur estimée de l'image  $u$ , l'intensité du signal bruit,  $\eta$ , est distribuée selon une loi exponentielle négative. La fonction de densité de probabilité est définie comme suit :

$$p_{\eta}(\eta) = \frac{1}{\eta} e^{-\frac{\eta}{\bar{\eta}}} \quad (\text{II.17})$$

$\bar{\eta}$  est la moyenne de  $\eta$ . On montre que cette moyenne est égale à 1, son écart type égal à sa moyenne, et en considérant l'aspect multiplicatif du bruit, nous constatons que la variation du signal est importante. Une manière de réduire la variance est de prendre la moyenne. L'intensité du bruit est distribuée selon une loi Gamma avec une moyenne égale à 1, et une variance égale à  $\frac{1}{L}$ . La fonction de densité de probabilité est définie comme suit :

$$p_{\eta}(\eta) = \frac{L^L}{\Gamma(L)} e^{-L\eta} \eta^{L-1} \quad (\text{II.18})$$

Le coefficient de variation du bruit est :

$$C_{\eta} = \frac{1}{\sqrt{L}} \quad (\text{II.19})$$

Dans une région homogène, c'est-à-dire avec une image constante, l'intensité d'une image,  $u$  est distribuée selon une loi Gamma. La fonction de densité de probabilité conditionnelle est donnée par :

$$p(f/u) = \frac{L^L}{\Gamma(L) u^L} e^{-L\frac{f}{u}} f^{L-1} \quad (\text{II.20})$$

La moyenne  $\bar{f}$ , et la variance  $Var(f)$  de  $f$  sont données par :

$$\bar{f} = u \quad (\text{II.21})$$

$$Var(f) = \frac{u^2}{L} \quad (\text{II.22})$$

Ainsi, dans une région homogène le coefficient de variation de l'intensité  $f$  est donné par :

$$C_f = \frac{\sqrt{Var(f)}}{\bar{f}} \quad (\text{II.23})$$

ou encore

$$C_f = \frac{1}{\sqrt{L}} \quad (\text{II.24})$$

Les méthodes de filtrage des images dégradées par un bruit multiplicatif reposent sur ce modèle mathématique, et visent à détecter la nature d'une région (homogène ou hétérogène) avant de la filtrer. Ces filtres se basent surtout sur le résultat de l'équation II.24. Il s'agit donc de calculer localement le coefficient de variation sur une fenêtre, et de comparer cette valeur avec  $\frac{1}{\sqrt{L}}$  afin de déterminer la nature de la fenêtre. Le processus de filtrage tient compte de la nature de cette fenêtre dans l'estimation de l'image originale pour chaque pixel. Ce type de filtrage est dit filtrage adaptatif. Le filtrage adaptatif sera décrit avec plus de détails dans la section II.6 où nous présenterons quelques méthodes de ce type de filtrage.

## II.4 Filtre moyennneur

Le filtre moyennneur (Box Filter) consiste à calculer pour chaque pixel la valeur moyenne sur une fenêtre de  $N \times N$  centrée sur le pixel :

$$\hat{u} = \frac{1}{N \times N} \sum_{x=1}^N \sum_{y=1}^N f(x, y) \quad (\text{II.25})$$

- Si les pixels inclus dans la fenêtre appartiennent tous à la même région homogène c'est-à-dire ils proviennent tous de la même population, dans ce cas la moyenne des pixels est le meilleur estimé de l'image  $u$ .
- Par contre, si la fenêtre est à cheval sur deux régions distinctes, les pixels vont appartenir à deux populations différentes, la valeur moyenne calculée sera donc un estimé de la moyenne des pixels des deux régions. Ce type de lissage fait disparaître les lignes et les petites régions isolées, ce qui peut rendre les contours flous et provoquer une perte au niveau de la conservation de la structure de l'image.

La Figure II.2 montre le résultat par un filtre moyennneur avec des fenêtres de tailles différentes pour l'image 'clock.tiff', Nous remarquons bien que ce filtre rend les contours flous.

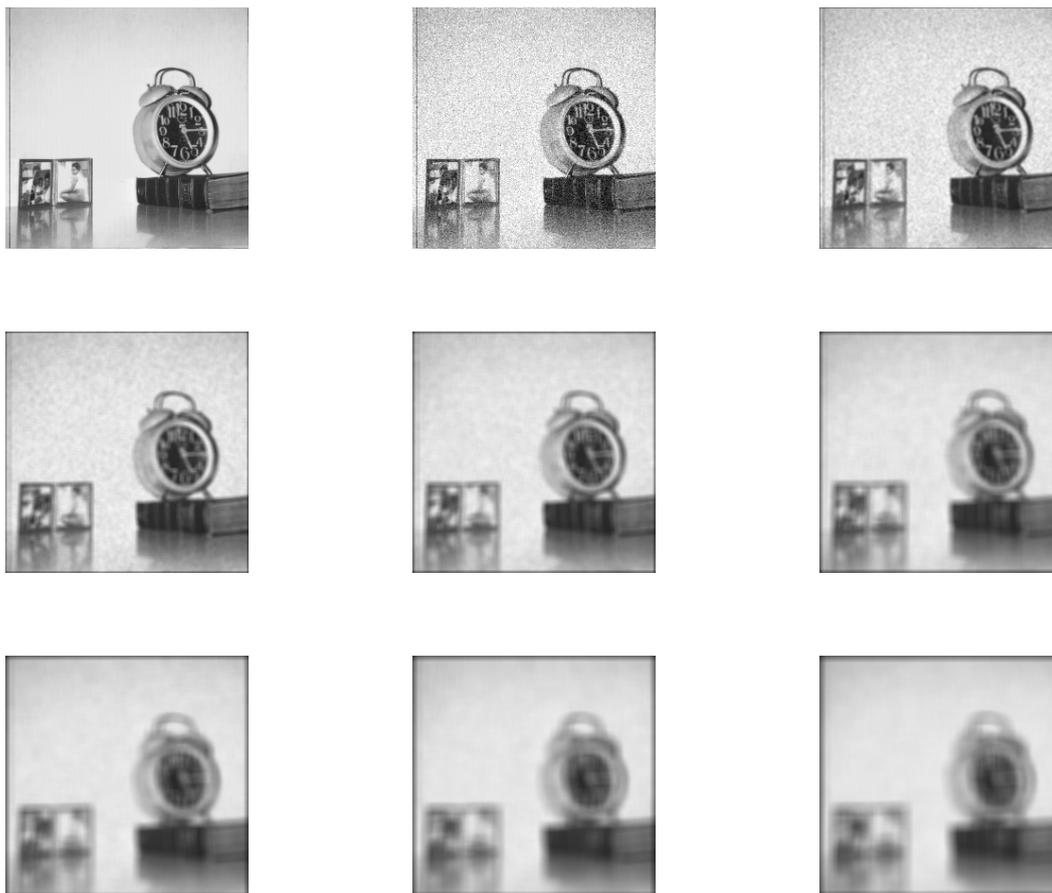


FIGURE II.2 – Restauration par le filtre moyenneur de l'image Clock.tiff' : de gauche à droite : originale, corrompue par un bruit multiplicatif, restaurée pour les fenêtres :  $W = 3$ ,  $W = 5$ ,  $W = 7$ ,  $W = 9$ ,  $W = 11$ ,  $W = 13$ ,  $W = 15$

## II.5 Filtre median

La technique de filtre médian remplace l'intensité de chaque pixel par l'intensité médiane prise parmi toutes les intensités des pixels voisins.

La Figure [II.3](#) montre le résultat du filtrage median avec des fenêtres de tailles différentes pour l'image 'clock.tiff'.

Le filtre médian donne des résultats non satisfaisants.

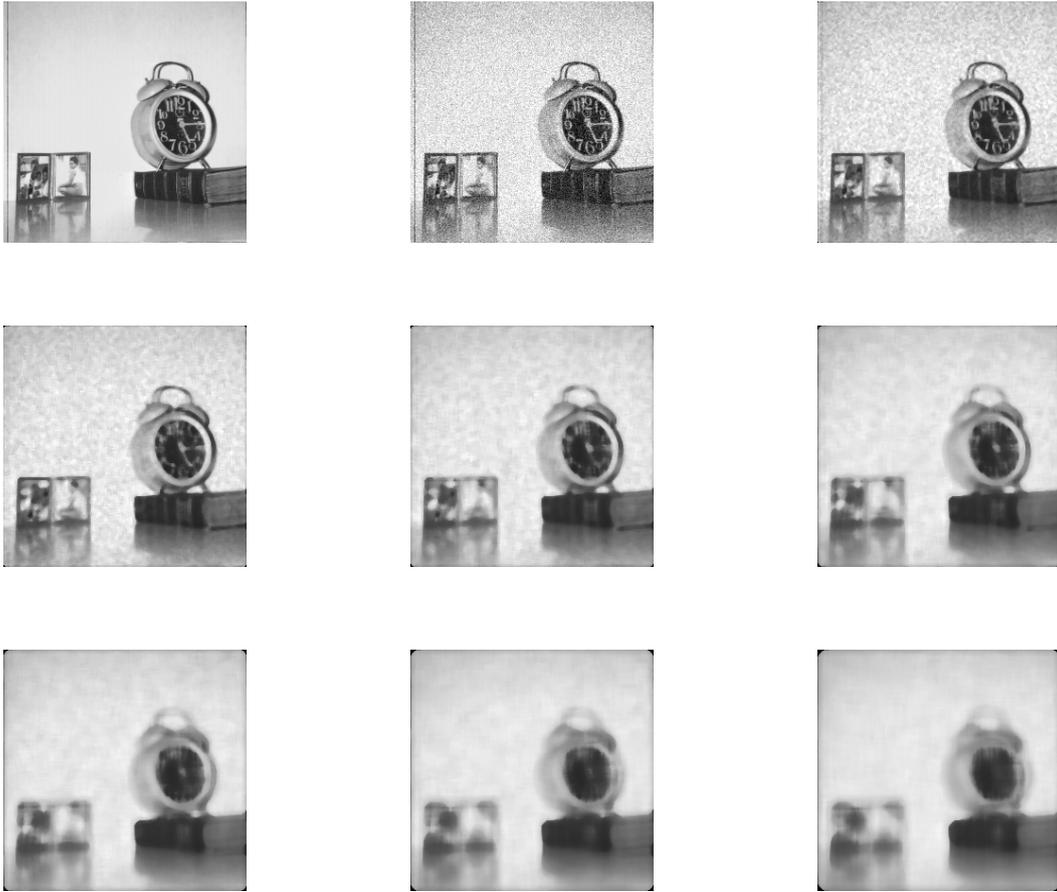


FIGURE II.3 – Restauration par le filtre Median de l'image Clock.tiff' : de gauche à droite : originale, corrompue par un bruit multiplicatif, restaurée pour les fenêtres :  $W = 3$ ,  $W = 5$ ,  $W = 7$ ,  $W = 9$ ,  $W = 11$ ,  $W = 13$ ,  $W = 15$

## II.6 Filtrage par les méthodes adaptatives : Filtres basés sur la minimisation de l'erreur quadratique moyenne

### II.6.1 Filtre de Lee

Le filtre de Lee [Lee, 1980] est donné par la formulation suivante :

$$\hat{u} = \bar{f} + k(f - \bar{f}) \quad (\text{II.26})$$

où  $f$  est l'image observée,  $\bar{f}$  la moyenne de  $f$  et  $k$  est appelé gain du filtre, il est donné par l'expression II.27 :

$$k = 1 - \left( \frac{\gamma_g}{\gamma_\ell} \right)^2 \quad (\text{II.27})$$

Ce paramètre  $k$  est évalué en comparant le  $CV$  local au  $CV$  global :

- Si  $\gamma_g \simeq \gamma_\ell \Rightarrow \frac{\gamma_g}{\gamma_\ell} \simeq 1 \Rightarrow k \simeq 0$  : sur une zone homogène le  $CV$  local et le  $CV$  global

sont proches ; la valeur de  $k$  est donc très faible et le filtre se comporte comme un filtre moyennneur.

- Si  $\gamma_\ell \gg \gamma_g \Rightarrow \frac{\gamma_g}{\gamma_\ell} \simeq 0 \Rightarrow k \simeq 1$  : sur un fort contour, le  $CV$  local est très supérieur au  $CV$  global. La valeur de  $k$  s'approche donc de l'unité et le filtrage est arrêté.
- Si  $0 < k < 1 \Rightarrow 0 < 1 - \left(\frac{\gamma_g}{\gamma_\ell}\right)^2 < 1 \Rightarrow 0 < \left(\frac{\gamma_g}{\gamma_\ell}\right)^2 < 1$ , le filtre lisse l'image à la manière d'un filtre passe-bas.

La Figure II.4 montre le résultat du filtrage de Lee avec des fenêtres de tailles différentes pour l'image 'clock.tiff'.

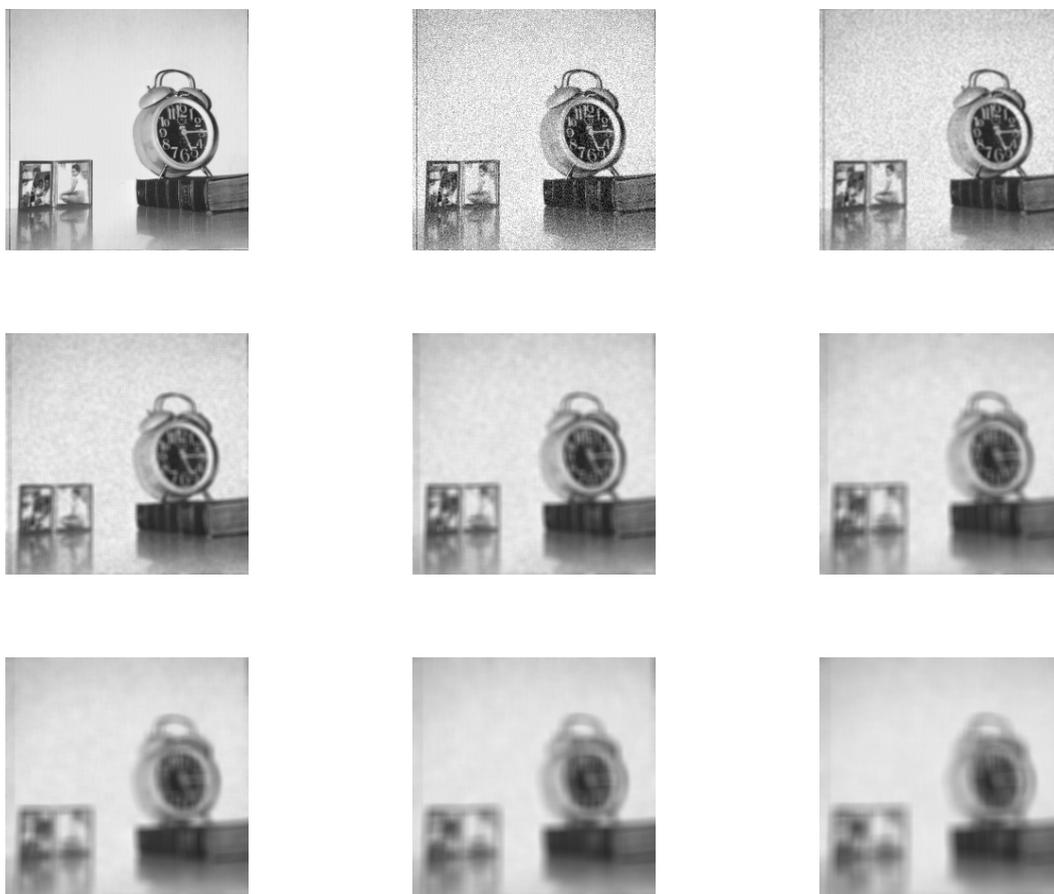


FIGURE II.4 – Restauration par le filtre de Lee de l'image 'Clock.tiff' : de gauche à droite : originale, corrompue par un bruit multiplicatif, restaurée pour les fenêtres :  $W = 3$ ,  $W = 5$ ,  $W = 7$ ,  $W = 9$ ,  $W = 11$ ,  $W = 13$ ,  $W = 15$

### II.6.2 Filtre de Frost

Le filtre de Frost [Frost et al., 1982] peut être considéré comme une variante du filtre de Wiener [Wiener, 1949], adapté au bruit multiplicatif. L'image filtrée est obtenue en

convoluant l'image par la réponse impulsionnelle  $h$  du filtre de Frost :

$$f = u \times \eta - h \quad (\text{II.28})$$

La réponse impulsionnelle du filtre de Frost se calcule comme suit :

$$m(t_0) = K_1 \cdot \exp(-K_2 \cdot \gamma_\ell^2(t_0) \cdot |t|) \quad (\text{II.29})$$

où  $t_0 = (i_0, j_0)$  est le centre de la fenêtre de calcul,  $t = (i - i_0, j - j_0)$  les coordonnées du pixel courant exprimées dans le repère centré sur  $t_0$ ,  $|t|$  la distance au pixel  $t_0$ <sup>4</sup>,  $K_1$  est une constante de normalisation introduite pour ne pas biaiser la moyenne et  $K_2$  le paramètre du filtre.

Le principe de ce filtre est que tous les pixels de la fenêtre participent au filtrage mais avec un poids qui diminue en fonction de :

- la distance  $|t|$  au pixel central  $t_0$  et
- l'hétérogénéité mesurée par le  $CV$  local  $\gamma_\ell$ .

$K_2$  permet de régler la quantité de filtrage :

- lorsque  $K_2$  est grand, le filtrage est plus faible,
- lorsque  $K_2$  est faible, la fonction exponentielle décroît plus lentement : le filtrage est plus fort.

$K_2$  est choisi de telle sorte que le produit  $K_2 \gamma_\ell^2$  tend vers zéro sur les zones homogènes et devient suffisamment grand aux abords des contours pour que le filtrage s'arrête.

---

4.  $|t| = \sqrt{(i - i_0)^2 + (j - j_0)^2}$

La figure II.5 montre les résultats du filtrage de Frost sur l'image 'Clock.tiff' :

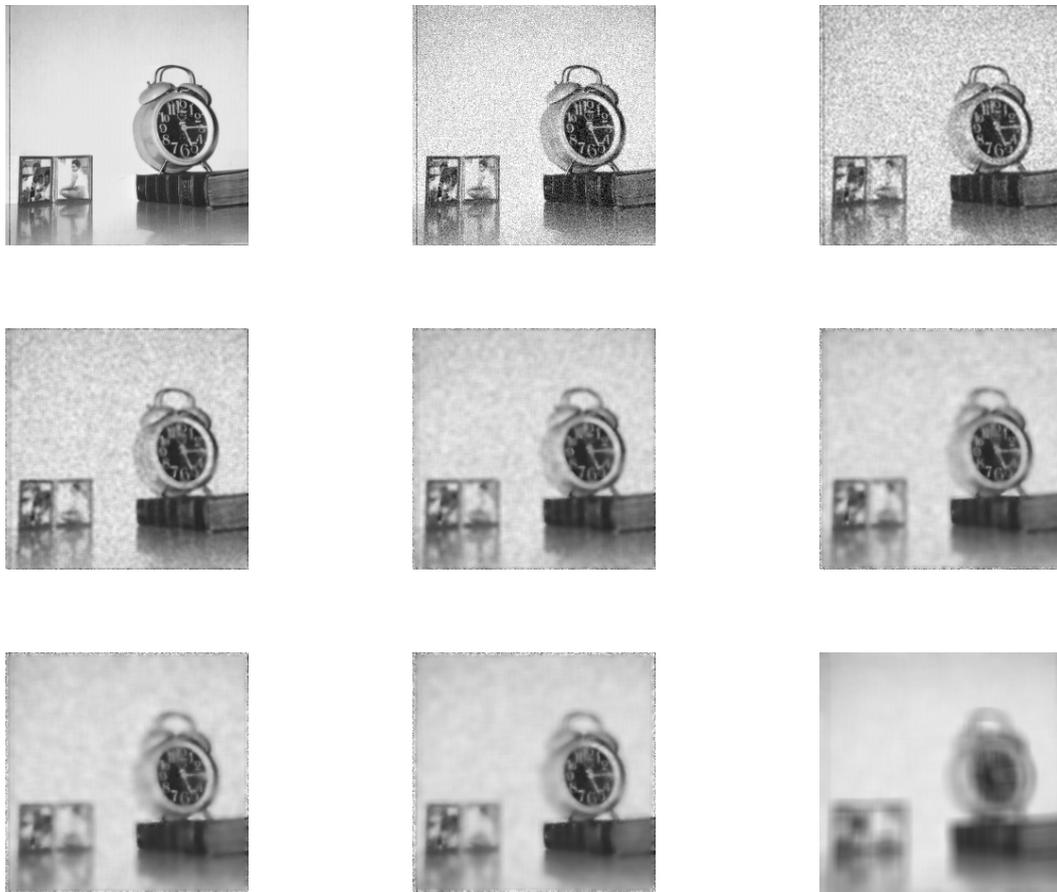


FIGURE II.5 – Restauration par le filtre de Frost de l'image 'Clock.tiff' : de gauche à droite : originale, corrompue par un bruit multiplicatif, restaurée pour les fenêtres :  $W = 3$ ,  $W = 5$ ,  $W = 7$ ,  $W = 9$ ,  $W = 11$ ,  $W = 13$ ,  $W = 15$

### II.6.3 Filtre de Kuan

Dans l'approche proposée par Kuan [Kuan et al., 1985], le modèle de bruit multiplicatif II.16 est d'abord transformé en un modèle additif dépendant du signal<sup>5</sup> :

$$f = u + u(\eta - 1) \quad (\text{II.30})$$

Le filtre de Kuan est semblable au filtre de Lee, mais avec un gain qui se calcule comme suit :

$$k = \frac{1 - \frac{\gamma_g^2}{\gamma_l^2}}{1 + \gamma_g^2} \quad (\text{II.31})$$

Selon la relation II.31, nous pouvons distinguer deux cas :

5.  $f = u \times \eta = u + u \times \eta - u = u + u(\eta - 1)$

- Si  $\gamma_g \simeq \gamma_\ell \Rightarrow \frac{\gamma_g}{\gamma_\ell} \simeq 1 \Rightarrow k \simeq 0$  : le gain est faible dans les zones homogènes.
- Si  $\gamma_g^2 \ll \gamma_\ell \Rightarrow k \simeq 1$  : le gain est proche de l'unité sur les contours.

La figure II.6 illustre les résultats du filtrage de Kuan sur l'image 'Clock.tiff' :

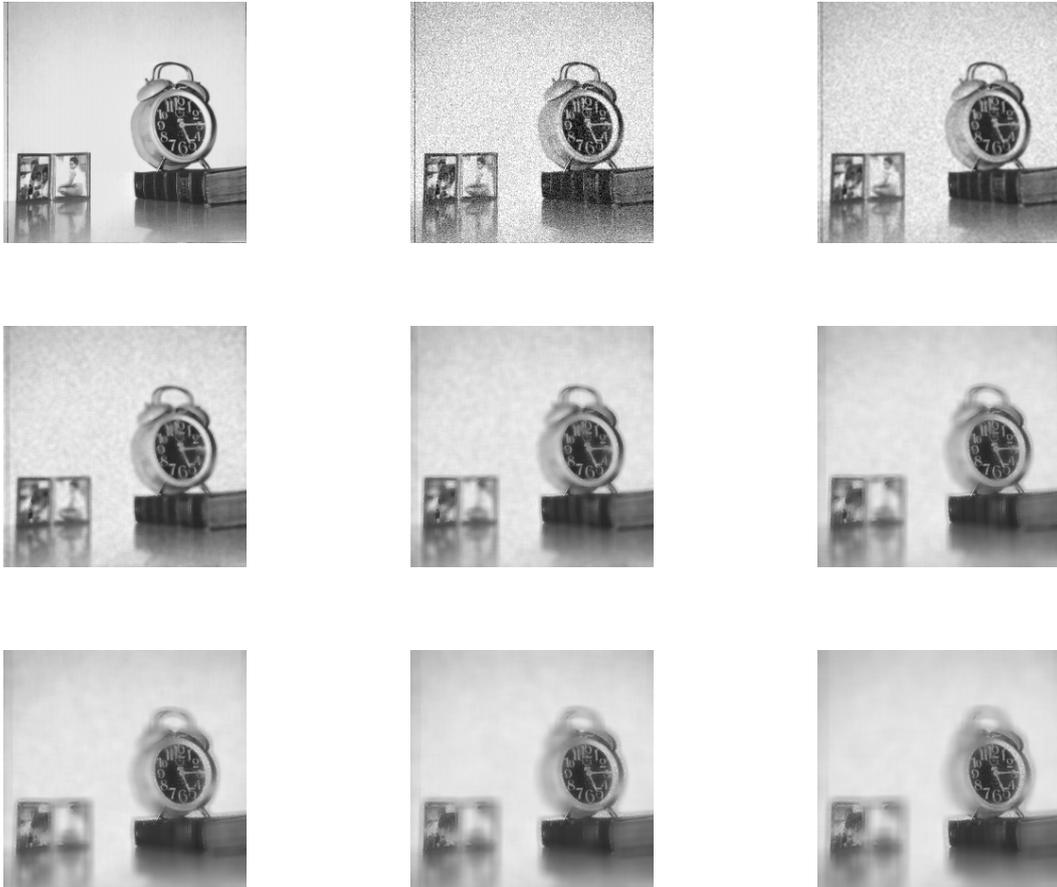


FIGURE II.6 – Restauration par le filtre de Kuan de l'image 'Clock.tiff' : de gauche à droite : originale, corrompue par un bruit multiplicatif, restaurée pour les fenêtres :  $W = 3$ ,  $W = 5$ ,  $W = 7$ ,  $W = 9$ ,  $W = 11$ ,  $W = 13$ ,  $W = 15$

#### II.6.4 Filtres de Lee et Frost modifiés

Dans [Lopes et al., 1990], les auteurs proposent une méthode pour améliorer le filtre de Lee et le filtre de Frost. Cette méthode se base sur une subdivision de l'image en différentes zones de trois classes :

- La première classe correspond aux zones homogènes dans lesquelles l'utilisation d'un filtre moyen est justifiée.
- La deuxième correspond aux zones hétérogènes dans lesquelles le filtrage doit réduire le bruit en préservant les contours.

- La troisième classe correspond aux zones où il faut éviter de filtrer. Cette classe dite préservée, le filtre doit fonctionner comme un filtre passe-tout et laisser identique l'intensité observée du pixel.

Le filtre de Lee amélioré est défini comme suit :

$$\hat{u} = \begin{cases} \bar{f} & \text{si } \gamma_\ell < \gamma_g \\ \bar{f} + k(f - \bar{f}) & \text{si } \gamma_g < \gamma_\ell < C_{\max} \\ f & \text{si } \gamma_\ell \geq C_{\max} \end{cases} \quad (\text{II.32})$$

avec

$$k = \exp\left(-K \frac{\gamma_\ell - \gamma_g}{C_{\max} - \gamma_\ell}\right) \quad (\text{II.33})$$

$K$  étant une constante,  $C_{\max} = 3\gamma_g$ ,

Pour le filtre de Frost amélioré, l'expression de la fonction réponse impulsionnelle est la suivante :

$$m(t) = K_1 \exp(-K_2 \hbar(\gamma_\ell)(t_0) |t|) \quad (\text{II.34})$$

$\hbar(\gamma_\ell)$  est une fonction hyperbolique du CV local défini comme suit :

$$\hbar(\gamma_\ell) = \begin{cases} 0 & \text{si } \gamma_\ell < \gamma_g \\ \frac{\gamma_\ell - \gamma_g}{C_{\max} - \gamma_\ell} & \text{si } \gamma_g < \gamma_\ell < C_{\max} \\ \infty & \text{si } \gamma_\ell \geq C_{\max} \end{cases} \quad (\text{II.35})$$

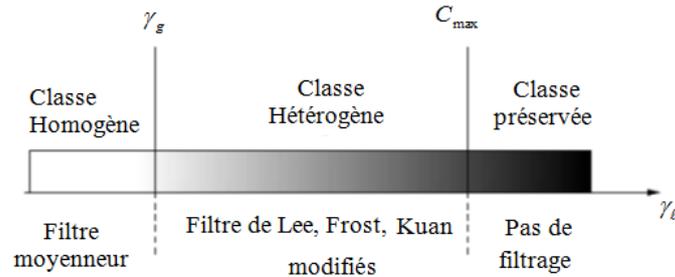


FIGURE II.7 – Classification des pixels selon la valeur du coefficient de variation local, pour les filtres de Lee, Frost et Kuan modifiés.

Notons que :

- dans le cas où  $\hbar(\gamma_\ell)(t_0) = 0$ , le filtre de Frost se comporte comme le filtre moyen ;
- alors que si  $\hbar(\gamma_\ell)(t_0) \rightarrow \infty$ , la réponse impulsionnelle tend vers zéro, ce qui correspond à garder la valeur observée du pixel.

Nous pouvons remarquer que le gain  $k$  de la version modifiée des filtres de Lee et Kuan est similaire à la fonction  $\hbar$  de la version modifiée de Frost lorsque la distance  $|t|$  est constante.

## II.7 Filtrage par les méthodes de diffusion

Dans cette section, nous allons étudier et analyser quelques modèles basés sur l'application des équations aux dérivées partielles (*EDP*) afin de minimiser le phénomène de bruitage de l'image.

Nous commençons par l'équation de la chaleur, c'est un filtrage de diffusion isotrope pour homogénéiser l'intensité de l'image. Ce filtre ne distingue pas entre le bruit et les contours de l'image, l'image résultante après filtrage serait souvent floue. Pour remédier à cet inconvénient, nous introduisons ensuite le filtrage anisotrope basé sur le modèle de Pérona-Malik, c'est un filtre adaptatif qui consiste à lisser les zones homogènes et de rehausser les contours. Nous ajoutons par la suite deux autres modèles basés sur la régularisation du modèle de Perona Malik, nous décrivons en [Catté et al., 1992] un modèle fondé sur la régularisation du gradient avec une gaussienne, et le modèle de Weickert [Weickert, 1998] qui introduit un tenseur de diffusion régit le type de diffusion voulue suivant les structures présentes dans l'image.

### II.7.1 Diffusion isotrope

La première méthode historique de débruitage est la diffusion isotrope d'une image, qui s'effectue à l'aide d'un produit de convolution de l'image par un opérateur Gaussien de moyenne nulle et d'écart-type  $\sigma$ , ceci revient à résoudre l'équation de propagation de la chaleur à une échelle  $t = \frac{\sigma^2}{2}$  [Koenderink, 1984] :

$$\frac{\partial u}{\partial t} = \text{div}(c \cdot \nabla u) \quad (\text{II.36})$$

$c$  est la constante de diffusivité,  $\nabla$  l'opérateur gradient et  $\text{div}$  est l'opérateur divergence. La solution de l'équation II.36 est donnée par :

$$u = G_{\sqrt{2t}} * f \quad (\text{II.37})$$

où  $G_\sigma$  est le noyau gaussien de dimension 2 :

$$G_\sigma(x) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{|x|^2}{2\sigma^2}\right) \quad (\text{II.38})$$

Cela correspond à un filtrage passe-bas.

Nous discrétisons les différentes dérivées par différences finies selon la section II.2.3, le schéma numérique est donné par la formule suivante :

$$u_{i,j}^{n+1} = u_{i,j}^n + \delta t \cdot c \left( u_{i,j+1}^n + u_{i+1,j}^n + u_{i-1,j}^n + u_{i,j-1}^n - 4u_{i,j}^n \right) \quad (\text{II.39})$$

Le déroulement de l'algorithme de l'équation de la chaleur II.1 est montré comme suit :

---

### Algorithme II.1 Diffusion isotrope : Equation de la chaleur

---

**1: Initialisation des paramètres**

- Donner la valeur du pas temporel  $0 \leq \delta t \leq \frac{1}{4}$
- Fixer la valeur du coefficient de diffusion  $c$
- Fixer le nombre d'itérations  $N$

**2: Calcul des différences avec les plus proches voisins dans les quatre directions Nord Sud Est Ouest NEWS**

- $u_N = u_{i,j+1} - u_{i,j}$
- $u_E = u_{i+1,j} - u_{i,j}$
- $u_W = u_{i-1,j} - u_{i,j}$
- $u_S = u_{i,j-1} - u_{i,j}$

**3: Détermination du nouveau pixel**

$$u_{i,j}^{n+1} = u_{i,j}^n + \delta t \cdot c (u_N^n + u_E^n + u_W^n + u_S^n)$$


---

Le lissage par un opérateur Gaussien permet de réduire le bruit dans les images de faible complexité, il opère de manière identique dans toutes les directions et ne préserve pas les discontinuités ni le contraste des images. La figure II.8 est un exemple de débruitage de l'image 'Clock.tiff' par l'EDP de la chaleur avec le nombre d'itérations fixé à 2, 20 et 100 :

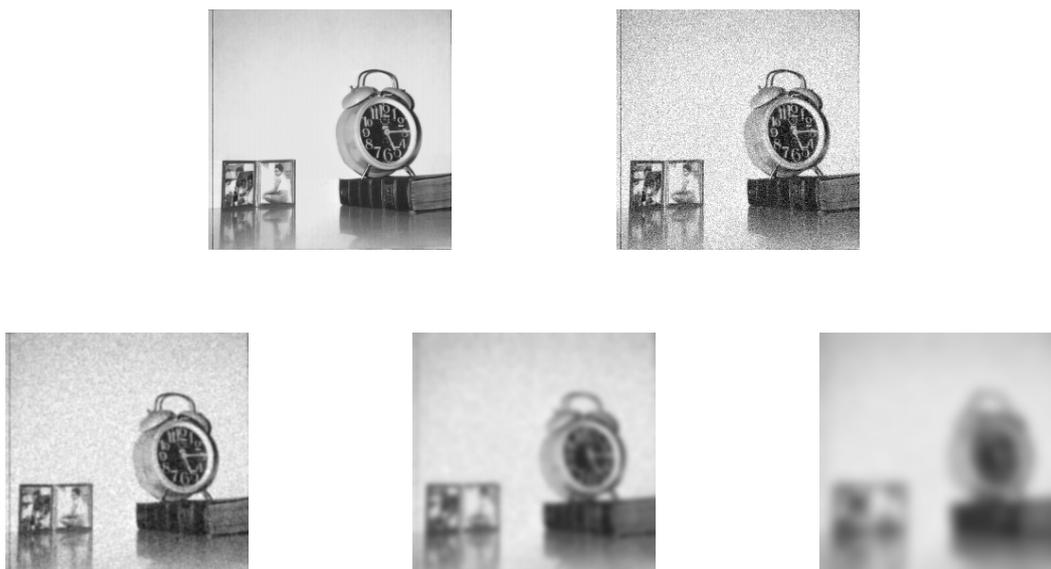


FIGURE II.8 – Restauration par équation de la chaleur de l'image 'Clock.tiff' : de gauche à droite : originale, dégradée par un bruit additif gaussien, restaurée pour :  $t = 2$ ,  $t = 20$ ,  $t = 100$

La restauration par l'EDP de la chaleur ne prend pas en compte la structure de l'image et engendre une disparition des contours, cette méthode rend l'image de plus en plus floue. Les zones homogènes sont bien restaurées par-contre les bords sont érodés.

Le problème d'un filtrage par diffusion isotropique est le lissage non selectif de toute l'image. Bien que le lieu des contours soit conservé, ces derniers deviennent flous. Pour éviter d'effacer progressivement les contours, les auteurs dans [Perona and Malik, 1990] ont proposé de modifier l'équation de diffusion pour faire de la diffusion anisotrope sur les principes suivants :

- Diffusion et donc homogénéisation maximale loin des contours.
- Diffusion minimale au niveau des contours.

## II.7.2 Diffusion anisotrope

Afin de résoudre les problèmes du lissage, la solution dans l'utilisation de la diffusion anisotrope [Perona and Malik, 1990] , [Perona et al., 1994], [Niessen et al., 1994] qui s'inspire de la diffusion isotrope. Il s'agit d'introduire dans l'équation une fonction de diffusion dépendant de la norme du gradient de l'image choisie de manière à mieux préserver le contraste de l'image :

$$\frac{\partial u}{\partial t} = \text{div} (g(|\nabla u|) \cdot \nabla u) = \nabla g \cdot \nabla u + g \Delta u \quad (\text{II.40})$$

$g$  est la fonction de diffusion ou fonction de conductance,  $\Delta u$  est l'opérateur laplacien et  $|\nabla u|$  est la norme du gradient de l'image  $u$ .

Deux formes classiques proposées par Pérona Malik :

1. fonction lorentzienne :

$$g(|\nabla u|) = \frac{1}{1 + \frac{|\nabla u|^2}{k^2}} \quad (\text{II.41})$$

2. ou bien fonction gaussienne :

$$g(|\nabla u|) = e^{-\frac{|\nabla u|^2}{k^2}} \quad (\text{II.42})$$

$g$  est une fonction décroissante. Si  $g(\cdot)$  est égale à une constante, on revient au processus de diffusion isotrope.  $k$  est appelé barrière ou seuil de diffusion, est un paramètre permettant de sélectionner l'amplitude seuil sur les gradients. , il sert à fixer la limite entre les forts gradients correspondant aux transitions à maintenir, et les faibles gradients correspondant au bruit. La diffusion anisotrope présente un lissage conditionnel.

La diffusivité réduit le lissage dans les zones à fort gradient. L'effet est un lissage non-uniforme, ralenti sur les contours donc une meilleure préservation des contours, mais les zones d'intensité croissante se retrouvent 'en escalier', la diffusion a tendance à créer

une image constante par morceau. Le choix de  $k$  n'est pas toujours facile, pas de méthode universelle. Pour trouver des informations sur le choix de  $k$  et  $t$ , on peut consulter respectivement les références [Whitaker, 1993], [Li and Chen, 1994],[Weickert, 1999a], [Weickert, 1999b] [Gilboa et al., 2001]

Une boucle de l'algorithme de diffusion anisotropique II.2 peut être décomposée comme suit :

---

### Algorithme II.2 Diffusion anisotropique de Pérona-Malik

---

**1: Initialisation des paramètres**

- Donner la valeur du pas temporel  $0 \leq \delta t \leq \frac{1}{4}$
- Fixer le nombre d'itérations  $N$

**2: Calcul des différences avec les plus proches voisins dans les quatre directions Nord Sud Est Ouest NEWS**

- $u_N = u_{i,j+1} - u_{i,j}$
- $u_E = u_{i+1,j} - u_{i,j}$
- $u_W = u_{i-1,j} - u_{i,j}$
- $u_S = u_{i,j-1} - u_{i,j}$

**3: Calcul des coefficients de diffusion dans les différentes directions du gradient**

- $g_N = \frac{1}{1+(\frac{u_N}{k})^2}$
- $g_E = \frac{1}{1+(\frac{u_E}{k})^2}$
- $g_W = \frac{1}{1+(\frac{u_W}{k})^2}$
- $g_S = \frac{1}{1+(\frac{u_S}{k})^2}$

**4: Modification de la valeur du pixel**

$$u_{i,j}^{n+1} = u_{i,j}^n + \delta t (g_N u_N^n + g_E u_E^n + g_W u_W^n + g_S u_S^n)$$


---

La figure II.9 montre le débruitage par Pérona-Malik pour un nombre d'itérations égal à 10, 20, 30, 50, 100, 200 et 400 :

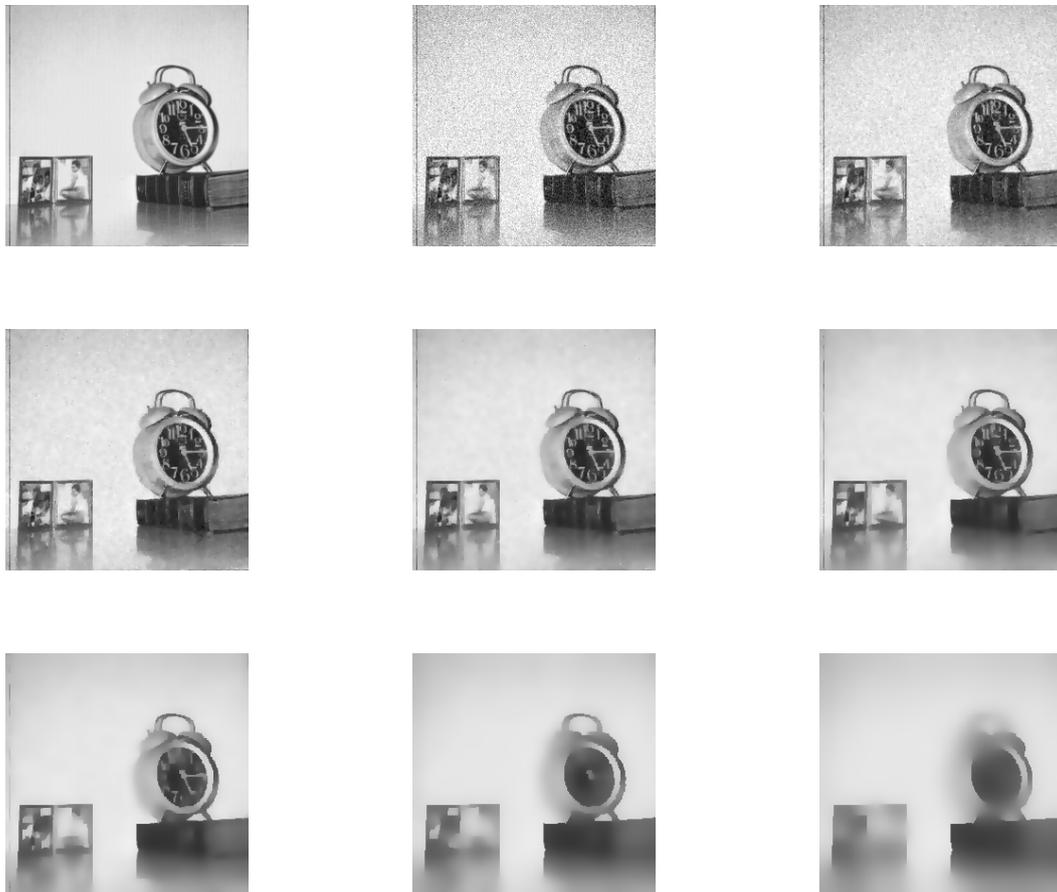


FIGURE II.9 – Restauration par Pérona-Malik de l'image Clock.tiff' : de gauche à droite : originale, dégradée par un bruit additif gaussien, restaurée pour :  $t = 10$  ,  $t = 20$ ,  $t = 30$ ,  $t = 50$ ,  $t = 100$  ,  $t = 200$ ,  $t = 400$

Des améliorations ont été proposées dans les sous sections II.7.3 et II.7.4 pour rendre le problème bien posé.

### II.7.3 Régularisation du gradient avec une Gaussienne

Dans l'article [Catté et al., 1992], les auteurs démontrent le caractère mal posé de la diffusion anisotrope de Perona et Malik [Perona and Malik, 1990] et proposent une version régularisée tout en donnant des résultats de stabilité en présence de bruit ainsi que des preuves d'existence et d'unicité de la solution.

L'équation continue proposée par Perona Malik est mal posée, le bruit est caractérisé par de forts gradients et risque de ne pas être lissé. La solution est de régulariser le gradient avec une gaussienne [Catté et al., 1992], [Nitzberg and Shiota, 1992]. L'équation

II.43 utilise ainsi une version lissée de l'image pour calculer les gradients à injecter dans la fonction de diffusion.

$$\frac{\partial u}{\partial t} = \text{div}(g(|\nabla(G_\sigma * u)|) \cdot \nabla u) \quad (\text{II.43})$$

où  $*$  est l'opérateur de convolution. l'opération  $G_\sigma * u$  consiste à convoluer  $u$  avec une gaussienne d'écart type  $\sigma$ .

Le terme  $g(|\nabla(G_\sigma * u)|)$  est un terme d'accentuation des bords :

- Si  $|\nabla u|$  est petit au voisinage d'un point  $x$ , alors ce point est considéré comme appartenant à une zone homogène, et la diffusion est donc forte.
- Si  $|\nabla u|$  est grand au voisinage d'un point  $x$ , alors ce point est considéré comme appartenant à un bord, et la diffusion est faible.

Observons l'effet de l'approche de Catté sur l'image 'Clock.tif' bruitée (figure II.10) :

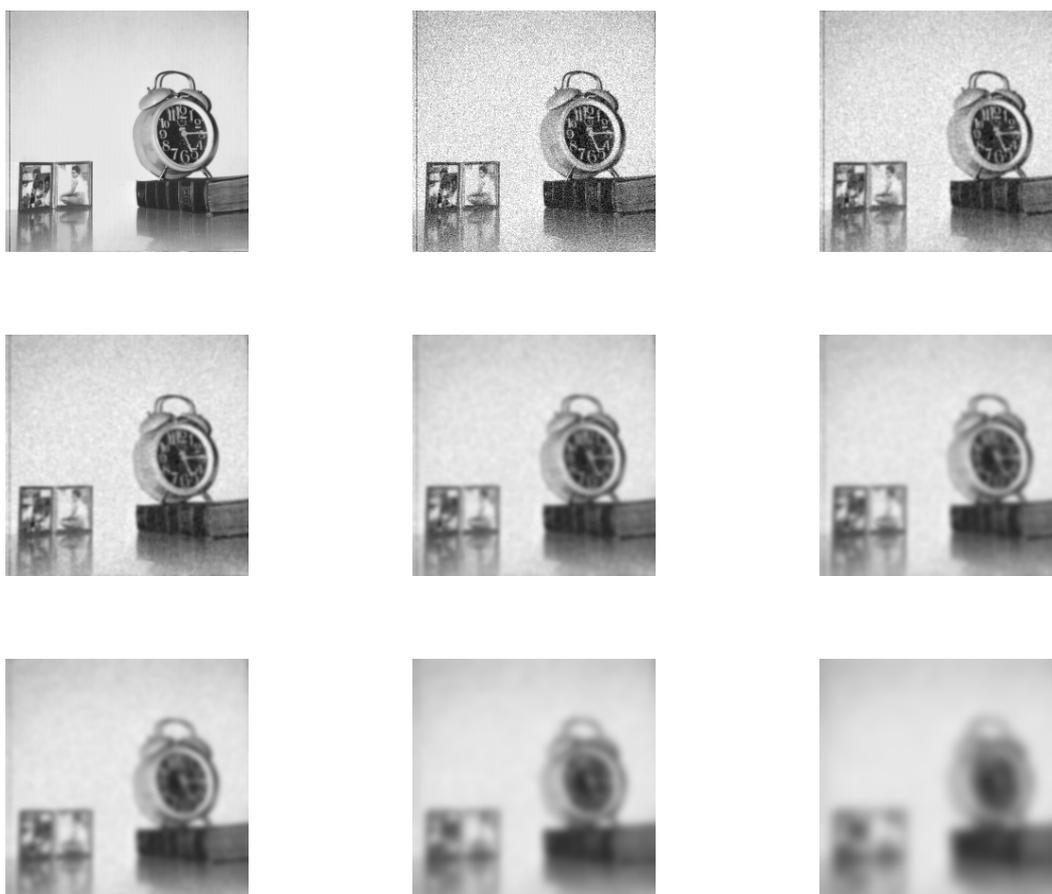


FIGURE II.10 – Restauration par Catté de l'image Clock.tif' : de gauche à droite : originale, dégradée par un bruit additif gaussien, restaurée pour :  $t = 2$  ,  $t = 4$  ,  $t = 10$  ,  $t = 15$  ,  $t = 20$  ,  $t = 50$  ,  $t = 100$

Le choix de  $\sigma$  influence directement sur la taille des détails à préserver. En effet, le lissage de l'image induit des gradients moins forts pour les petits détails. Les gradients correspondant à des structures de taille significative se trouveront donc plus souvent dans l'intervalle de rehaussement de la fonction  $g$ , elles seront ainsi conservées. Cette solution rend le problème bien posé, le gradient du bruit est réduit par la régularisation mais celui des contours ne l'est pas.

La méthode de Catté [Catté et al., 1992] est basée sur la convolution gaussienne, ce type de processus conduit à des instabilités, et laisse une part d'effet de bruit. L'inefficacité de la méthode dans les zones où le bruit présente de grosses discontinuités et même l'absence de solutions dans certains cas.

Afin de prendre en compte l'orientation du gradient et pas seulement la norme, Weickert dans [Weickert, 1998] propose un schéma capable de prendre en compte ces orientations.

#### II.7.4 Diffusion tensorielle

Pour effectuer un lissage dans les zones homogènes et le long des contours avec un rehaussement, Weickert [Weickert, 1999a] a proposé la diffusion pour le rehaussement de contours (Edge Enhancing Diffusion 'EED'). Le modèle de diffusion anisotrope tensorielle qui est une évolution du modèle classique et qui s'appuie sur une matrice de diffusion dont les directions de diffusion sont fixées par les vecteurs propres de cette matrice est donné par :

$$\frac{\partial u(x, y; t)}{\partial t} = \text{div}(D \cdot \nabla u(x, y; t)) \quad (\text{II.44})$$

La figure II.11 nous montre l'illustration de ce type de diffusion sur l'image 'Clock.tiff' : L'EED produit un arrondissement plus fort des coins anguleux dans ces zones à forte courbure, les pré-lissages isotropes avec des noyaux gaussiens produisent de fausses directions de diffusion.

#### II.7.5 Filtrage d'image par méthode de Yu et Acton

Yu et Acton [Yu and Acton, 2002] modifient le schéma classique de Pérona-Malik en introduisant le coefficient de variation instantanée  $q$ .

$$\begin{cases} \frac{\partial u(x, y; t)}{\partial t} = \text{div}[c(x, y; t) \cdot \nabla u(x, y; t)] \\ u(x, y; t) = u_0(x, y), \left. \frac{\partial u(x, y; t)}{\partial \vec{n}} \right|_{\partial \Omega} = 0 \end{cases} \quad (\text{II.45})$$

$c$  est le coefficient de variation défini comme suit :

$$c(x, y; t) = \frac{1}{1 + \frac{q^2(x, y; t) - q_0^2(t)}{q_0^2(t)(1 + q_0^2(t))}} \quad (\text{II.46})$$

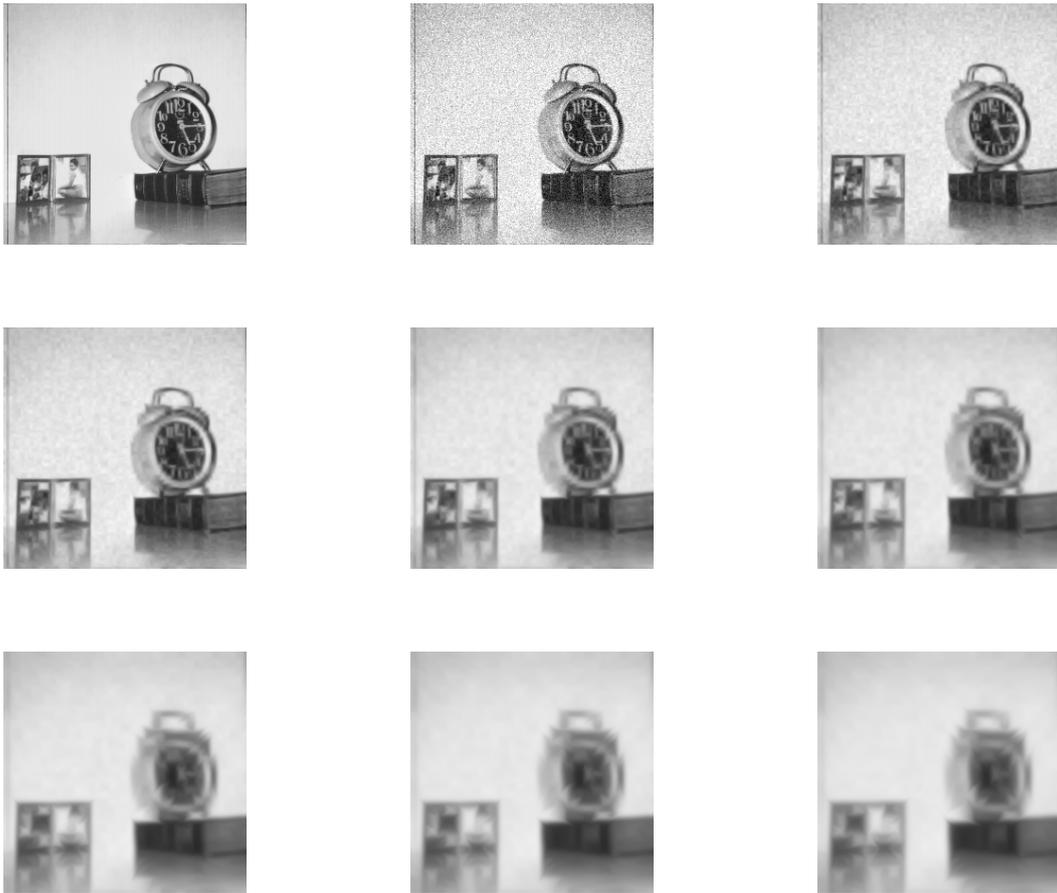


FIGURE II.11 – Restauration par EED de l’image Clock.tiff : de gauche à droite : originale, dégradée par un bruit additif gaussien, restaurée pour :  $t = 5$  ,  $t = 10$ ,  $t = 20$ ,  $t = 30$ ,  $t = 60$  ,  $t = 80$ ,  $t = 100$

où  $q(x, y; t)$  est le coefficient de variation instantanée, et est défini ainsi :

$$q(x, y; t) = \sqrt{\frac{\frac{1}{2} \left( \left( \frac{|\nabla u(x,y;t)|}{u(x,y;t)} \right)^2 - \left( \frac{1}{4^2} \right) \left( \frac{\nabla^2 u(x,y;t)}{u(x,y;t)} \right)^2 \right)}{\left( 1 + \frac{1}{4} \left( \frac{\nabla^2 u(x,y;t)}{u(x,y;t)} \right) \right)^2}} \quad (\text{II.47})$$

avec  $q_0(t)$  une fonction estimée comme suit :

$$q_0 = \frac{\sqrt{\text{var}(z(t))}}{\bar{z}(t)} \quad (\text{II.48})$$

où  $\text{var}(z(t))$  et  $\bar{z}(t)$  sont respectivement la variance et la moyenne à l’intérieur d’une zone homogène à l’instant  $t$ .

Le coefficient de variation instantanée est souvent utilisé comme détecteur de contours dans un contexte d’images bruitées. Dans les zones homogènes ( $q(x, y; t) \simeq q_0(t)$ ), nous

retombons sur un schéma de diffusion isotrope ( $c(x, y; t) \simeq 1$ ).

1.  $\lim_{q \rightarrow 0} c(x, y; t) = \frac{1+q_0^2(t)}{q_0^2(t)}$
2.  $\lim_{q \rightarrow +\infty} c(x, y; t) = 0$
3.  $\lim_{q(x,y;t) \rightarrow q_0(t)} c(x, y; t) = 1$

La diffusion de Yu et Acton est appliquée par un algorithme analogue [Diffellah et al., 2012] selon la fonction de diffusion  $c$  :

---

### Algorithme II.3 Diffusion de Yu et Acton SRAD

---

**1: Initialisation des paramètres**

- Déclaration de l'image bruitée  $u(x, y; t)$
- Fixer la valeur du pas temporel  $\delta t$
- Fixer le nombre d'itération  $N$

**2: Calcul du gradient dans les différentes directions**

- $U_N = u_{i,j+1}^n - u_{i,j}^n$
- $U_E = u_{i+1,j}^n - u_{i,j}^n$
- $U_W = u_{i,j-1}^n - u_{i,j}^n$
- $U_S = u_{i-1,j}^n - u_{i,j}^n$

**3: Calcul du coefficient de variation instantannée (zone homogène)**

$$q_0 = \frac{\sqrt{\text{var } z}}{\bar{z}(t)}$$

**4: Calcul du coefficient de variation  $q_{i,j}^n$**

- Calcul du module de Laplacien  $|\nabla u_{i,j}^n|$

$$Mod = \sqrt{U_N^2 + U_E^2 + U_W^2 + U_S^2}$$

- Calcul du Laplacien  $\nabla^2 u_{i,j}^n$

$$Lap = U_N + U_E + U_W + U_S$$

- Calcul du terme  $\frac{|\nabla u_{i,j}^n|}{u_{i,j}^n}$

$$Mod1 = \frac{Mod(\nabla u_{i,j}^n)}{u}$$

- Calcul du terme  $\frac{\nabla^2 u_{i,j}^n}{u_{i,j}^n}$

$$lap1 = \frac{Lap(u_{i,j}^n)}{u}$$

$$Q = \sqrt{\frac{\frac{1}{2}(Mod1^2 - (\frac{1}{4^2})Lap1^2)}{(1 + \frac{1}{4}Lap1)^2}}$$

**5: Calcul des coefficients de variation  $C_{i,j}^n$**

- $c = \frac{Q^2 - q_0^2}{q_0^2(1 + q_0^2)}$

- $C_{i,j}^n = \frac{1}{1+c}$

**6: Modification de la valeur de pixel**

$$u_{i,j}^{n+1} = u_{i,j}^n + \frac{1}{4}\delta t (C_N \cdot U_N + C_E \cdot U_E + C_W \cdot U_W + C_S \cdot U_S)$$


---

La figure II.12 montre l'implémentation de l'algorithme de SRAD II.3 sur l'image 'clock' :

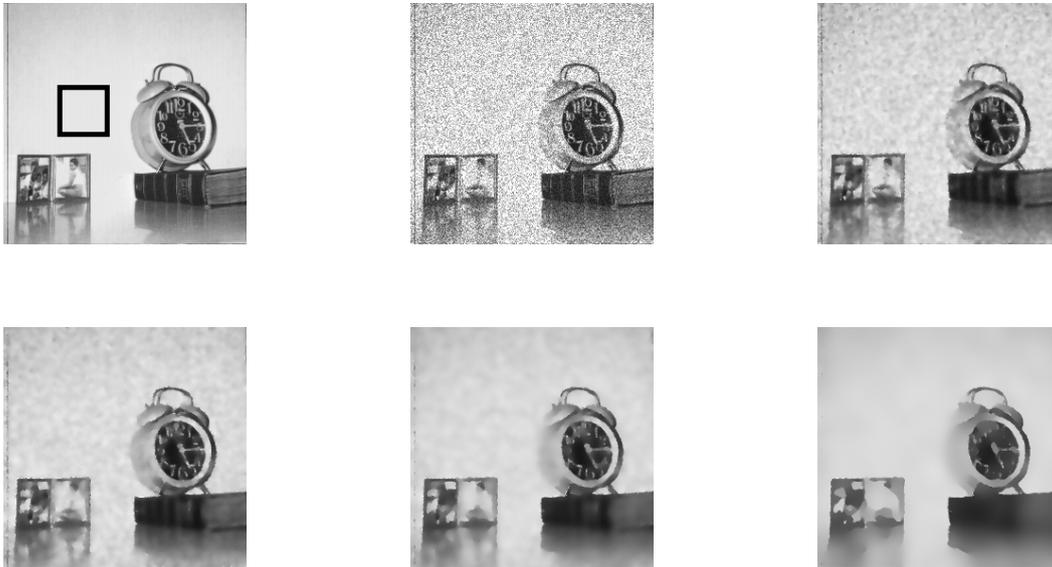


FIGURE II.12 – Restauration par SRAD de l'image 'Clock.tiff' : de gauche à droite : originale avec la zone homogène, bruitée par speckle, restaurée pour :  $t = 5$ ,  $t = 10$ ,  $t = 20$ ,  $t = 100$

## II.8 Filtrage par approche variationnelle

L'approche variationnelle permettant de réécrire les *EDPs* les plus classiquement utilisées sous la forme d'une minimisation d'une fonctionnelle appelée énergie. Cette énergie est décomposée en un terme d'attache aux données lié aux bruit et un terme de régularisation [Deriche and Faugeras, 2006]<sup>6</sup>, [Charbonnier et al., 1997], [Aubert et al., 1999] :

$$E(u) = E_{\text{données}}(u) + \lambda E_{\text{régularisation}}(u) \quad (\text{II.49})$$

### II.8.1 Méthode des moindres carrés

Le problème qui se pose alors est de pouvoir reconstruire  $u$  à partir de l'image dégradée  $f$  :

$$f = Ku + \eta \quad (\text{II.50})$$

où  $f$  est l'image observée,  $K$  la matrice de dégradation,  $u$  l'image originale et  $\eta$  un bruit de mesure.

6. <https://hal.inria.fr/inria-00073993/document>

Le problème de restauration consiste à résoudre l'équation II.50, c'est-à-dire à estimer l'image originale  $u$  ou une image de bonne qualité proche de la réalité, connaissant les données  $f$ , le modèle de dégradation  $K$ , des informations a priori sur la solution recherchée et, éventuellement, la distribution statistique du bruit  $\eta$  :

$$\eta = f - Ku \quad (\text{II.51})$$

Nous sommes alors amené à chercher  $u$  comme solution du problème :

$$(P) : \min_u \|f - Ku\|^2 \quad (\text{II.52})$$

La résolution de l'équation II.52 est un problème inverse mal-posé lié au mauvais conditionnement de la matrice  $K$  et à la présence de bruit.

Rappelons que selon la définition d'Hadamard [Hadamard, 1923] :

**Définition II.7** *Un problème est bien posé lorsque ces trois conditions sont remplies :*

1. *il existe une solution au problème,*
2. *si une solution existe, alors elle est unique,*
3. *une faible perturbation des données entraîne une faible perturbation de la solution du problème non perturbé.*

*Si au moins une de ces conditions n'est pas vérifiée le problème est dit mal posé.*

On propose la méthode des moindres carrés pour résoudre le problème mal posé II.53 :

$$f - Ku = 0 \quad (\text{II.53})$$

La norme du bruit s'écrit comme suit :

$$\eta^2 = \|f - Ku\|^2 = (Ku - f)^T (Ku - f) \quad (\text{II.54})$$

$u$  minimise le bruit  $f$  si et seulement si :

$$\nabla (\eta^2) = 0 \quad (\text{II.55})$$

$$2K^T Ku - 2K^T f = 0 \quad (\text{II.56})$$

L'image estimée :

$$\hat{u} = (K^T K)^{-1} K^T f \quad (\text{II.57})$$

La non-continuité de la solution II.57 en fonction des données aboutit à une solution instable. Dans le cas général, l'opérateur  $K$  n'est pas nécessairement inversible, il est

souvent de grande taille et mal conditionné. Trouver une solution numérique nécessite l'introduction d'un terme de régularisation :

$$(P) : \min_{u \in \ell_2} \|f - Ku\|^2 + R(u) \quad (\text{II.58})$$

$R(u)$  est le terme de régularisation. Notons que  $R$  est fonction de  $\|\nabla u\|$ .

Dans l'expression II.58, le premier terme sert à rapprocher la solution de l'image observée et le deuxième terme est un terme de régularisation qui est différent selon la valeur donnée à  $R$ . Lorsque l'image est bruitée, la valeur  $\|\nabla u\|$  aura tendance à avoir des pics au point où il y a du bruit, ce terme sert ainsi à réduire ces pics. De plus, l'utilisation de ce terme permet de réaliser un processus anisotrope (qui dépend de la direction et de la valeur du gradient), ce qui est important pour permettre de garder les contours mais d'éliminer le bruit<sup>7</sup>

Les méthodes de régularisation basées sur l'introduction d'informations a priori sur la solution recherchée, conduisent à une solution stable.

## II.8.2 Le modèle de Tykhonov

Nous allons présenter une première méthode de restauration proposée par Tykhonov. Le problème II.58 s'appuie sur la définition d'un critère à optimiser, qui peut s'écrire sous la forme [Demoment, 1989] :

$$J(u) : J_1(u, f) + \lambda J_2(u) \quad (\text{II.59})$$

où  $J_1(u, f)$ <sup>8</sup> est une distance mesurant la fidélité de la solution aux données ( $J_1(u, f)$ , qui est en général une distance quadratique) :

$$J_1(u) : \|f - Ku\|^2 \quad (\text{II.60})$$

$J_2(u)$ <sup>9</sup> est une mesure de la fidélité de la solution a priori :

$$J_2(u) = \|Du\|^2 \quad (\text{II.61})$$

$\lambda$  est un paramètre de régularisation qui traduit l'importance que l'on donne à l'information a priori par rapport aux données. L'estimée de  $u$  selon Tykhonov s'exprime par :

$$\hat{u} = (K^T K + \lambda D^T D)^{-1} K^T f \quad (\text{II.62})$$

avec  $D$ , un opérateur utilisé pour traduire l'information a priori.

7. Humbert Florent, Techniques de débruitage d'images, 23 janvier 2008. <https://humbert-florent.developpez.com/algorithmique/traitement/debruitage/>

8. dit terme d'attache aux données

9. dit terme de fidélité

Dans la suite, nous ne considérerons que le débruitage des images, ce qui correspond au cas simplifié où  $K$  est l'opérateur identité.

Remplaçons d'abord les équations II.60 et II.61 dans l'équation II.59, nous obtenons la formule II.63 :

$$\min_u J(u) = \min_u \left\{ \frac{1}{2} \|u - f\|^2 + \lambda \frac{1}{2} \|\nabla u\|^2 \right\} \quad (\text{II.63})$$

Nous discrétisons l'équation II.59 par la méthode des différences finies pour obtenir la solution à implémenter. Chercher le  $\min_u J(u)$  revient à résoudre  $\nabla J(u) = 0$ .

Nous allons résoudre le problème II.63 avec un algorithme de descente de gradient, cet algorithme utilise le gradient de l'itération précédente :

$$\frac{du}{dt} = -\nabla J(u) \quad (\text{II.64})$$

Les approximations de la première dérivée conduisent à la relation II.65 :

$$\frac{u(x, y; t+1) - u(x, y; t)}{\Delta(t+1) - \Delta t} = -(u(x, y; t) - f(x, y; t) + \lambda \Delta u(x, y; t)) \quad (\text{II.65})$$

Maintenant, nous passons à la discrétisation par différences finies qui donne le résultat suivant :

$$\frac{u_{i,j}^{n+1} - u_{i,j}^n}{\delta t} = - \left( u_{i,j}^n - f_{i,j}^n + \lambda \left( u_{i-1,1}^n + u_{i+1,1}^n + u_{i,j+1}^n + u_{i,j-1}^n - 4u_{i,j}^n \right) \right) \quad (\text{II.66})$$

Finalement, l'équation à implémenter est donnée par :

$$u_{i,j}^{n+1} = u_{i,j}^n - \delta t \left( u_{i,j}^n - f_{i,j}^n + \lambda \left( u_{i-1,1}^n + u_{i+1,1}^n + u_{i,j+1}^n + u_{i,j-1}^n - 4u_{i,j}^n \right) \right) \quad (\text{II.67})$$

Ce modèle n'est pas vraiment adapté, la minimisation de la norme  $\ell_2$  du gradient entraînant que l'image reconstruite est beaucoup trop floutée.

Un test de l'équation II.67 sur l'image 'Clock.tiff'<sup>10</sup> donne les résultats de la figure II.13 :

Nous pouvons noter que la méthode de Tykhonov a tendance à rendre flou l'image et faire disparaître les détails.

### II.8.3 Modèle de Rudin, Osher et Fatemi (ROF)

Le modèle de Rudin, Osher et Fatemi (ROF) suppose qu'une image peut se décomposer en une partie oscillante, représentant la texture et le bruit, et une partie «cartoon». Le but de cette méthode est de trouver une image qui minimise la longueur des contours, ce qui tend à se débarrasser du bruit, tout en restant proche des données initiales.

10. <http://sipi.usc.edu/database/database.php?volume=misc&image=17#top>

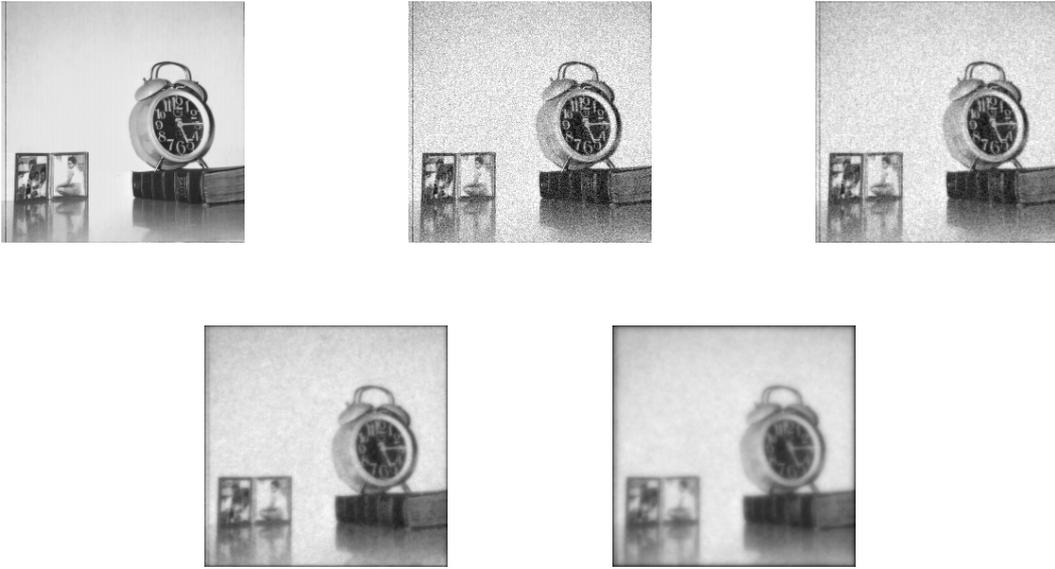


FIGURE II.13 – Restauration de l’image Clock.tiff’ par Tykhonov : de gauche à droite : originale, dégradée par un bruit additif gaussien, restaurée pour :  $\lambda = 0.3$ ,  $\lambda = 1$ ,  $\lambda = 5$

Cependant la minimisation de la longueur des contours engendre des perturbations numériques et la solution a tendance à être constante par morceaux. Les images débruitées forment des «plateaux» de niveaux de gris (ou de couleurs) qui sont observées en tant que signal 1D ressemblent à des marches d’escalier : c’est l’effet staircasing.

Nous allons chercher à débruiter l’image en minimisant la fonctionnelle suivante :

$$\min_u J(u) = \min_u \left\{ \frac{1}{2} \|u - f\|^2 + \lambda \frac{1}{2} \|\nabla u\| \right\} \quad (\text{II.68})$$

Le terme d’attache aux données et le terme de fidélité sont donnés respectivement par les equations II.69 et II.70 :

$$J_1(u) = \frac{1}{2} \|u - f\|^2 \quad (\text{II.69})$$

$$J_2(u) = \frac{1}{2} \|\nabla u\| \quad (\text{II.70})$$

La solution de la minimisation de  $J(u)$  de l’équation II.68 peut être calculée directement en résolvant  $\nabla J(u) = 0$ , et pour minimiser cette fonctionnelle nous utiliserons une méthode de descente de gradient à pas fixe. Nous procédons comme suit :

$$\frac{du}{dt} = -\nabla J(u) \quad (\text{II.71})$$

telle que :

$$\frac{du}{dt} = - \left( \frac{1}{2} \|u - f\| + \text{div} \frac{\Delta u}{\|\nabla u\|} \right) \quad (\text{II.72})$$

La discrétisation par différence finie a donné le résultat suivant :

$$\operatorname{div} \left( \frac{\nabla u}{\|\nabla u\|} \right) \xrightarrow{\text{discretization}} D_{i,j}^n = \frac{1}{h^2} \left[ \frac{u_{i+1,j}^n - u_{i,j}^n}{c_{1,i,j}^n} - \frac{u_{i,j}^n - u_{i-1,j}^n}{c_{2,i,j}^n} + \frac{u_{i,j+1}^n - u_{i,j}^n}{c_{3,i,j}^n} - \frac{u_{i,j}^n - u_{i,j-1}^n}{c_{4,i,j}^n} \right] \quad (\text{II.73})$$

Les différents coefficients des dénominateurs sont donnés par les relations suivantes :

$$c_{1,i,j}^n = \sqrt{\varepsilon^2 + \left( \frac{u_{i+1,j}^n - u_{i,j}^n}{h^2} \right)^2 + \left( \frac{u_{i,j+1}^n - u_{i,j-1}^n}{2h} \right)^2} \quad (\text{II.74})$$

$$c_{2,i,j}^n = \sqrt{\left( \frac{u_{i,j}^n - u_{i-1,j}^n}{h^2} \right)^2 + \left( \frac{u_{i-1,j+1}^n - u_{i-1,j-1}^n}{2h} \right)^2} \quad (\text{II.75})$$

$$c_{3,i,j}^n = \sqrt{\left( \frac{u_{i+1,j}^n - u_{i-1,j}^n}{2h} \right)^2 + \left( \frac{u_{i,j+1}^n - u_{i,j}^n}{h^2} \right)^2} \quad (\text{II.76})$$

$$c_{4,i,j}^n = \sqrt{\left( \frac{u_{i+1,j-1}^n - u_{i-1,j-1}^n}{2h} \right)^2 + \left( \frac{u_{i,j}^n - u_{i,j-1}^n}{h^2} \right)^2} \quad (\text{II.77})$$

Finalement, nous obtenons comme solution de minimisation du problème [II.68](#) l'équation suivante :

$$u_{i,j}^{n+1} = u_{i,j}^n - \tau (u_{i,j}^n - f_{i,j}^n + \lambda D_{i,j}^n) \quad (\text{II.78})$$

**Remarque II.8.1** *En fait, le terme TV n'étant pas différentiable, on le remplace dans la pratique par une autre approximation de la variation totale :*

$$TV_\varepsilon(u) = \sqrt{\varepsilon^2 + TV^2(u)} \quad (\text{II.79})$$

où  $\varepsilon$  est un réel strictement positif.

Le test de l'équation trouvée en [II.78](#) sur l'image 'Clock.tiff' montre les résultats suivants : ( figure [II.14](#) ) :

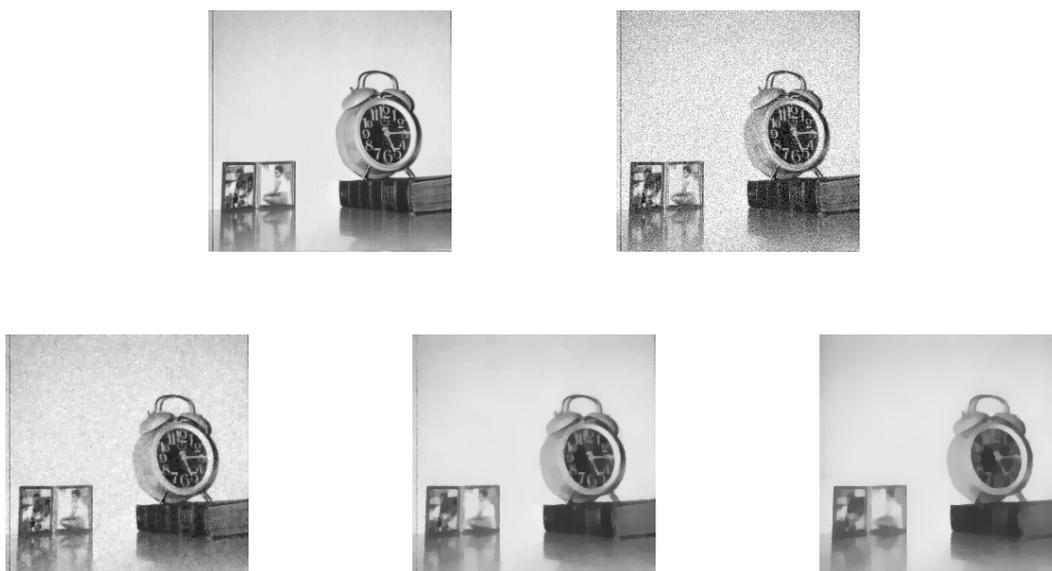


FIGURE II.14 – Restauration par TV de l'image Clock.tiff' : de gauche à droite : originale, dégradée par un bruit additif gaussien, restaurée pour :  $\rho = 10$ ,  $\rho = 30$ ,  $\rho = 50$

L'image semble meilleure qu'avec la méthode de la variation totale. Une approche possible pour s'attaquer à la minimisation du problème de ROF consiste à utiliser la méthode dite Implicite-Explicite ou encore Forward-Backward splitting, Il s'agit en fait de minimiser  $(\varphi + \psi)$  où  $\varphi$  et  $\psi$  sont toutes les deux des fonctions convexes avec certaines propriétés de régularité. Généralement dans le domaine du traitement du signal, étant donné un signal  $u$ ,  $\varphi(u)$  est le terme de fidélité aux données et est égal à  $\frac{1}{2}\|Ku - f\|^2$  où  $f$  est un signal bruité qui a également subi une perturbation linéaire  $K$ . Le second terme  $\psi$  reflète une connaissance a priori sur le bruit par exemple. Lorsque  $\psi = VT$  comme dans le problème de ROF, il est en général difficile de calculer (ou d'approcher) le minimiseur à savoir l'opérateur proximal du signal perturbé  $prox_{TV}f$  [Moreau, 1965], [Combettes and Wajs, 2005],[Jalalzai and Chambolle, 2009].

## II.9 Conclusion

Dans ce chapitre, nous avons présenté les méthodes de filtrage classiques telles que : le filtre moyennneur, le filtre median, les filtres adaptatifs basés sur la minimisation de l'erreur quadratique moyenne comme le filtre de Lee, le filtre de Frost, le filtre de Kuan et les filtres de Lee et Frost modifiés, puis nous avons étudié le filtrage par diffusion basé sur les équations aux dérivées partielles en partant de la diffusion isotrope linéaire suivi par le modèle non linéaire anisotrope de Pérona Malik. Aussi, nous avons montré l'efficacité de deux approches la régularisation du gradient avec une gaussienne et la diffusion tensorielle, et nous avons également terminé cette partie de diffusion par le modèle de Yu et Acton. A

la fin de ce chapitre, le filtrage par l'approche variationnelle a été développé, nous avons étudié la méthode des moindres carrés , le modèle de Tykhonov et le modèle de Rudin Osher et Fatemi.

Au chapitre 3, nous nous intéressons à l'application du proximal au problème de débruitage d'images.

## CHAPITRE III

# MÉTHODES PROXIMALES POUR LA RESTAURATION D'IMAGES

Une mer calme n'a jamais fait un bon marin

---

### III.1 Introduction

La résolution de problèmes inverses est présente dans la restauration d'images ; elle passe généralement par la minimisation d'un critère. De récentes méthodes d'optimisation convexe, nommées approches proximales, permettent de minimiser des critères non différentiables.

Les approches proximales possèdent des propriétés théoriques de convergence qui assurent la fiabilité de la solution.

### III.2 Rappels de topologie dans $\mathbb{R}^n$

Cette section est consacrée à un rappel rapide de quelques notions de topologie dans  $\mathbb{R}^n$  pour faciliter la lecture de certains passages.

#### III.2.1 Boule, ouvert, fermé

**Définition III.1 (Boule ouverte)** Soient  $x \in \mathbb{R}^n$  et  $r > 0$ . On appelle boule ouverte de centre  $x$  et de rayon  $r$  l'ensemble :

$$B(x, r) = \{y \in \mathbb{R}^n \text{ tel que } \|y - x\| < r\} \quad (\text{III.1})$$

**Définition III.2 (Ouvert, fermé)**

- Un ensemble  $O$  de  $\mathbb{R}^n$  est dit ouvert si pour tout  $x \in O$ , il existe une boule ouverte  $B(x, r)$  de centre  $x$  et de rayon  $r$  incluse dans  $O$ .
- Un ensemble  $F$  de  $\mathbb{R}^n$  est dit fermé si son complémentaire est un ouvert.

**Remarque III.2.1 (Interprétation géométrique)** L'interprétation géométrique de 'ouvert' et 'fermé' est qu' :

- un ensemble est ouvert si et seulement si il ne contient aucun point de sa frontière.
- Un ensemble est fermé si et seulement si il contient tous les points de sa frontière.
- Les ensembles qui contiennent une partie de leur frontière mais pas l'intégralité ne sont ni ouverts ni fermés.

### III.2.2 Minorant/Majorant, Infimum/Supremum, Minimum/-Maximum

Nous définissons les notions de minorant/majorant, infimum/supremum et minimum/-maximum comme suit :

**Définition III.3 (Minorant/Majorant)** Soit  $E$  un sous-ensemble de  $\mathbb{R}$ ,

- $m \in \mathbb{R} \cup \{-\infty, +\infty\}$  est un minorant de  $E$  si et seulement si  $m$  est inférieur ou égal à tous les éléments de  $E$ . Ainsi

$$m \text{ minorant} \Leftrightarrow \forall x \in E \quad m \leq x \quad (\text{III.2})$$

- tandis que  $M \in \mathbb{R} \cup \{-\infty, +\infty\}$  est un majorant de  $E$  si et seulement si il est supérieur ou égal à tous les éléments de  $E$ . Ainsi

$$M \text{ majorant} \Leftrightarrow \forall x \in E \quad M \geq x \quad (\text{III.3})$$

**Remarque III.2.2** Si  $E$  admet un minorant (respectivement majorant) fini alors il est dit minoré (respectivement majoré).

**Définition III.4 (Infimum/Supremum)**  $E \subset \mathbb{R}$ .

- L'infimum :  $\inf(E) \in \mathbb{R} \cup \{-\infty, +\infty\}$  de  $E$  est le plus grand des minorants. On le note :  $\inf(E) = \inf_{x \in E} (x)$
- Le supremum :  $\sup(E) \in \mathbb{R} \cup \{-\infty, +\infty\}$  de  $E$  est le plus petit des majorants. On le note :  $\sup(E) = \sup_{x \in E} (x)$

**Définition III.5 (Minimum, maximum)** Soit  $E \subset \mathbb{R}$ .

- L'infimum de  $E$  est appelé minimum si et seulement si  $\inf(E) \in E$ , noté  $\min(E)$ .

- Le supremum de  $E$  est appelé maximum si et seulement si  $\sup(E) \in E$  : Dans ce cas, noté  $\max(E)$ .

**Application III.1** Si on se donne  $E = ]0, 1]$ , alors l'ensemble des minorants de  $E$  est  $[-\infty, 0]$  et l'ensemble des majorants est  $[1, +\infty]$ . On en déduit que le l'infimum vaut 0 et le supremum vaut 1. Comme 1 appartient à  $E$  alors c'est aussi le maximum de  $E$ . Cependant 0 n'appartient pas à  $E$  et donc  $E$  n'a pas de minimum. On dit souvent que l'infimum de  $E$  n'est pas atteint.

**Proposition III.2.1**  $F \subset \mathbb{R}$  un fermé non vide de  $\mathbb{R}$ .

- Si  $F$  est minoré alors  $F$  admet un minimum et ainsi  $\inf(F) = \min(F) \in \mathbb{R}$ .
- Si  $F$  est majoré alors  $F$  admet un maximum et ainsi  $\sup(F) = \max(F) \in \mathbb{R}$ .

### III.3 Quelques rappels d'analyse convexe

Cette section est consacrée à un rappel rapide d'analyse convexe pour faciliter la lecture de certains passages, Pour plus de détails, nous pourrions se référer aux ouvrages : [Rockafellar, 1970], [Laurent, 1972], [Ekeland and Temam, 1972],[Auslender, 1976], [Polyak, 1987], [Bertsekas, 1999], [Zalinescu, 2002], [Bonnans et al., 2003],[Nesterov, 2004], [Boyd and Vandenberghe, 2004], [Bertsekas, 2006], [Bertsekas, 2009]. La référence [Bauschke and Combettes, 2018] propose une très bonne introduction à l'analyse convexe dans des espaces de Hilbert. Pour découvrir les nombreux détails de l'analyse convexe en dimension finie nous pourrions accéder aux références [Urruty et al., 1993b], [Urruty et al., 1993a], [Rockafellar, 2015]. Quelques notes aussi sont très largement inspirées des cours des sites :

[https://www.ljll.math.upmc.fr/~cerf/Techniques\\_Optimisation\\_2018.pdf](https://www.ljll.math.upmc.fr/~cerf/Techniques_Optimisation_2018.pdf),

[https://www.math.univ-toulouse.fr/~rondep/CoursTD/poly4GMM\\_nondiff.pdf](https://www.math.univ-toulouse.fr/~rondep/CoursTD/poly4GMM_nondiff.pdf)

et [https://www.math.univ-toulouse.fr/~weiss/Docs/EcolePrintemps2011/Partie3\\_2.pdf](https://www.math.univ-toulouse.fr/~weiss/Docs/EcolePrintemps2011/Partie3_2.pdf).

Dans toute la suite  $H$  désigne l'espace euclidien  $\mathbb{R}^n$ , muni du produit scalaire usuel noté  $\langle \cdot, \cdot \rangle$  et de la norme euclidienne associée  $\|x\| = \langle x, x \rangle^{\frac{1}{2}}$ . On note par  $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$  muni d'une structure algébrique déduite de celle de  $\mathbb{R}$  avec la convention que  $-\infty - (+\infty) = +\infty$  [Rockafellar, 1970].

### III.3.1 Domaine, fonction propre, fonction indicatrice, épigraphe, lignes de niveau

**Définition III.6 (Domaine)** Soit  $f : H \rightarrow ]-\infty, +\infty]$ , le domaine de la fonction  $f$  est donné par :

$$\text{dom } f = \{x \in H \mid f(x) < +\infty\} \quad (\text{III.4})$$

**Application III.2** La figure III.4 illustre la représentation graphique de la fonction valeur absolue :

$$f(x) = |x| = \begin{cases} +x & \text{si } x \geq 0 \\ -x & \text{ailleurs} \end{cases}$$

et son domaine

$$\text{dom } f = \mathbb{R}$$

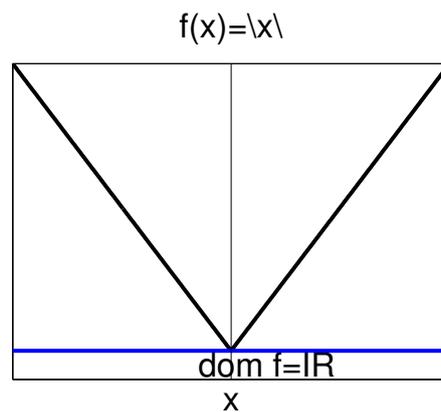


FIGURE III.1 – Fonction valeur absolue et son domaine

**Définition III.7 (Fonction indicatrice)** Soit  $E$  un sous-ensemble de  $H$ , la fonction indicatrice  $i_E : H \rightarrow ]-\infty, +\infty]$  de cet ensemble est définie par :

$$\forall x \in H \quad i_E(x) = \begin{cases} +\infty & \text{si } x \notin E \\ 0 & \text{si } x \in E \end{cases} \quad (\text{III.5})$$

**Application III.3** La figure III.2 illustre la représentation graphique de la fonction indicatrice :

$$f(x) = i_{[-\delta, \delta]}(x) = \begin{cases} 0 & \text{si } x \in [-\delta, \delta] \\ +\infty & \text{ailleurs} \end{cases}$$

et son domaine :

$$\text{dom } f = [-\delta, +\delta]$$

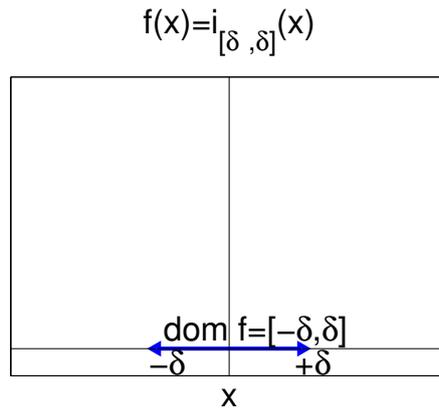


FIGURE III.2 – Fonction indicatrice et son domaine

**Définition III.8 (Infimale convolution)** Soit  $g : H \rightarrow ]-\infty, +\infty]$ . L'infimale convolution de  $h$  et  $g$  est définie par  $g : H \rightarrow ]-\infty, +\infty]$  :  $x \rightarrow \inf_{y \in H} (h(y) + g(x - y))$ . L'élément neutre de l'infimale convolution est la fonction indicatrice de l'ensemble  $\{0\}$ , où  $0$  représente le vecteur nul de  $H$  :

$$\forall x \in H, \quad h \square i_{\{0\}}(x) = h(x) \tag{III.6}$$

**Propriété III.1 (Fonction propre)** Si  $\text{dom } f = \emptyset$ , la fonction  $f$  est dite propre.

**Définition III.9 (Épigraphe)** L'épigraphe d'une fonction est l'ensemble des points situés au-dessus de la courbe représentative de la fonction  $f$ .

**Application III.4** La figure III.3 ci dessous illustre les épigraphes des fonctions valeur absolue et fonction indicatrice :

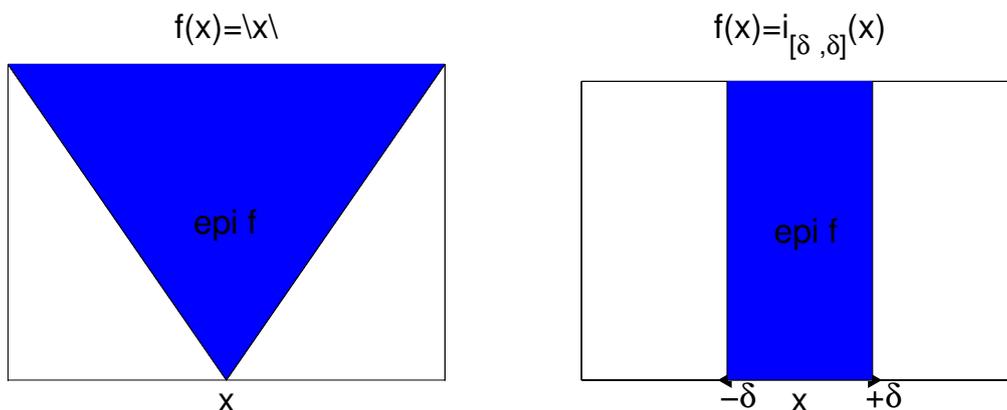


FIGURE III.3 – Epigraphe des fonctions valeur absolue et indicatrice

**Définition III.10 (Lignes de niveau ou section inférieure)** L'ensemble des lignes de niveau ou section inférieure de la fonction  $f : H \rightarrow ]-\infty, +\infty]$  à la hauteur  $\delta \in \mathbb{R}$  est

notée  $lev_{\leq \delta} f$  et se définit par :

$$lev_{\leq \delta} f = \{x \in H \mid h(x) \leq \delta\} \quad (\text{III.7})$$

**Application III.5** Une représentation graphique des définitions III.6 et III.10 est donnée dans la figure III.4 :

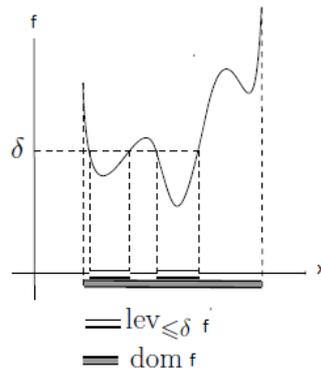


FIGURE III.4 – Illustration du domaine et l'ensemble des lignes de niveau  $lev_{\leq \delta} f$  de la fonction  $f : H \rightarrow ]-\infty, +\infty]$

### III.3.2 Fonction semi-continue inférieurement, fonction coercive

**Définition III.11 (Fonction semi-continue inférieurement)** Soit  $f : H \rightarrow ]-\infty, +\infty]$ , la fonction  $f$  est dite semi-continue inférieurement (s.c.i) sur  $H$  si, pour tout  $x \in H$ , pour toute suite  $(x_k)_{k \in \mathbb{N}}$  de  $H$ ,

$$x_k \rightarrow x \Rightarrow \liminf f(x_k) \geq f(x) \quad (\text{III.8})$$

**Remarque III.3.1** Toute fonction continue est s.c.i.

**Remarque III.3.2** La fonction indicatrice d'un ensemble convexe fermé est convexe et semi-continue inférieurement s.c.i.

**Définition III.12 (fonction coercive)** Soit  $f : H \rightarrow ]-\infty, +\infty]$ , la fonction  $f$  est dite coercive si :

$$\lim_{\|x\| \rightarrow \infty} f(x) = +\infty \quad (\text{III.9})$$

### III.3.3 Ensemble convexe, fonction convexe, fonction concave

**Définition III.13 (Ensemble convexe)** Un ensemble  $C$  est dit convexe lorsque, pour tous  $x$  et  $y$  de  $C$ , le segment  $[x, y]$  est tout entier contenu dans  $C$ , c'est-à-dire :

$$C \subseteq \mathbb{R}^n \Leftrightarrow \forall x, y \in C, \forall \lambda \in [0, 1], \lambda x + (1 - \lambda) y \in C \quad (\text{III.10})$$

**Application III.6** La figure III.5 montre un exemple d'application des fonctions convexes et non convexes :

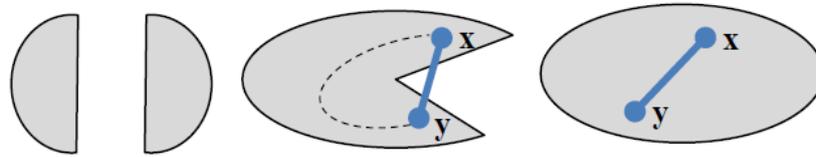


FIGURE III.5 – Exemple de convexité

*Interprétation géométrique de la figure III.5 : Si le segment  $[x, y]$  est inclus dans  $C$ , l'ensemble est dit convexe (3ème figure à droite), sinon l'ensemble est dit non-convexe (les deux premières figures).*

**Définition III.14 (Fonction convexe)** Une fonction  $f$  de  $\mathbb{R}^n$  dans  $\mathbb{R}$  est dite convexe lorsque, pour tous  $x$  et  $y$  de  $\mathbb{R}$  et tout  $\lambda$  dans  $[0, 1]$  on a :

$$\forall x, y \in \mathbb{R}^n, \forall \lambda \in [0, 1], \quad f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \quad (\text{III.11})$$

La figure III.6 montre un exemple illustratif d'une fonction convexe.

**Théorème III.3.1**  $f$  est convexe si et seulement si son épigraphe est convexe.

**Remarque III.3.3** La fonction indicatrice d'un ensemble convexe est convexe et s.c.i.

**Définition III.15 (Fonction strictement convexe)** Une fonction  $f$  de  $\mathbb{R}^n$  dans  $\mathbb{R}$  est dite convexe lorsque, pour tous  $x$  et  $y$  de  $\mathbb{R}^n$  et tout  $\lambda$  dans  $[0, 1]$  on a :

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y) \quad (\text{III.12})$$

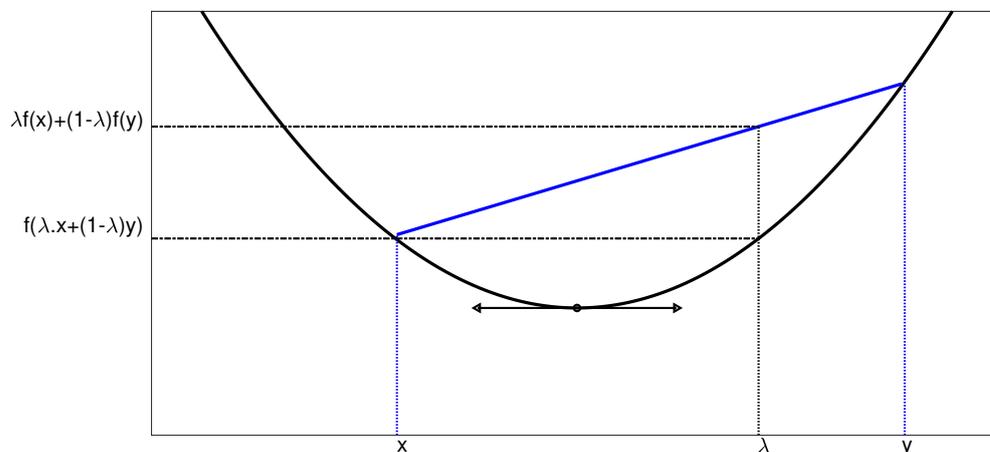


FIGURE III.6 – Exemple d'une fonction convexe

**Application III.7** Si la sécante de la figure III.6 est au dessus de la courbe, la concavité est vers le haut.

**Remarque III.3.4 (Fonction concave)** Une fonction est dite concave si sa fonction opposée est convexe. Si  $f$  est concave,  $-f$  est convexe.

**Définition III.16 (Fonction fortement convexe)** Soit  $(H, \|\cdot\|)$  un espace normé, on dit qu'une fonction  $f : H \rightarrow \mathbb{R} \cup \{+\infty\}$  est fortement convexe, de module  $a > 0$  si, pour tous  $x$  et  $y$  dans  $\text{dom } f$  et tout  $\lambda$  dans  $[0, 1]$ , on a :

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) - \frac{1}{2}a\lambda(1 - \lambda)\|x - y\|^2 \quad (\text{III.13})$$

**Remarque III.3.5** Dans l'équation III.13, on retrouve la notion de fonction convexe lorsque  $a = 0$ .

### III.3.4 Différentiabilité

**Définition III.17 (Dérivée directionnelle)** Soit  $x \in \text{dom}(f)$ , on appelle dérivée directionnelle au point  $x$  dans la direction  $p$  la limite suivante :

$$f'(x, p) = \lim_{\alpha \rightarrow 0^+} \frac{1}{\alpha} (f(x + \alpha p) - f(x)) \quad (\text{III.14})$$

Si cette limite existe, on dit que  $f'(x, p)$  est la dérivée directionnelle de  $f$  en  $x$ .

**Théorème III.3.2 (Fonction convexe différentiable)** Une fonction convexe est différentiable dans toutes les directions sur tous les points de l'intérieur de son domaine.

**Définition III.18 (Sous-gradient et sous-différentiel)** Soit  $f$  une fonction convexe. Un vecteur  $v \in \mathbb{R}^n$  est appelé sous-gradient de  $f$  au point  $x_0 \in \text{dom}(f)$  si :

$$\forall x \in \text{dom}(f), \quad f(x) \geq f(x_0) + \langle v, x - x_0 \rangle \quad (\text{III.15})$$

**Application III.8** Un exemple d'application est illustré dans la figure III.7 :

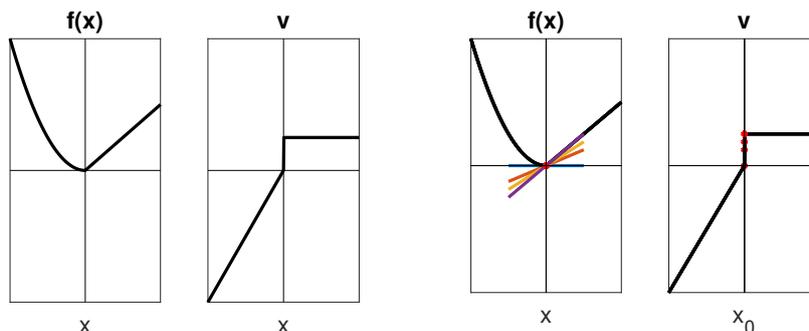


FIGURE III.7 – Exemple de sous-différentiel d'une fonction

**Remarque III.3.6** L'ensemble de tous les sous-gradients en  $x_0$  est appelé sous-différentiel de  $f$ . Il est noté  $\partial f(x_0)$ .

**Théorème III.3.3** Le sous-différentiel

$$\partial f(x_0) = \{v \in \mathbb{R}^n, \forall x \in \mathbb{R}^n, f(x) \geq f(x_0) + \langle v, x - x_0 \rangle\} \quad (\text{III.16})$$

est un ensemble convexe fermé.

**Application III.9** Soit  $f : x \in \mathbb{R} \rightarrow |x|$ . Calculons les sous-gradients de  $f$  en tout point de  $x$ . Le sous gradient différentiel est représenté par la figure III.8 :

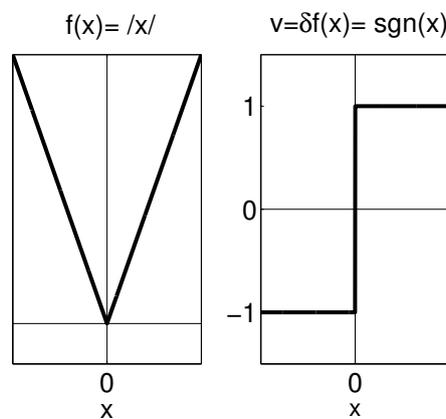


FIGURE III.8 – La fonction  $f(x) = |x|$  et son sous-différentiel  $\partial f(x) = \text{sgn}(x)$

Dans  $\mathbb{R}^2$ , les hyperplans d'appui sont des droites et les sous-gradients associés sont leurs coefficients directeurs :

$$\partial f(x) = \begin{cases} \{-1\} & \text{si } x < 0 \\ [-1, +1] & \text{si } x = 0 \\ \{+1\} & \text{si } x > 0 \end{cases}$$

**Théorème III.3.4 (Caractérisation différentielle de la convexité)** Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction différentiable. Les propositions suivantes sont équivalentes :

- $f$  est convexe.
- Ses hyperplans tangents sont des minorants (Voir figure III.9), avec  $v = \nabla f(x_0)$   
 $\forall (x_0, x) \in \mathbb{R}^n \times \mathbb{R}^n, f(x) \geq f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle$
- Le gradient de  $f$  est un opérateur monotone :  
 $\forall (x_0, x) \in \mathbb{R}^n \times \mathbb{R}^n, \langle \nabla f(x) - \nabla f(x_0), x - x_0 \rangle \geq 0$

**Application III.10** La figure III.9 est un exemple concret de la notion de sous-différentiabilité :

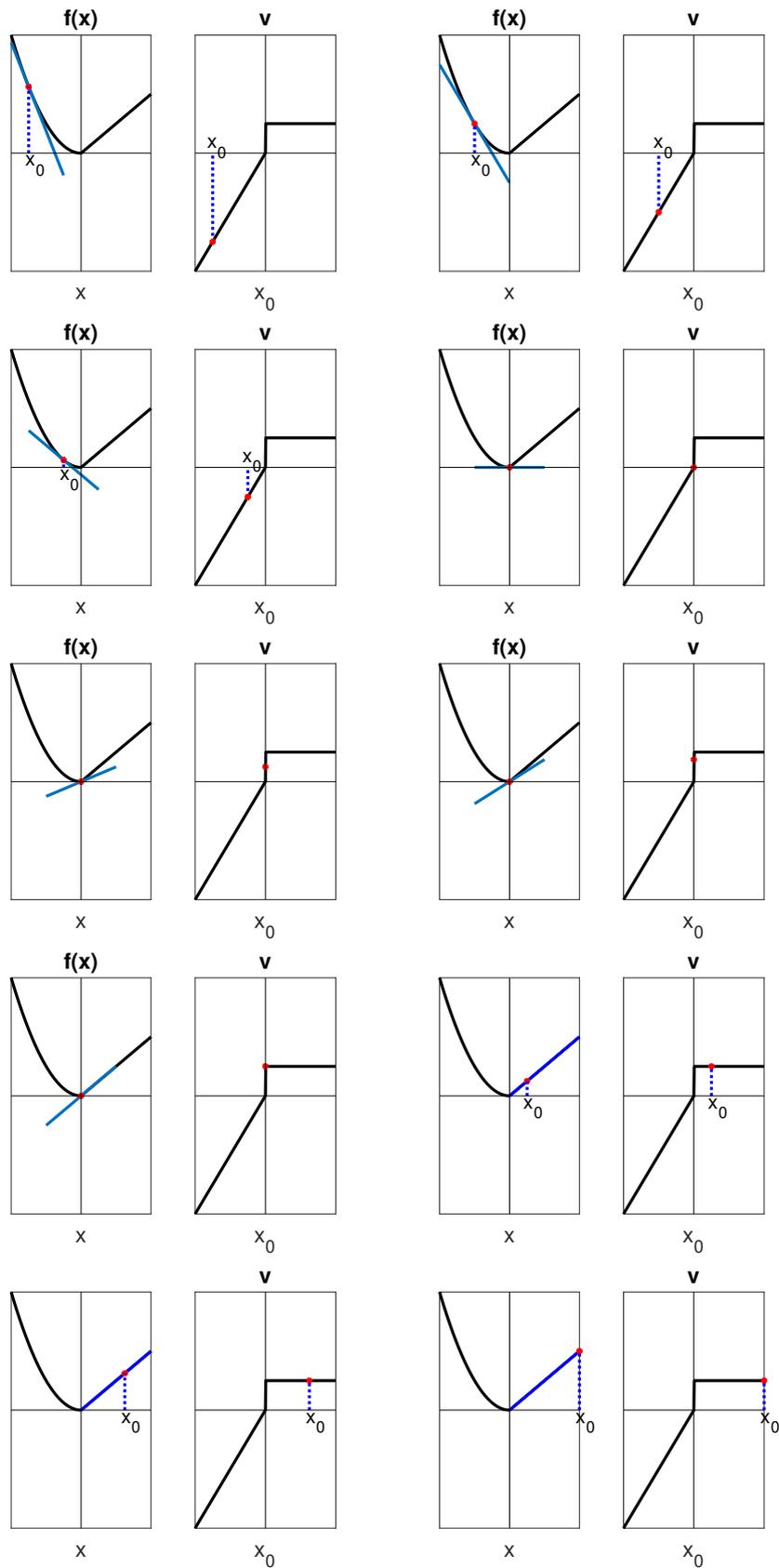


FIGURE III.9 – Exemple de sous-différentiel et hyperplans d'une fonction  $f(x) \geq f(x_0) + \langle v, x - x_0 \rangle$

Le sous-différentiel au point  $x_0$  a une interprétation géométrique simple. C'est l'ensemble des hyperplans qui passent par le point  $(x_0, f(x_0))$  et dont le graphe passe sous le graphe de  $f$ .  $\partial f(x_0)$  est réduit à un singleton si et seulement si  $f$  est dérivable en  $x$ , dans ce cas,  $\partial f(x) = \nabla f(x)$ .

**Proposition III.3.1** Si  $f : H \rightarrow ]-\infty, +\infty]$  est une fonction propre, alors nous avons :

$$x \in \text{Arg min } f \quad \Rightarrow \quad 0 \in \partial f(x) \quad (\text{III.17})$$

**Remarque III.3.7 (Gradient)** Lorsque  $f : H \rightarrow ]-\infty, +\infty]$  et  $f$  non convexe est différentiable au point  $x \in H$ , nous avons :

$$\partial f(x) = \nabla f(x) \quad (\text{III.18})$$

**Définition III.19 (Fonction de gradient  $\beta$ -Lipschitz)** Soit  $f : H \rightarrow ]-\infty, +\infty]$ , la fonction  $f$  est dite de gradient  $\beta$ -Lipschitz si elle est différentiable sur  $H$  et si son gradient vérifie :

$$\forall (x, y) \in H^2, \beta \in ]0, +\infty[ \quad \|\nabla f(x) - \nabla f(y)\| \leq \beta \|x - y\| \quad (\text{III.19})$$

Lorsqu'une fonction  $f : H \rightarrow ]-\infty, +\infty]$  est la somme d'une fonction différentiable et d'une fonction non différentiable, la règle de différentiation est alors donnée par la proposition III.3.2 :

**Proposition III.3.2 [Rockafellar et al., 1998]** Soit  $f : H \rightarrow ]-\infty, +\infty]$  une fonction propre telle que, pour tout  $x \in H$ ,  $f(x) = f_1(x) + f_2(x)$ , où  $f_1$  est une fonction propre différentiable sur  $\text{dom } f$  et  $f_2$  est une fonction propre. Nous avons alors :

$$\forall x \in \text{dom } f, \quad \partial f(x) = \nabla f_1(x) + \partial f_2(x)$$

**Définition III.20 (Opérateur monotone)** Un opérateur  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  est dit monotone si :

$$\forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^n, \quad \langle F(y) - F(x), y - x \rangle \geq 0 \quad (\text{III.20})$$

**Théorème III.3.5 (Minimiseurs d'une fonction)** Le sous-différentiel permet notamment de caractériser les minimiseurs d'une fonction. On considère le problème suivant :

(P) : trouver  $x^*$  tel que  $f(x^*) = \min_{x \in \mathbb{R}^n} f(x)$ , où  $f$  est convexe et s.c.i.  
 $x^*$  est solution du problème (P) si et seulement si  $0 \in \partial f(x^*)$

### III.3.5 Fonction affine, Enveloppe supérieure (fonction $\Gamma^\circ(H)$ )

**Définition III.21 (Fonction affine continue)** *Les fonctions affines continues sur  $H$  sont les fonctions de type  $x \rightarrow l(x) + a$*

*où  $l$  est une forme linéaire continue sur  $H$  et  $a$  est un réel.*

**Définition III.22 (Enveloppe supérieure)** *On note  $\Gamma^\circ(H)$  l'ensemble des fonctions  $f$  sur  $H$  ( $f \neq +\infty$ ), qui sont l'enveloppe supérieure d'une famille quelconque de fonctions affines continues.*

*Il découle de cette définition que  $\Gamma^\circ(H)$  est l'ensemble des fonctions propres convexes semi-continues inférieurement sur  $H$ .*

**Remarque III.3.8** *Pour  $f$  l'élément de  $\Gamma^\circ(H)$ , on remarque que  $\text{epi}(f)$  est un ensemble convexe fermé non vide dans  $H \times \mathbb{R}^n$ , la projection sur cet ensemble est donc bien définie.*

## III.4 Algorithmes d'optimisation

Il existe de multiples algorithmes dédiés à la résolution des problèmes d'optimisation [Rockafellar, 1970], [Nocedal and Wright, 2006]. Dans cette section, nous présentons un certain nombre d'algorithmes d'optimisation, et nous donnons leurs conditions d'application. Pour tous ces algorithmes, l'objectif est de minimiser la contrainte d'une fonction  $f(x)$  par rapport à la variable  $x \in \mathbb{R}^n$ .

Les algorithmes que nous présentons sont des algorithmes itératifs, la convergence de l'algorithme, est donc l'arrêt des itérations, est décidée par des conditions d'arrêt. Ces algorithmes peuvent être décomposés en deux grandes classes : les algorithmes basés sur le gradient de la fonction à optimiser et les algorithmes sans calcul de gradient.

Nous pourrions nous référer aux ouvrages : [Gill et al., 1982], [Lee, 2004],[Fletcher, 2009], [Snyman and Wilke, 2018] etc, si nous désirions plus d'informations sur l'optimisation.

### III.4.1 Conditions d'optimalité

**Théorème III.4.1 (Conditions nécessaires d'optimalité locale)** *Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  une application différentiable. Si  $x^* \in \mathbb{R}^n$  réalise un minimum local (respectivement maximum local) de  $f$ , alors (conditions nécessaires du premier ordre) :*

$$\nabla f(x^*) = 0 \tag{III.21}$$

*Si, de plus,  $f$  est deux fois différentiable dans un voisinage ouvert de  $x^*$ , alors (condition nécessaire d'optimalité du deuxième ordre) :  $H[f](x^*)$  semi définie positive (respectivement semi définie négative)*

La condition du premier ordre joue un rôle central en optimisation numérique : elle permet

de sélectionner un certain nombre de points à être des extrema locaux, appelés points critiques ou points stationnaires. Parmi eux, figurent des minima locaux, des maxima locaux et d'autres qui ne sont ni l'un, ni l'autre, appelés points selle.

**Définition III.23 (Points critiques)** *Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  une application différentiable. Tout point  $x \in \mathbb{R}^n$  vérifiant  $\nabla f(x) = 0$  est appelé point critique (ou point stationnaire) de  $f$ .*

**Remarque III.4.1** *Tout point où le gradient est nul n'est pas nécessairement un extremum.*

**Théorème III.4.2 (Condition suffisante d'optimalité locale)** *Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  une application supposée de classe  $C^2$  vérifie :*

$\nabla f(\bar{x}) = 0$  et  $H[f](\bar{x})$  symétrique, définie positive (respectivement définie négative).

Alors  $\bar{x}$  est un point de minimum local (respectivement maximum local) de  $f$ .

Si de plus la fonctionnelle à optimiser est convexe ou strictement convexe, en appliquant le théorème III.4.2 au convexe  $C = \mathbb{R}^n$ , on obtient :

**Théorème III.4.3 (Condition suffisante d'optimalité globale)** *Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  une application différentiable et  $\bar{x}$  un point critique de  $f$ .*

- Si  $f$  est convexe, alors  $\bar{x}$  est un point de minimum global de  $f$ .
- Si  $f$  est strictement convexe, alors  $\bar{x}$  est l'unique point de minimum global de  $f$ .

## III.4.2 Algorithmes avec calcul de gradient

C'est une optimisation sans contraintes.

### III.4.2.1 Algorithme d'optimisation convexe 'algorithmes à directions de descente'

Partant d'un point  $x_0$  arbitrairement choisi, un algorithme de descente III.1 va chercher à générer une suite d'itérés  $(x_k)_{k \in \mathbb{N}}$  telle que :  $\forall k \in \mathbb{N} \quad f(x_{k+1}) \leq f(x_k)$  et on connaît  $f(x)$  et  $\nabla f(x_k)$ .

Le paramètre  $\tau_k$  est le pas à effectuer le long de la direction de descente  $d_k$  au point courant  $x_k$ ,  $-\nabla f(x_k)$  indique la direction avec le plus grand taux de décroissance de  $f$  au point  $x_k$ .

Un algorithme à directions de descente III.1 est donc déterminé par les paramètres  $\tau_k$  et  $d_k$ , la façon dont la direction de descente  $d_k$  est calculée donne son nom à l'algorithme et la façon dont le pas  $\tau_k$  est déterminé, est appelée recherche linéaire et peut se déterminer de différentes façon à préciser.

---

**Algorithme III.1 Directions de descente**


---

- 1: Choisir un point de départ  $x_0$  et  $k = 0$
- 2: Choisir la direction de descente  $d_k$ ,  $d_k \in \mathbb{R}^n$  telle que  $\nabla f(x_k)^T d_k < 0$
- 3: Choisir le pas de descente  $\tau_k > 0$  sachant :  $f(x_k + \tau_k d_k) < f(x_k)$  ( $f$  est une fonction décroissante)
- 4: Calculer le nouveau vecteur  $x_{k+1}$   

$$x_{k+1} = x_k + d_k \cdot \tau_k$$
- 5: Incrémenter  $k$  :  $k = k + 1$
- 6: Critère d'arrêt du calcul :
  - $\|x_{k+1} - x_k\| < \varepsilon \|x_k\|$
  - $\|f(x_{k+1}) - f(x_k)\| < \varepsilon \|f(x_k)\|$ .

Stop s'il y a convergence sinon reprendre de 2

---

**III.4.2.2 Algorithme à descente de gardient**

La formule de Taylor-Young assure qu'une fonction  $f$ , dérivable 2 fois au point  $x_0$ , admet un Développement Limité 'DL' en ce point :

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + o((x - x_0)^2)$$

On pose :  $x_0 = x_k$  et  $x = x_k + \tau_k$

$$f(x_k + \tau_k) = f(x_k) + \nabla f(x_k)(\tau_k) + \frac{1}{2}(\nabla^2 f)(x_k)(\tau_k)^2 + o((\tau_k)^2)$$

$$\nabla f(x_k + \tau_k) = \nabla f(x_k) + (\nabla^2 f)(x_k)(\tau_k)$$

Minimiser signifie « annuler la dérivée » :  $\nabla f(x_k + \tau_k) = 0$

$$\nabla f(x_k) + \tau_k \cdot \nabla^2 f(x_k) = 0$$

Nous choisissons  $\tau_k$  de telle sorte :  $\tau_k = -\frac{\nabla f(x_k)}{\|\nabla^2 f(x_k)\|}$  et nous prenons comme direction de descente  $d_k = -\nabla f(x_k)$ .

Le pas  $\tau_k$  est, soit fixé, éventuellement le même pour toutes les étapes, on parle alors de méthode à pas variable, soit calculé à chaque étape de façon à minimiser  $f$  dans la direction  $d_k$ , nous parlons alors de méthode à pas optimal.

Ci-joint l'algorithme de la méthode descente du gradient III.2 :

---

**Algorithme III.2 Descente de gradient**


---

- 1: Choisir un point de départ  $x_0$  et  $k = 0$
- 2: Choisir la direction de descente  $d_k = -\nabla f(x_k)$  .
- 3: Choisir le pas de descente  $\tau_k > 0$  sachant :  $f(x_k + \tau_k d_k) < f(x_k)$  ( $f$  est une fonction décroissante)
- 4: Calculer le nouveau vecteur  $x_{k+1}$   

$$x_{k+1} = x_k + \tau_k d_k = x_k - \tau_k \nabla f(x_k)$$
- 5: incrémenter  $k$  :  $k = k + 1$
- 6: Critère d'arrêt du calcul :
  - $\|x_{k+1} - x_k\| < \varepsilon \|x_k\|$
  - $\|f(x_{k+1}) - f(x_k)\| < \varepsilon \|f(x_k)\|$

stop s'il y a convergence sinon reprendre de 2

---

## 1. Méthode du gradient à pas fixe :

La descente de gradient avec un pas  $\tau$  constant est un problème simple, la descente à pas fixe ne marche que si la fonction est douce, et que si le point de départ est proche de la solution. L'organigramme de cette méthode est montré par la figure III.10 suivi par son algorithme III.3 :

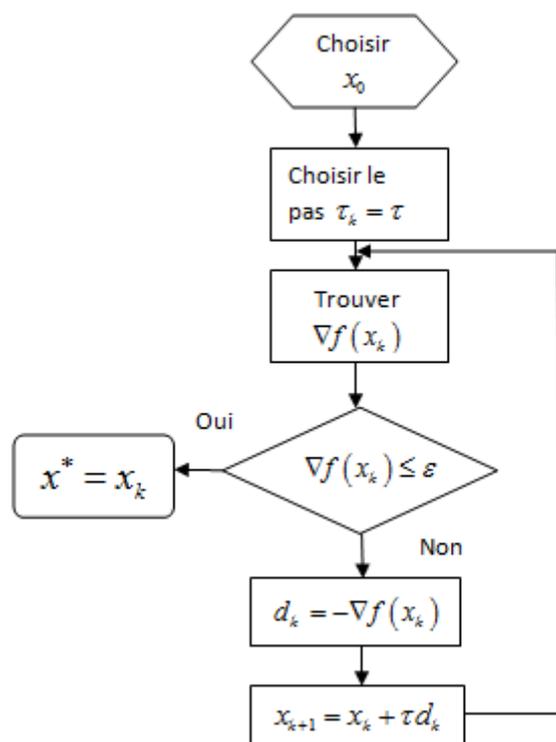


FIGURE III.10 – Organigramme de la méthode de la descente de gradient à pas fixe

**Algorithme III.3** Descente de gradient à pas fixe

- 1: Choisir un point de départ  $x_0$  et  $k = 0$
- 2: Choisir la direction de descente  $d_k = -\nabla f(x_k)$  .
- 3: Choisir le pas de descente  $(\tau_k = \tau = cste) > 0$  sachant :  $f(x_k + \tau d_k) < f(x_k)$  ( $f$  est une fonction décroissante)
- 4: Calculer le nouveau vecteur  $x_{k+1}$   

$$x_{k+1} = x_k - \tau \nabla f(x_k)$$
- 5: Incrémenter  $k : k = k + 1$
- 6: Critère d'arrêt du calcul :
  - $\|x_{k+1} - x_k\| < \varepsilon \|x_k\|$
  - $\|f(x_{k+1}) - f(x_k)\| < \varepsilon \|f(x_k)\|$

Stop s'il y a convergence sinon reprendre de 2

- Avantage : C'est la méthode de gradient la plus simple à utiliser.
- Inconvénient : la convergence peut être très lente.

## 2. Méthode du gradient à pas optimal ou méthode de plus profonde descente :

L'idée consiste à suivre la direction de plus forte descente et à faire un pas qui rende la fonction à minimiser la plus petite possible dans cette direction.

L'organigramme de cette méthode est montré par la figure III.11 suivi par son algorithme III.4

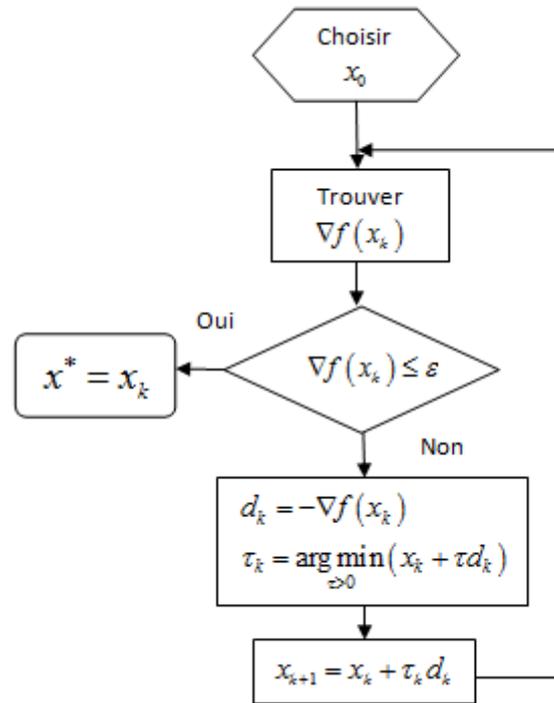


FIGURE III.11 – Organigramme de la méthode de la descente de gradient à pas optimal

---

### Algorithme III.4 Descente de gradient à pas optimal

---

- 1: Choisir un point de départ  $x_0$  et  $k = 0$
- 2: Choisir la direction de descente  $d_k = -\nabla f(x_k)$  .
- 3: Calculer un pas optimal  $\tau_k$  solution de  $\min_{\tau > 0} f(x_k - \tau d_k)$  sachant :  $f(x_k + \tau d_k) < f(x_k)$  ( $f$  est une fonction décroissante)
- 4: Calculer le nouveau vecteur  $x_{k+1}$   
 $x_{k+1} = x_k - \tau \nabla f(x_k)$
- 5: Incrémenter  $k : k = k + 1$
- 6: Critère d'arrêt du calcul :
  - $\|x_{k+1} - x_k\| < \varepsilon \|x_k\|$
  - $\|f(x_{k+1}) - f(x_k)\| < \varepsilon \|f(x_k)\|$

Stop s'il y a convergence sinon reprendre de 2

---

Propriétés :

- $(x_{k+1} - x_k) \perp (x_k - x_{k-1})$
- Si  $\nabla f(x_k) \neq 0$  alors  $f(x_{k+1}) < f(x_k)$

La méthode de plus profonde descente est une sorte d'idéalisation : d'une part, nous ne savons pas en pratique calculer de façon exacte un point minimum  $\tau_k$  de l'objectif

dans une direction donnée et le problème n'est en général pas trivial, d'autre part, la résolution du problème de minimisation même de façon approchée, coûte cher en temps de calcul. Pour ces raisons, nous pouvons lui préférer parfois l'algorithme de gradient à pas constant.

- Avantage : convergence admise.
- Inconvénient : difficile de trouver le pas  $\tau_k$ .

### III.4.2.3 Méthode de gradient conjugué

La figure III.12 montre de la méthode du gradient conjugué :

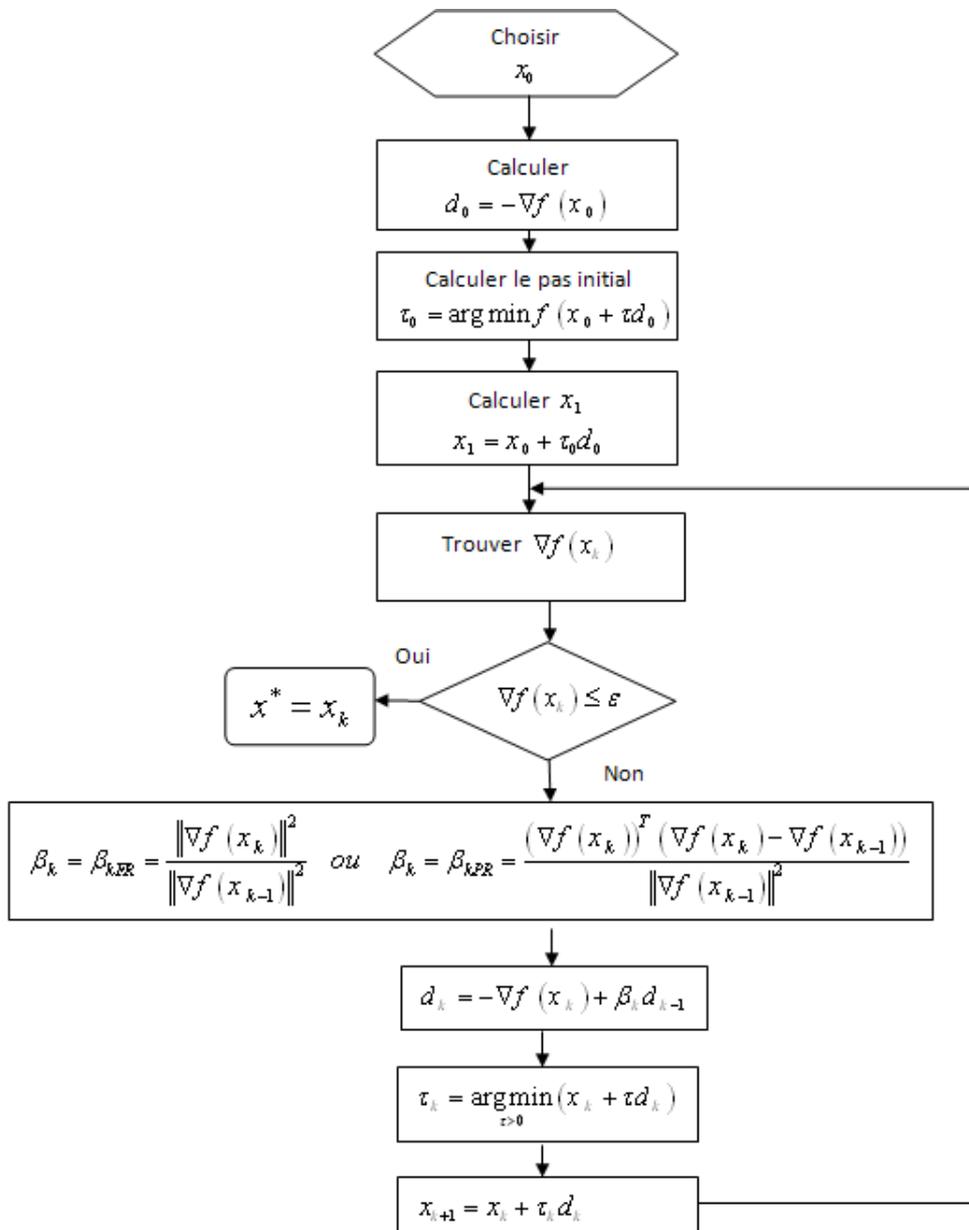


FIGURE III.12 – Organigramme du gradient conjugué

La méthode du gradient conjugué [Hager and Zhang, 2006] linéaire est une technique très utile pour résoudre de grands systèmes linéaires et elle peut même être adaptée de telle manière à ce qu'elle résolve des problèmes d'optimisation non-linéaires. La direction de descente à l'itération  $k + 1$  est une combinaison entre le gradient au point  $d = -\nabla f(\cdot)$  et de la direction de descente précédente pondérée par un coefficient  $\beta$ , de la forme :  $d_{k+1} = -\nabla f(x_{k+1}) + \beta d_k$ . On minimise donc sur un espace plus "grand" que dans la méthode de gradient à pas optimal, on s'attend alors à un "meilleur" minimum.

Le coefficient  $\beta$  selon [Fletcher and Reeves, 1964] est donné par l'équation (III.22) :

$$\beta_{kFR} = \frac{\|\nabla f(x_k)\|^2}{\|\nabla f(x_{k-1})\|^2} \quad (\text{III.22})$$

avec :  $\|\nabla f(x_{k-1})\|^2 = \nabla f(x_{k-1})^T \nabla f(x_{k-1})$  et  $\|\nabla f(x_k)\|^2 = \nabla f(x_k)^T \nabla f(x_k)$

Une autre extension de ce coefficient donnée par l'équation (III.23) est celle proposée par [Polak and Ribiere, 1969] :

$$\beta_{kPR} = \frac{(\nabla f(x_k))^T (\nabla f(x_k) - \nabla f(x_{k-1}))}{\|\nabla f(x_{k-1})\|^2} \quad (\text{III.23})$$

La figure III.12 illustre l'organigramme de la méthode du gradient conjugué selon Fletcher Reeves et Polak Rivière suivi par son algorithme montré en III.5 .

La méthode du gradient conjugué est facile à implémenter et elle donne le résultat attendu, elle résout les problèmes quadratiques en  $n$  itérations , et convergence plus rapide que le gradient.

Un avantage de l'algorithme de Fletcher et Reeves est d'abord, le faible coût en mémoire car un nombre réduit de vecteurs doit être stocké en mémoire, de plus, à chaque itération, seul le gradient de la fonction doit être calculé. L'utilisation du coefficient  $\beta$  selon l'extension de Polak et Rivière mène à une optimisation plus robuste qu'avec celui de Fletcher. La variante de Polack-Ribiere donne de meilleurs résultats en pratique.

### III.4.3 Algorithmes sans calcul du gradient

Les méthodes sans calcul du gradient sont intéressantes lorsque les méthodes classiques utilisant le gradient de la fonction de coût ne peuvent être utilisées ou fournissent des résultats médiocres (convergence locale). Ci-dessous quelques points expliquant pourquoi les méthodes classiques sont souvent écartées pour la résolution de problèmes complexes :

- Pour certains problèmes, la fonction de coût est non-différentiable par nature ou présente des discontinuités (par exemple les problèmes de type min-max).
- Le calcul du gradient de la fonction de coût peut être délicat.
- La fonction de coût peut présenter de multiples minima locaux.

---

**Algorithme III.5 Gradient conjugué selon Fletcher et Reeves ou Polak et Ribière**


---

- 1: Choisir un point de départ  $x_0$  et  $k = 0$
  - 2: Calculer la direction de descente au point 0,  $d_0 = -\nabla f(x_0)$
  - 3: Calculer le pas  $\tau_0$  : solution de  $\min_{\tau > 0} f(x_0 + \tau d_0)$  sachant :  $f(x_k + \tau d_k) < f(x_k)$  ( $f$  est une fonction décroissante)
  - 4: Calculer le coefficient  $x_1$   
 $x_1 = x_0 + \tau_0 d_0$
  - 5: Calculer le coefficient  $\beta_k$  selon Fletcher et Reeves Polak et Ribière
    - $\beta_{kFR} = \frac{\|\nabla f(x_k)\|^2}{\|\nabla f(x_{k-1})\|^2}$
    - ou
    - $\beta_{kPR} = \frac{(\nabla f(x_k))^T (\nabla f(x_k) - \nabla f(x_{k-1}))}{\|\nabla f(x_{k-1})\|^2}$
  - 6: Calculer la descente  $d_k$ 
    - $d_k = -\nabla f(x_k) + \beta_{kFR} d_{k-1}$
    - ou
    - $d_k = -\nabla f(x_k) + \beta_{kPR} d_{k-1}$
  - 7: Calculer le pas  $\tau_k$   
 $\tau_k = \arg \min_{\tau} f(x_k + \tau d_k)$
  - 8: Calculer le nouveau vecteur  $x_{k+1}$   
 $x_{k+1} = x_k + \tau_k d_k$
  - 9: Incrémenter  $k : k = k + 1$
  - 10: Critère d'arrêt du calcul :
    - $\|x_{k+1} - x_k\| < \varepsilon \|x_k\|$
    - $\|f(x_{k+1}) - f(x_k)\| < \varepsilon \|f(x_k)\|$
- Stop s'il y a convergence sinon reprendre de 5**
- 

Le but de cette sous section est de donner des algorithmes numériques pour la résolution du problème de minimisation avec contraintes. La difficulté du problème dépend de la nature des contraintes.

### III.4.3.1 Rappel sur les multiplicateurs de Lagrange

La méthode des multiplicateurs de Lagrange permet de trouver les points stationnaires (maximum, minimum...) d'une fonction dérivable d'une ou plusieurs variables, sous contraintes.

Le problème à résoudre est de trouver le minimum suivant :

$$\begin{aligned}
 & \min_x f(x) \\
 & \text{s.c.} \\
 & h_i(x) = 0, \quad i = 1, \dots, m \\
 & g_j(x) \leq 0, \quad j = 1, \dots, r
 \end{aligned}
 \tag{III.24}$$

avec  $f(x), h(x), g(x) \in C^1$ , ces fonctions admettent des dérivées partielles par rapport à chaque variable.

Le Lagrangien permet de transformer le problème avec contraintes en un problème sans contraintes. A chaque contrainte d'égalité  $h_i(x) = 0$  ou d'inégalité  $g_j(x) \leq 0$  est associée respectivement un paramètre scalaire  $\lambda_i, i = 1, \dots, m$  et  $\mu_j, j = 1, \dots, r$  appelé multiplicateur de Lagrange.

**Définition III.24 (Lagrangien)** On appelle lagrangien du problème (III.24) la fonction  $L$  définie par :

$$L(x, \lambda, \mu) \triangleq f(x) + \sum_{i=1}^m \lambda_i h_i(x) + \sum_{j=1}^r \mu_j g_j(x) \quad (\text{III.25})$$

Les paramètres  $\lambda_i$  et  $\mu_j$  sont appelés variables duales, ils sont des inconnus du problème qu'il faut aussi déterminer.  $x^*$  est une solution du problème (III.25) si et seulement si il existe  $\lambda_i^* (i = 1, \dots, m)$  et  $\mu_j^* (j = 1, \dots, r)$  satisfaisant les conditions Karush-Kuhn-Tucker (KKT) (voir la sous section III.4.3.2)

**Définition III.25 (Fonction duale lagrangienne)** On appelle la fonction duale lagrangienne du problème (III.24) la fonction  $d$  définie par :

$$d(\lambda, \mu) = \min_x L(x, \lambda, \mu) \quad (\text{III.26})$$

**Remarque III.4.2** Pour la contrainte d'égalité le multiplicateur est quelconque par contre pour la contrainte d'inégalité le multiplicateur doit être positif.

### III.4.3.2 Conditions d'optimalité du premier ordre : conditions de Karush-Kuhn-Tucker KKT

$x^*$  est une solution du problème (III.25) s'il existe  $\mu^* = [\mu_1^*, \mu_2^*, \dots, \mu_r^*]^T$  et  $\lambda^* = [\lambda_1^*, \lambda_2^*, \dots, \lambda_m^*]^T$  tels que :

- Stationarité :  $\nabla L(x, \lambda, \mu) = 0$ ,  $\nabla f(x) + \sum_{i=1}^m \lambda_i \nabla h_i(x) + \sum_{j=1}^r \mu_j \nabla g_j(x) = 0$ .
- Admissibilité primale :  $h_i(x) = 0$ ,  $i = 1, \dots, m$  et  $g_j(x) \leq 0$ ,  $j = 1, \dots, r$
- Admissibilité duale :  $\mu_j \geq 0$ ,  $j = 1, \dots, r$
- Complémentarité :  $\mu_j g_j(x) = 0$ ,  $\forall j = 1, \dots, r$

Condition de Slater :

$$\begin{aligned} f(x^*) &= d(\lambda^*, \mu^*) \\ f(x^*) &= \min_x f(x) + \sum_{i=1}^m \lambda_i^* h_i(x) + \sum_{j=1}^r \mu_j^* g_j(x) \\ &\leq f(x^*) + \sum_{i=1}^m \lambda_i^* h_i(x^*) + \sum_{j=1}^r \mu_j^* g_j(x^*) \\ &\leq f(x^*) \end{aligned}$$

**Théorème III.4.4 (Condition KKT)** *Sous certaines conditions de régularité, une solution  $x^*$  du problème d'optimisation vérifie nécessairement les conditions de Karush, Kuhn et Tucker.*

## III.5 Algorithme proximal

Dans cette section, nous allons rappeler certaines définitions, propriétés et résultats dans [Moreau, 1965], qui a été le premier à étudier l'application  $prox_f$  associée à une fonction  $f$ , où  $f \in \Gamma_0$  sans considérer l'aspect itératif. Cet aspect itératif a été abordé plus tard pour la première fois par [Martinet, 1972] et puis [Rockafellar, 1976].

Dans [Combettes and Wajs, 2005], [Combettes and Pesquet, 2007b], [Chaux et al., 2007], [Combettes and Pesquet, 2011], [Parikh and Boyd, 2014], les auteurs ont largement étudié l'opérateur proximal pour des fonctions convexes usuelles du traitement du signal. Par contre dans références [Urruty et al., 1993b], [Urruty et al., 1993a], [Attouch et al., 2013], [Combettes and Vu, 2013], [Emilie et al., 2014], nous trouvons un exemple d'utilisation de cet opérateur si les fonctions ne sont pas nécessairement convexes dans le cas des normes pondérées.

Cette application proximale est définie par la relation suivante :

$$prox_f x = \arg \min_{x_0} \left\{ f(x_0) + \frac{1}{2} \|x_0 - x\|^2 \right\} \quad (\text{III.27})$$

Où  $\arg \min f(x)$  désigne une solution du problème de minimisation de  $f$ .

**Proposition III.5.1** *Soient  $x \in H$  et  $f \in \Gamma^\circ(H)$ . La fonction définie sur  $H$  par :*

$$x_0 \rightarrow g(x, x_0) = f(x_0) + \frac{1}{2} \|x_0 - x\|^2 \quad (\text{III.28})$$

*est fortement convexe par rapport à  $x_0$ , donc elle possède une valeur minimale, qu'elle atteint en un unique point.*

**Remarque III.5.1** *La fonction à minimiser dans la définition de l'opérateur proximal équation III.28 est strictement convexe et possède donc un unique minimiseur pour tout  $x \in \mathbb{R}$ .*

**Proposition III.5.2** *Soient  $x \in H$  et  $f \in \Gamma^\circ(H)$ . La fonction  $f_1$  définie sur  $H$  par :*

$$f_1(x_0) = \inf \{g(x, x_0); \quad x_0 \in H\} \quad (\text{III.29})$$

*est une minorante différentiable de  $f$ .*

**Définition III.26 (Point proximal)** *L'unique point qui réalise le minimum de l'équation III.27 est appelé le point proximal de  $f$  en  $x$  et est noté  $prox_f(x)$  ou  $prox_f x$ .*

**Définition III.27 (Approximation de Moreau-Yosida)** La fonction  $f_1$  est appelée l'approximation de Moreau Yosida de  $f$ , de rapport 1 [Moreau, 1967].

**Définition III.28 (Opérateur proximal ou résolvente)** On appelle opérateur proximal ou résolvente d'un opérateur maximal monotone  $T$  l'opérateur suivant :

$$J^T = (1 + T)^{-1} \quad (\text{III.30})$$

On appelle aussi résolvente de  $T$ , de l'opérateur proximal de  $\lambda T$  pour  $\lambda > 0$  et on le note par :

$$J_\gamma^T = (1 + \gamma T)^{-1} \quad (\text{III.31})$$

L'approximation de Yosida de l'opérateur  $T$  est définie par :

$$A_\gamma^T = \frac{I - J_\gamma^T}{\gamma} \quad (\text{III.32})$$

La méthode du point proximal consiste à appliquer l'itération :

$$x^{k+1} = J_{\gamma_k}^T x^k \quad (\text{III.33})$$

### III.5.1 Algorithme du point proximal

Le problème général est de trouver la solution de l'équation III.34 :

$$0 \in Tx \quad (\text{III.34})$$

C'est-à-dire trouver un zéro de l'opérateur  $T$ .

**Proposition III.5.3**  $x$  est un zéro de  $T$  si et seulement si  $x$  est un point fixe de  $J_\gamma^T$ .  
(Voir Preuve C.1)

#### Preuve III.1

$$0 \in Tx$$

$$\Leftrightarrow \gamma 0 \in \gamma Tx \Rightarrow 0 \in \gamma Tx$$

$$\Leftrightarrow x \in x + \gamma Tx \Rightarrow x \in (I + \gamma T)x$$

$$\Leftrightarrow (I + \gamma T)^{-1}x = \underbrace{(I + \gamma T)^{-1}(I + \gamma T)x}_I$$

$$\Leftrightarrow (I + \gamma T)^{-1}x = x$$

$$\Leftrightarrow x = \underbrace{(I + \gamma T)^{-1}x}_{J_\gamma^T}$$

$$x = J_\gamma^T x$$

Nous pouvons voir clairement que la méthode du point proximal consiste à appliquer l'itération :

$$x^{k+1} = J_{\gamma_k}^T x^k \quad (\text{III.35})$$

L'algorithme III.6 peut être vu comme la discrétisation de la méthode d'Euler appliquée

---

### Algorithme III.6 Point proximal

---

- 1: **Initialisation**  $x^0 \in H$  et  $k = 0$
  - 2: **Itération proximale**  $x^{k+1} = J_{\gamma_k}^T x^k$
  - 3: **Test** :
    - $0 \in Tx^k$  alors fin et on a une solution
    - Sinon  $k := k + 1$  et reprendre de 2
- 

à l'inclusion différentielle ci-dessous :

$$\begin{aligned} \frac{dx}{dt} &\in Tx(t) \\ x(0) &= x^0 \end{aligned} \quad (\text{III.36})$$

L'équation III.36 peut être discrétisée par deux manières :

- La première est une discrétisation d'Euler explicite donnée par l'équation :

$$\frac{x^k - x^{k+1}}{\lambda_k} \in T(x^k) \quad (\text{III.37})$$

Pour  $y^k \in Tx^k$  et  $\lambda_k > 0$ , sa discrétisation donne : (Voir preuve III.2)

$$x^{k+1} = x^k - \gamma_k y^k \quad (\text{III.38})$$

#### Preuve III.2

$$\begin{aligned} \frac{x^k - x^{k+1}}{\gamma_k} &\in T(x^k) \\ \Leftrightarrow \frac{x^k - x^{k+1}}{\gamma_k} &= y^k \\ x^{k+1} &= x^k - \gamma_k y^k \end{aligned}$$

Qui n'est rien d'autre que l'algorithme du sous-gradient lorsque  $T = \partial f$  et  $f \in \Gamma_0(H)$

- La deuxième est une discrétisation d'Euler implicite donnée par :

$$\frac{x^{k-1} - x^k}{\gamma_k} \in T(x^k) \quad (\text{III.39})$$

La solution de l'équation III.39 est donnée par : (Voir preuve III.3 )

$$x^{k+1} = J_k^T x^k \quad (\text{III.40})$$

### Preuve III.3

$$\begin{aligned} \frac{x^{k-1} - x^k}{\gamma_k} &\in T(x^k) \\ \Leftrightarrow x^{k-1} &\in x^k + \gamma_k T(x^k) = (I + \gamma_k T)(x^k) \\ \Leftrightarrow x^k &\in (I + \gamma_k T)(x^{k+1}) \\ \Leftrightarrow (I + \lambda_k T)^{-1} x^k &\in (I + \gamma_k T)^{-1} (I + \gamma_k T)(x^{k+1}) \\ \Leftrightarrow x^{k+1} &= (I + \gamma_k T)^{-1} x^k \\ x^{k+1} &= J_k^T x^k \end{aligned}$$

## III.5.2 Application du point proximal à l'optimisation convexe

Soit  $f$  un élément de  $\Gamma_0(H)$ , on veut résoudre le problème suivant :

$$\min_{x \in H} \text{imiser } f(x) \quad (\text{III.41})$$

La condition d'optimalité étant de trouver  $x^*$  tel que :  $0 \in \partial f(x^*)$ , qui est un cas particulier de l'équation III.34 si on pose  $T = \partial f$ , on peut donc itérer la suite III.36 et appliquer l'algorithme III.6 au problème III.41.

La suite III.36 peut être interprétée comme étant une régularisation proximale, et elle prendra la forme ci-dessous :

$$x^{k+1} = \arg \min_{x_0 \in H} \left\{ f(x_0) + \frac{1}{2\gamma_k} \|x_0 - x^k\| \right\} \quad (\text{III.42})$$

On pose :

$$f_\gamma(x) = \min_{x_0 \in H} \left\{ f(x_0) + \frac{1}{2\gamma_k} \|x_0 - x^k\| \right\} \quad (\text{III.43})$$

Ce minimum est atteint en un point unique qu'on notera  $\text{prox}_{\lambda f} x$ .

$f_\gamma(x)$  est appelée fonction régularisée ou approximation de Moreau-Yosida de  $f$ .

**Proposition III.5.4** Soit  $f$  un élément de  $\Gamma_0(H)$ , la fonction régularisée  $f_\gamma(x)$  de  $f(x)$  convexe et différentiable, son gradient est donné par :

$$\nabla f_\gamma(x) = \frac{x - \text{prox}_{\gamma f} x}{\gamma} \quad (\text{III.44})$$

de plus on a :  $\|\nabla f_\gamma(x)\| \leq \inf \{\|y\| : y \in \partial f(x)\}$

**Remarque III.5.2** Notons que :

- $f_\gamma$  est différentiable.
- $f_\gamma$  est l'infconvolution de  $f$  et  $\|\cdot\|^2$ .

**Proposition III.5.5** Soit  $f$  un élément de  $\Gamma_0(H)$ ,  $0 < \gamma_1 < \gamma_2$  et  $\alpha = \inf f$ , alors pour tout  $x \in \text{dom } f$ , on a :

—

$$\alpha \leq f_{\gamma_2}(x) \leq f_{\gamma_1}(x) \leq f(x) \tag{III.45}$$

—

$$\lim_{\gamma \rightarrow 0} f_\gamma(x) = f(x) \tag{III.46}$$

—

$$\lim_{\gamma \rightarrow +\infty} f_\gamma(x) = \alpha = \inf f(x) \tag{III.47}$$

**Application III.11 (Approximation de Moreau-Yosida de  $f(x) = |x|$ )**

L'approximation de Moreau-Yosida  $f_\gamma(x)$  de  $f(x) = |x|$ , est définie pour  $\gamma > 0$  par :

$$f_\gamma(x) = \begin{cases} \frac{x^2}{2\gamma} & \text{si } |x| \leq \gamma \\ |x| - \frac{\gamma}{2} & \text{si } |x| > \gamma \end{cases} \tag{III.48}$$

La figure III.13 représente la courbe de  $f(x)$  ainsi que celles de  $f_\gamma(x)$  pour certaines valeurs de  $\gamma > 0$  ( $\gamma = 0.25, 0.50, 1, 2, 4$ )

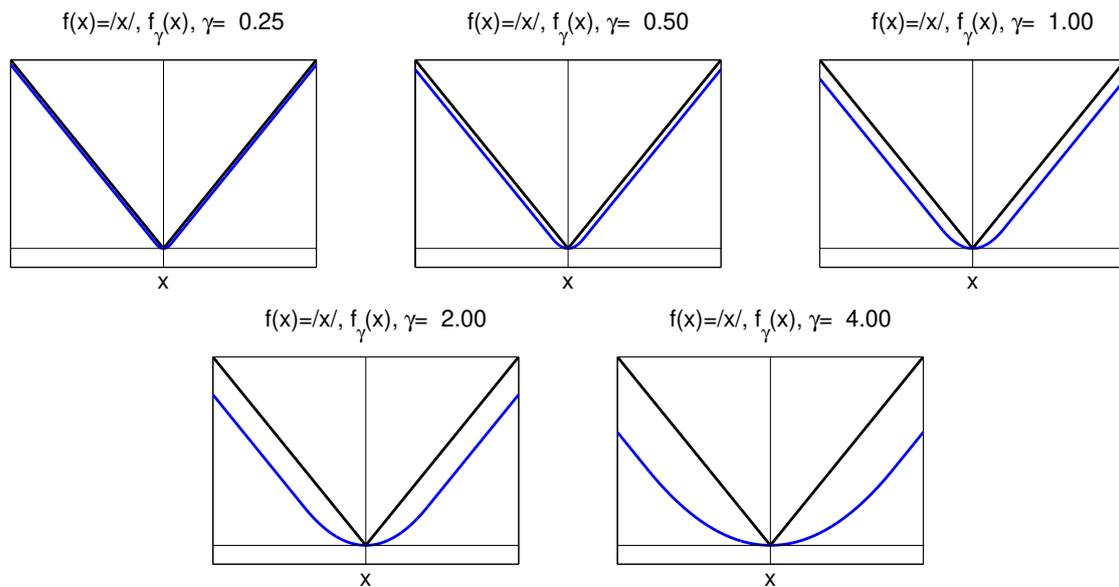


FIGURE III.13 – Approximation de Moreau-Yosida  $f_\gamma(x)$  de  $f(x) = |x|$  pour  $\gamma = 0.25, 0.50, 1, 2, 4$

**Application III.12 (Proximal de  $f(x) = i_C x$ )** Lorsque  $f$  est la fonction indicatrice d'un ensemble convexe fermé non vide :

$$i_C x = \begin{cases} 0 & \text{si } y \in C \\ +\infty & \text{si } y \notin C \end{cases} \quad (\text{III.49})$$

L'opérateur proximal correspondant est la projection euclidienne sur  $C$  :

$$\text{prox}_{i_C} x = P_C x = \min_{x_0 \in C} \|x - x_0\| \quad (\text{III.50})$$

rappelons que :  $\text{prox}_f x = \arg \min_{x_0 \in H} \{f(x_0) + \frac{1}{2}\|x_0 - x\|^2\}$ . La fonction indicatrice a comme opérateur proximal l'équation III.51 :

$$\text{prox}_{i_C} x = \arg \min_{x_0 \in C} \left\{ \frac{1}{2} \|x_0 - x\|^2 \right\} = \text{Pr } o j_C \quad (\text{III.51})$$

**Application III.13 (Proximal  $f(x) = |x|$ )** Soit  $\gamma \in ]0, +\infty[$  et  $f(x) = |x|$ , on a pour tout  $x \in \mathbb{R}$  :

$$\text{prox}_{\gamma f}(x) = \arg \min_{x_0 \in H} \left\{ \|x_0\| + \frac{1}{2\gamma} \|x_0 - x\|^2 \right\} \quad (\text{III.52})$$

$$\text{prox}_{\gamma f}(x) = \max \left( 0, 1 - \frac{\gamma}{\|x\|} \right) \quad (\text{III.53})$$

La figure III.14 illustre la représentation graphique du proximal de  $f(x) = |x|$  donné par la relation III.53 :

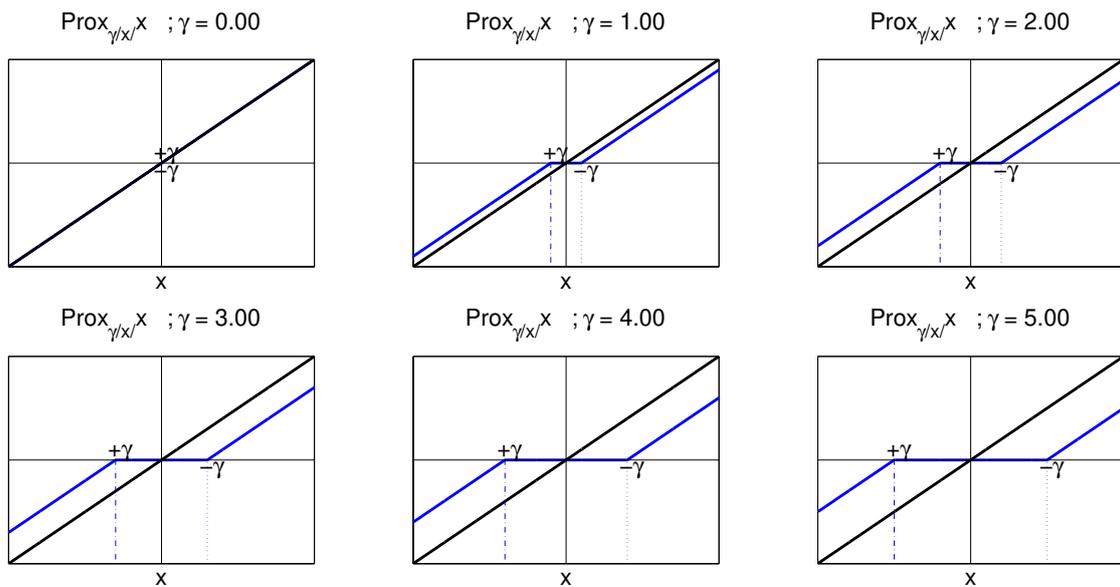


FIGURE III.14 – Proximal de  $f(x) = |x|$  pour  $\gamma = 0, 1, 3, 3, 4, 5$

**Application III.14 (Enveloppe de Moreau Yoside et opérateur proximal de  $f(x) = |x|$ )**  
 Pour  $x_0$  varie de  $-3$  à  $3$  avec un pas de  $0.5$  et  $\gamma = 1$ , la figure III.16 montre les tracés<sup>1</sup>  
 des fonctions suivantes :

$$g(x) = f(x_0) + \frac{1}{2\gamma} \|x_0 - x\|^2 \Rightarrow g(x) = \|x_0\| + \frac{1}{2} \|x_0 - x\|^2$$

$$f_\gamma(x) = \inf_{x_0 \in H} \left\{ f(x_0) + \frac{1}{2\gamma} \|x_0 - x\|^2 \right\} \Rightarrow f_\gamma(x) = \inf_{x_0 \in H} \left\{ \|x_0\| + \frac{1}{2} \|x_0 - x\|^2 \right\}$$

$$f_\gamma(x) = \begin{cases} \frac{x^2}{2\gamma} & \text{si } |x| \leq \gamma \\ |x| - \frac{\gamma}{2} & \text{si } |x| > \gamma \end{cases} \Rightarrow f_1(x) = \begin{cases} \frac{x^2}{2} & \text{si } |x| \leq 1 \\ |x| - \frac{1}{2} & \text{si } |x| > 1 \end{cases}$$

$$\text{prox}_{\gamma f}(x) = \max\left(0, 1 - \frac{\gamma}{\|x\|}\right) \Rightarrow \text{prox}_{1f}(x) = \max\left(0, 1 - \frac{1}{\|x\|}\right)$$

La liason entre l'enveloppe de moreau et l'opérateur proximal est esquissée par la figure III.15 :

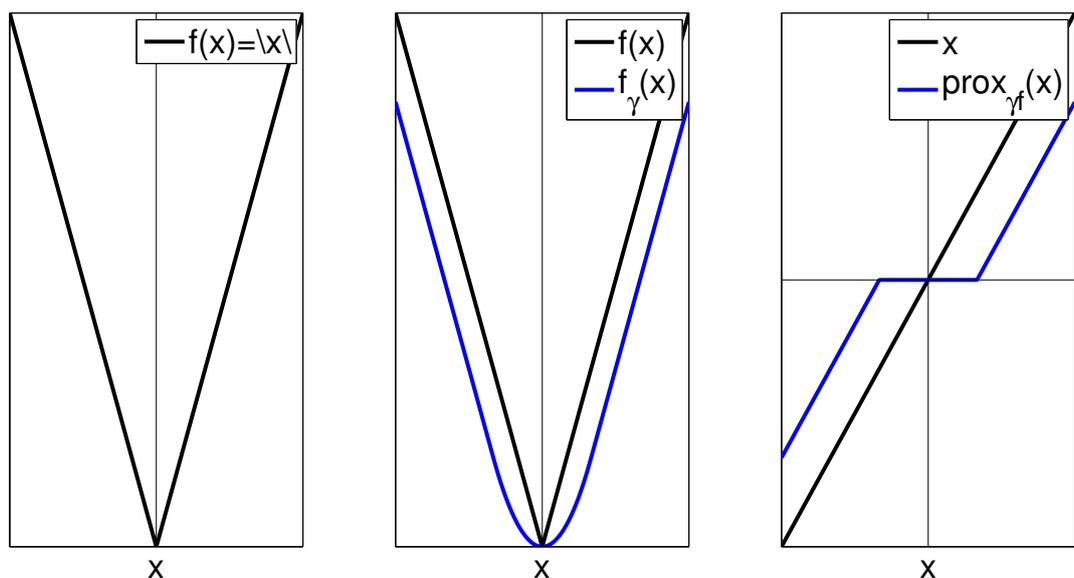


FIGURE III.15 – Représentation de  $f(x) = |x|$ , son enveloppe de Moreau Yosida et son opérateur proximal pour  $\gamma = 1.5$

Pour mieux comprendre, plus de détails sont donnés dans les figures III.16 : La fonction  $g(x) = f(x_0) + \frac{1}{2} \|x_0 - x\|^2$  est fortement convexe par rapport à  $x_0$ , donc elle possède une valeur minimale qu'elle atteint en un point unique qui est le point proximal noté  $\text{prox}_f x$ . Les deux fonctions  $f(x)$  et  $g(x)$  se coupent au point  $x_0$ , elles sont égales en ce point, c'est à dire  $f(x_0) = g(x_0)$ .

1. Voir annexe B

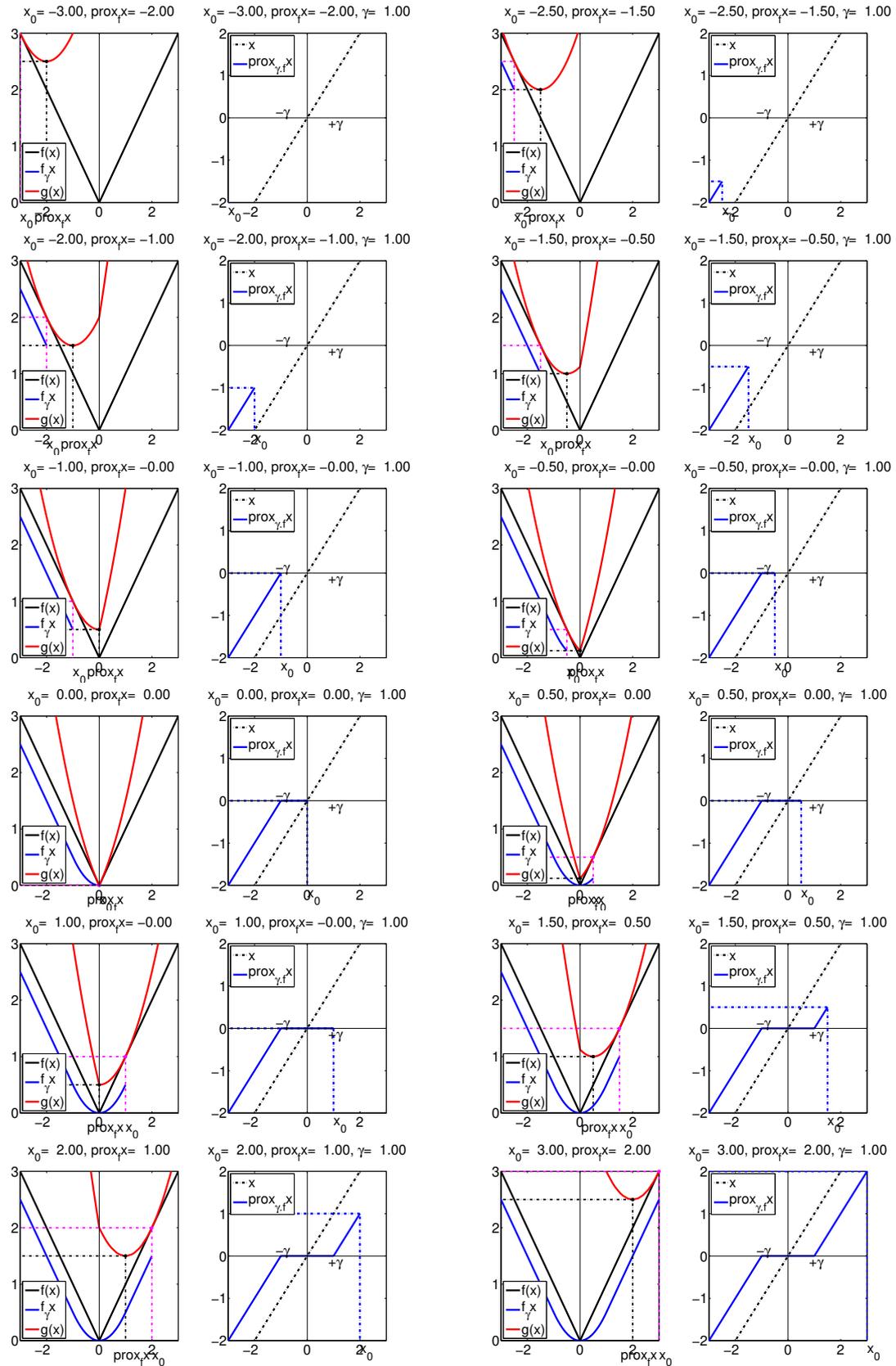


FIGURE III.16 – Visualisation de l'enveloppe de Moreau Yosida et le proximal de  $f(x) = |x|$  pour  $\gamma = 1$

**Proposition III.5.6** *Si  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction convexe fermée,  $\text{prox}_f x$  existe et est unique pour tout  $x \in \mathbb{R}^n$ . De plus il est caractérisé par les équations suivantes :*

$$\begin{aligned} u &= \text{prox}_f x \\ &\Leftrightarrow 0 \in \partial f(u) + u - x \\ &\Leftrightarrow u = (I + \partial f)^{-1} x \end{aligned} \tag{III.54}$$

Ci-joint la démonstration III.4 de la proposition III.5.6 :

#### Preuve III.4

$$\begin{aligned} u &= \text{prox}_f x \\ &\Leftrightarrow x - u \in \partial f(u) \quad (0 \in \partial f(u) + u - x) \\ &\Leftrightarrow x \in u + \partial f(u) \quad (x \in (I + \partial f)(u)) \\ &\Leftrightarrow x \in (I + \partial f)(u) \\ &\Leftrightarrow (I + \partial f)^{-1} x = \underbrace{(I + \partial f)^{-1} (I + \partial f)}_I (u) \\ &\Leftrightarrow (I + \partial f)^{-1} x = u \\ u = \text{prox}_f x &\Leftrightarrow u = (I + \partial f)^{-1} x \end{aligned}$$

**Proposition III.5.7**  $f \in \Gamma_0$  alors :

- $f \in \Gamma_0 \Leftrightarrow f^* \in \Gamma_0$  dans ce cas on a  $f = f^{**}$  ,
- $y \in \partial f(x) \Leftrightarrow f(x) + f^*(y) = \langle x, y \rangle$  et  $y \in \partial f(x) \Leftrightarrow x \in \partial f(y^*)$ ,
- $\partial f(x)$  est une partie convexe fermée,
- Si  $\partial f(x) = \{y\}$  alors  $f$  est différentiable en  $x$  et  $\nabla f(x) = y$  ,
- $f(x^0) = \min \{f(x), x \in H\} \Leftrightarrow 0 \in \partial f(x^0)$

## III.6 Application du proximal au débruitage d'images

Dans cette section et dans la cadre de l'application du proximal au débruitage des images, nous étudions et implémentons deux algorithmes Forward-Backward (FB) et Douglas-Rachford (DR), qui permettent de résoudre le problème décrit par l'équation III.60.

### III.6.1 Problématique

On commence à donner le modèle utilisé pour la restauration d'images. Soit l'image observée  $f$  telle que :

$$f = u + \eta \tag{III.55}$$

Rappelons que  $u$  est l'image originale et  $\eta$  le bruit gaussien de moyenne nulle et de variance  $\sigma^2$ . On cherche donc à calculer une image proche de  $u$  présentant une texture moins bruitée. Le problème de débruitage revient à minimiser la fonction objectif qui n'est que la variation  $\min_u \int_{\Omega} |\nabla u|$  sous la contrainte  $Ku = f$  :

$$\begin{aligned} \min_u \int_{\Omega} |\nabla u| \\ \text{sous contrainte} \\ Ku = f \end{aligned} \quad (\text{III.56})$$

$K$  est un opérateur matriciel. Pour qu'on puisse utiliser le proximal, on ramène le problème de débruitage III.56 à un modèle sans contrainte différentiable convexe. Ceci revient alors à minimiser une énergie composée de deux termes :

$$\min_u E = \min_u \{E_{reg}(u) + \lambda.E_{data}(u, f)\} \quad (\text{III.57})$$

— Le premier terme est le terme de régularisation donné par une variation totale norme  $\ell_1$  :

$$E_{reg}(u) = \int_{\Omega} |\nabla u| \quad (\text{III.58})$$

— Le deuxième terme est le terme de fidélité ou terme d'attache aux données, lié au bruit :

$$E_{data}(u) = \int_{\Omega} \|Ku - f\|^2 \quad (\text{III.59})$$

Le multiplicateur de Lagrange  $\lambda$  est important pour la qualité finale de la restauration. L'équation III.57 s'écrit alors comme suit :

$$\min_u E(u) = \min_u \left( |\nabla u| + \frac{\lambda}{2} \|Ku - f\|^2 \right) \quad (\text{III.60})$$

Par la suite, nous prenons  $K = I$ ,  $I$  est la matrice identité.

## III.6.2 Algorithme Forward-Backward (FB) appliqué au débruitage d'images (bruit additif)

### III.6.2.1 Algorithme explicite-implicite (FB : Forward-Backward)

Cet algorithme [Daubechies et al., 2004],[Bect et al., 2004],[Combettes and Wajs, 2005], [Weiss et al., 2009] permet de gérer efficacement un critère composé d'un terme quadratique (plus généralement une fonction de gradient Lipschitz) et d'une fonction pouvant être non différentiable.

Soient  $f_1$  et  $f_2$  deux fonctions de  $\Gamma_0(H)$  telles que  $\arg \min f_1 + f_2 \neq \emptyset$  où  $f_2$  est une

fonction différentiable sur  $H$ , de gradient  $\beta$ -Lipschitz.

On souhaite trouver :

$$\hat{x} = \arg \min_x f_1(x) + f_2(x) \quad (\text{III.61})$$

$x$  est solution si et seulement si :

$$x = \text{prox}_{\gamma f_1}(x - \gamma \nabla f_2(x)) \quad (\text{III.62})$$

l'équation III.62 peut s'écrire encore sous la forme ci-dessous :

$$\begin{cases} x = \text{prox}_{\gamma f_1}(y) \\ y = x - \gamma \nabla f_2(x) \end{cases} \quad (\text{III.63})$$

La démonstration de l'équation III.63 est décrite ci-dessous :

### Preuve III.5

$f_2 : \mathbb{R}^n \rightarrow \mathbb{R}$ , différentiable, propre, convexe

$f_1 : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ , nondifférentiable, propre, convexe

$f_1(x)$  nondifférentiable  $\Rightarrow \partial f_1(x) \neq \nabla f_1(x) \Rightarrow$  chercher  $\text{prox}_{f_1} x$

$f_2(x)$  différentiable  $\Rightarrow \partial f_2(x) = \nabla f_2(x)$

$\min_x f_1(x) + f_2(x)$

$0 \in \nabla f_2(x) + \partial f_1(x)$

$\gamma \cdot 0 \in \gamma \cdot \nabla f_2(x) + \gamma \cdot \partial f_1(x)$ ,  $\gamma > 0$

$0 \in (\gamma \cdot \nabla f_2(x) - x) + (x + \gamma \cdot \partial f_1(x))$

$0 \in (\gamma \cdot \nabla f_2 - I)(x) + (I + \gamma \cdot \partial f_1)(x)$

$(I - \gamma \cdot \nabla f_2)(x) \in (I + \gamma \cdot \partial f_1)(x) \quad (I + \gamma \cdot \partial f_1)(x) \ni (I - \gamma \cdot \nabla f_2)(x)$

$(I + \gamma \cdot \partial f_1)^{-1}(I + \gamma \cdot \partial f_1)(x) \ni (I + \gamma \cdot \partial f_1)^{-1}(I - \gamma \cdot \nabla f_2)(x)$

$x = (I + \gamma \cdot \partial f_1)^{-1}(I - \gamma \cdot \nabla f_2)(x) \quad x = \text{prox}_{\gamma f_1}(I - \gamma \cdot \nabla f_2)(x)$

$x = \text{Pr } \text{ox}_{\gamma f_1}(x - \gamma \cdot \nabla f_2(x))$  (\*)....ce qu'il faut démontrer.

(\*) peut s'écrire aussi sous la forme :  $\begin{cases} x = \text{Pr } \text{ox}_{\gamma f_1} y \dots \text{Forward} \\ y = x - \gamma \nabla f_2(x) \dots \text{Backward} \end{cases}$

Les auteurs dans [Combettes and Pesquet, 2011] montrent que la méthode FB exige le déroulement des deux étapes suivantes :

— Etape explicite (en anglais Forward step) utilise la descente de gradient :

$$\frac{x_{n+1} - x_n}{\tau_n} = -\nabla f_2(x) : f_1 = 0$$

— Etape implicite (en anglais Backward step) nécessite la calcul du proximal :

$$x_{n+1} = \text{prox}_{\gamma_n f_1} x_n : f_2 = 0$$

Les itérations de l'algorithme III.7 explicite-implicite deviennent :

---

**Algorithme III.7 Forward-Backward (Explicite-Implicite)**


---

- 1: Fixer  $\varepsilon \in ]0, \min\{1, \frac{1}{\beta}\}]$  et  $x_0 \in H$
  - 2: Pour  $n = 0, 1, \dots$  :
    - $\gamma_n \in [\varepsilon, \frac{2}{\beta} - \varepsilon]$
    - **Etape Forward (implicite)** :  $y_n = x_n - \gamma_n \nabla f_2(x_n)$
    - $\tau_n \in [\varepsilon, 1]$
    - **Etape Backward (explicite)** :  $x_{n+1} = x_n + \tau_n (\text{prox}_{\gamma_n f_1} y_n - x_n)$
- 

### III.6.2.2 Résolution du problème

Considérons le problème sans contraintes :  $\min_u E(u) = \min_u \{|\nabla u| + \frac{\lambda}{2} \|Ku - f\|^2\}$ .  
 $E_{reg}(u) \in \Gamma_0(H) \rightarrow E_{reg}(u) : \mathbb{R}^N \rightarrow ]-\infty, +\infty]$   
 $E_{data}(u) \in \Gamma_0(H) \rightarrow E_{data}(u) : \mathbb{R}^N \rightarrow ]-\infty, +\infty]$   
 $E_{data}(u)$  est une fonction différentiable sur  $\mathbb{R}^N$ , de gradient  $\beta$ -Lipschitz :

$$\|E_{data}(u_2) - E_{data}(u_1)\| \leq \beta \|u_2 - u_1\|; \quad \beta > 0$$

$$\arg \min_u \{E_{data}(u) + E_{reg}(u)\} \neq \emptyset \Rightarrow \exists! \hat{u} \quad |\hat{u} = \min_u \{E_{data}(u) + E_{reg}(u)\}$$

Il faut chercher :  $\text{Pr } ox_{\gamma E_{reg}}(u)$  et  $\nabla E_{data}(u)$ .

Comme  $E_{reg}(u) = |\nabla u|_1$  et  $E_{data}(u) = \frac{\lambda}{2} \|Ku - f\|^2$ , il est facile de montrer que :

$$\text{prox}_{\gamma E_{reg}}(\nabla u) = \max\left(0, 1 - \frac{\gamma}{|\nabla u|}\right) \nabla u \quad (\text{III.64})$$

$$\nabla E_{data}(u) = \partial E_{data}(u) = \lambda K^T (Ku - f) \quad (\text{III.65})$$

D'après l'équation III.63, on aboutit à :

$$\begin{cases} u = \text{Pr } ox_{\gamma E_{reg}} y \\ y = u - \gamma \cdot \partial E_{data}(u) \end{cases} \quad (\text{III.66})$$

En appliquant la méthode de sous-gradient implicite et après discrétisation, on trouve l'équation III.67 :

$$\frac{u_{i,j}^{n+1} - u_{i,j}^n}{\Delta t} = - (\text{prox}_{\gamma E_{reg}}(y))_{i,j}^{n+1} \quad (\text{III.67})$$

La solution est :

$$u_{i,j}^{n+1} = u_{i,j}^n - \tau_n \cdot \text{prox}_{\gamma E_{reg}} \left( u_{i,j}^n - \gamma (\nabla E_{data}(u))_{i,j}^{n+1} \right) \quad (\text{III.68})$$

Les figures III.17, III.18, III.19, III.20 illustrent la technique Forward-Backward appliquée respectivement au débruitage des images 'Cameraman', 'Clock', 'Parrort' et 'Satellite' pour différentes valeurs de  $\sigma$  :



(a) Originale

(b) Bruitée  $\sigma = 10$ 

(c) Masquée



(d) Restaurée

(e) Bruitée  $\sigma = 20$ 

(f) Masquée



(g) Restaurée

(h) Bruitée  $\sigma = 30$ 

(i) Masquée

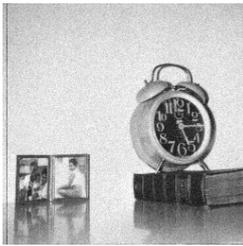
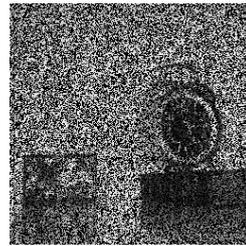


(j) Restaurée

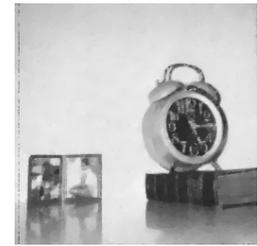
FIGURE III.17 – Débruitage de l'image Cameraman de dimension  $256 \times 256$  par l'algorithme Forward-Backward lorsque l'image est corrompue par un bruit gaussien de variance  $\sigma = 10$ ,  $\sigma = 20$ ,  $\sigma = 30$



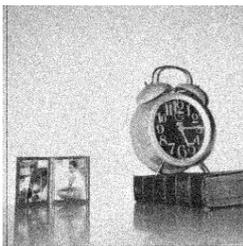
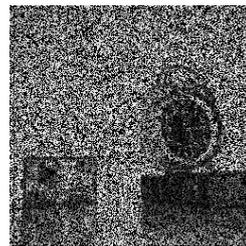
(a) Originale

(b) Bruitée  $\sigma = 10$ 

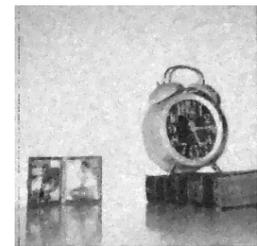
(c) Masquée



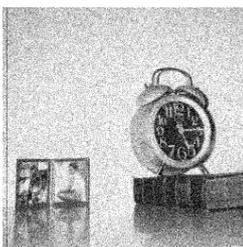
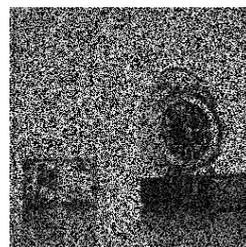
(d) Restaurée

(e) Bruitée  $\sigma = 20$ 

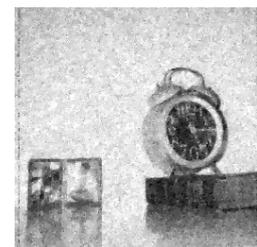
(f) Masquée



(g) Restaurée

(h) Bruitée  $\sigma = 30$ 

(i) Masquée

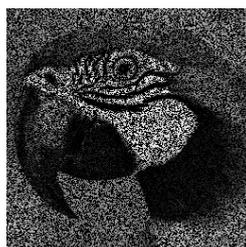


(j) Restaurée

FIGURE III.18 – Débruitage de l'image Clock de dimension  $256 \times 256$  par l'algorithme Forward-Backward lorsque l'image est corrompue par un bruit gaussien de variance  $\sigma = 10$ ,  $\sigma = 20$ ,  $\sigma = 30$



(a) Originale

(b) Bruitée  $\sigma = 10$ 

(c) Masquée



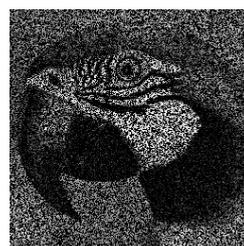
(d) Restaurée

(e) Bruitée  $\sigma = 20$ 

(f) Masquée



(g) Restaurée

(h) Bruitée  $\sigma = 30$ 

(i) Masquée



(j) Restaurée

FIGURE III.19 – Débruitage de l'image parrot de dimension  $256 \times 256$  par l'algorithme Forward-Backward lorsque l'image est corrompue par un bruit gaussien de variance  $\sigma = 10$ ,  $\sigma = 20$ ,  $\sigma = 30$

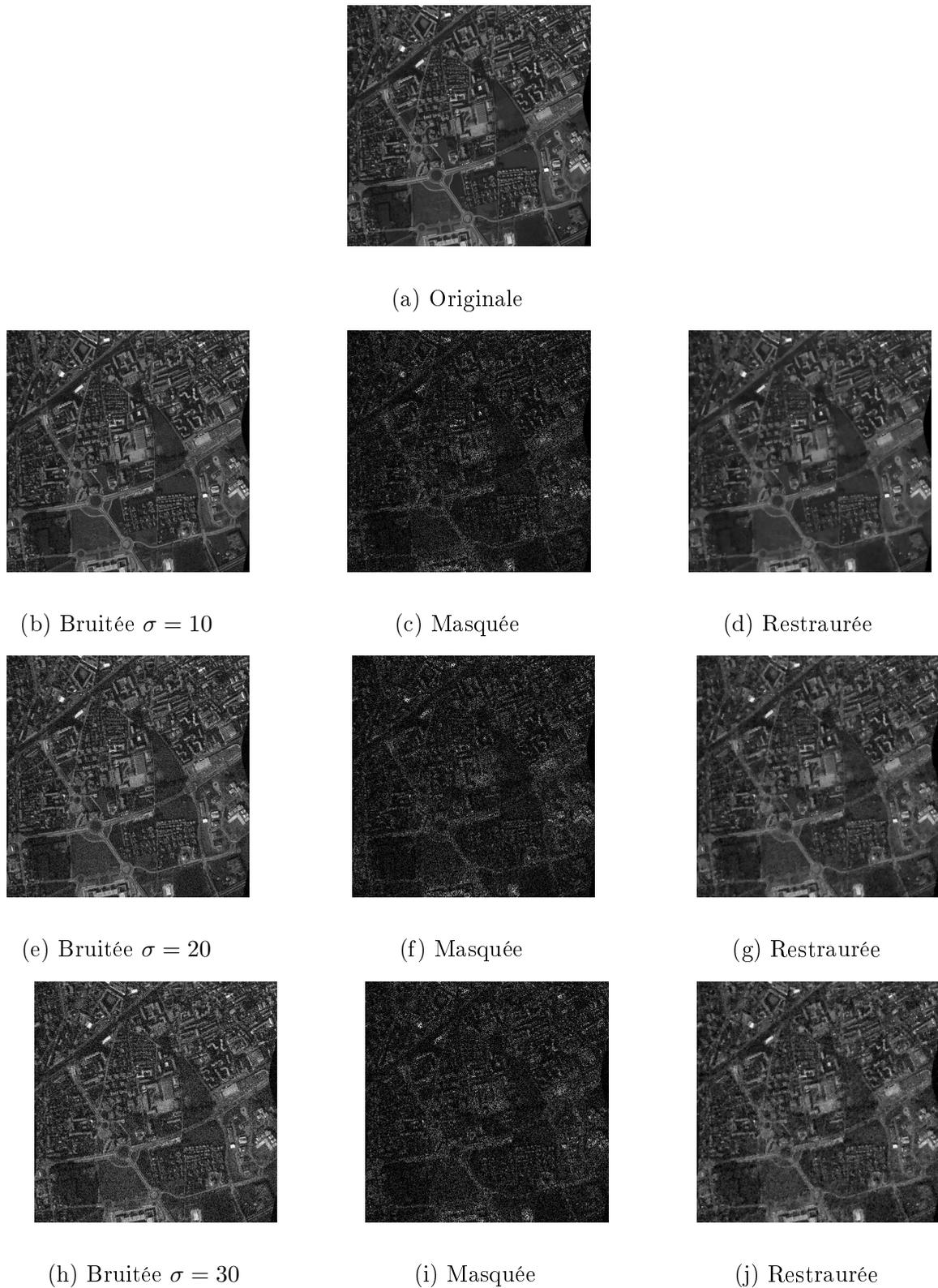


FIGURE III.20 – Débruitage de l'image Satellite de dimension  $512 \times 512$  par l'algorithme Forward-Backward lorsque l'image est corrompue par un bruit gaussien de variance  $\sigma = 10$ ,  $\sigma = 20$ ,  $\sigma = 30$

### III.6.3 Algorithme Douglas-Rachford (DR) appliqué au débruitage d'images bruit additif

#### III.6.3.1 Algorithme de Douglas-Rachford (DR)

Pour l'algorithme Douglas-Rachford 'DR' [Combettes and Pesquet, 2007a], [Eckstein and Bertsekas, 1992], la contrainte de différentiabilité au sens de Lipschitz formulée par l'algorithme Forward-Backward est relâchée.

Soient  $f_1$  et  $f_2$  deux fonctions de  $\Gamma_0(H)$  telles que  $\arg \min f_1 + f_2 \neq \emptyset$ .

On souhaite trouver :  $\hat{x} = \arg \min_x f_1(x) + f_2(x)$

$x$  est solution si et seulement si :

$$\begin{cases} x = \text{prox}_{\gamma f_2} y \\ \text{prox}_{\gamma f_2} y = \text{prox}_{\gamma f_1} (2\text{prox}_{\gamma f_2} y - y) \end{cases} \quad (\text{III.69})$$

La démonstration de l'équation III.69 est décrite ci-dessous :

#### Preuve III.6

$$\min_x \{f_1(x) + f_2(x)\}$$

$$0 \in \partial(f_1 + f_2)(x)$$

$$\gamma \cdot 0 \in \partial(\gamma f_1 + \gamma f_2)(x), \quad \gamma > 0$$

$$0 \in \gamma \partial(f_1)(x) + \gamma \partial(f_2)(x)$$

$$2x \in x + \gamma \partial(f_1)(x) + x + \gamma \partial(f_2)(x)$$

$$2x \in (I + \gamma \partial f_1)(x) + (I + \gamma \partial f_2)(x) \dots \dots (*)$$

Supposons :  $y \in (I + \gamma \partial f_2)(x)$

$$y \in (I + \gamma \partial f_2)(x) \Rightarrow (I + \gamma \partial f_2)^{-1} y \in \underbrace{(I + \gamma \partial f_2)^{-1} (I + \gamma \partial f_2)(x)}_I$$

$$\underbrace{(I + \gamma \partial f_2)^{-1} y}_{\text{prox}_{\gamma f_2}} = x \Rightarrow x = \text{prox}_{\gamma f_2} y \dots \dots (**)$$

(\*) dans (\*\*):  $2x \in (I + \gamma \partial f_1)(x) + \underbrace{(I + \gamma \partial f_2)(x)}_y \Rightarrow 2x \in (I + \gamma \partial f_1)(x) + y$

$$2x - y = (I + \gamma \partial f_1)(x)$$

$$(I + \gamma \partial f_1)^{-1} (2x - y) = \underbrace{(I + \gamma \partial f_1)^{-1} (I + \gamma \partial f_1)(x)}_I \Rightarrow \underbrace{(I + \gamma \partial f_1)^{-1} (2x - y)}_{\text{prox}_{\gamma f_1}} = x$$

$$x = \text{prox}_{\gamma f_1} (2x - y) \dots \dots (***)$$

(\*\*) dans (\*\*\*) :  $x = \text{prox}_{\gamma f_1} (2\text{prox}_{\gamma f_2} y - y) \dots \dots$  ce qu'il faut démontrer.

Finalement :

$$0 \in \partial f_1(x) + \partial f_2(x) \Leftrightarrow \begin{cases} x = \text{prox}_{\gamma f_2} y \\ x = \text{prox}_{\gamma f_1} (2\text{prox}_{\gamma f_2} y - y) \end{cases}$$

### III.6.3.2 Résolution du problème

On procède comme suit :

$$y = \text{prox}_{\gamma f_2} x \quad (\text{III.70})$$

$$\text{prox}_{\gamma f_1} (2y - x) = y \quad (\text{III.71})$$

$$-\text{prox}_{\gamma f_1} (2y - x) + y = 0 \quad (\text{III.72})$$

Nous appliquons la méthode de la descente du gradient, nous obtenons :

$$-\frac{dx}{dt} = -\text{prox}_{\gamma f_1} (2y - x) + y \quad (\text{III.73})$$

Après discrétisation de la relation III.73, nous obtenons la relation III.74 :

$$\frac{x_{n+1} - x_n}{\tau_n} = \text{prox}_{\gamma f_1} (2y_n - x_n) - y_n \quad (\text{III.74})$$

$$x_{n+1} = x_n + \tau_n \cdot (\text{prox}_{\gamma f_1} (2y_n - x_n) - y_n) \quad (\text{III.75})$$

avec :

$$y_n = \text{prox}_{\gamma f_2} x_n \quad (\text{III.76})$$

Les auteurs dans [Combettes and Pesquet, 2011] ont proposé un algorithme III.8 explicatif de la méthode DR :

---

#### Algorithme III.8 Douglas Rachford

---

- 1: **Fixer**  $\varepsilon \in ]0, 1[$ ,  $\gamma > 0$  et  $x_0 \in H$
  - 2: **Pour**  $n = 0, 1, \dots$  :
    - $y_n = \text{prox}_{\gamma f_2} x_n$
    - $\tau_n \in [\varepsilon, 2 - \varepsilon]$
    - $x_{n+1} = x_n + \tau_n (\text{prox}_{\gamma f_1} (2y_n - x_n) - y_n)$
- 

- Pour trouver une solution au problème de débruitage sous contrainte, nous transformons l'équation III.56 comme suit :

$$\begin{aligned} & \arg \min_u |\nabla u| \\ & \text{sous - contrainte} \\ & \|Ku - f\|^2 \leq \varepsilon \end{aligned} \quad (\text{III.77})$$

et nous considérons la contrainte comme une fonction indicatrice et  $\varepsilon$  est le rayon  $r$

de la boule de l'équation III.1, puis nous appliquons l'algorithme III.8, cette solution est appelée 'DR1'.

Les opérateurs proximaux de la contrainte et de la fonction objectif sont donnés respectivement par les équations III.51 et III.53.

- Par contre, pour trouver une solution au problème III.60, nous suivons les étapes de l'algorithme III.8 qui nécessitent le calcul des opérateurs proximaux des fonctionnelles  $E_{data}(u) = \frac{\lambda}{2}\|Ku - f\|^2$  et  $E_{reg}(u) = |\nabla u|$ , ces opérateurs sont donnés respectivement par :

$$\text{Pr } ox_{\gamma E_{data}} u = \frac{\lambda}{2} K^*(I + \gamma K^* K)^{-1} \quad (\text{III.78})$$

et

$$\text{Pr } ox_{\gamma E_{reg}} u = \max \left( 0, 1 - \frac{\gamma}{|\nabla u|} \right) \nabla u \quad (\text{III.79})$$

Cette solution est appelée 'DR2'.

Pour une meilleure appréciation des résultats théoriques, nous renforçons notre théorie par quelques résultats de simulations issus de la méthode DR.

Une comparaison des résultats obtenus par DR1 et DR2 sera effectuée.

**Remarque III.6.1** *L'algorithme de Douglas-Rachford permet de gérer des problèmes de débruitage en présence de bruit non nécessairement additif gaussien.*

Les figures III.21, III.22, III.23, III.24 montrent le débruitage des images Cameraman', 'Clock', 'Parrort' et 'Satellite' en appliquant la première méthode 'DR1', c'est à dire lorsque le problème est avec contraintes.

L'illustration de la deuxième méthode 'DR2', c'est à dire lorsque le problème est sans contraintes, est démontrée dans les figures III.25, III.26, III.27, III.28 appliquée sur les mêmes images.



(a) Originale

(b) Bruitée  $\sigma = 10$ 

(c) Masquée



(d) Restaurée

(e) Bruitée  $\sigma = 20$ 

(f) Masquée



(g) Restaurée

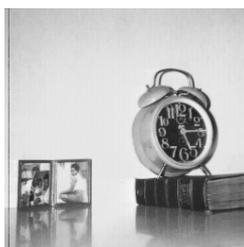
(h) Bruitée  $\sigma = 30$ 

(i) Masquée

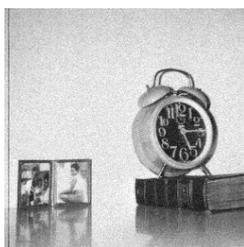
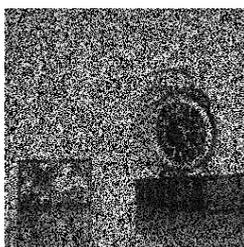


(j) Restaurée

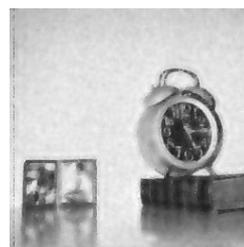
FIGURE III.21 – Débruitage de l'image Cameraman de dimension  $256 \times 256$  par l'algorithme Douglas-Rachford avec contraintes lorsque l'image est corrompue par un bruit gaussien de variance  $\sigma = 10$ ,  $\sigma = 20$ ,  $\sigma = 30$



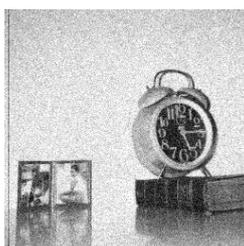
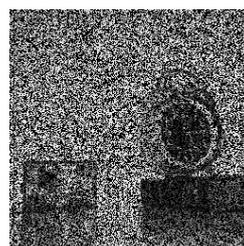
(a) Originale

(b) Bruitée  $\sigma = 10$ 

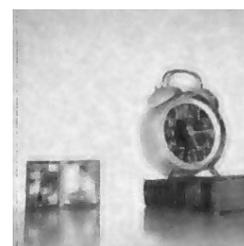
(c) Masquée



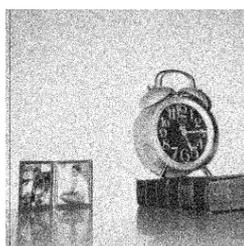
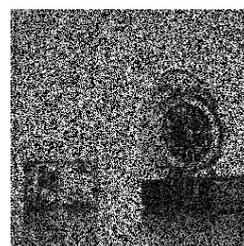
(d) Restaurée

(e) Bruitée  $\sigma = 20$ 

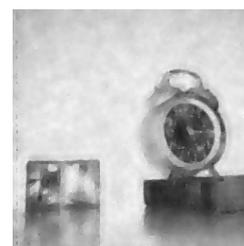
(f) Masquée



(g) Restaurée

(h) Bruitée  $\sigma = 30$ 

(i) Masquée

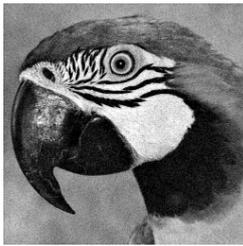
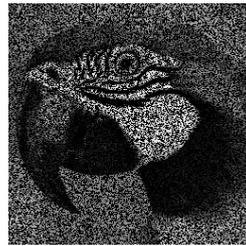


(j) Restaurée

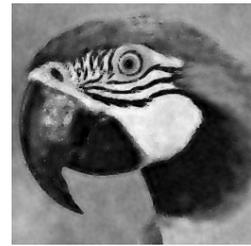
FIGURE III.22 – Débruitage de l'image Clock de dimension  $256 \times 256$  par l'algorithme Douglas-Rachford avec contraintes lorsque l'image est corrompue par un bruit gaussien de variance  $\sigma = 10$ ,  $\sigma = 20$ ,  $\sigma = 30$



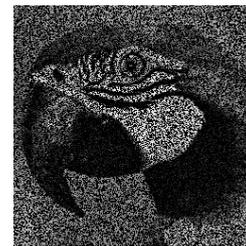
(a) Originale

(b) Bruitée  $\sigma = 10$ 

(c) Masquée



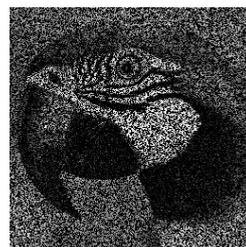
(d) Restaurée

(e) Bruitée  $\sigma = 20$ 

(f) Masquée



(g) Restaurée

(h) Bruitée  $\sigma = 30$ 

(i) Masquée



(j) Restaurée

FIGURE III.23 – Débruitage de l'image parrot de dimension  $256 \times 256$  par l'algorithme Douglas-Rachford avec contraintes lorsque l'image est corrompue par un bruit gaussien de variance  $\sigma = 10$ ,  $\sigma = 20$ ,  $\sigma = 30$

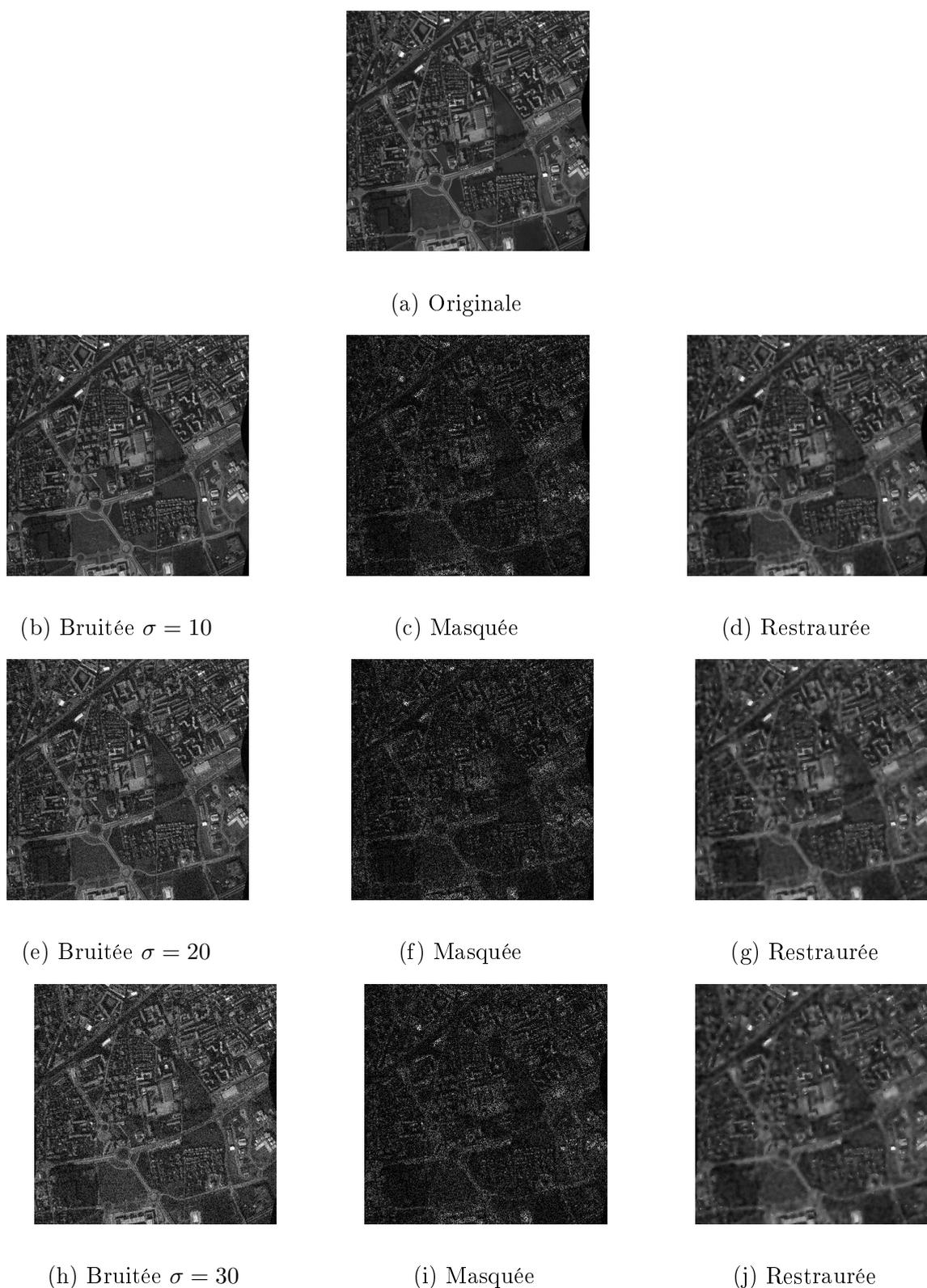


FIGURE III.24 – Débruitage de l'image Satellite de dimension  $512 \times 512$  par l'algorithme Douglas-Rachford avec contraintes lorsque l'image est corrompue par un bruit gaussien de variance  $\sigma = 10$ ,  $\sigma = 20$ ,  $\sigma = 30$



(a) Originale

(b) Bruitée  $\sigma = 10$ 

(c) Masquée



(d) Restaurée

(e) Bruitée  $\sigma = 20$ 

(f) Masquée



(g) Restaurée

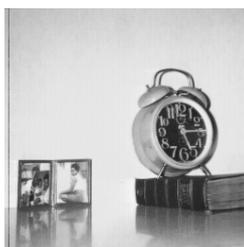
(h) Bruitée  $\sigma = 30$ 

(i) Masquée

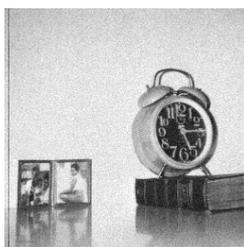
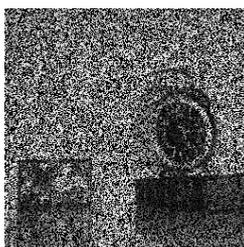


(j) Restaurée

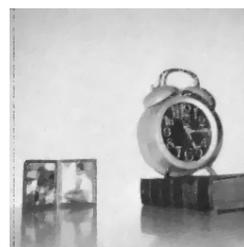
FIGURE III.25 – Débruitage de l'image Cameraman de dimension  $256 \times 256$  par l'algorithme Douglas-Rachford sans contraintes lorsque l'image est corrompue par un bruit gaussien de variance  $\sigma = 10$ ,  $\sigma = 20$ ,  $\sigma = 30$



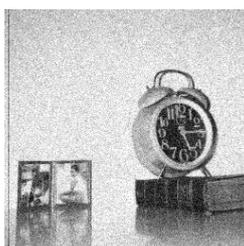
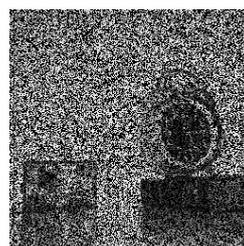
(a) Originale

(b) Bruitée  $\sigma = 10$ 

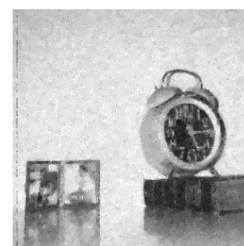
(c) Masquée



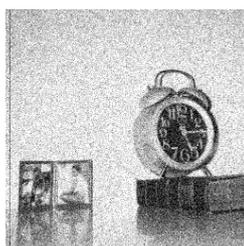
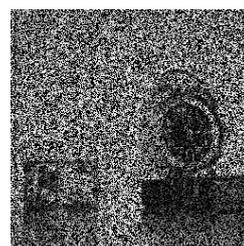
(d) Restaurée

(e) Bruitée  $\sigma = 20$ 

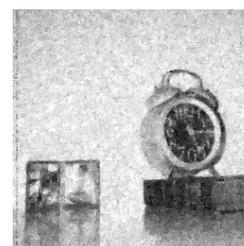
(f) Masquée



(g) Restaurée

(h) Bruitée  $\sigma = 30$ 

(i) Masquée

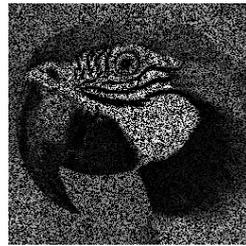


(j) Restaurée

FIGURE III.26 – Débruitage de l'image Clock de dimension  $256 \times 256$  par l'algorithme Douglas-Rachford sans contraintes lorsque l'image est corrompue par un bruit gaussien de variance  $\sigma = 10$ ,  $\sigma = 20$ ,  $\sigma = 30$



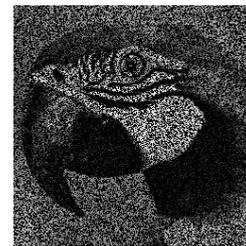
(a) Originale

(b) Bruitée  $\sigma = 10$ 

(c) Masquée



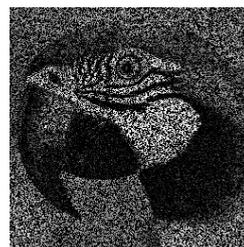
(d) Restaurée

(e) Bruitée  $\sigma = 20$ 

(f) Masquée



(g) Restaurée

(h) Bruitée  $\sigma = 30$ 

(i) Masquée



(j) Restaurée

FIGURE III.27 – Débruitage de l'image parrot de dimension  $256 \times 256$  par l'algorithme Douglas-Rachford sans contraintes lorsque l'image est corrompue par un bruit gaussien de variance  $\sigma = 10$ ,  $\sigma = 20$ ,  $\sigma = 30$

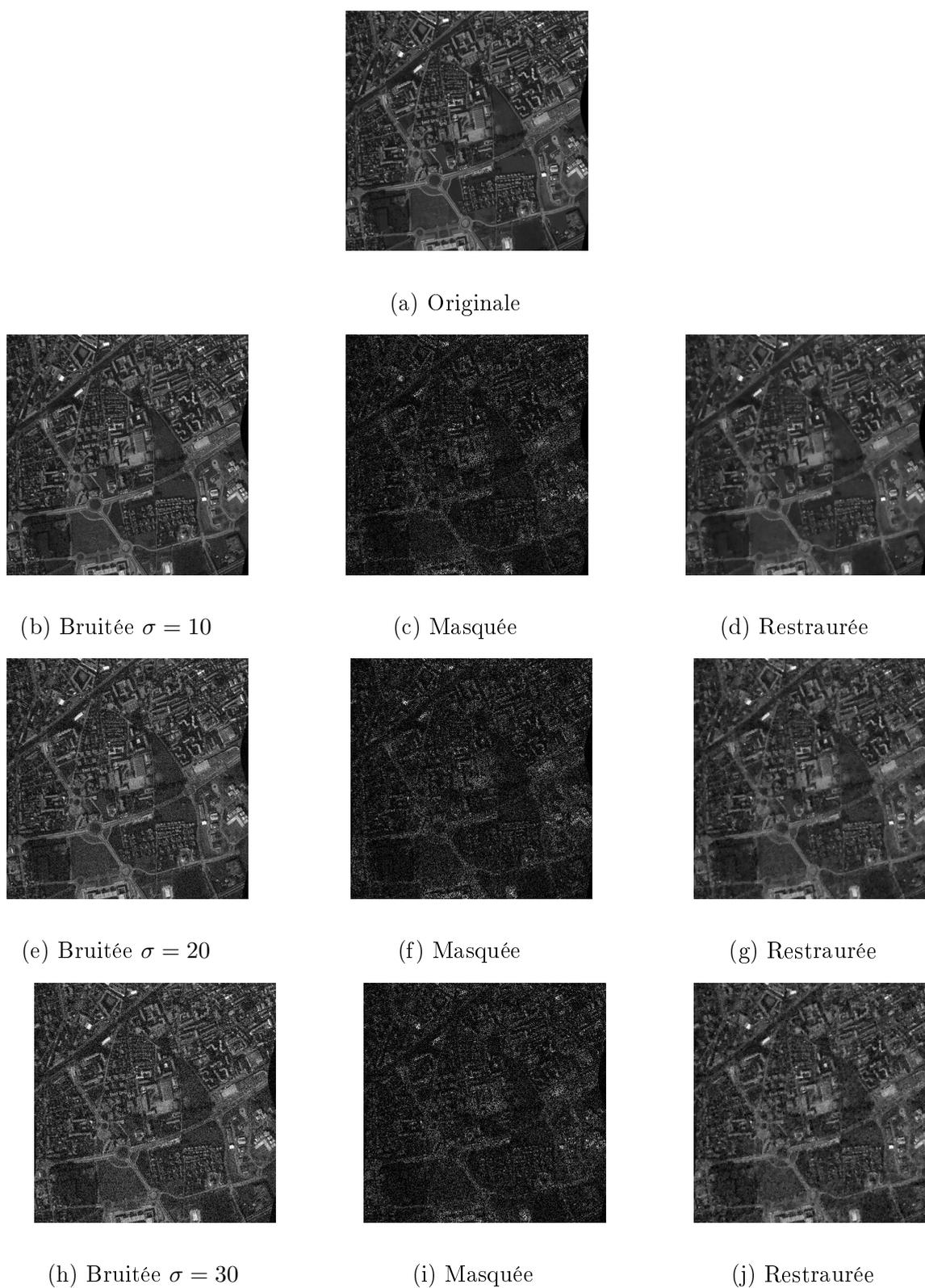


FIGURE III.28 – Débruitage de l'image Satellite de dimension  $512 \times 512$  par l'algorithme Douglas-Rachford sans contraintes lorsque l'image est corrompue par un bruit gaussien de variance  $\sigma = 10$ ,  $\sigma = 20$ ,  $\sigma = 30$

## III.7 Conclusion

Les approches proximales ont jusqu'à présent montré leur efficacité sur des problèmes spécifiques de débruitage en présence de bruit de Gauss. Une interprétation des deux algorithmes basés sur la méthode de point proximal  $FB$  et  $DR$  a donc été proposée et appuyée par des résultats de simulation. Dans le prochain chapitre nous nous intéressons à la résolution du problème de débruitage en présence d'un bruit multiplicatif de type Gamma par les algorithmes proximaux.

Une heure de lecture est le souverain  
remède contre les dégoûts de la vie

---

Montesquieu

## IV.1 Introduction

Les algorithmes proximaux ont montré leurs intérêts pour la résolution des problèmes inverses, ces algorithmes permettent de minimiser un critère composé de plusieurs fonctions convexes. Dans le chapitre III, nous avons montré l'efficacité des approches proximales Forward Backward et Douglas Rachford sur des problèmes spécifiques de restauration par filtrage en présence du bruit de Gauss. La question qui se pose c'est comment adapter ces approches à la restauration d'image bruitée par un bruit non additif?

Dans cette thèse, notre intérêt portera sur les approches variationnelles, le problème est de chercher à minimiser des critères convexes composés d'une somme de deux fonctions, la variation totale comme étant une fonction de régularisation et le terme d'attache aux données lié au bruit de type Gamma. Le problème provient du calcul de l'opérateur proximal associé au terme d'attache aux données.

Notons que l'algorithme de Douglas-Rachford permet de gérer des problèmes de débruitage en présence du bruit, non nécessairement additif gaussien. Nous proposons de calculer l'opérateur proximal du terme d'attache aux données lié au bruit multiplicatif Gamma, puis montrer que l'utilisation de cet opérateur dans l'algorithme de DR permet de gérer efficacement des problèmes de débruitage en présence de bruit multiplicatif.

Dans un premier temps, cinq méthodes basées que ce soit sur les EDP ou le variationnel, sont développées et implémentées, ce sont les travaux de [Yu and Acton, 2002],

[Shi and Osher, 2008], [Bernardes et al., 2010], [Dong and Zeng, 2013] et [Zhao et al., 2014]. Dans la suite de ce chapitre, nous allons donc présenter l’algorithme développé au cours de cette thèse, ainsi que son application à des images réelles.

## IV.2 Méthodes de comparaison

Afin d’évaluer les performances de notre méthode, nous allons la comparer à cinq autres algorithmes : l’algorithme inspiré des travaux de [Yu and Acton, 2002] SRAD cité dans l’état de l’art présenté dans le chapitre II section II.7.5, l’algorithme SO [Shi and Osher, 2008] basé sur le modèle de ROF décrit dans [Rudin et al., 2003], l’algorithme de la diffusion complexe adaptative [Bernardes et al., 2010], l’algorithme de la méthode variationnelle convexe [Dong and Zeng, 2013] et enfin l’algorithme de la méthode d’optimisation convexe [Zhao et al., 2014].

### IV.2.1 Filtrage d’image par la méthode de Yu et Acton

Dans [Yu and Acton, 2002], Yu et Acton ont proposé une méthode combinant la diffusion anisotrope et le coefficient de variation, intitulée SRAD, pour Speckle Reducing Anisotropic Diffusion. Pour plus de détails, on se réfère à l’algorithme II.3 qui se trouve dans la sous section II.7.5 du chapitre II. (Voir aussi la référence [Diffellah et al., 2012]). La version proposée dans [Diffellah et al., 2019] consiste à utiliser la distribution gamma au lieu de celle de Rayleigh proposée par les auteurs [Yu and Acton, 2002].

Cette méthode effectue un lissage des zones homogènes tout en renforçant les contours.

### IV.2.2 Filtrage d’image par la méthode de ROF multiplicative

Dans le chapitre II section II.8.3, nous avons expliqué et implémenté l’algorithme de débruitage d’images corrompues par un bruit additif selon le modèle de ROF. Rudin, Lions et Osher ‘RLO’ dans [Rudin et al., 2003] ont proposé une méthode de débruitage d’images corrompues par un bruit multiplicatif en appliquant la descente du gradient, ce sont les premiers qui ont utilisé la variation totale comme terme de régularisation. L’algorithme C.1 dans l’annexe C relate les étapes à suivre lors du déroulement de la méthode dite variationnelle, suivi par un développement mathématique cité dans la preuve C.1.

- La fonction objectif est nonconvexe et elle est très onéreuse pour trouver un minimiseur global au problème C.5.
- Le modèle C.6 proposé dans [Rudin et al., 2003] pose le problème de l’unicité et le problème de la convergence de l’algorithme numérique.

Dans [Shi and Osher, 2008], les auteurs ont considéré l’image observée comme étant  $\log f = \log u + \log \eta$ , ci dessous l’algorithme IV.1 qui permet de récupérer l’image restaurée : Les

**Algorithme IV.1 Méthode de ROF multiplicative**

1: Transformer le problème multiplicatif à un problème additif  $\underbrace{f = u \cdot \eta}_{\text{Produit}} \Rightarrow \underbrace{F = U + N}_{\text{Somme}}$  en appliquant la fonction logarithmique :

$$— f = u \cdot \eta \Rightarrow \log f = \log(u \cdot \eta) \Rightarrow \underbrace{\log f}_F = \underbrace{\log u}_U + \underbrace{\log \eta}_N \Rightarrow F = U + N$$

2: Appliquer ROF additif pour chercher  $U$  (Voir la section II.8.3)

3: Passage de la fonction logarithmique à la fonction l'exponentielle pour trouver l'image recherchée  $u$  :

$$u = e^U$$

hypothèses de ce modèle ne sont pas cohérentes, ci dessous une preuve démonstrative :

**Preuve IV.1**

— Pour la multiplication :  $f = u \times \eta$

Les conditions sont :

— La valeur moyenne du bruit est égale à 1 donc  $E(\eta) = 1$

—  $u$  et  $f$  sont indépendants :  $f = u \times \eta \Rightarrow E(f) = E(u) \times E(\eta)$

$$\begin{cases} E(f) = E(u) \times E(\eta) \\ E(\eta) = 1 \end{cases} \Rightarrow E(f) = E(u)$$

L'hypothèse trouvée est :  $E(f) = E(u) \dots (a)$

— Pour l'addition :  $F = U + B$  avec  $U = \log u$ ,  $F = \log f$ ,  $B = \log \eta$

— La valeur moyenne du bruit est égale à 0 donc  $E(B) = 0$

—  $F = U + B \Rightarrow E(F) = E(U) + E(B)$

$$\begin{cases} E(f) = E(U) + E(B) \\ E(B) = 0 \end{cases} \Rightarrow E(F) = E(U) \text{ L'hypothèse trouvée est : } E(F) = E(U)$$

— de L'inégalité de Jensen on a :  $e^{E(B)} \leq E(e^B)$

$$— \begin{cases} e^{E(B)} \leq E(e^B) \\ E(B) = 0 \end{cases} \Rightarrow e^0 \leq E(e^B) \Rightarrow 1 \leq E(e^B)$$

$$— \begin{cases} B = \log \eta \Rightarrow \eta = e^B \\ 1 \leq E(e^B) \end{cases} \Rightarrow 1 \leq E(\eta) \Rightarrow E(\eta) \geq 1$$

Pour le cas où  $E(\eta) > 1$  :

$$\begin{cases} E(\eta) > 1 \\ E(\eta) = \frac{E(f)}{E(u)} \end{cases} \Rightarrow \frac{E(f)}{E(u)} > 1 \Rightarrow E(f) > E(u) \dots (b)$$

Conclusion : on voit clairement la non-cohérence entre les équations (a) et (b).

### IV.2.3 Diffusion non linéaire complexe améliorée : diffusion complexe adaptative

Dans [Bernardes et al., 2010], la diffusion non linéaire complexe est donnée par :

$$\frac{\partial u}{\partial t} = \nabla \cdot (D \nabla u) = \text{div} (D \nabla u) \quad (\text{IV.1})$$

$\text{div}$  ou  $\nabla \cdot$  est la divergence,  $\nabla$  le gradient et selon les auteurs,  $D$  le coefficient de diffusion donné par la formule suivante :

$$D = \frac{e^{j\theta}}{1 + \left( \frac{\text{Im}(u)}{\kappa\theta} \right)^2} \quad (\text{IV.2})$$

$\theta$  est l'angle de phase proche de zéro,  $\text{Im}(u)$  est la partie imaginaire de  $u$ ,  $\kappa$  est le paramètre de seuil donné par l'équation :

$$\kappa = \kappa_{\max} + (\kappa_{\min} - \kappa_{\max}) \frac{g - \min(g)}{\max(g) - \min(g)} \quad (\text{IV.3})$$

$\min(g)$  et  $\max(g)$  représentent le minimum et le maximum de  $g$ .  $g$  est définie par :

$$g = G_{N,\sigma} * \text{Re}(u) \quad (\text{IV.4})$$

$\text{Re}(u)$  est la partie réelle de  $u$ ,  $*$  est l'opérateur de convolution,  $G_{N,\sigma}$  une gaussienne de moyenne de kernel de taille  $N \times N$ ,  $\kappa_{\min}$  et  $\kappa_{\max}$  sont des constantes réelles telles que  $\kappa_{\min} < \kappa_{\max}$ .

Selon les auteurs dans [Bernardes et al., 2010], le coefficient de diffusion peut être donné approximativement par l'équation :

$$D \simeq \frac{1}{1 + \left( \frac{\Delta u}{k} \right)^2} \quad (\text{IV.5})$$

La figure IV.1 illustre les allures du coefficient de diffusion  $D$  pour  $k = 0.5, 2, 5, 10$  :

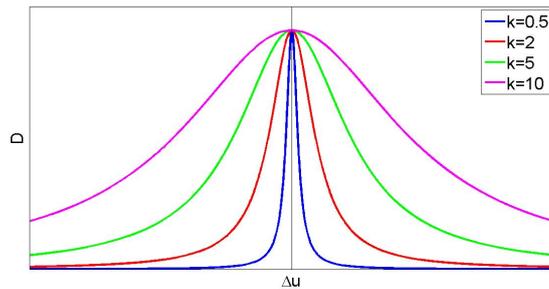


FIGURE IV.1 – Coefficient de diffusion  $D$  en fonction de  $\Delta u$  de selon les valeurs de  $k$

Après discrétisation par la méthode des différences finies, l'équation à implémenter est la suivante :

$$u_{i,j}^{n+1} = u_{i,j}^n + \Delta t^n \left( \bar{D}_{i,j}^n \Delta_h u_{i,j}^n + \nabla_h D_{i,j}^n \nabla_h u_{i,j}^n \right) \quad (\text{IV.6})$$

avec :

$$\bar{D}_{i,j}^n = \frac{4D_{i,j}^n + D_{i-1,j}^n + D_{i+1,j}^n + D_{i,j-1}^n + D_{i,j+1}^n}{8} \quad (\text{IV.7})$$

$\Delta_h$  et  $\nabla_h$  sont le laplacien et l'opérateur gradient discret respectivement.

#### IV.2.4 Modèle variationnel convexe

On a vu dans la sous-section IV.2.2, que les auteurs dans [Rudin et al., 2003] ont démontré que la récupération de l'image  $u$  était basée sur la résolution du problème d'optimisation sous contraintes suivant :

$$\begin{aligned} \min_u \quad & \int_{\Omega} |Du| \\ \text{Sous - contraintes} \quad & \int_{\Omega} \frac{f}{u} dx \\ & \int_{\Omega} \left( \frac{f}{u} - 1 \right)^2 dx = \sigma^2 \end{aligned} \quad (\text{IV.8})$$

La variation totale est utilisée comme fonction objectif afin de conserver des bords significatifs dans les images et seules les propriétés statistiques de base, la moyenne et la variance du bruit sont utilisées comme contraintes, ce qui limite en quelque sorte les résultats restaurés. Pour ces raisons, les auteurs dans [Aubert and Aujol, 2008] ont proposé un modèle variationnel en se référant à la règle de Bayes et à la distribution gamma de moyenne égale à 1 en appliquant l'estimation par Maximum A Posteriori MAP :

$$\min_u \quad \int_{\Omega} \left( \log(u) + \frac{f}{u} \right) dx + \lambda \int_{\Omega} |Du| \quad (\text{IV.9})$$

$\lambda > 0$  le paramètre de régularisation.

Les deux formules IV.8 et IV.9 sont non-convexes.

Le modèle IV.8 proposé dans [Rudin et al., 2003] pose le problème de l'unicité et le problème de la convergence de l'algorithme numérique. Pour remédier à ce problème, les auteurs [Dong and Zeng, 2013] ont proposé un nouveau modèle convexe en ajoutant un terme de pénalité quadratique basé sur les propriétés statistiques du bruit Gamma multiplicatif.

Ils ont proposé le modèle suivant :

$$\min_u \int_{\Omega} \left( \log(u) + \frac{f}{u} \right) dx + \alpha \int_{\Omega} \left( \sqrt{\frac{u}{f}} - 1 \right)^2 dx + \lambda \int_{\Omega} |Du| \quad (\text{IV.10})$$

Le modèle IV.10 est convexe, les auteurs [Dong and Zeng, 2013] ont proposé l'algorithme prima-dual pour résoudre ce problème de minimisation.

### IV.2.5 Modèle variationnel convexe multiplicatif

Le modèle convexe proposé par les auteurs dans [Zhao et al., 2014] est le suivant :

$$\min_{w,u} \frac{1}{2} \|w - \mu e\|_2^2 + \alpha_1 \|Gw - Hu\|_1 + \alpha_2 \|Du\|_2 \quad (\text{IV.11})$$

avec  $f = NKu$ ,  $N$  matrice diagonale du bruit  $\eta_i$ ,  $w_i = \frac{1}{\eta_i}$ ,  $F$  matrice diagonale de l'image bruitée  $f_i$ , donc  $Fw = Ku$ ,  $\mu$  peut être défini comme la valeur moyenne de  $w$ ,  $e$  est un vecteur dont tous les éléments sont égaux à 1 et  $\alpha_1$ ,  $\alpha_2$  sont deux paramètres de régularisation positifs pour contrôler l'équilibre entre les trois termes de la fonction objectif.

- Le premier terme du modèle IV.11 permet de mesurer la variance de  $w$ , et par conséquent, nous aimerions déterminer  $w$  et  $f$  dans le modèle de telle sorte que la variance de  $w$  puisse être minimisée.
- Le deuxième terme est le terme d'attache aux données entre l'image observée  $u$  et  $w$ .
- Le troisième terme est le terme de régularisation de  $u$ .

La résolution du problème IV.11 peut être faite par la méthode de direction alternée des multiplicateurs [Gabay and Mercier, 1976], les méthodes de Douglas–Rachford dans [Combettes and Wajs, 2005], la méthode de Bregman [Goldstein and Osher, 2009], les méthodes primal-dual [Esser et al., 2010] - [Chambolle and Pock, 2011] et les méthodes de division proximale [Combettes and Pesquet, 2011].

Les techniques de la direction alternée des multiplicateurs [Gabay and Mercier, 1976] ont été utilisées par les auteurs [Zhao et al., 2014].

## IV.3 Méthode proposée

Avant de décrire et implémenter la méthode proposée, il est nécessaire de commencer d'abord à donner un rappel sur la loi Gamma afin de tirer la densité de probabilité du

bruit utilisé pour la contamination de l'image, puis passer à l'estimateur de MAP pour extraire le modèle de débruitage.

### IV.3.1 Rappel de la loi Gamma

Une variable aléatoire continue  $X$  suit une loi Gamma de paramètres  $\alpha, \beta \in \mathbb{R}_+^*$  ce que l'on la note aussi  $X \sim \Gamma(\alpha, \beta)$ , si elle admet pour densité de probabilité la fonction :

$$g(x; \alpha, \beta) = \begin{cases} x^{\alpha-1} \frac{\beta^\alpha e^{-\beta x}}{\Gamma(\alpha)} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases} \quad (\text{IV.12})$$

Dans la formule IV.12,  $\Gamma$  est la fonction Gamma d'Euler définie par :

$$\Gamma(x) = \int_0^{+\infty} e^{-u} u^{x-1} du \quad (\text{IV.13})$$

$g(x; \alpha, \beta)$  est la première formulation de la densité de probabilité de la loi Gamma,  $x$  est une variable aléatoire qui suit la loi Gamma,  $\alpha$  est le paramètre de forme strictement positif et  $\beta$  est le paramètre d'intensité strictement positif.

Alternativement, la distribution Gamma peut être paramétrée d'un paramètre de forme  $k = \alpha$  et d'un paramètre d'échelle  $\theta = \frac{1}{\beta}$  :

$$g(x; k, \theta) = \begin{cases} x^{k-1} \frac{e^{-\frac{x}{\theta}}}{\theta^k \Gamma(k)} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases} \quad (\text{IV.14})$$

ce que l'on la note aussi  $X \sim \Gamma(k, \theta)$ . L'équation IV.14 décrit la deuxième formule de la distribution Gamma.

Une récapitulation des deux formulations de la distribution Gamma est donnée dans le tableau IV.1 :

Tableau IV.1 – Les deux formulations de la distribution Gamma

|                              | Première formulation   | Deuxième formulation  |
|------------------------------|--|---|
| Densité $g(x; k, \theta)$    | $\begin{cases} x^{k-1} \frac{e^{-\frac{x}{\theta}}}{\theta^k \Gamma(k)} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$ | $\begin{cases} x^{\alpha-1} \frac{\beta^\alpha e^{-\beta x}}{\Gamma(\alpha)} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$ |
| Paramètres                   | $k, \theta$  | $\alpha, \beta$   |
| Espérance $E[X]$             | $k \cdot \theta$   | $\frac{\alpha}{\beta}$  |
| Variance $\sigma^2 = Var[X]$ | $k \cdot \theta^2$   | $\frac{\alpha}{\beta^2}$  |

**Remarque IV.3.1** ( Densité de probabilité  $g(x; L)$  ) Si  $E[X] = 1$  alors :

- $k = \frac{1}{\theta} = L$
- $\alpha = \beta = L$ ,
- la variance devient  $\sigma^2 = \text{Var}[X] = \frac{1}{L}$
- et la densité de probabilité sera noté  $g(x; L)$  telle que :

$$g(x; L) = \begin{cases} x^{L-1} \frac{L^L e^{-Lx}}{\Gamma(L)} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases} \quad (\text{IV.15})$$

**Remarque IV.3.2** La famille des distributions Gamma inclut entre autres les lois exponentielles et Poisson :

- La loi Gamma de paramètres  $\alpha > 0$  et  $\beta = 1$ , s'appelle la loi exponentielle de paramètre 1, sa densité de probabilité est :

$$f(x) = \lambda e^{-\lambda x}; \quad x > 0 \quad (\text{IV.16})$$

- La loi Gamma de paramètres  $\alpha = 1$  et  $\beta = n+1$ , s'appelle loi d'Erlang de paramètre  $n$ , elle est liée à la loi de Poisson et sa densité de probabilité est :

$$f(x) = \frac{1}{n!} x^n e^{-x}; \quad x \geq 0 \quad (\text{IV.17})$$

### IV.3.2 Bruit multiplicatif de la loi Gamma

L'image bruitée s'écrit en général sous forme :

$$f = u \times \eta \quad (\text{IV.18})$$

$f$  est l'image bruitée (observée),  $u$  est l'image originale (estimée),  $\eta$  est le bruit multiplicatif de loi Gamma de moyenne égale à 1. On note le bruit comme suit :

$$\eta \sim \Gamma \left( \sigma = \frac{1}{\sqrt{L}}, \mu = 1 \right) \quad (\text{IV.19})$$

$\Gamma$  est la loi Gamma,  $\sigma$  est l'écart type et  $\mu$  est la valeur moyenne de loi Gamma.

Pour  $x = \eta$ , l'équation IV.15 devient :

$$p(\eta; L) = \begin{cases} \eta^{L-1} \frac{L^L e^{-L\eta}}{\Gamma(L)} & \text{si } \eta > 0 \\ 0 & \text{si } \eta \leq 0 \end{cases} \quad (\text{IV.20})$$

de l'équation IV.18 :

$$\eta = \frac{f}{u}, \quad \eta \geq 0 \quad (\text{IV.21})$$

Remplacer IV.21 dans l'équation IV.20 donne ce qui suit :

$$p\left(\frac{f}{u}; L\right) = \frac{L^L}{(u)^{L-1}\Gamma(L)} f^{L-1} \cdot e^{-L\left(\frac{f}{u}\right)} \quad (\text{IV.22})$$

l'équation IV.22 est appelée densité de probabilité de la loi Gamma.

**Théorème IV.3.1 (Théorème de Bayes)** *Le théorème de Bayes permet de déterminer la probabilité de  $u$  sachant  $f$ , si on connaît les probabilités de  $u$ ,  $f$  et  $f$  sachant  $u$ . Pour aboutir au théorème de Bayes, on part de la définition de la probabilité conditionnelle :*

$$p\left(\frac{u}{f}\right) \cdot p(f) = p\left(\frac{f}{u}\right) \cdot p(u) \quad (\text{IV.23})$$

En divisant les deux membres par  $p(f)$ , on obtient :

$$p\left(\frac{u}{f}\right) = \frac{p\left(\frac{f}{u}\right) \cdot p(u)}{p(f)} \quad (\text{IV.24})$$

$p\left(\frac{u}{f}\right)$  est la probabilité a posteriori de  $u$  sachant  $f$  ou probabilité de  $u$  sous condition de  $f$ , elle est postérieure au sens qu'elle dépende directement de  $f$ .

$p\left(\frac{f}{u}\right)$  étant la fonction de vraisemblance de  $u$ , pour un  $f$  connu.

$p(u)$  est la probabilité a priori de  $u$ , elle est antérieure au sens qu'elle précède toute information sur  $f$ , appelée aussi la probabilité marginale de  $u$ .

$p(f)$  est la probabilité marginale ou a priori de  $f$ .

### IV.3.3 Estimateur du Maximum A Posteriori (MAP)

L'image inconnue  $u$ , peut être calculée à l'aide de l'estimateur du Maximum A Posteriori (MAP), on souhaite trouver un  $\hat{u} \in \arg \max_u p\left(\frac{u}{f}\right)$ .

$$\hat{u} \in \arg \max_u p\left(\frac{u}{f}\right) \Leftrightarrow \hat{u} \in \arg \min_u \left\{ -\log \left( p\left(\frac{u}{f}\right) \right) \right\} \quad (\text{IV.25})$$

L'équation IV.24 dans IV.25

$$\hat{u} \in \arg \min_u \left\{ -\log \left[ p\left(\frac{f}{u}\right) \right] - \log [p(u)] + \underbrace{\log [p(f)]}_{\text{indépendante de } u} \right\} \quad (\text{IV.26})$$

$$\hat{u} \in \arg \min_u \left\{ -\log \left[ p\left(\frac{f}{u}\right) \right] - \log [p(u)] \right\} \quad (\text{IV.27})$$

Le premier terme de l'équation IV.27 correspond au terme d'attache aux données  $f$  et dépend du modèle de dégradation (équation IV.22), le second terme correspond à un

terme d'a priori sur l'image originale.

— Développement du terme d'attache aux données  $\arg \min_u \{-\log [p(\frac{f}{u})]\}$  :

$$\arg \min_u \left\{ -\log \left[ p \left( \frac{f}{u} \right) \right] \right\} = \arg \min_u \left\{ -\log \left[ \frac{L^L}{\Gamma(L)} f^{L-1} \right] + L \frac{f}{u} + (L-1) \log u \right\} \quad (\text{IV.28})$$

$$\arg \min_u \left\{ -\log \left[ p \left( \frac{f}{u} \right) \right] \right\} \simeq \arg \min_u \left\{ L \frac{f}{u} + (L-1) \log u \right\} \quad (\text{IV.29})$$

$$\arg \min_u \left\{ -\log \left[ p \left( \frac{f}{u} \right) \right] \right\} \simeq \arg \min_u \left\{ L \left( \frac{f}{u} + \log u \right) \right\} \quad (\text{IV.30})$$

— Développement du deuxième terme  $\arg \min_u \{-\log [p(u)]\}$  :  
la probabilité a priori de  $u$ , est donnée par :

$$p(u) = \frac{1}{Z} e^{-\Gamma \phi(u)} \quad (\text{IV.31})$$

$Z$  est le terme de stabilisation,  $\Gamma$  un paramètre positif et  $\phi(u)$  une fonction convexe.

$$\arg \min_u \{-\log [p(u)]\} = \arg \min_u \{\log(Z) + \Gamma \phi(u)\} \quad (\text{IV.32})$$

$$\arg \min_u \{-\log [p(u)]\} = \arg \min_u \{\Gamma \phi(u)\} \quad (\text{IV.33})$$

— Remplaçons les équations IV.30 et IV.33 dans l'équation IV.27, nous obtenons :

$$\max_u \left\{ p \left( \frac{u}{f} \right) \right\} = \arg \min_u \left\{ -\log \left( p \left( \frac{u}{f} \right) \right) \right\} = \arg \min_u \left\{ L \left( \log u + \frac{f}{u} \right) + \Gamma \phi(u) \right\} \quad (\text{IV.34})$$

$$\max_u \left\{ p \left( \frac{u}{f} \right) \right\} = \arg \min_u \left\{ -\log \left( p \left( \frac{u}{f} \right) \right) \right\} = \arg \min_u \left\{ \left( \log u + \frac{f}{u} \right) + \frac{\Gamma}{L} \phi(u) \right\} \quad (\text{IV.35})$$

**Remarque IV.3.3** Pour notre cas le terme  $\phi(u)$  présente une régularisation de type variation totale  $TV$ .

Si on pose  $\lambda = \frac{\Gamma}{L}$  et  $\phi(u) = TV(u)$ , le critère à minimiser s'écrira sous la forme :

$$\hat{u} \in \arg \min_u \left\{ \int \left( \log u + \frac{f}{u} \right) + \alpha.TV(u) \right\} \quad (\text{IV.36})$$

Ci-dessous le tableau IV.2 regroupe les avantages et les inconvénients de l'utilisation de la variation totale en débruitage :

Tableau IV.2 – Avantages et inconvénients de l'utilisation de la variation totale

|               |  |
|---------------|--|
| Avantages     | La variation totale permet d'obtenir des images débruitées de bonne qualité, principalement si l'image originale est de type « bande-dessinée », c'est à dire contenant beaucoup d'aplats. |
| Inconvénients | Pour des images naturelles, la variation totale présente des artefacts donnant un effet « de marche d'escalier » (staircasing effect).   |

### IV.3.4 Algorithme proposé

Dans cette sous section, nous formulons le problème d'optimisation pour restaurer l'image dégradée par un bruit multiplicatif de type Gamma, puis nous proposons les étapes à suivre pour minimiser la fonctionnelle énergie et nous montrons l'intérêt du proximal pour assurer l'unicité et la convergence de notre problème qui s'écrit sous forme :

$$\hat{u} \in \arg \min_u \left\{ \int \left( \log u + \frac{f}{u} \right) + \lambda \int |\nabla u| \right\} \quad (\text{IV.37})$$

Rappelons que :

$$E_{data}(u) = \int \left( \log u + \frac{f}{u} \right) \quad (\text{IV.38})$$

$$E_{reg}(u) = \int |\nabla u| \quad (\text{IV.39})$$

$E_{reg}(u)$  and  $E_{data}(u)$  sont convexes et semi continues inférieurement. Nous devons chercher le proximal des deux fonctions puis appliquer l'algorithme de Douglas Rachford [Diffellah et al., 2019].

#### IV.3.4.1 Première formulation de débruitage par une regularization variationnelle lorsque le bruit multiplicatif suit la loi Gamma

L'idée principale est de subdiviser le problème en deux fonctions :

—  $E_{reg}(u)$  est la variation totale.

$$E_{reg}(u) = \|u\|_{TV} = \|u\|_1 \quad (\text{IV.40})$$

—  $E_{data}(u)$  est la fonction indicatrice de l'ensemble convexe  $C$  définie comme suit :

$$E_{data}(u) = i_C(u) = \begin{cases} 0 & \text{if } \left\| \log u + \frac{f}{u} \right\|_2 \leq \varepsilon \\ +\infty & \text{ailleurs} \end{cases} \quad (\text{IV.41})$$

La première formulation pour écrire le problème décrit par l'expression IV.37 est la suivante :

$$\begin{aligned} & \arg \min_u \|u\|_{TV} \\ & \text{sous - contraintes} \\ & \left\| \log u + \frac{f}{u} \right\|_2 \leq \varepsilon \end{aligned} \tag{IV.42}$$

Pour résoudre le problème IV.42, nous appliquons l'algorithme IV.2 ci-dessous :

---

#### Algorithme IV.2 Squelette de l'algorithme proposé : Formulation1

---

1: Définir le proximal de  $E_{reg}(u)$  :

$$Prox_{\gamma E_{reg}} u = \arg \min_z \frac{1}{2} \|u - z\|_2^2 + \gamma \|z\|_1$$

2: Définir le proximal de  $E_{data}(u)$  :

$$Prox_{\gamma E_{data}} u = \arg \min_z \frac{1}{2} \|u - z\|_2^2 + \gamma i_C(z)$$

$Prox_{\gamma E_{data}} u \equiv$  c'est la projection sur la boule  $B2$  de rayon  $\varepsilon$

( $\varepsilon$  est le rayon de la boule  $B2$  'B2 ball')

$$\arg \min_z \|u - z\|_2^2 \text{ subject to } \left\| \log z + \frac{f}{z} \right\|_2 \leq \varepsilon$$

3: Appliquer la méthode de Douglas-Rachford.

---

#### IV.3.4.2 Deuxième formulation de débruitage par une regularization variationnelle lorsque le bruit multiplicatif suit la loi Gamma

Pour ce cas, l'expression IV.42 s'écrit comme suit :

$$\arg \min_u \left\{ \|u\|_{TV} + \lambda \left\| \log u + \frac{f}{u} \right\|_2 \right\} \tag{IV.43}$$

Les deux fonctions s'écrivent respectivement :

$$— E_{reg}(u) = \|u\|_{TV} = \|u\|_1$$

$$— \text{et } E_{data}(u) = \left\| \log u + \frac{f}{u} \right\|_2$$

Pour résoudre le problème IV.43, nous proposons l'algorithme IV.3 suivant :

---

#### Algorithme IV.3 Squelette de l'algorithme proposé : Formulation2

---

1: Définir le proximal de  $E_{reg}(u)$  :

$$Prox_{\gamma E_{reg}} u = \arg \min_z \frac{1}{2} \|u - z\|_2^2 + \gamma \|z\|_1$$

2: Définir le proximal de  $E_{data}(u)$  :

$$Prox_{\gamma E_{data}} u = \arg \min_z \frac{1}{2} \|u - z\|_2^2 + \gamma \left\| \log z + \frac{f}{z} \right\|_2$$

3: Appliquer la méthode de Douglas-Rachford.

---

## IV.3.4.3 L'algorithme de Douglas-Rachford

La méthode de Douglas-Rachford est utilisée pour résoudre un problème d'optimisation convexe de la forme :

$$x \in \arg \min_x f_1(x) + f_2(x)$$

$$f_1(x) : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}, f_2(x) : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$$

$$(\text{ridom} f_1) \cap (\text{ridom} f_2) \neq \emptyset$$

$f_1(x), f_2(x)$  sont convexes et non-lisses.

$$\text{Prox}_{f_1} x : \mathbb{R}^N \rightarrow \mathbb{R}^N$$

$$\text{Prox}_{f_1} x = \min_{y \in \mathbb{R}^N} f_1(y) + \frac{1}{2} \|x - y\|^2 \quad (\text{IV.44})$$

$$\text{Prox}_{f_2} x : \mathbb{R}^N \rightarrow \mathbb{R}^N$$

$$\text{Prox}_{f_2} x = \min_{y \in \mathbb{R}^N} f_2(y) + \frac{1}{2} \|x - y\|^2 \quad (\text{IV.45})$$

Pour décrire les étapes de la méthode de Douglas-Rachford, nous proposons l'algorithme IV.4 ci-dessous :

---

**Algorithme IV.4 Algorithme détaillé de Douglas-Rachford**


---

- 1: **Initialisation :** Fixer  $\varepsilon \in ]0, 1[$ ,  $\gamma > 0$ ,  $u_0 \in \mathbb{R}^N$   
 $n = 0$
- 2: **Itérations de Douglas Rachford :**  
 $y_n = \text{prox}_{\gamma E_{data}} u_n$   
 $\tau_n \in [\varepsilon, 2 - 2\varepsilon]$   
 $u_{n+1} = u_n + \tau_n (\text{prox}_{\gamma E_{reg}} (2y_n - u_n) - y_n)$
- 3: **Calcul de la tolérance :**  
 $Tol = u_{n+1} - u_n$
- 4: **Test :**  
**Si**  $u_{n+1} - u_n < \varepsilon$  **alors**  $u_n$  **est solution**  $\hat{u} = u_{n+1}$ .  
**Sinon**  $n := n + 1$  **retour à 2.**

Avec :

$$\text{prox}_{\gamma E_{reg}}(u) = \arg \min_z \frac{1}{2} \|u - z\|^2 + \gamma E_{reg}(z)$$

$$\text{prox}_{\gamma E_{data}}(u) = \arg \min_z \frac{1}{2} \|u - z\|^2 + \gamma E_{data}(z)$$


---

## IV.3.4.4 La fonction Edata

Comme  $\nabla E_{data} = \frac{u-f}{u^2}$  et  $\nabla^2 E_{data} = \frac{2f-u}{u^3}$  et si  $0 < u < 2f$ , nous pouvons démontrer que  $E_{data}$  est strictement convexe ; le minimiseur est unique :  $u = f$ , la fonction  $\log u + \frac{f}{u}$  passe par un minimum égal à  $1 + \log f$ . Voir les tracés IV.2 ci dessous des courbes représentatives :

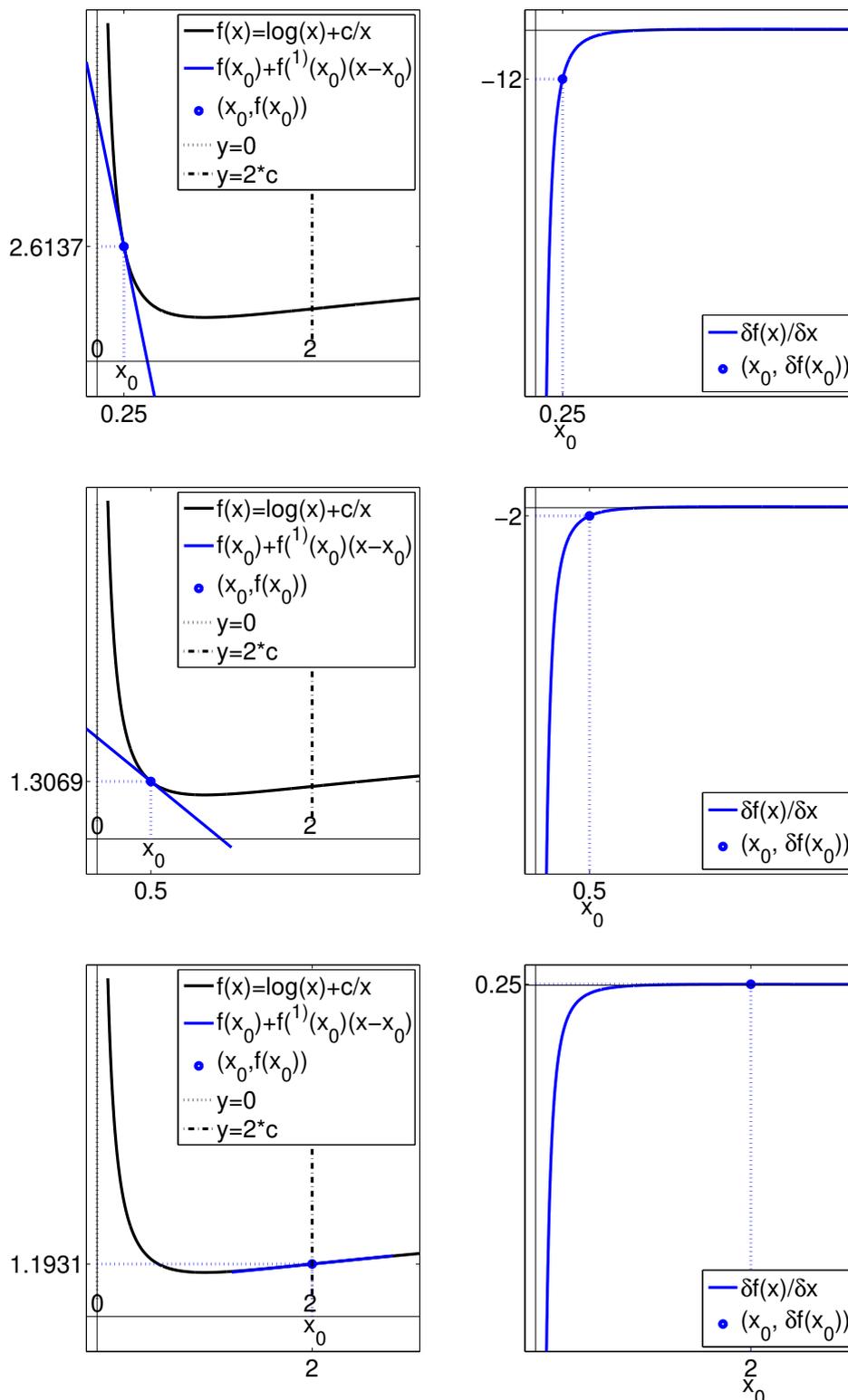


FIGURE IV.2 – Tracés de la fonction  $\log x + \frac{c}{x}$  pour  $c = 1$ , sa dérivée et sa tangente

- si  $0 < f < 2c$  la tangente est au dessous de la fonction  $f$ ,
- mais si  $f$  dépasse  $2c$  alors la tangente sera au dessus de la fonction  $f$ , dans ce cas  $f$  ne sera plus convexe.

- La tangente et la fonction  $f$  se coupent au point  $(x_0, f(x_0))$ ,
- Le nombre dérivé de  $f$  en  $x_0$  correspond au coefficient directeur de la tangente en  $(x_0, f(x_0))$

#### IV.3.4.5 Existence de ProxEdata

D'après [Combettes and Wajs, 2005], [Combettes and Pesquet, 2007a], [Chaux et al., 2007] et [Combettes and Pesquet, 2011], les auteurs ont montré que lorsque  $\phi \in \Gamma_0$ ,  $\alpha \in \mathbb{R}$ ,  $\tau \geq 0$ ,  $w > 0$  et si

$$\phi(x) = \begin{cases} -\kappa \ln(x) + \frac{\tau x^2}{2} + \alpha x & \text{si } x > 0 \\ 0 & \text{ailleurs} \end{cases} \quad (\text{IV.46})$$

alors

$$\text{prox}_{\phi} x = \frac{1}{2(1+\tau)} \left( x - \alpha + \sqrt{|x - \alpha|^2 + 4\kappa(1+\tau)} \right) \quad (\text{IV.47})$$

Si  $\tau = 0$  les équations IV.46 et IV.47 deviennent respectivement :

$$\phi(x) = \begin{cases} -\kappa \ln(x) + \alpha x & \text{si } x > 0 \\ 0 & \text{ailleurs} \end{cases} \quad (\text{IV.48})$$

$$\text{prox}_{\phi} x = \frac{1}{2} \left( x - \alpha + \sqrt{|x - \alpha|^2 + 4\kappa} \right) \quad (\text{IV.49})$$

Si on pose  $x = \frac{1}{u}$ , alors l'équation IV.46 s'écrit :

$$\phi(u) = \begin{cases} -\kappa \ln\left(\frac{1}{u}\right) + \frac{\alpha}{u} & \text{si } \frac{1}{u} > 0 \\ 0 & \text{ailleurs} \end{cases} = \begin{cases} \kappa \ln(u) + \frac{\alpha}{u} & \text{si } u < 1 \\ 0 & \text{ailleurs} \end{cases} \quad (\text{IV.50})$$

Si on prend  $\kappa = 1$  et  $\alpha = f$  dans l'équation IV.50, on obtient l'équation ci dessous :

$$\phi(u) = \ln(u) + \frac{f}{u}, \quad u < 1 \quad (\text{IV.51})$$

On déduit de l'équation IV.49 que le proximal de l'équation IV.51 est le suivant :

$$\text{Prox}_{E_{data}}(u) = \frac{\frac{1}{u} - f + \sqrt{\left|\frac{1}{u} - f\right|^2 + 4}}{2} \quad (\text{IV.52})$$

#### IV.3.4.6 Recherche du ProxEreg

Montrons que  $E_{reg}$  est semi continue inférieurement et non différentiable. La figure IV.3 représente les courbes de la fonction  $y = |x|$ , sa dérivée et sa tangente :

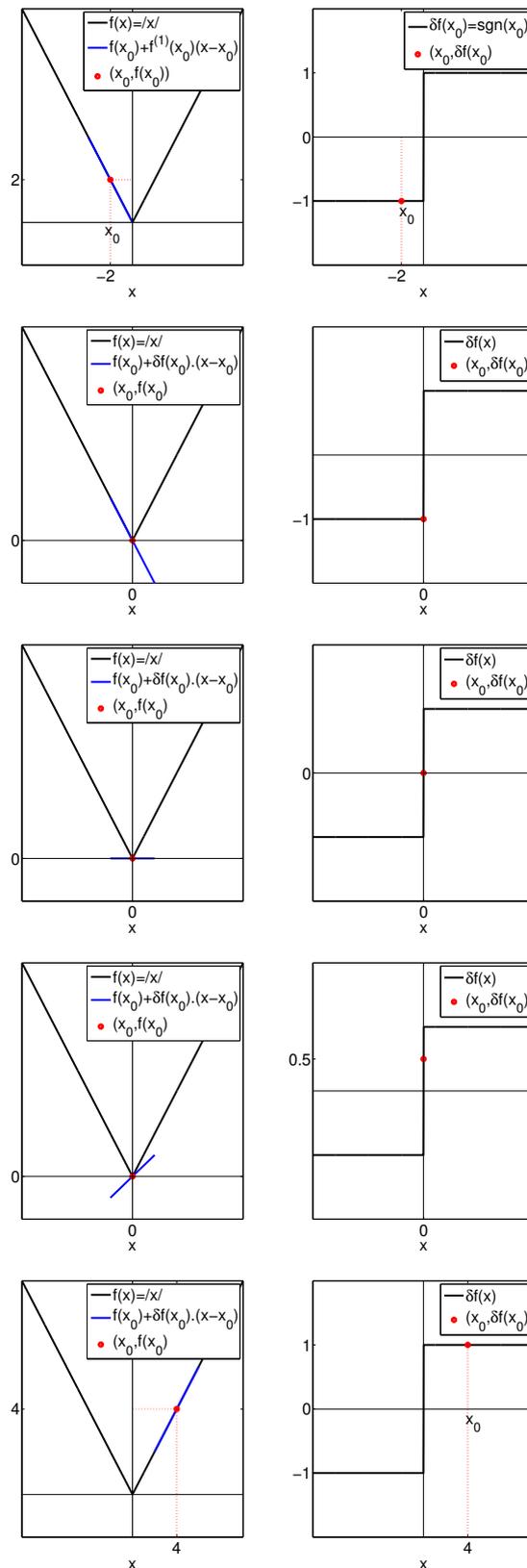


FIGURE IV.3 – Tracés de la fonction  $y = |x|$ , sa dérivée et sa tangente

- La fonction  $f(x) = |x|$  est continue donc elle est semi continue inférieurement s.c.i.
- La limite étudiée pour savoir si  $f$  est dérivable en  $x_0 = 0$  est exprimée de la façon

suivante :

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} = \frac{|h|}{h} \tag{IV.53}$$

— Limite à droite  $\lim_{h \rightarrow 0^+} \frac{f(x_0+h)-f(x_0)}{h} = \lim_{h \rightarrow 0^+} \frac{|h|}{h} = \frac{h}{h} = 1$

— Limite à gauche  $\lim_{h \rightarrow 0^-} \frac{f(x_0+h)-f(x_0)}{h} = \lim_{h \rightarrow 0^-} \frac{|h|}{h} = \frac{-h}{h} = -1$

Les limites à droite et à gauche en 0 du rapport ne sont pas égales, on ne peut parler de limite en 0. La fonction  $f(x) = |x|$  n'est donc pas dérivable en 0.

Par contre :

- La courbe admet deux demi-tangentes en 0.
- Une demi-tangente à gauche du coefficient directeur  $-1$ .
- Une demi-tangente à droite du coefficient directeur  $1$ .
- Cas particulier : si le nombre dérivé est nul, la tangente, dont le coefficient directeur est alors nul, est horizontale.

La figure IV.4 montre que  $y = |x|$  est continue et non différentiable en 0.

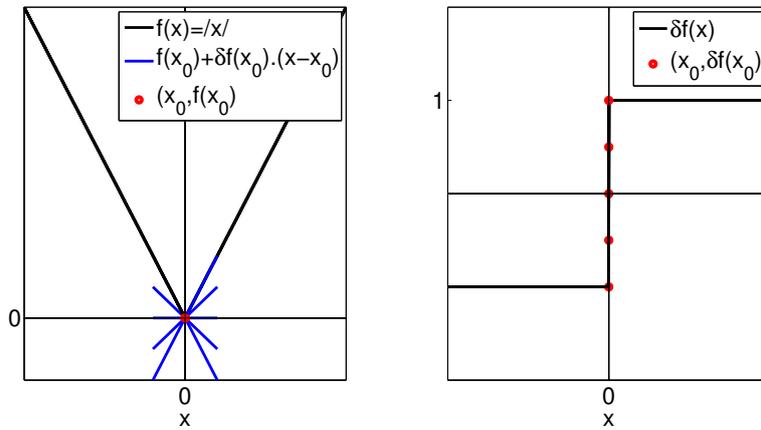


FIGURE IV.4 – Tracés de la fonction  $y = |x|$  , sa dérivée et sa tangente en  $x_0 = 0$

$E_{reg}u$  est une fonction convexe non-différentiable, cherchons son proximal  $\text{Pr } ox_{E_{reg}}u$  :

$$\text{Pr } ox_{E_{reg}}u = \text{Pr } ox_{TV}u = \text{Pr } ox_{\|u\|_1}u \tag{IV.54}$$

$$\text{Pr } ox_{\gamma E_{reg}}(x) = \arg \min_z \frac{1}{2} \|x - z\|^2 + \gamma E_{reg}(z) \tag{IV.55}$$

avec  $\gamma > 0$

$$\text{Pr } ox_{\gamma E_{reg}}x = \max \left( 0, 1 - \frac{\gamma}{|x|} \right) x \tag{IV.56}$$

Le principe général de la méthode proposée peut être résumé par l'algorithme IV.5 :

---

**Algorithme IV.5 Algorithme de la méthode proposée**


---

1: **Initialisation** : Fixer  $\varepsilon \in ]0, 1[$ ,  $\gamma > 0$ ,  $u_0 \in \mathbb{R}^N$

$$n = 0$$

2: **Définir le prox of  $E_{reg}$**  :

$$\text{Prox}_{\gamma E_{reg}} u_n = \max\left(0, 1 - \frac{\gamma}{|u_n|}\right) u_n$$

3: **Définir le prox of  $E_{data}$**  :

$$\text{Prox}_{E_{data}}(u_n) = \frac{1}{2} \left( \frac{1}{u_n} - f + \sqrt{\left| \frac{1}{u_n} - f \right|^2 + 4} \right)$$

4: **Itérations de Douglas Rachford** :

$$y_n = \text{prox}_{\gamma E_{data}} u_n$$

$$\tau_n \in [\varepsilon, 2 - 2\varepsilon]$$

$$u_{n+1} = u_n + \tau_n (\text{prox}_{\gamma E_{reg}}(2y_n - u_n) - y_n)$$

5: **Calcul de la tolérance** :

$$\text{Tol} = u_{n+1} - u_n$$

6: **Test** :

Si  $u_{n+1} - u_n < \varepsilon$  alors  $u_n$  est solution  $\hat{u} = u_{n+1}$ .

Sinon  $n := n + 1$  retour à 2.

---

## IV.4 Résultats de simulation et interprétations

Pour tester la méthode de filtrage proposée et afin d'utiliser le formalisme proximal pour la résolution des problèmes de filtrage en présence de bruit additif ou multiplicatif, nous avons utilisé le UNLocBoX : 'Matlab convex optimization toolbox'<sup>1</sup>, un toolbox implémenté sous Matlab. Les autres méthodes de comparaison ont été testées en les implémentant également sous Matlab.

Nous procédons comme suit :

- Dégrader l'image en générant numériquement un bruit multiplicatif de type Gamma avec différentes valeurs du nombre de look :  $L = 1$ ,  $L = 4$  et  $L = 10$ ,
- débruiter l'image dégradée par notre propre algorithme ainsi que les cinq méthodes de comparaison,
- mesurer les performances des différentes méthodes, en comparant l'image résultante estimée avec l'image d'entrée originale de bonne qualité.

Le tableau IV.3 donne l'appellation des images utilisées ainsi que leurs dimensions.

Pour la simulation, nous utilisons des images réelles en niveau de gris (Voir figure IV.5).

---

1. <https://epfl-lts2.github.io/unlocbox-html/>

Tableau IV.3 – Dimension des images test

| Image       | Dimension en pixels |
|-------------|---------------------|
| 'Cap'       | $512 \times 768$    |
| 'Girl'      | $512 \times 512$    |
| 'Subway'    | $310 \times 450$    |
| 'Cameraman' | $256 \times 256$    |
| 'Parrot'    | $256 \times 256$    |



(a) Girl



(b) Cap



(c) Parrot



(d) Subway



(e) Cameraman

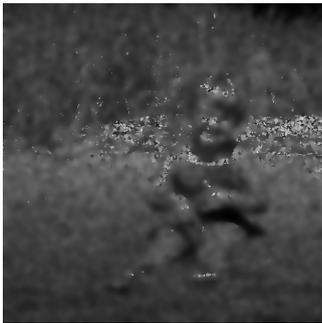
FIGURE IV.5 – Collection des images utilisées.

Les résultats de simulation sont présentés dans les figures [IV.6](#), [IV.7](#), [IV.8](#), [IV.9](#), [IV.10](#), [IV.11](#), [IV.12](#), [IV.13](#), [IV.14](#), [IV.15](#), [IV.16](#), [IV.17](#), [IV.18](#), [IV.19](#), [IV.20](#), chaque figure illustre le résultat de débruitage par la méthode proposée et les différentes méthodes de comparaison pour différentes valeurs de  $L$ ,  $L = 1$ ,  $L = 2$  et  $L = 4$ .

En utilisant les critères présentés dans le chapitre I section [I.5](#), l'évaluation quantitative des résultats obtenus par les différentes méthodes est présentée dans le tableau [IV.4](#).

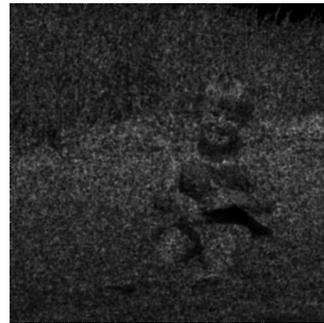


Image originale

Image bruitée  $L = 1$ 

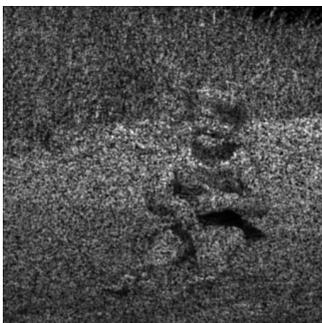
Méthode de ROFM

$PSNR = 16.97dB, SSIM = 0.32, VSNR = 9.81dB$



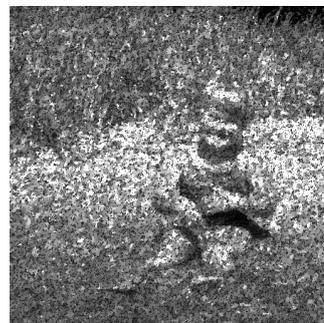
Méthode de SRAD

$PSNR = 10.98dB, SSIM = 0.29, VSNR = 10.80dB$



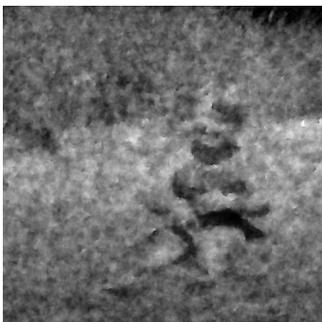
Méthode ACD

$PSNR = 9.94dB, SSIM = 0.26, VSNR = 10.66dB$



Méthode CV

$PSNR = 13.81dB, SSIM = 0.29, VSNR = 4.88dB$



Méthode CVM

$PSNR = 9.86dB, SSIM = 0.04, VSNR = 11.70dB$



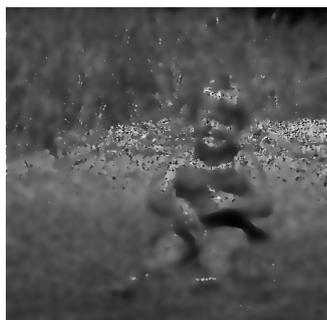
Méthode proposée

$PSNR = 23.66dB, SSIM = 0.82, VSNR = 15.75dB$

FIGURE IV.6 – Résultats de restauration de l'image 'Girl' pour  $L = 1$



Image originale

Image bruitée  $L = 4$ 

Méthode de ROFM

$PSNR = 21.64dB, SSIM = 0.57, VSNR = 12.25dB$



Méthode de SRAD

$PSNR = 12.68dB, SSIM = 0.35, VSNR = 10.86dB$



Méthode ACD

$PSNR = 11.52dB, SSIM = 0.35, VSNR = 10.66dB$



Méthode CV

$PSNR = 20.86dB, SSIM = 0.57, VSNR = 11.22dB$



Méthode CVM

$PSNR = 20.14dB, SSIM = 0.57, VSNR = 6.71dB$



Méthode proposée

$PSNR = 24.69dB, SSIM = 0.88, VSNR = 18.55dB$

FIGURE IV.7 – Résultats de restauration de l'image 'Girl' pour  $L = 4$



Image originale

Image bruitée  $L = 10$ 

Méthode de ROFM

$PNSR = 22.62dB, SSIM = 0.67, VSNR = 13.41dB$



Méthode de SRAD

$PNSR = 13.94dB, SSIM = 0.38, VSNR = 10.91dB$



Méthode ACD

$PNSR = 12.70dB, SSIM = 0.40, VSNR = 10.79dB$



Méthode CV

$PNSR = 22.06dB, SSIM = 0.60, VSNR = 12.34dB$



Méthode CVM

$PNSR = 22.82dB, SSIM = 0.70, VSNR = 12.24dB$



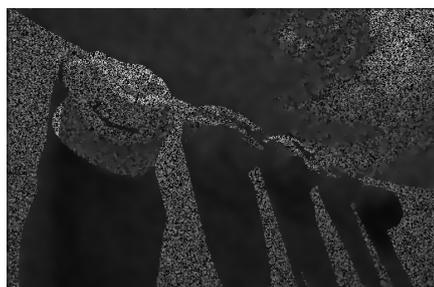
Méthode proposée

$PNSR = 25.76dB, SSIM = 0.91, VSNR = 20.20dB$

FIGURE IV.8 – Résultats de restauration de l'image 'Girl' pour  $L = 10$

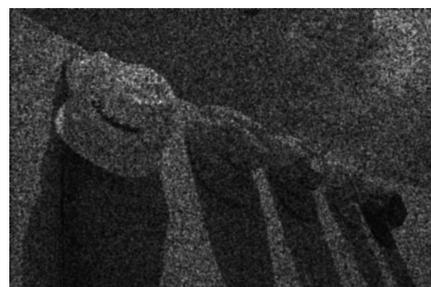


Image originale

Image bruitée  $L = 1$ 

Méthode de ROFM

$PSNR = 15.99dB, SSIM = 0.35, VSNR = 13.46dB$



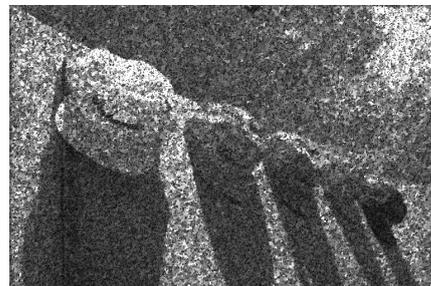
Méthode de SRAD

$PSNR = 11.56dB, SSIM = 0.42, VSNR = 13.06dB$



Méthode ACD

$PSNR = 11.66dB, SSIM = 0.32, VSNR = 14.74dB$



Méthode CV

$PSNR = 15.15dB, SSIM = 0.21, VSNR = 6.11dB$



Méthode CVM

$PSNR = 7.51dB, SSIM = 0.01, VSNR = infdB$



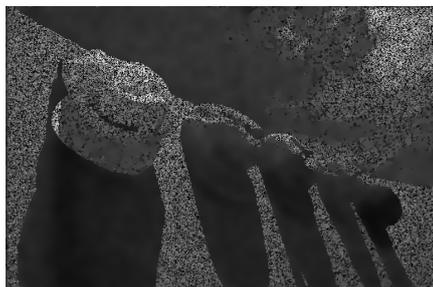
Méthode proposée

$PSNR = 31.77dB, SSIM = 0.93, VSNR = 24.72dB$

FIGURE IV.9 – Résultats de restauration de l'image 'Cap' pour  $L = 1$



Image originale

Image bruitée  $L = 4$ 

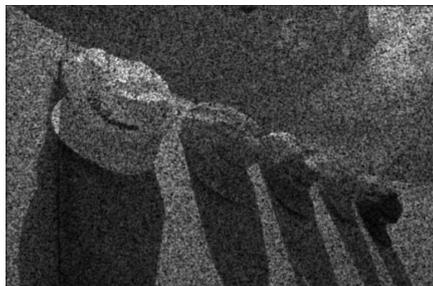
Méthode de ROFM

$PSNR = 23.48dB, SSIM = 0.61, VSNR = 16.78dB$



Méthode de SRAD

$PSNR = 13.09dB, SSIM = 0.49, VSNR = 12.98dB$



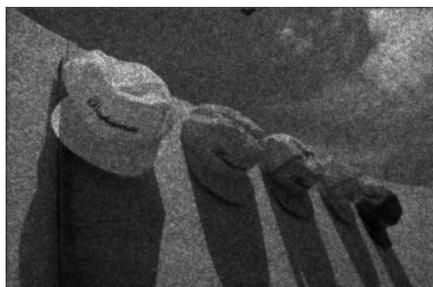
Méthode ACD

$PSNR = 13.33dB, SSIM = 0.38, VSNR = 14.36dB$



Méthode CV

$PSNR = 25.41dB, SSIM = 0.72, VSNR = 15.88dB$



Méthode CVM

$PSNR = 20.87dB, SSIM = 0.50, VSNR = 8.02dB$



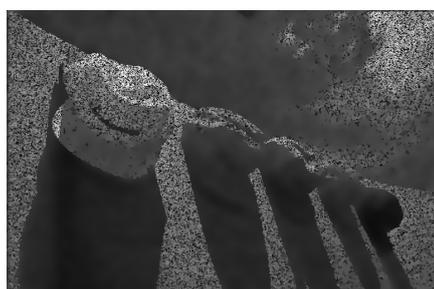
Méthode proposée

$PSNR = 31.55dB, SSIM = 0.94, VSNR = 27.48dB$

FIGURE IV.10 – Résultats de restauration de l'image 'Cap' pour  $L = 4$



Image originale

Image bruitée  $L = 10$ 

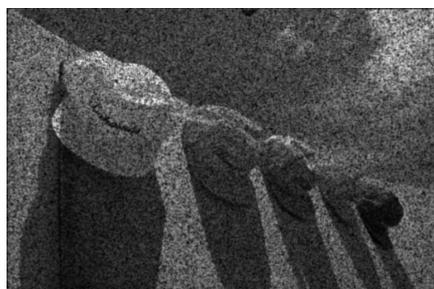
Méthode de ROFM

$PSNR = 27.19dB, SSIM = 0.75, VSNR = 19.25dB$



Méthode de SRAD

$PSNR = 14.53dB, SSIM = 0.57, VSNR = 14.08dB$



Méthode ACD

$PSNR = 14.61dB, SSIM = 0.43, VSNR = 14.32dB$



Méthode CV

$PSNR = 29.50dB, SSIM = 0.86, VSNR = 19.96dB$



Méthode CVM

$PSNR = 26.08dB, SSIM = 0.73, VSNR = 16.41dB$



Méthode proposée

$PSNR = 33.31dB, SSIM = 0.96, VSNR = 29.56dB$

FIGURE IV.11 – Résultats de restauration de l'image 'Cap' pour  $L = 10$



Image originale

Image bruitée  $L = 1$ 

Méthode de ROFM

$PSNR = 16.59dB, SSIM = 0.24, VSNR = 12.18dB$



Méthode de SRAD

$PSNR = 9.71dB, SSIM = 0.27, VSNR = 12.44dB$



Méthode ACD

$PSNR = 9.09dB, SSIM = 0.24, VSNR = 12.91dB$



Méthode CV

$PSNR = 13.70dB, SSIM = 0.23, VSNR = 8.59dB$



Méthode CVM

$PSNR = 8.45dB, SSIM = 0.05, VSNR = 10.57dB$



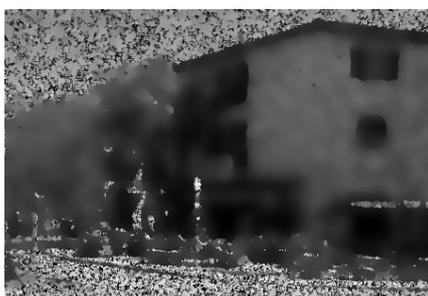
Méthode proposée

$PSNR = 24.51dB, SSIM = 0.80, VSNR = 15.06dB$

FIGURE IV.12 – Résultats de restauration de l'image 'Subway' pour  $L = 1$



Image originale

Image bruitée  $L = 4$ 

Méthode de ROFM

$PSNR = 21.73dB$ ,  $SSIM = 0.41$ ,  $VSNR = 13.76dB$



Méthode de SRAD

$PSNR = 11.45dB$ ,  $SSIM = 0.32$ ,  $VSNR = 12.41dB$



Méthode ACD

$PSNR = 10.68dB$ ,  $SSIM = 0.31$ ,  $VSNR = 12.81dB$



Méthode CV

$PSNR = 20.83dB$ ,  $SSIM = 0.59$ ,  $VSNR = 12.72dB$



Méthode CVM

$PSNR = 19.07dB$ ,  $SSIM = 0.46$ ,  $VSNR = 7.24dB$



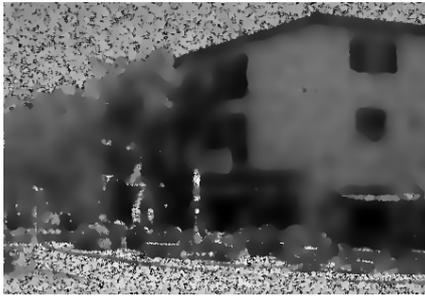
Méthode proposée

$PSNR = 25.27dB$ ,  $SSIM = 0.79$ ,  $VSNR = 16.38dB$

FIGURE IV.13 – Résultats de restauration de l'image 'Subway' pour  $L = 4$



Image originale

Image bruitée  $L = 10$ 

Méthode de ROFM

$PSNR = 22.20dB, SSIM = 0.49, VSNR = 14.34dB$



Méthode de SRAD

$PSNR = 12.86dB, SSIM = 0.35, VSNR = 12.43dB$



Méthode ACD

$PSNR = 11.92dB, SSIM = 0.35, VSNR = 12.91dB$



Méthode CV

$PSNR = 22.72dB, SSIM = 0.69, VSNR = 14.10dB$



Méthode CVM

$PSNR = 22.53dB, SSIM = 0.57, VSNR = 12.44dB$



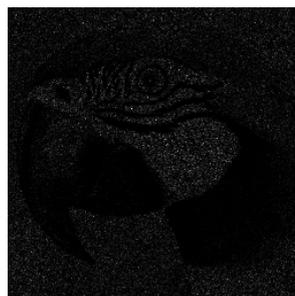
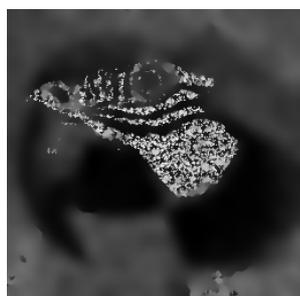
Méthode proposée

$PSNR = 26.54dB, SSIM = 0.84, VSNR = 16.66dB$

FIGURE IV.14 – Résultats de restauration de l'image 'Subway' pour  $L = 10$

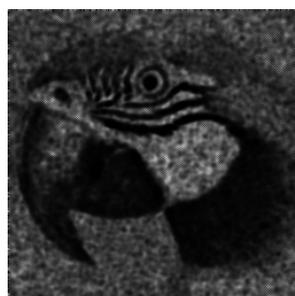


Image originale

Image bruitée  $L = 1$ 

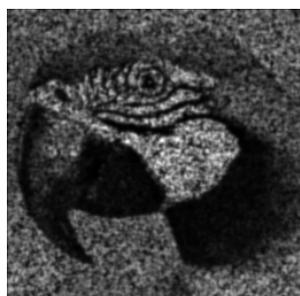
Méthode de ROFM

$PSNR = 15.11dB, SSIM = 0.31, VSNR = 13.52dB$



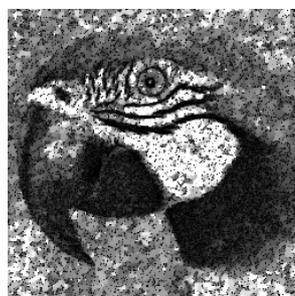
Méthode de SRAD

$PSNR = 10.86dB, SSIM = 0.43, VSNR = 12.56dB$



Méthode ACD

$PSNR = 10.08dB, SSIM = 0.31, VSNR = 12.75dB$



Méthode CV

$PSNR = 14.44dB, SSIM = 0.26, VSNR = 10.26dB$



Méthode CVM

$PSNR = 9.34dB, SSIM = 0.08, VSNR = 5.85dB$



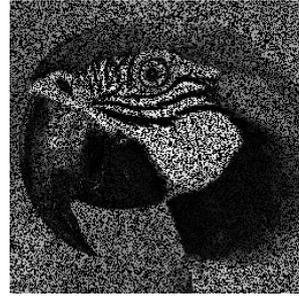
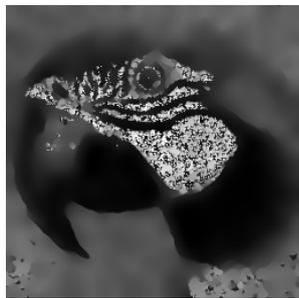
Méthode proposée

$PSNR = 24.96dB, SSIM = 0.83, VSNR = 19.48dB$

FIGURE IV.15 – Résultats de restauration de l'image 'Parrot' pour  $L = 1$



Image originale

Image bruitée  $L = 4$ 

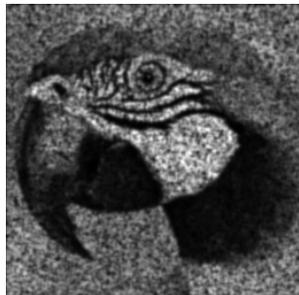
Méthode de ROFM

$PSNR = 20.13dB, SSIM = 0.49, VSNR = 16.14dB$



Méthode de SRAD

$PSNR = 12.57dB, SSIM = 0.50, VSNR = 13.46dB$



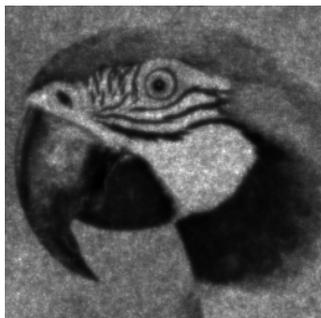
Méthode ACD

$PSNR = 11.61dB, SSIM = 0.38, VSNR = 12.92dB$



Méthode CV

$PSNR = 22.75dB, SSIM = 0.68, VSNR = 17.97dB$



Méthode CVM

$PSNR = 19.87dB, SSIM = 0.51, VSNR = 9.16dB$



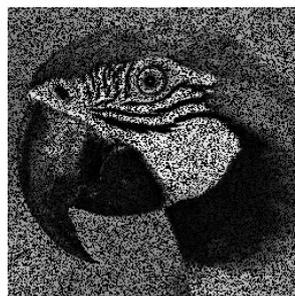
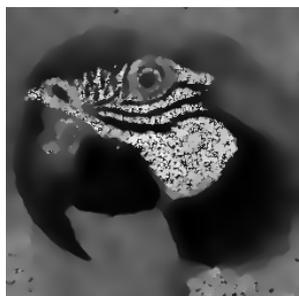
Méthode proposée

$PSNR = 26.58dB, SSIM = 0.83, VSNR = 24.30dB$

FIGURE IV.16 – Résultats de restauration de l'image 'Parrot' pour  $L = 4$

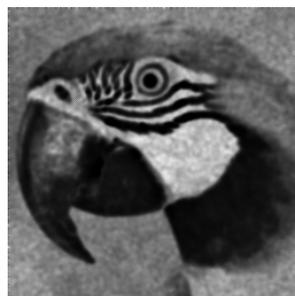


Image originale

Image bruitée  $L = 10$ 

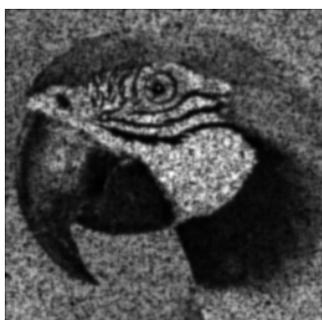
Méthode de ROFM

$PSNR = 21.03dB, SSIM = 0.59, VSNR = 17.61dB$



Méthode de SRAD

$PSNR = 13.99dB, SSIM = 0.54, VSNR = 13.95dB$



Méthode ACD

$PSNR = 12.72dB, SSIM = 0.43, VSNR = 13.14dB$



Méthode CV

$PSNR = 25.60dB, SSIM = 0.79, VSNR = 21.72dB$



Méthode CVM

$PSNR = 23.34dB, SSIM = 0.65, VSNR = 15.23dB$



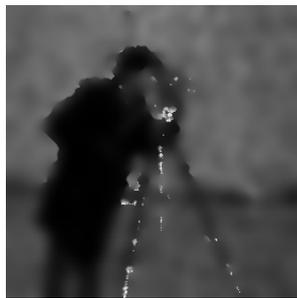
Méthode proposée

$PSNR = 28.12dB, SSIM = 0.87, VSNR = 26.86dB$

FIGURE IV.17 – Résultats de restauration de l'image 'Parrot' pour  $L = 10$



Image originale

Image bruitée  $L = 1$ 

Méthode de ROFM

$PSNR = 17.91dB, SSIM = 0.27, VSNR = 10.62dB$



Méthode de SRAD

$PSNR = 10.90dB, SSIM = 0.47, VSNR = 9.31dB$



Méthode ACD

$PSNR = 9.71dB, SSIM = 0.27, VSNR = 10.31dB$



Méthode CV

$PSNR = 13.78dB, SSIM = 0.21, VSNR = 5.90dB$



Méthode CVM

$PSNR = 7.28dB, SSIM = 0.03, VSNR = 12.08dB$



Méthode proposée

$PSNR = 29.87dB, SSIM = 0.93, VSNR = 23.89dB$

FIGURE IV.18 – Résultats de restauration de l'image 'Cameraman' pour  $L = 1$



Image originale

Image bruitée  $L = 4$ 

Méthode de ROFM

$PSNR = 19.47dB, SSIM = 0.41, VSNR = 13.45dB$



Méthode de SRAD

$PSNR = 12.75dB, SSIM = 0.54, VSNR = 10.37dB$



Méthode ACD

$PSNR = 11.34dB, SSIM = 0.34, VSNR = 10.59dB$



Méthode CV

$PSNR = 22.72dB, SSIM = 0.63, VSNR = 15.56dB$



Méthode CVM

$PSNR = 17.98dB, SSIM = 0.43, VSNR = 6.69dB$



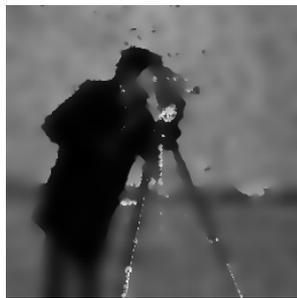
Méthode proposée

$PSNR = 30.14dB, SSIM = 0.93, VSNR = 29.82dB$

FIGURE IV.19 – Résultats de restauration de l'image 'Cameraman' pour  $L = 4$



Image originale

Image bruitée  $L = 10$ 

Méthode de ROFM

$PSNR = 21.58dB, SSIM = 0.50, VSNR = 14.89dB$



Méthode de SRAD

$PSNR = 13.97dB, SSIM = 0.57, VSNR = 10.33dB$



Méthode ACD

$PSNR = 12.46dB, SSIM = 0.37, VSNR = 10.87dB$



Méthode CV

$PSNR = 25.26dB, SSIM = 0.76, VSNR = 18.51dB$



Méthode CVM

$PSNR = 21.72dB, SSIM = 0.57, VSNR = 13.74dB$



Méthode proposée

$PSNR = 32.25dB, SSIM = 0.95, VSNR = 33.92dB$

FIGURE IV.20 – Résultats de restauration de l'image 'Cameraman' pour  $L = 10$

La comparaison des différentes méthodes de filtrage en présence du bruit Gamma pour différentes valeurs de  $L$ , nous a conduit aux remarques suivantes :

- Si le bruit est trop élevé  $L = 1$  : pour une bonne illustration, nous avons pris l'image de 'cap', en effet, nous remarquons que les textures de l'ombre de la casquette ne sont pas clairvoyants. Idem pour le cas des détails et des bords, si nous prenons l'exemple des chaises et les feuilles des arbres, dans l'image 'subway'.
- Si  $L = 10$  : notre méthode restaure les images d'une manière très satisfaisante
  - le texte des casquettes est clair et net après la restauration.
- Le fond : élimination du bruit quelque soit la valeur de  $L$ .

Nous récapitulons que quelque soit la valeur de  $L$ , la méthode proposée surpasse les méthodes de comparaison et elle est beaucoup plus meilleure, effectivement, nous pouvons observer aussi que notre approche permet de retrouver la plupart des détails qui ne sont pas perceptibles à cause de la présence du bruit. Nous avons donc trouvé un juste équilibre entre élimination du bruit et conservation des détails importants.

Le tableau IV.4 présente les valeurs de SNR et SSIM obtenues lorsque la restauration des images est effectuée par la méthode proposée et les cinq méthodes de comparaison . Nous pouvons remarquer que les résultats obtenus tant en PSNR , tant en SSIM, qu'en VSNR nous montrent que la méthode proposée donne le meilleur résultat pour toutes les images de test et quel que soit le bruit.

Nous tenons en compte lors de notre évaluation que le filtre le plus performant est celui qui exige conjointement ce qui suit :

- un  $PSNR$  le plus grand possible.
- un  $SSIM$  le plus proche de 1.
- et un  $VSNR$  le plus grand possible.

Tableau IV.4 – Mesures de qualité des différentes images

| Image     | Methode        | L = 1        |             |              | L = 4        |             |              | L = 10       |             |              |
|-----------|----------------|--------------|-------------|--------------|--------------|-------------|--------------|--------------|-------------|--------------|
|           |                | PSNR         | SSIM        | VSNR         | PSNR         | SSIM        | VSNR         | PSNR         | SSIM        | VSNR         |
| Girl      | "SRAD "        | 10.98        | 0.29        | 10.80        | 12.68        | 0.35        | 10.86        | 13.94        | 0.38        | 10.91        |
|           | " ROF "        | 16.97        | 0.32        | 9.81         | 21.64        | 0.57        | 12.25        | 22.62        | 0.67        | 13.41        |
|           | "Adaptative"   | 9.94         | 0.26        | 10.66        | 11.52        | 0.35        | 10.66        | 12.70        | 0.40        | 10.79        |
|           | "Convex "      | 13.81        | 0.29        | 4.88         | 20.86        | 0.57        | 11.22        | 22.06        | 0.60        | 12.34        |
|           | "Convex mult " | 9.86         | 0.04        | 11.70        | 20.14        | 0.57        | 6.71         | 22.82        | 0.70        | 12.24        |
|           | "Proposée "    | <b>23.66</b> | <b>0.82</b> | <b>15.75</b> | <b>24.69</b> | <b>0.88</b> | <b>18.55</b> | <b>25.76</b> | <b>0.91</b> | <b>20.20</b> |
| Cap       | "SRAD "        | 11.56        | 0.42        | 13.06        | 13.09        | 0.49        | 12.98        | 14.53        | 0.57        | 14.08        |
|           | "ROF "         | 15.99        | 0.35        | 13.46        | 23.48        | 0.61        | 16.78        | 27.19        | 0.75        | 19.25        |
|           | "Adaptative "  | 11.66        | 0.32        | 14.74        | 13.33        | 0.38        | 14.36        | 14.61        | 0.43        | 14.32        |
|           | "Convex"       | 15.15        | 0.21        | 6.11         | 25.41        | 0.72        | 15.88        | 29.50        | 0.86        | 19.96        |
|           | "Convex mult " | 7.51         | 0.01        | inf          | 20.87        | 0.50        | 8.02         | 26.08        | 0.73        | 16.41        |
|           | "Proposée"     | <b>31.77</b> | <b>0.93</b> | <b>24.72</b> | <b>31.55</b> | <b>0.94</b> | <b>27.48</b> | <b>33.31</b> | <b>0.96</b> | <b>29.56</b> |
| Subway    | "SRAD "        | 9.71         | 0.27        | 12.44        | 11.45        | 0.32        | 12.41        | 12.86        | 0.35        | 12.43        |
|           | "ROF "         | 16.59        | 0.24        | 12.18        | 21.73        | 0.41        | 13.76        | 22.20        | 0.49        | 14.34        |
|           | "Adaptative "  | 9.09         | 0.24        | 12.91        | 10.68        | 0.31        | 12.81        | 11.92        | 0.35        | 12.91        |
|           | "Convex "      | 13.70        | 0.23        | 8.59         | 20.83        | 0.59        | 12.72        | 22.72        | 0.69        | 14.10        |
|           | "Convex mult " | 8.45         | 0.05        | 10.57        | 19.07        | 0.46        | 7.24         | 22.53        | 0.57        | 12.44        |
|           | "Proposée "    | <b>24.51</b> | <b>0.80</b> | <b>15.06</b> | <b>25.27</b> | <b>0.79</b> | <b>16.38</b> | <b>26.54</b> | <b>0.84</b> | <b>16.66</b> |
| Parrot    | "SRAD "        | 10.86        | 0.43        | 12.56        | 12.57        | 0.50        | 13.46        | 13.99        | 0.54        | 13.95        |
|           | "ROF"          | 15.11        | 0.31        | 13.52        | 20.13        | 0.49        | 16.14        | 21.03        | 0.59        | 17.61        |
|           | "Adaptative "  | 10.08        | 0.31        | 12.75        | 11.61        | 0.38        | 12.92        | 12.72        | 0.43        | 13.14        |
|           | "Convex "      | 14.44        | 0.26        | 10.26        | 22.75        | 0.68        | 17.97        | 25.60        | 0.79        | 21.72        |
|           | "Convex mult " | 9.34         | 0.08        | 5.85         | 19.87        | 0.51        | 9.16         | 23.34        | 0.65        | 15.23        |
|           | "Proposée "    | <b>24.96</b> | <b>0.83</b> | <b>19.48</b> | <b>26.58</b> | <b>0.83</b> | <b>24.30</b> | <b>28.12</b> | <b>0.87</b> | <b>26.86</b> |
| Cameraman | "SRAD "        | 10.90        | 0.47        | 9.31         | 12.75        | 0.54        | 10.37        | 13.97        | 0.57        | 10.33        |
|           | "ROF "         | 17.91        | 0.27        | 10.62        | 19.47        | 0.41        | 13.45        | 21.58        | 0.50        | 14.89        |
|           | "Adaptative "  | 9.71         | 0.27        | 10.31        | 11.34        | 0.34        | 10.59        | 12.46        | 0.37        | 10.87        |
|           | "Convex "      | 13.78        | 0.21        | 5.90         | 22.72        | 0.63        | 15.56        | 25.26        | 0.76        | 18.51        |
|           | "Convex mult " | 7.28         | 0.03        | 12.08        | 17.98        | 0.43        | 6.69         | 21.72        | 0.57        | 13.74        |
|           | "Proposée"     | <b>29.87</b> | <b>0.93</b> | <b>23.89</b> | <b>30.14</b> | <b>0.93</b> | <b>29.82</b> | <b>32.25</b> | <b>0.95</b> | <b>33.92</b> |

Les mesures de qualité appliquées aux images d'origines et les images restaurées montrent clairement la supériorité et l'efficacité de notre approche comparativement aux méthodes implémentées existantes en littérature, ce qui nous permet de dire que notre modèle est le plus performant.

## IV.5 Conclusion

Après un fort labeur, nous avons pu atteindre l'objectif de notre travail, qui est la création d'un modèle énergétique convexe de débruitage des images dégradées par un bruit multiplicatif de type Gamma en se basant sur l'estimateur du Maximum A Posteriori MAP. Cette énergie, est composée de deux termes le premier est le terme d'attache aux données lié au bruit et le second est le terme de régularisation qui n'est que la variation totale. Pour trouver la solution nous avons minimisé l'énergie en se basant sur les algorithmes proximaux. La contribution réside dans la recherche et la démonstration de l'existence et l'unicité du proximal du terme d'attache aux données, puis l'intégration du résultat trouvé dans l'algorithme de Douglas-Rachford pour trouver la solution et garantir sa convergence. Des critères quantitatifs ont été utilisés pour mesurer les performances tels que le rapport signal sur bruit (PSNR), l'indice de similarité structurelle (SSIM) et le rapport visuel signal sur bruit (VSNR) et la qualité visuelle. L'évaluation de l'approche proposée montre qu'elle fournit des résultats meilleurs que ceux obtenus par les autres méthodes existantes en littérature.

---

## CONCLUSION ET PERSPECTIVES

Celui qui se perd dans sa passion a  
moins à perdre que celui qui perd sa  
passion.

---

Saint Augustin

La réduction du bruit dans les images est un problème délicat et largement traité par les scientifiques chercheurs. La problématique primordiale est de réussir à éliminer le maximum de bruit, tout en préservant la structure et les détails de l'image.

Une image peut subir des dégradations additives, multiplicatives ou convolutives :

- Quant aux dégradations additives ou multiplicatives, elles sont dues généralement soit aux capteurs ou durant la transmission, qui se traduisent par un flou causé par le système optique, sous forme d'un bruit quantique ou un bruit de mesures ou bien quelques défauts de transmission, comme on peut citer aussi les dégradations suite à une quantification ou une compression de l'information.
- Par contre les dégradations convolutives sont dues aux conditions de prise de vues comme le flou de bougé et les perturbations atmosphériques.

Le speckle apparaît lors de l'éclairage d'une scène par une lumière cohérente, il a suscité un regain d'intérêt lors de la commercialisation des premiers lasers dans les années 1960 [Goodman, 2007], il n'est pas spécifique au seul domaine de l'optique mais apparaît également dans des domaines tels que l'imagerie radar à synthèse d'ouverture (RSO) ou l'échographie. Le speckle donne aux images un aspect granuleux et en modifie la statistique et les caractéristiques texturales, rendant difficile tout traitement de débruitage. Le développement de méthodes de filtrage adaptées en vue de faciliter tout processus de filtrage semble donc nécessaire.

Avant d'entamer le problème de la réduction du bruit multiplicatif, il nous paraît impor-

tant de passer par le bruit additif, c'est pourquoi, dans un premier temps, nous avons présenté les algorithmes classiquement utilisés en filtrage d'images corrompues par un bruit additif, puis une classe spécifique de méthodes a été décrite spécialement pour le bruit multiplicatif.

Quoique les algorithmes développés dans la littérature ont subi un succès remarquable dans l'amélioration de la qualité des images débruitées, néanmoins ils restent loin des résultats souhaités. Cela a relancé une course concurrentielle dans la recherche scientifique, et dans ce contexte que nous avons proposé notre contribution dans cette thèse de Doctorat.

L'insuffisance de la qualité de filtrage des méthodes citées dans la littérature, nous a amené à développer une nouvelle approche qui se base sur les méthodes d'optimisation convexe, ces dernières ont montré leur efficacité pour la résolution de problèmes inverses.

Cette approche consiste à améliorer le débruitage des images dégradées par un bruit multiplicatif de type Gamma. Nos intérêts étant axés sur l'application des algorithmes proximaux, ces algorithmes se basent sur la minimisation d'un critère convexe composé d'une somme de deux fonctions non-nécessairement différentiables et possèdent des garanties de convergence.

L'algorithme proposé permet de gérer efficacement un critère composé d'un terme lié au type de bruit dit attache aux données et d'une fonction non différentiable qui est la variation totale dite terme de régularisation. La recherche de l'opérateur proximal du terme d'attache aux données lié au bruit multiplicatif Gamma puis son application à l'algorithme de Douglas Rachford constitue l'essentiel de notre contribution dans la restauration d'images par filtrage variationnel.

Afin d'évaluer les performances de la méthode proposée, cinq approches à base des EDPs ou de variationnel ont été implémentées, nous citons les travaux de :

- Yu et Acton [Yu and Acton, 2002] où le bruit de Rayleigh a été remplacé par un bruit de type Gamma.
- Shi et Osher [Shi and Osher, 2008] ont transformé le modèle multiplicatif en un modèle additif afin de pouvoir appliquer le modèle de ROF.
- Bernardes et al [Bernardes et al., 2010] qui ont travaillé sur la diffusion complexe adaptative.
- Dong et Zeng [Dong and Zeng, 2013] ont traité le débruitage par une méthode variationnelle convexe.
- Zhao et al [Zhao et al., 2014] ont proposé une méthode d'optimisation convexe.

La comparaison qui a été effectuée, nous a conduit à tirer la conclusion que la méthode proposée donne des résultats satisfaisants de débruitage d'images par rapport aux autres méthodes. L'évaluation de ces résultats appuyés sur l'aspect visuel, rapport signal sur

bruit (PSNR), l'indice de similarité structurelle (SSIM) et le rapport visuel du signal sur bruit VSNR, démontre la pertinence et l'efficacité de l'approche proposée. Les métriques utilisées sont toutes calculées entre une image originale et l'image filtrée.

L'algorithme proposé est implémenté sous UNLocbox<sup>2, 3</sup>. Les résultats obtenus montrent une très bonne amélioration de la restauration par rapport à l'état de l'art. Ceci ouvre de nouvelles perspectives de recherche telles que :

- Extension de l'application de l'algorithme sur des images couleurs.
- Proposition d'autres subdivisions du modèle énergétique en plusieurs fonctionnelles au lieu de deux.
- Diversification et extension de l'algorithme sur d'autres types de bruit.
- Reflexion à implémenter d'autres algorithmes du proximal.
- Accentuation des travaux futurs sur les images SAR ou médicales.

Nous espérons que ce modeste travail sera une très bonne référence, et bénéfique aux chercheurs intéressés par l'optimisation convexe, le traitement d'images, les mathématiques appliquées,...

---

2. <https://epfl-lts2.github.io/unlocbox-html/>

3. UNLocbox est une boîte à outils dédiée aux méthodes d'optimisation convexe, plus précisément approches proximales

## ANNEXE A

### EULER EXPLICITE-EULER IMPLICITE

On suppose que le problème de Cauchy<sup>1</sup>

$$\begin{cases} y'(t) = F(t, y(t)) \\ y(t_0) = y_0, \quad t_0 \in \mathbb{R}, \quad y_0 \in \mathbb{R}^n \end{cases} \quad (\text{A.1})$$

admet une unique solution  $y$  définie sur l'intervalle  $I = [t_0, t_0 + T]$  comme le montre la figure A.1 :

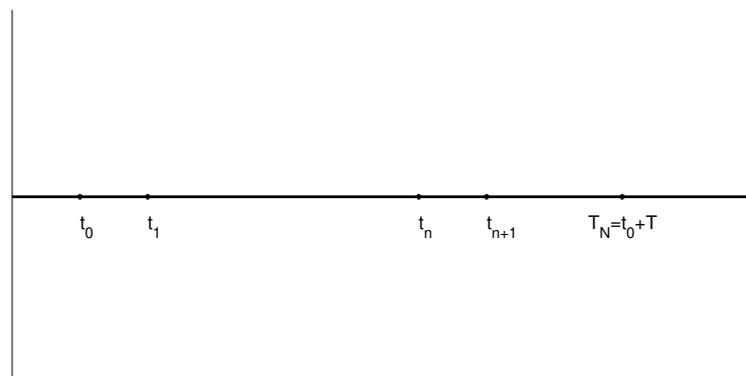


FIGURE A.1 – Subdivision de l'intervalle de temps

avec

$$\Delta t_n = t_{n+1} - t_n \text{ et } \Delta t = \max_{0 \leq n \leq N} \Delta t_n$$

L'objectif est de calculer des valeurs  $(Y_n)_{0 \leq n \leq N}$ , qui soient de bonnes approximations de  $y(t_n)_{0 \leq n \leq N}$ .

1. [http://math.univ-lille1.fr/~debievre/M45website/chapterequadiff\\_toprint.pdf](http://math.univ-lille1.fr/~debievre/M45website/chapterequadiff_toprint.pdf)

L'équation suivante montre le lien avec l'intégration numérique :

$$\begin{cases} \int_{t_n}^{t_{n+1}} y'(t) dt = \int_{t_n}^{t_{n+1}} F(t, y(t)) dt \\ y(t_0) = y_0, t_0 \in \mathbb{R}, y_0 \in \mathbb{R}^n \end{cases} \quad (\text{A.2})$$

Le premier membre de l'équation A.2 s'exprime comme suit :

$$\int_{t_n}^{t_{n+1}} y'(t) dt = y(t) \Big|_{t_n}^{t_{n+1}} = y(t_{n+1}) - y(t_n) \quad (\text{A.3})$$

L'application de la formule quadratique au deuxième membre de l'équation A.2 donne les approximations suivantes :

— Approximation à gauche

$$\int_{t_n}^{t_{n+1}} F(t, y(t)) dt = (t_{n+1} - t_n) F(t_n, y(t_n)) \quad (\text{A.4})$$

— Approximation à droite

$$\int_{t_n}^{t_{n+1}} F(t, y(t)) dt = (t_{n+1} - t_n) F(t_{n+1}, y(t_{n+1})) \quad (\text{A.5})$$

— Approximation du trapèze

$$\int_{t_n}^{t_{n+1}} F(t, y(t)) dt = (t_{n+1} - t_n) \frac{F(t_n, y(t_n)) + F(t_{n+1}, y(t_{n+1}))}{2} \quad (\text{A.6})$$

Les méthodes numériques correspondant aux trois approximations sont :

— Méthode d'Euler explicite :

$$\begin{cases} y(t_{n+1}) = y(t_n) + (t_{n+1} - t_n) F(t_n, y(t_n)) \\ y(t_0) = y_0 \end{cases} \quad (\text{A.7})$$

Après discrétisation :

$$\begin{cases} Y_{n+1} = Y_n + (t_{n+1} - t_n) F(t_n, Y_n) \\ Y_0 = y_0. \end{cases} \quad (\text{A.8})$$

— Méthode d'Euler implicite :

$$\begin{cases} y(t_{n+1}) = y(t_n) + (t_{n+1} - t_n) F(t_{n+1}, y(t_{n+1})) \\ y(t_0) = y_0 \end{cases} \quad (\text{A.9})$$

Après discrétisation :

$$\begin{cases} Y_{n+1} = Y_n + (t_{n+1} - t_n) F(t_{n+1}, Y_{n+1}) \\ Y_0 = y_0 \end{cases} \quad (\text{A.10})$$

— Méthode de Crank-Nicolson :

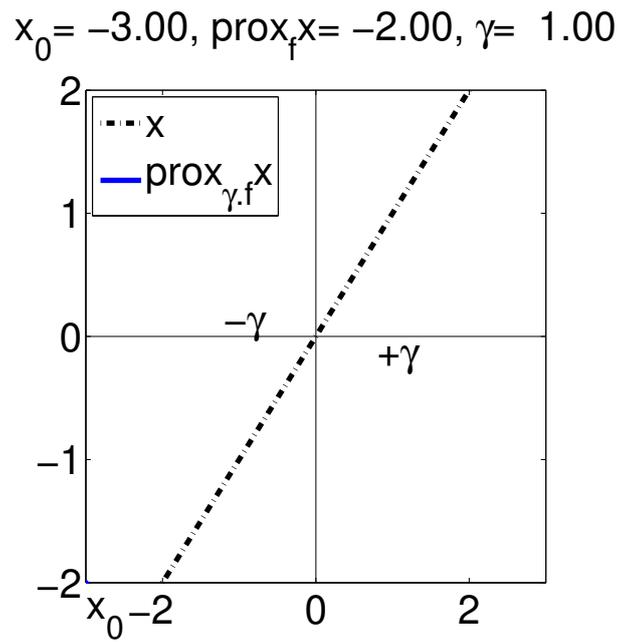
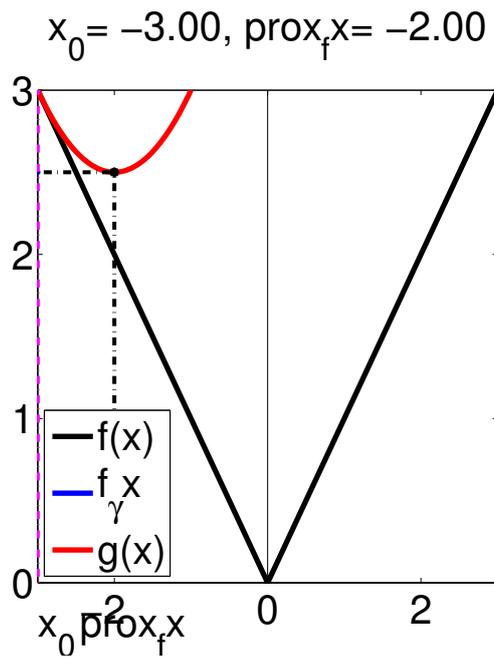
$$\begin{cases} y(t_{n+1}) = y(t_n) + (t_{n+1} - t_n) \frac{F(t_n, y(t_n)) + F(t_{n+1}, y(t_{n+1}))}{2} \\ y(t_0) = y_0 \end{cases} \quad (\text{A.11})$$

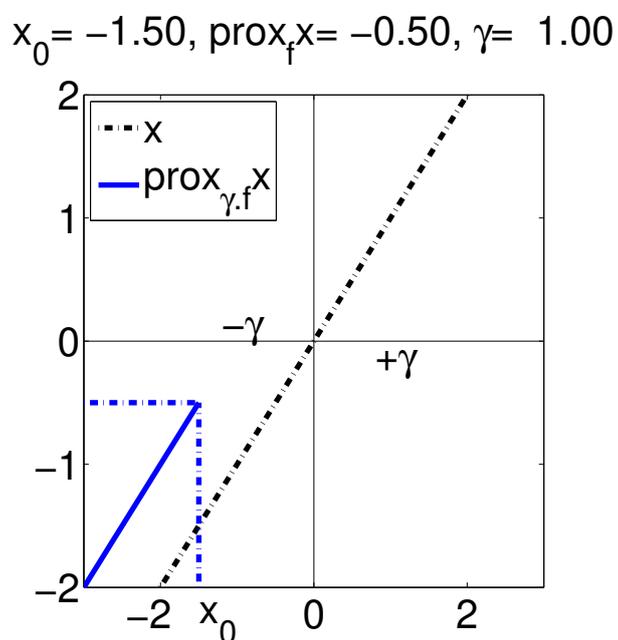
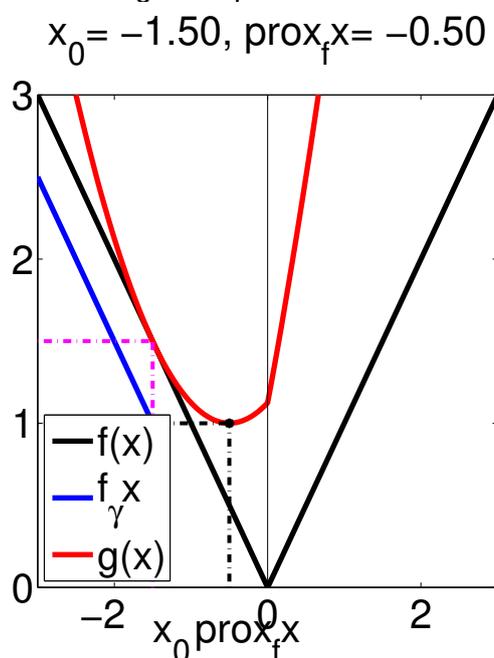
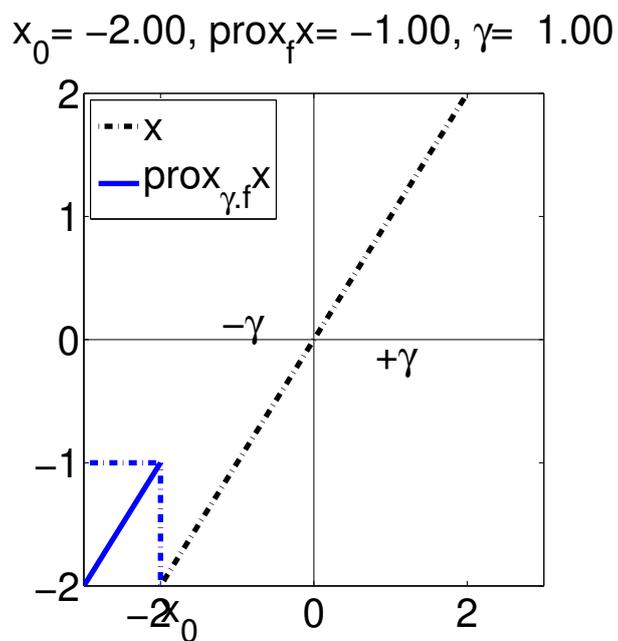
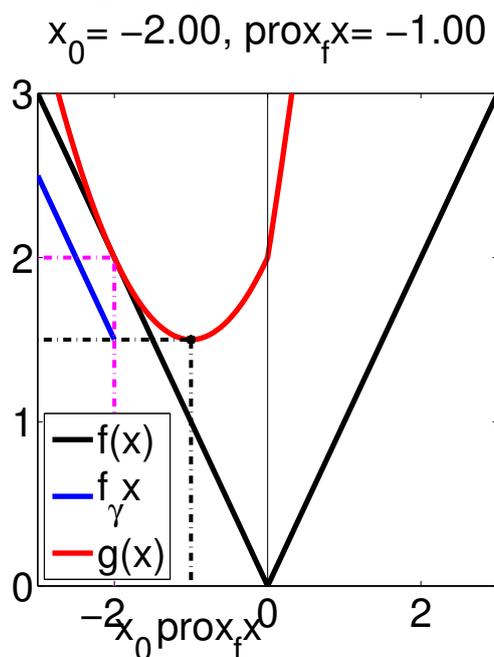
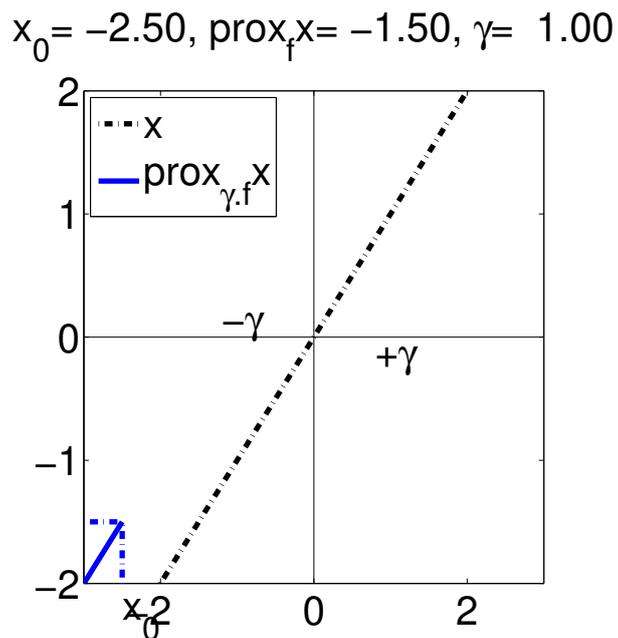
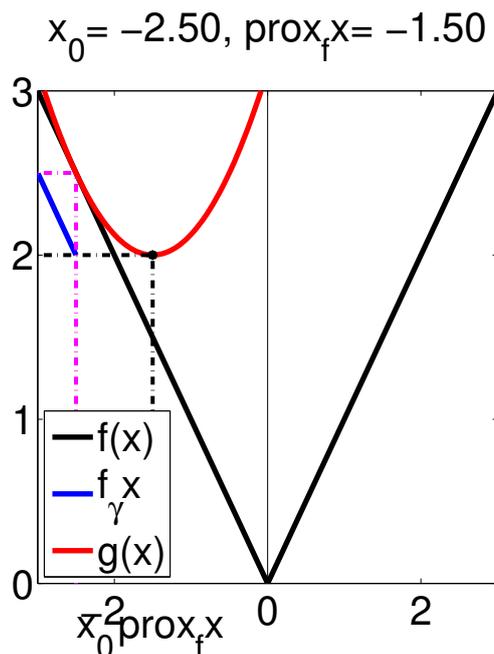
Après discrétisation :

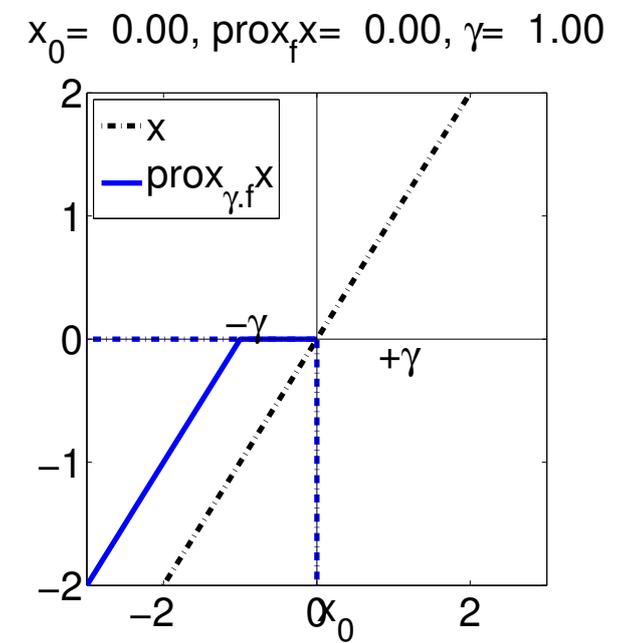
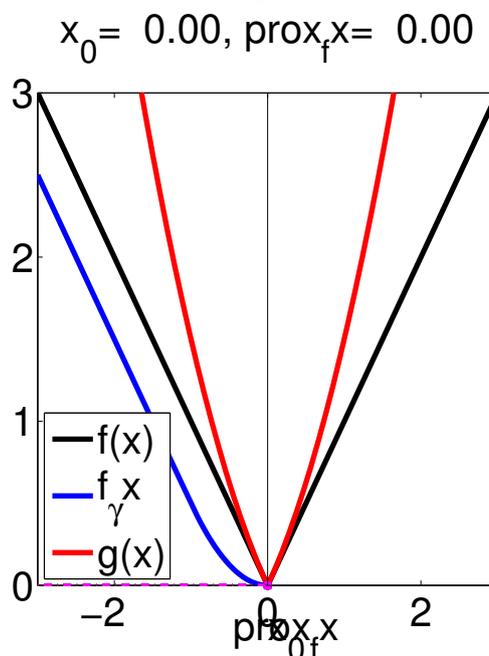
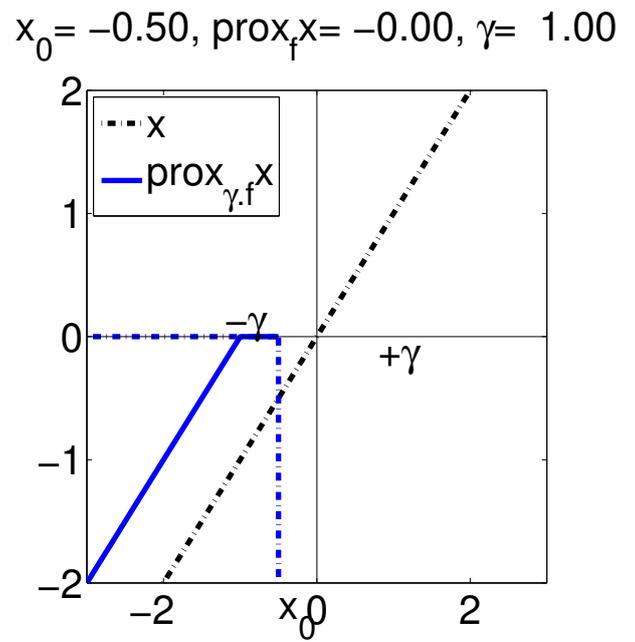
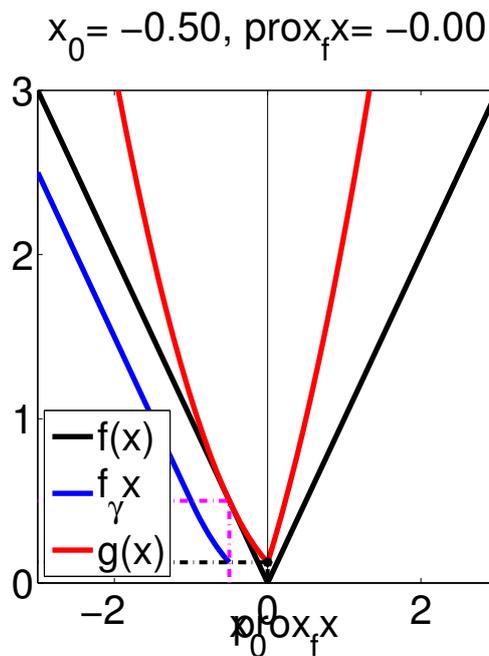
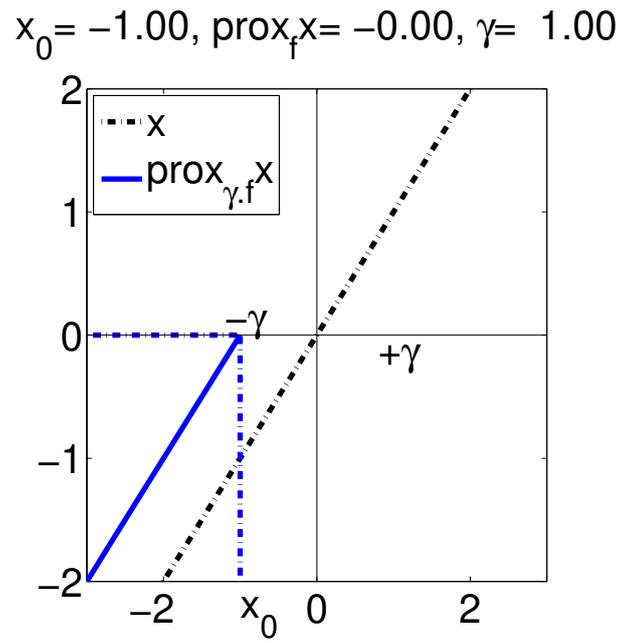
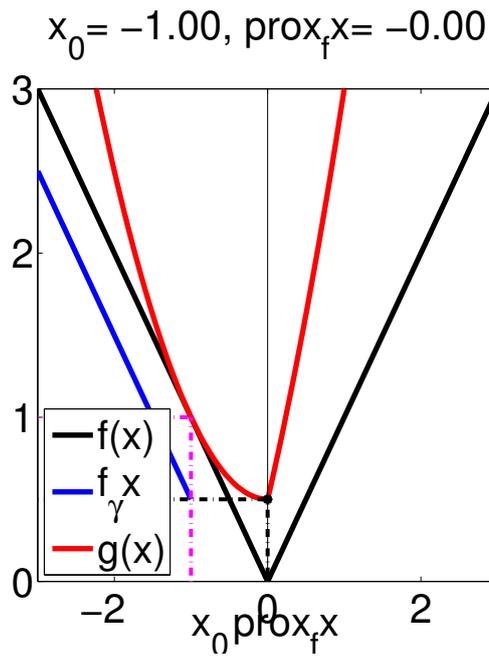
$$\begin{cases} Y_{n+1} = Y_n + (t_{n+1} - t_n) \frac{F(t_n, Y_n) + F(t_{n+1}, Y_{n+1})}{2} \\ Y_0 = y_0 \end{cases} \quad (\text{A.12})$$

# ANNEXE B

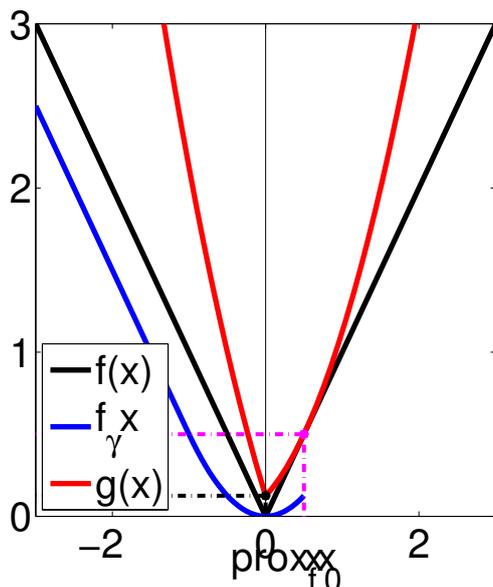
## VISUALISATION DE L'ENVELOPPE DE MOREAU YOSIDA ET LE PROXIMAL DE $F(X) = |X|$ POUR $\lambda = 1$



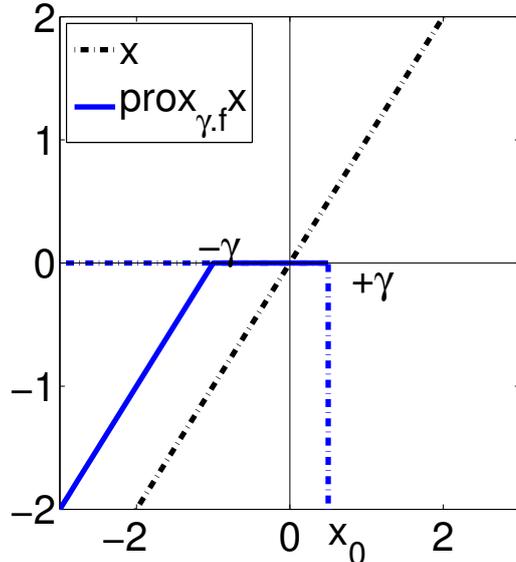




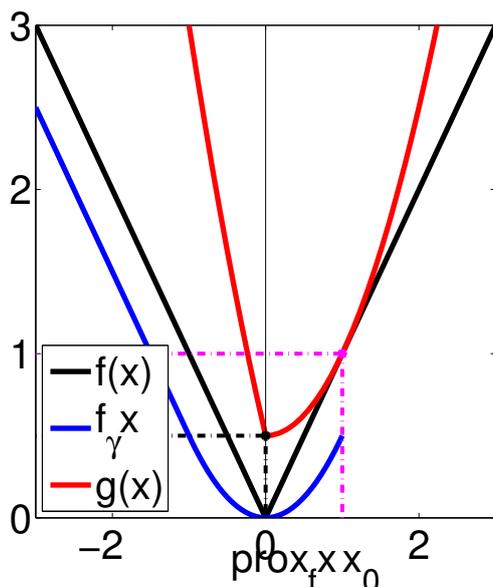
$x_0 = 0.50, \text{prox}_f x = 0.00$



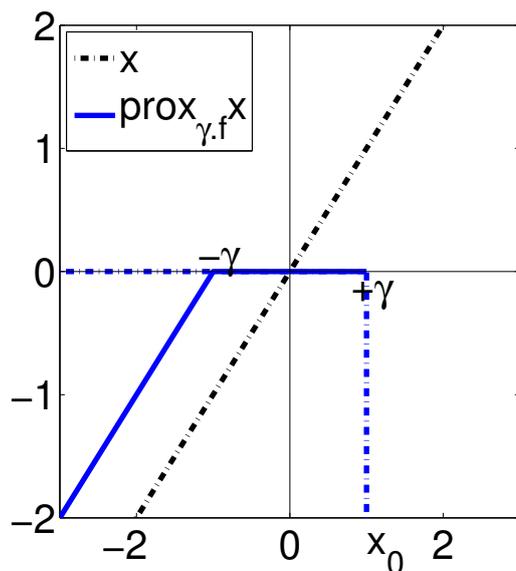
$x_0 = 0.50, \text{prox}_f x = 0.00, \gamma = 1.00$



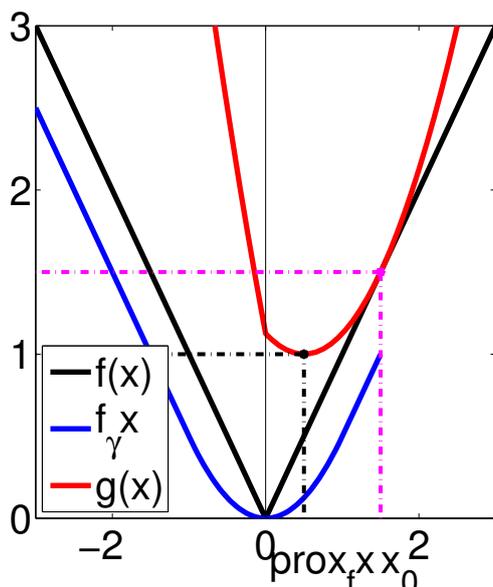
$x_0 = 1.00, \text{prox}_f x = -0.00$



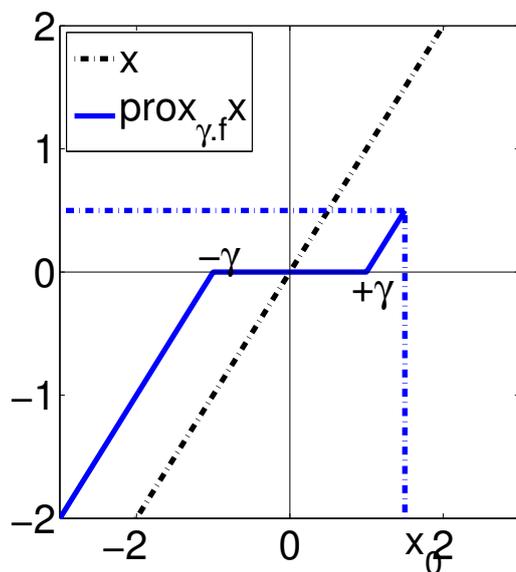
$x_0 = 1.00, \text{prox}_f x = -0.00, \gamma = 1.00$

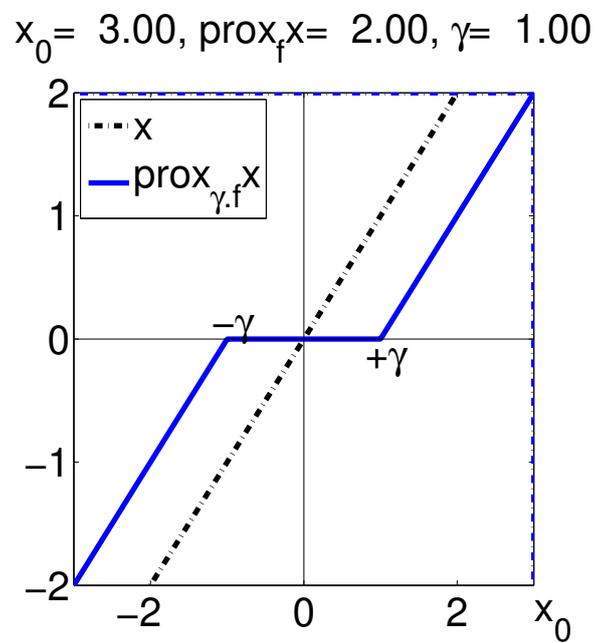
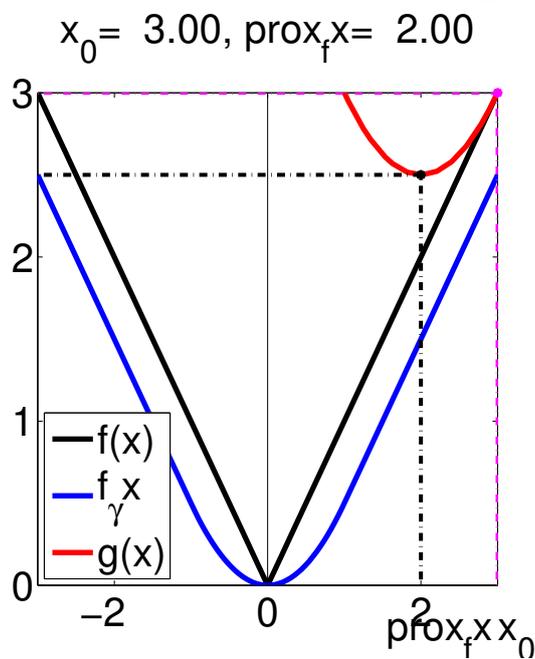
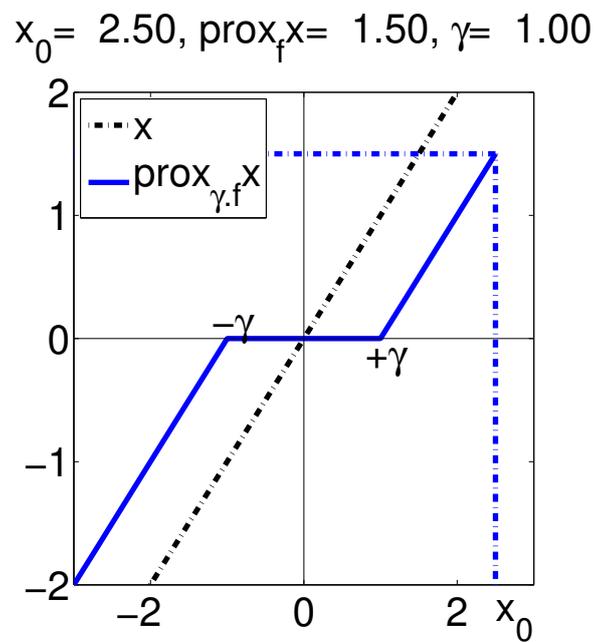
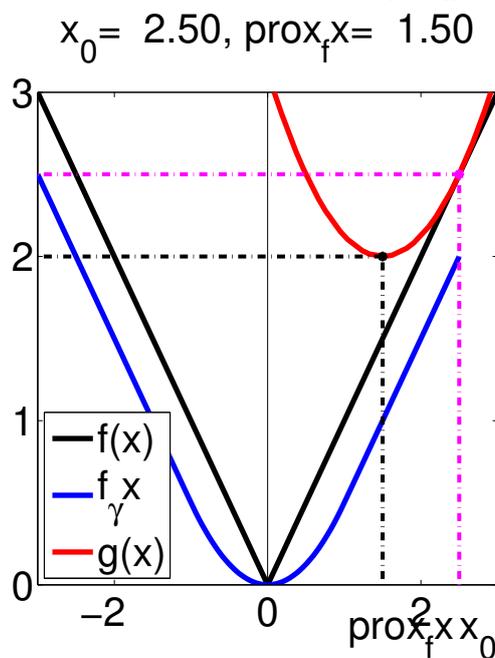
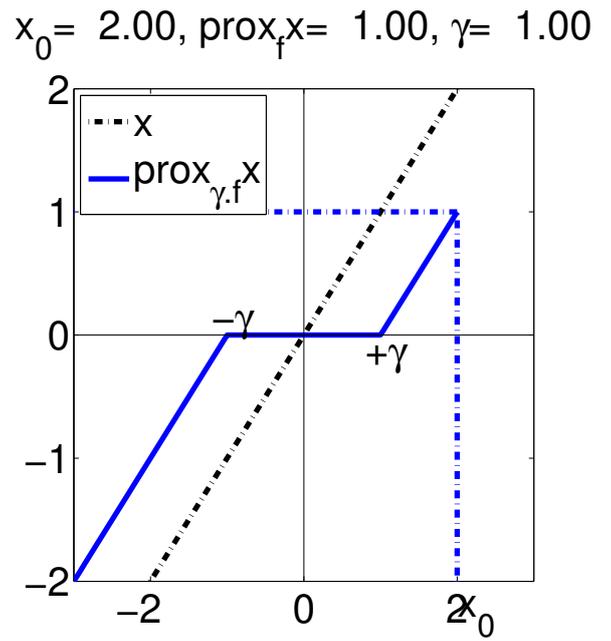
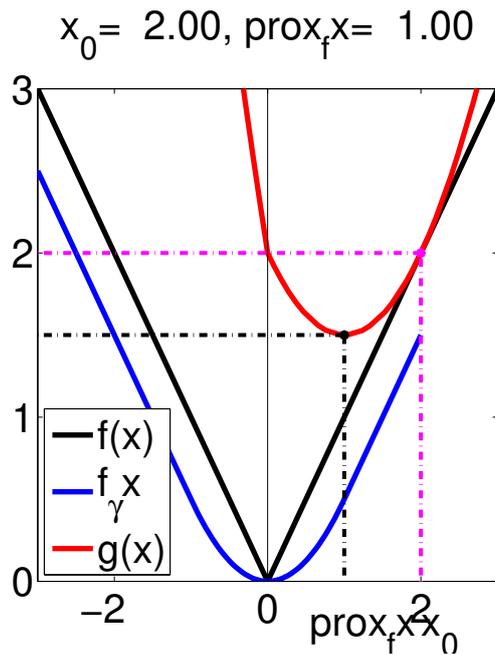


$x_0 = 1.50, \text{prox}_f x = 0.50$



$x_0 = 1.50, \text{prox}_f x = 0.50, \gamma = 1.00$





Rudin, Lions et Osher 'RLO' dans [Rudin et al., 2003] ont proposé une méthode de débruitage d'images corrompues par un bruit multiplicatif en appliquant la descente du gradient.

L'algorithme C.1 ci-dessous relate les étapes à suivre lors du déroulement de la méthode dite variationnelle :

---

**Algorithme C.1 Méthode variationnelle**


---

- 1: **Première étape : Création d'une énergie  $E$  qui décrit la qualité de l'image  $u$  :**

$$E(u) = \int_{\Omega} F(u, u_x, u_y)$$

- 2: **Deuxième étape : Calcul de la première variation d'énergie  $\nabla E$  :**

$$\nabla E = \frac{\partial F}{\partial u} - \frac{\partial}{\partial x} \frac{\partial F}{\partial u_x} - \frac{\partial}{\partial y} \frac{\partial F}{\partial u_y}$$

- 3: **Troisième étape : Application de la descente du gradient  $\nabla E$  :**

$$\frac{\partial u}{\partial t} = -\nabla E$$

- 4: **Quatrième étape : Discrétisation de l'EDP pour avoir le minimum de  $E$  :**

$$u_{i,j}^{n+1} = u_{i,j}^n + \tau_n (-\nabla E(u_{i,j}^n))$$


---

Étant donné une image  $f$  corrompue par un bruit multiplicatif d'écart-type  $\sigma^2$ , de valeur moyenne égal à 1 :

$$f = u \cdot \eta \tag{C.1}$$

Notre problème est la minimisation de la fonction objectif mettant en jeu le terme de variation totale :

$$\min_u \int_{\Omega} \sqrt{u_x^2 + u_y^2} dx dy \tag{C.2}$$

sous les deux contraintes :

La moyenne qui devra être égale à 1 :

$$\int \eta = 1 \quad (\text{C.3})$$

et la variance :

$$\int (\eta - 1)^2 = \sigma^2 \quad (\text{C.4})$$

Nous transformons le problème comme suit :

$$\left\{ \begin{array}{l} \min_u \int_{\Omega} \sqrt{u_x^2 + u_y^2} dx dy \\ \text{Sous contraintes} \\ \int \frac{f}{u} = 1 \\ \int \left( \frac{f}{u} - 1 \right)^2 = \sigma^2 \end{array} \right. \quad (\text{C.5})$$

La solution du problème de minimisation C.5 est donnée par l'équation C.6 :

$$u_t = \frac{u_{xx}u_y^2 - 2u_xu_yu_{xy} + u_{yy}u_x^2}{(u_x^2 + u_y^2)^{\frac{3}{2}}} - \lambda_1 \frac{f}{u^2} - \lambda_2 \left( \frac{f^2}{u^3} \right) \quad (\text{C.6})$$

$\lambda_1$  et  $\lambda_2$  sont appelés multiplicateurs de Lagrange.

Nous décrivons ci-dessous la démonstration et l'aboutissement au résultat trouvé dans l'équation C.6 :

### Preuve C.1

*Nous écrivons la deuxième contrainte sous forme :*

$$\int \frac{f}{u} = \frac{1}{2} \int \left( \left( \frac{f}{u} \right)^2 + 1 \right) = \frac{1}{2} \sigma^2$$

*puis le problème C.5 devient :*

$$\begin{array}{l} \min_u \int_{\Omega} \sqrt{u_x^2 + u_y^2} dx dy \\ \int \frac{f}{u} = 1 \\ \frac{1}{2} \int \left( \left( \frac{f}{u} \right)^2 + 1 \right) = \frac{1}{2} \sigma^2 \end{array}$$

nous utilisons Euler-Lagrange pour rendre le problème sans contraintes :

$$P : \min_u \left( \int_{\Omega} \sqrt{u_x^2 + u_y^2} dx dy + \lambda_1 \int \frac{f}{u} + \lambda_2 \frac{1}{2} \int \left( \left( \frac{f}{u} \right)^2 + 1 \right) \right)$$

Notons :

$$F(u, u_x, u_y) = \sqrt{u_x^2 + u_y^2} + \lambda_1 \frac{f}{u} + \lambda_2 \frac{1}{2} \left( \left( \frac{f}{u} \right)^2 + 1 \right)$$

puis procédons comme suit :

$$\frac{\partial F(u, u_x, u_y)}{\partial u} - \left( \frac{\partial}{\partial x} \frac{\partial F(u, u_x, u_y)}{\partial u_x} + \frac{\partial}{\partial y} \frac{\partial F(u, u_x, u_y)}{\partial u_y} \right) = 0$$

$$\frac{\partial F(u, u_x, u_y)}{\partial u} = \frac{\partial}{\partial u} \left( \sqrt{u_x^2 + u_y^2} + \lambda_1 \frac{f}{u} + \lambda_2 \frac{1}{2} \left( \left( \frac{f}{u} \right)^2 + 1 \right) \right)$$

$$\frac{\partial F(u, u_x, u_y)}{\partial u} = -\lambda_1 \frac{f}{u^2} + \lambda_2 \frac{1}{2} \left( 2 \left( \frac{f}{u} \right) \left( -\frac{f}{u^2} \right) \right) = -\lambda_1 \frac{f}{u^2} - \lambda_2 \left( \frac{f^2}{u^3} \right)$$

$$\frac{\partial F(u, u_x, u_y)}{\partial u_x} = \frac{\partial}{\partial u_x} \left( \sqrt{u_x^2 + u_y^2} + \lambda_1 \frac{f}{u} + \lambda_2 \frac{1}{2} \left( \left( \frac{f}{u} \right)^2 + 1 \right) \right) = \frac{\partial}{\partial u_x} \left( \sqrt{u_x^2 + u_y^2} \right)$$

$$\frac{\partial F(u, u_x, u_y)}{\partial u_x} = \frac{\partial}{\partial u_x} (u_x^2 + u_y^2)^{\frac{1}{2}} = \frac{1}{2} (2u_x) (u_x^2 + u_y^2)^{-\frac{1}{2}} = \frac{u_x}{\sqrt{u_x^2 + u_y^2}}$$

$$\frac{\partial F(u, u_x, u_y)}{\partial u_y} = \frac{\partial}{\partial u_y} \left( \sqrt{u_x^2 + u_y^2} + \lambda_1 \frac{f}{u} + \lambda_2 \frac{1}{2} \left( \left( \frac{f}{u} \right)^2 + 1 \right) \right) = \frac{\partial}{\partial u_y} \left( \sqrt{u_x^2 + u_y^2} \right)$$

$$\frac{\partial F(u, u_x, u_y)}{\partial u_y} = \frac{\partial}{\partial u_y} (u_x^2 + u_y^2)^{\frac{1}{2}} = \frac{1}{2} (2u_y) (u_x^2 + u_y^2)^{-\frac{1}{2}} = \frac{u_y}{\sqrt{u_x^2 + u_y^2}}$$

$$\frac{\partial F(u, u_x, u_y)}{\partial u} - \left( \frac{\partial}{\partial x} \frac{\partial F(u, u_x, u_y)}{\partial u_x} + \frac{\partial}{\partial y} \frac{\partial F(u, u_x, u_y)}{\partial u_y} \right) = 0$$

$$-\frac{\partial}{\partial x} \frac{u_x}{\sqrt{u_x^2 + u_y^2}} - \frac{\partial}{\partial y} \frac{u_y}{\sqrt{u_x^2 + u_y^2}} + \lambda_1 \frac{f}{u^2} + \lambda_2 \left( \frac{f^2}{u^3} \right) = 0$$

Nous appliquons maintenant la méthode de la descente du gradient :

$$\frac{\partial u}{\partial t} = -\nabla E$$

$$\nabla E = \frac{\partial F(u, u_x, u_y)}{\partial u} - \left( \frac{\partial}{\partial x} \frac{\partial F(u, u_x, u_y)}{\partial u_x} + \frac{\partial}{\partial y} \frac{\partial F(u, u_x, u_y)}{\partial u_y} \right)$$

$$u_t = \frac{\partial}{\partial x} \frac{u_x}{\sqrt{u_x^2 + u_y^2}} + \frac{\partial}{\partial y} \frac{u_y}{\sqrt{u_x^2 + u_y^2}} - \lambda_1 \frac{f}{u^2} - \lambda_2 \left( \frac{f^2}{u^3} \right)$$

avec :

$$\frac{\partial}{\partial x} \frac{u_x}{\sqrt{u_x^2 + u_y^2}} = \frac{u_{xx} \sqrt{u_x^2 + u_y^2} - u_x \frac{\partial}{\partial x} \sqrt{u_x^2 + u_y^2}}{u_x^2 + u_y^2} = \frac{u_{xx} u_y^2 - u_x u_y u_{xy}}{(u_x^2 + u_y^2)^{\frac{3}{2}}}$$

$$\frac{\partial}{\partial y} \frac{u_y}{\sqrt{u_x^2 + u_y^2}} = \frac{u_{yy} \sqrt{u_x^2 + u_y^2} - u_y \frac{\partial}{\partial y} \sqrt{u_x^2 + u_y^2}}{u_x^2 + u_y^2} = \frac{u_{yy} u_x^2 - u_x u_y u_{xy}}{(u_x^2 + u_y^2)^{\frac{3}{2}}}$$

L'évolution de l'EDP devient :

$$u_t = \frac{u_{xx} u_y^2 - 2u_x u_y u_{xy} + u_{yy} u_x^2}{(u_x^2 + u_y^2)^{\frac{3}{2}}} - \lambda_1 \frac{f}{u^2} - \lambda_2 \left( \frac{f^2}{u^3} \right) \dots$$

ce qu'il faut démontrer

# ANNEXE D

## VISUALISATION DES TRACÉS DE LA FONCTION $Y = |X|$ , SA DÉRIVÉE ET SA TANGENTE

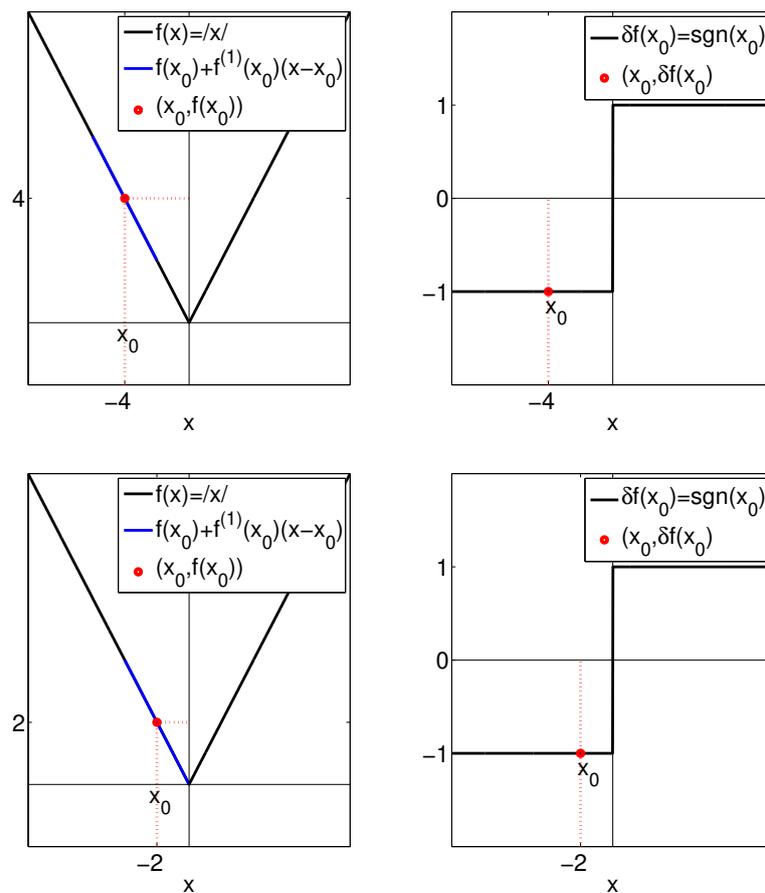


FIGURE D.1 – Tracés de la fonction  $y = |x|$  , sa dérivée et sa tangente lorsque  $x$  négatif

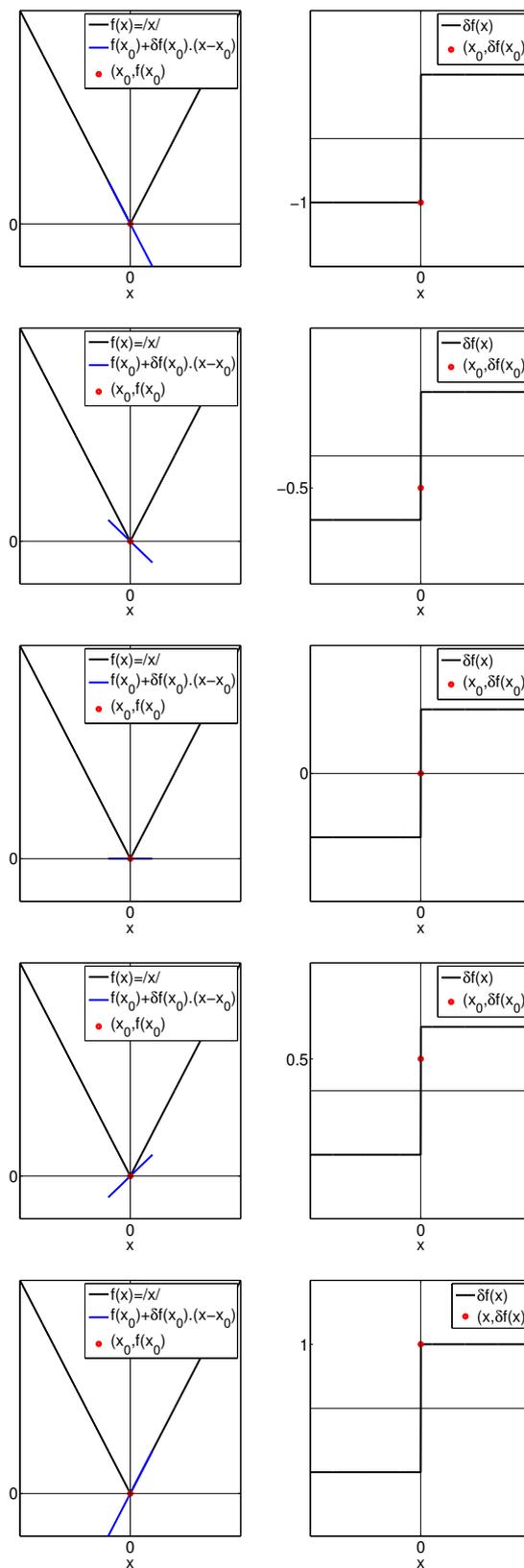


FIGURE D.2 – Tracés de la fonction  $y = |x|$ , sa dérivée et sa tangente au point  $x = 0$

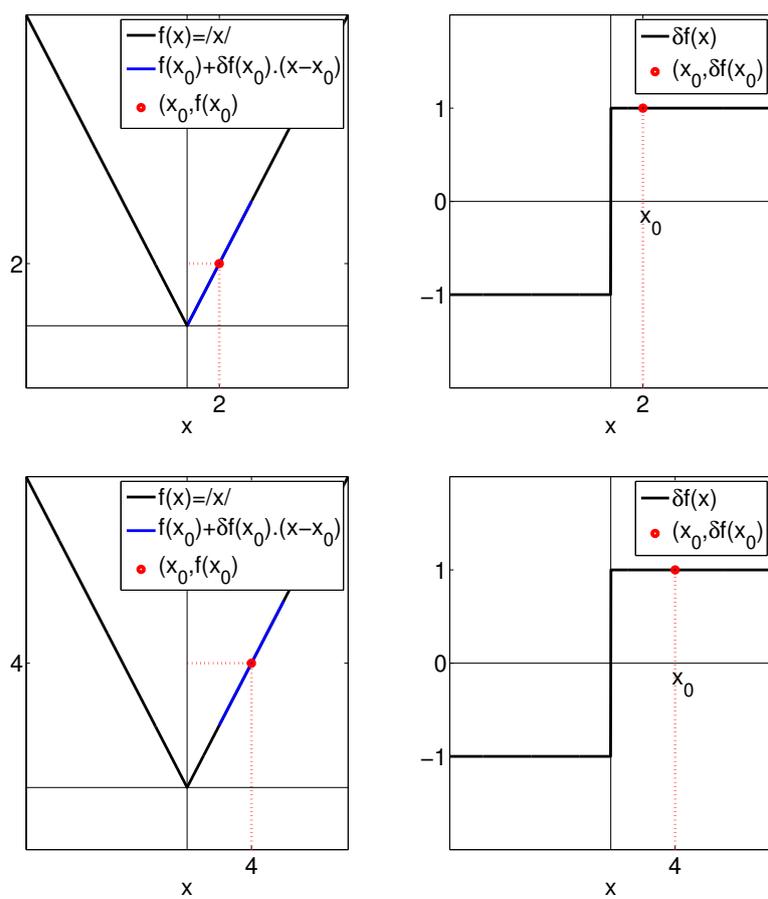


FIGURE D.3 – Tracés de la fonction  $y = |x|$ , sa dérivée et sa tangente lorsque  $x$  positif

---

## BIBLIOGRAPHIE

- [Attouch et al., 2013] Attouch, H., Bolte, J., and Svaiter, B. F. (2013). Convergence of descent methods for semi algebraic and tame problems : proximal algorithms, forward backward splitting, and regularized gauss seidel methods. *Mathematical Programming*, 137(1-2) :91–129.
- [Aubert and Aujol, 2008] Aubert, G. and Aujol, J.-F. (2008). A variational approach to removing multiplicative noise. *SIAM Journal on Applied Mathematics*, 68(4) :925–946.
- [Aubert et al., 1999] Aubert, G., Deriche, R., and Kornprobst, P. (1999). Computing optical flow via variational techniques. *SIAM Journal on Applied Mathematics*, 60(1) :156–182.
- [Aubert and Kornprobst, 2006] Aubert, G. and Kornprobst, P. (2006). *Mathematical Problems in Image Processing Partial Differential Equations and the Calculus of Variations*, volume 147 of *Applied Mathematical Sciences*. Springer-Verlag New York.
- [Auslender, 1976] Auslender, A. (1976). *Optimisation : Méthodes numériques (maîtrise de mathématiques et applications fondamentales)*. Masson, Paris.
- [Bauschke and Combettes, 2018] Bauschke, H. H. and Combettes, P. L. (2018). *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*,. Springer, New York, 2 edition.
- [Bect et al., 2004] Bect, J., Blanc-Féraud, L., Aubert, G., and Chambolle, A. (2004). A l1-unified variational framework for image restoration. In T. Pajdla, J. M., editor, *8th European Conference on Computer Vision (ECCV 2004*, volume 3024 of *Lecture Notes in Computer Science*, pages 1–13, Prague, Czech Republic. Springer.
- [Bernardes et al., 2010] Bernardes, R., Maduro, C., Serranho, P., Araújo, A., Barbeiro, S., and Cunha-Vaz, J. (2010). Improved adaptive complex diffusion despeckling filter. *Optics Express*, 18 :24048–24059.

- [Bertsekas, 2006] Bertsekas, D. (2006). *Convex Analysis and Optimization*. Athena Scientific.
- [Bertsekas, 2009] Bertsekas, D. (2009). *Convex Optimization Theory*. Athena Scientific.
- [Bertsekas, 1999] Bertsekas, D. P. (1999). *Nonlinear programming*. Athena Scientific, Belmont, Massachusetts, 2 edition.
- [Bloch et al., 2005] Bloch, I., Gousseau, Y., Maître, H., Matignon, D., Pesquet-Popescu, B., Schmitt, F., Sigelle, M., and Tupin, F. (2005). *Le traitement des images*, volume 5.
- [Bonnans et al., 2003] Bonnans, F., Lemarechal, C., Sagastizabal, C., and GILBERT, J. C. (2003). *Numerical Optimization : Theoretical and Practical Aspects*. Springer.
- [Boyd and Vandenberghe, 2004] Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization*. Cambridge University Press.
- [Catté et al., 1992] Catté, F., Lions, P.-L., Morel, J.-M., and Coll, T. (1992). Image selective smoothing and edge detection by nonlinear diffusion. *SIAM Journal on Numerical analysis*, 29(1) : :182–193.
- [Chambolle and Pock, 2011] Chambolle, A. and Pock, T. (2011). A first order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1) :120–145.
- [Chandler and Hemami, 2007] Chandler, D. M. and Hemami, S. S. (2007). Vsnr : A wavelet-based visual signal-to-noise ratio for natural images. *IEEE Transactions on Image Processing*, 16(9) :2284 – 2298.
- [Charbonnier et al., 1997] Charbonnier, P., Blanc-Feraud, L., Aubert, G., and Barlaud, M. (1997). Deterministic edge-preserving regularization in computed imaging. *IEEE Transactions on Image Processing*, 6(2) :298 – 311.
- [Chaux et al., 2007] Chaux, C., Combettes, P. L., Pesquet, J. L. C., and Wajs, V. R. (2007). A variational formulation for frame based inverse problems. *Inverse Problems*, 23(4) :1495–1518.
- [Combettes and Pesquet, 2007a] Combettes, P. L. and Pesquet, J.-C. (2007a). A Douglas–Rachford splitting approach to nonsmooth convex variational signal recovery. *IEEE Journal of Selected Topics in Signal Processing*, 1(4) :564 – 574.
- [Combettes and Pesquet, 2007b] Combettes, P. L. and Pesquet, J. C. (2007b). Proximal thresholding algorithm for minimization over orthonormal bases. *SIAM Journal on Optimization*, 18(4) :1351–1376.
- [Combettes and Pesquet, 2011] Combettes, P. L. and Pesquet, J.-C. (2011). *Proximal Splitting Methods in Signal Processing*, chapter Fixed Point Algorithms for Inverse Problems in Science and Engineering, pages 185–212. Springer-Verlag, New York.
- [Combettes and Vu, 2013] Combettes, P. L. and Vu, B. C. (2013). Variable metric quasi - fejer monotonicity. *Nonlinear Analysis : Theory, Methods & Applications*, 78 :17–31.

- [Combettes and Wajs, 2005] Combettes, P. L. and Wajs, V. (2005). Signal recovery by proximal forward-backward splitting. *Multiscale Modeling and Simulation : A SIAM Interdisciplinary Journal*, 4(4) :1164–1200.
- [Daubechies et al., 2004] Daubechies, I., Defrise, M., and Mol, C. D. (2004). An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics*, 57(11) :1413–1457.
- [Demoment, 1989] Demoment, G. (1989). Image reconstruction and restoration : overview of common estimation structures and problems. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37(12) :2024 – 2036.
- [Deriche and Faugeras, 2006] Deriche, R. and Faugeras, O. (2006). Les edp en traitement des images et vision par ordinateur. Technical report, Inria-00073993f.
- [Diffellah et al., 2019] Diffellah, N., Baarir, Z.-E., Derraz, F., and Ahmed-Taleb, A. (2019). A global variational filter for restoring noised images with gamma multiplicative noise. *Engineering, Technology & Applied Science*, 9(3) :4188–4195.
- [Diffellah et al., 2012] Diffellah, N., Derraz, F., Ahmed, A. T., Baarir, Z. E., and Hamadache, F. (2012). Image restoration with denoising using a split -bregman algorithm. In *The third International Conference on Image and Signal Processing and their Applications (ISPA'12)*.
- [Djenabou, 2018] Djenabou, B. (2018). Débruitage d'image par fusion de filtrage spatio-fréquentielle. Montréal.
- [Dong and Zeng, 2013] Dong, Y. and Zeng, T. (2013). A convex variational model for restoring blurred images with multiplicative noise. *SIAM Journal on Imaging Sciences*, 6(3) :1598–1625.
- [Eckstein and Bertsekas, 1992] Eckstein, J. and Bertsekas, D. P. (1992). On the Douglas—rachford splitting method and the proximal point algorithm for maximal monotone operators. *Mathematical Programming*, 55(1-3) :293–318.
- [Ekeland and Temam, 1972] Ekeland, I. and Temam, R. (1972). *Convex Analysis & Variational Problems1*. Dunod,Paris.
- [Emilie et al., 2014] Emilie, C., Christophe, P. J., and Audrey, R. (2014). Variable metric forward-backward algorithm for minimizing the sum of a differentiable function and a convex function. *Journal of Optimization Theory and Applications*, 162(1) :107–132.
- [Esser et al., 2010] Esser, E., Zhang, X., and Chan, T. F. (2010). A general framework for a class of first order primal-dual algorithms for convex optimization in imaging science. *SIAM Journal on Imaging Sciences*, 3(4) :1015–1046.
- [Fletcher, 2009] Fletcher, R. (2009). *Practical Methods of Optimization*. Wiley India, 2 edition.

- [Fletcher and Reeves, 1964] Fletcher, R. and Reeves, C. M. (1964). Function minimization by conjugate gradients. *The Computer Journal*, 7(2) :149–154.
- [Frost et al., 1982] Frost, V. S., Stiles, J. A., Shanmugan, K. S., and Holtzman, J. C. (1982). A model for radar images and its application to adaptive digital filtering of multiplicative noise. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 4(2) :157 – 166.
- [Gabay and Mercier, 1976] Gabay, D. and Mercier, B. (1976). A dual algorithm for the solution of nonlinear variational problems via finite element approximation. *Computers & Mathematics with Applications*, 2 :17–40.
- [Ghazel, 2004] Ghazel, M. (2004). *Adaptive fractal and wavelet image denoising*. PhD thesis, Waterloo, Ontario, Canada.
- [Gilboa et al., 2001] Gilboa, G., Zeevi, Y. Y., and Sochen, N. A. (2001). Image enhancement segmentation and denoising by time dependent nonlinear diffusion processes. *International Conference on Image Processing*, 3 :134–137.
- [Gill et al., 1982] Gill, P. E., Murray, W., and Wright, M. H. (1982). *Practical Optimization*. Academic Press, London.
- [Goldstein and Osher, 2009] Goldstein, T. and Osher, S. (2009). The split bregman method for l1 regularized problems. *SIAM Journal on Imaging Sciences*, 2(2) :323–343.
- [Goodman, 2007] Goodman, J. W. (2007). *Speckle Phenomena in Optics : Theory and Applications*. Roberts and Company Publishers.
- [Hadamard, 1923] Hadamard, J. (1923). Lectures on cauchy’s problem in linear partial differential equations. *Physics Today*, 6(3).
- [Hager and Zhang, 2006] Hager, W. W. and Zhang, H. (2006). A survey of nonlinear conjugate gradient methods. *Pacific journal of Optimization*, 2(1) :35–58.
- [Jalalzai and Chambolle, 2009] Jalalzai, K. and Chambolle, A. (2009). Restauration d’images floutées & bruitées par une variante originale de la variation totale. In *ORASIS’09 - Congrès des jeunes chercheurs en vision par ordinateur*, Trégastel, France.
- [Koenderink, 1984] Koenderink, J. (1984). The structure of images. *Biological Cybernetics*, 50(5) :363–370.
- [Kuan et al., 1985] Kuan, D. T., Sawchuk, A. A., Strand, T. C., and Chavel, P. (1985). Adaptive noise smoothing filter for images with signal-dependent noise. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7(2) :165 – 177.
- [Laurent, 1972] Laurent, J. (1972). *Approximation et Optimisation*. Hermann.
- [Lee, 2004] Lee, J. (2004). *A First Course in Combinatorial Optimization*, volume 36. Cambridge University Press.

- [Lee, 1980] Lee, J.-S. (1980). Digital image enhancement and noise filtering by use of local statistics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2(2) :165 – 168.
- [Li and Chen, 1994] Li, X. and Chen, T. (1994). Nonlinear diffusion with multiple edginess thresholds. *Pattern Recognition*, 27(8) :1029–1037.
- [Lopes et al., 1990] Lopes, A., Touzi, R., and Nezry, E. (1990). Adaptive speckle filters and scene heterogeneity. *IEEE Transactions on Geoscience and Remote Sensing*, 28(6).
- [Maher, 1997] Maher, N. (1997). Filtrage et analyse des images radars. Université Laval.
- [Martinet, 1972] Martinet, B. (1972). Détermination approchée d’un point fixe d’une application pseudo contractante. *Comptes Rendus Hebdomadaires des Sciences de l’Académie, Paris*, (274A) :163–165.
- [Moreau, 1965] Moreau, J. J. (1965). Proximité et dualité dans un espace hilbertien. *Bulletin de la Société Mathématique de France*, 93 :273–299.
- [Moreau, 1967] Moreau, J. J. (1966-1967). Fonctionnelles convexes. *Séminaire Jean Leray*, pages 1–108.
- [Nesterov, 2004] Nesterov, Y. (2004). *Introductory Lectures on Convex Optimization. A Basic Course*, volume 87 of *Applied Optimization*. Springer US, 1 edition.
- [Niessen et al., 1994] Niessen, ter Haar Romeny, Florack, Salden, and Viergever (1994). Nonlinear diffusion of scalar images using well-posed differential operators. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 92–97.
- [Nitzberg and Shiota, 1992] Nitzberg, M. and Shiota, T. (1992). Nonlinear image filtering with edge and corner enhancement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(8) :826 – 833.
- [Nocedal and Wright, 2006] Nocedal, J. and Wright, S. (2006). *Numerical Optimization*. Springer series in operations research. Springer, 2nd edition.
- [Parikh and Boyd, 2014] Parikh, N. and Boyd, S. (2014). Proximal algorithms. *Foundations and Trends in Optimization*, 1(3) :127–239.
- [Perona and Malik, 1990] Perona, P. and Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on pattern analysis and machine intelligence*, 12(7) :629–639, 1990.
- [Perona et al., 1994] Perona, P., Shiota, T., and Malik, J. (1994). *Geometry-Driven Diffusion in Computer Vision*, chapter Anisotropic Diffusion, pages 73–92. Springer, Dordrecht.
- [Polak and Ribiere, 1969] Polak, E. and Ribiere, G. (1969). Note sur la convergence de méthodes de directions conjuguées. *ESAIM : Mathematical Modelling and Numerical Analysis - Modélisation Mathématique et Analyse Numérique*, 3(R1) :35–43.

- [Polyak, 1987] Polyak, B. T. (1987). *Introduction to optimization*. New York : Optimization Software, Inc.
- [Rockafellar et al., 1998] Rockafellar, R., Tyrrell, W., and J-B, R. (1998). *Variational Analysis*, volume 1 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag Berlin Heidelberg, 1 edition.
- [Rockafellar, 1970] Rockafellar, R. T. (1970). *Convex Analysis*. Princeton University Press.
- [Rockafellar, 1976] Rockafellar, R. T. (1976). Augmented lagrangians and applications of the proximal point algorithm in convex programming. *Mathematics of Operations Research*, 1(2) :97–116.
- [Rockafellar, 2015] Rockafellar, R. T. (2015). *Convex analysis. Princeton*. Princeton university press.
- [Rudin et al., 2003] Rudin, L., Lions, P. L., and Osher, S. (2003). *Geometric Level Set Methods in Imaging, Vision, and Graphics*, chapter Multiplicative Denoising and Deblurring : Theory and Algorithms, pages 103–119. Springer, New York, NY.
- [Salmon, 2010] Salmon, J. (2010). *Agrégation d'estimateurs et méthodes à patch pour le débruitage d'images numériques*. PhD thesis, Paris-Diderot.
- [Shi and Osher, 2008] Shi, J. and Osher, S. (2008). A nonlinear inverse scale space method for a convex multiplicative noise model. *SIAM Journal on Imaging Sciences*, 1(3) :294–321.
- [Snyman and Wilke, 2018] Snyman, J. A. and Wilke, D. N. (2018). *Practical Mathematical Optimization : Basic Optimization Theory and Gradient-Based Algorithms*. Springer International Publishing, 2 edition.
- [Tauber, 2005] Tauber, C. (2005). *Filtrage anisotrope robuste et segmentation par B-spline snake : application aux images échographiques*. PhD thesis, Institut National Polytechnique de Toulouse.
- [Urruty et al., 1993a] Urruty, H., Baptiste, J., and Claude, L. (1993a). *Convex Analysis and Minimization Algorithms II Advanced Theory and Bundle Methods*, volume 306 of *Grundlehren der mathematischen Wissenschaften*. Springer Verlag Berlin Heidelberg, 1 edition.
- [Urruty et al., 1993b] Urruty, H., Baptiste, J. J., and Claude, L. (1993b). *Convex analysis and minimization algorithms I : fundamentals*, volume 305 of *Grundlehren der mathematischen Wissenschaften*. Springer Science & Business Media, 1 edition.
- [Wang et al., 2004] Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P. (2004). Image quality assessment : From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4) :600–612.

- [Weickert, 1998] Weickert, J. (1998). *Anisotropic Diffusion in Image Processing*. B.G. Teubner Stuttgart.
- [Weickert, 1999a] Weickert, J. (1999a). Coherence enhancing diffusion filtering. *International Journal of Computer Vision*, 31(2-3) :111–127.
- [Weickert, 1999b] Weickert, J. (1999b). Coherence enhancing diffusion of colour images. *Image and Vision Computing*, 17(3-4) :201–212.
- [Weiss et al., 2009] Weiss, P., Blanc-Féraud, L., and Aubert, G. (2009). Efficient schemes for total variation minimization under constraints in image processing. *SIAM Journal on Scientific Computing*, 31(3) :2047–2080.
- [Whitaker, 1993] Whitaker, R. (1993). *Geometry limited diffusion*. PhD thesis, Department of Computer Science, University of North Carolina.
- [Wiener, 1949] Wiener, N. (1949). *Extrapolation, Interpolation, and Smoothing of Stationary Time Series : With Engineering Applications*. Technology Press of the Massachusetts Institute of Technology.
- [Yu and Acton, 2002] Yu, Y. and Acton, S. (2002). Speckle reducing anisotropic diffusion. *IEEE Transactions on Image Processing*, 11(11) :1260 – 1270.
- [Zalinescu, 2002] Zalinescu, C. (2002). *Convex analysis in general vector spaces*. World Scientific.
- [Zhao et al., 2014] Zhao, X.-L., Wang, F., and Ng, M. K. (2014). A new convex optimization model for multiplicative noise and blur removal. *SIAM Journal on Imaging Sciences*, 7(1).

## المخلص

تمركز بحثنا في الدكتوراه حول ترشيح الصور المشوشة بواسطة ضوضاء مضاعفة من نوع جاما، باستعمال زيادة كثافة الاحتمال الخلفي MAP، حيث قمنا باستخراج نموذج تم تقسيمه إلى طرفين، الأول له علاقة مباشرة بالضوضاء، أما الثاني فهو يمثل التباعد أو الاختلاف الكلي VT. ولحلّ المشكل المطروح، قمنا بإيجاد الكمية الأقرب للطرف الاول، و برهنا أنّ حلّها موجود و وحيد، بعد ذلك استعملنا نظرية دوغلاس و راش فورد لإيجاد الحل الذي يضمن التقارب، ولإثبات نجاعة الطريقة المقترحة، قمنا بمقارنة نتائجنا مع النتائج المتحصّل عليها سابقا، فوجدنا أنّ طريقتنا جدّ فعالة لنزع الضوضاء من الصور، سواء بالعين المجردة أو استعمال مقاييس الجودة مثل PSNR, SSIM و VSNR.

## الكلمات المفتاحية:

ضوضاء جاما المضاعفة، الترشيح، التباين الكلي، الخوارزمية القريبية، مقياس الجودة...

## Abstract

This thesis considers the restoration problem of images, which are corrupted by gamma multiplicative noise. Our problem is assumed as a regularization problem in total variation (TV) framework with the data fitting term which is deduced by maximizing the a-posteriori probability density (MAP estimation). We need to evaluate the proximal operator of a data fitting term then we adapt numerically the Douglas - Rachford (DR) splitting method to solve the problem.

To validate the effectiveness of the proposed method, we compare it with other variational models. The measures of image quality such as: PSNR (Peak Signal-to-Noise Ratio), VSNR (Visual Signal-to-Noise Ratio) and SSIM (Structural SIMilarity index) explain that the proposed model has good performance. The proposed method is implemented under UNLocBoX which is, a MATLAB convex optimization toolbox based on proximal splitting methods.

Key words:

Gamma multiplicative noise, filtering, total variation, proximal algorithm, quality measure,...

## Résumé

L'axe primordial de notre thèse est le filtrage des images dégradées par un bruit multiplicatif de type Gamma. L'application de l'estimateur du Maximum A Posteriori MAP nous a permis de créer un modèle basé sur un critère énergétique, cette énergie est la somme de deux termes le premier est le terme d'attache aux données lié au bruit et le second est le terme de régularisation qui n'est que la variation totale, le minimiseur de l'énergie correspond à l'image recherchée. La méthode proposée pour résoudre ce problème est liée à l'optimisation convexe fondée sur les algorithmes proximaux, nous avons d'abord recherché et montré l'existence et l'unicité du proximal du terme d'attache aux données, puis nous avons appliqué l'algorithme de Douglas-Rachford pour trouver la solution et garantir sa convergence. Afin de comparer notre approche, nous avons présenté cinq méthodes récemment introduites qui utilisent les Equations aux Dérivées Partielles EDP et les approches variationnelles. Cette étude comparative nous a conduit à tirer la conclusion que la méthode proposée donne des résultats satisfaisants de filtrage d'images par rapport aux autres méthodes. Des critères quantitatifs ont été utilisés pour mesurer les performances tels que: le rapport signal sur bruit (PSNR), l'indice de similarité structurelle (SSIM) et le rapport visuel signal sur bruit (VSNR) et la qualité visuelle. L'évaluation de l'approche proposée montre qu'elle fournit des résultats meilleurs que ceux obtenus par les autres méthodes testées. La méthode proposée est programmée sous UNLocBoX qui est une boîte à outils d'optimisation convexe MATLAB basée sur les méthodes de fractionnement proximal.

Mots clés:

bruit multiplicatif Gamma, filtrage, variation totale, algorithme proximal, mesures de qualité,...