



HAL
open science

Décomposition d'image par modèles variationnels : Débruitage et extraction de texture.

Loïc Piffet

► **To cite this version:**

Loïc Piffet. Décomposition d'image par modèles variationnels: Débruitage et extraction de texture..
Mathématiques [math]. Université d'Orléans, 2010. Français. NNT: . tel-00598289v1

HAL Id: tel-00598289

<https://theses.hal.science/tel-00598289v1>

Submitted on 6 Jun 2011 (v1), last revised 28 Jun 2011 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ÉCOLE DOCTORALE SCIENCES ET TECHNOLOGIES
LABORATOIRE DE MATHÉMATIQUES ET APPLICATIONS,
PHYSIQUE MATHÉMATIQUE D'ORLÉANS

THÈSE présentée par :
Loïc PIFFET

Soutenue le : **23 novembre 2010**

pour obtenir le grade de **Docteur de l'université d'Orléans**

Discipline : **MATHÉMATIQUES**

**Décomposition d'image par modèles
variationnels -
Débruitage et extraction de texture.**

THÈSE dirigée par :

Maitine BERGOUNIOUX

Professeur, Université d'Orléans

RAPPORTEURS :

Jean-François AUJOL

Chargé de recherche CNRS, LATP Aix-Marseille

Pierre MARECHAL

Professeur, Université Paul Sabatier

JURY :

Romain ABRAHAM

Professeur, Université d'Orléans

Jean-François AUJOL

Chargé de recherche CNRS, LATP Aix-Marseille

Maitine BERGOUNIOUX

Professeur, Université d'Orléans

Anne-Marie JOLLY

Professeur, Université d'Orléans

Pierre MARECHAL

Professeur, Université Paul Sabatier

Simon MASNOU

Professeur, Université Lyon 1



UNIVERSITÉ D'ORLÉANS



ÉCOLE DOCTORALE SCIENCES ET TECHNOLOGIES
LABORATOIRE DE MATHÉMATIQUES ET APPLICATIONS,
PHYSIQUE MATHÉMATIQUE D'ORLÉANS

THÈSE présentée par :
Loïc PIFFET

Soutenue le : **23 novembre 2010**

pour obtenir le grade de **Docteur de l'université d'Orléans**

Discipline : **MATHÉMATIQUES**

**Décomposition d'image par modèles
variationnels -
Débruitage et extraction de texture.**

THÈSE dirigée par :

Maitine BERGOUNIOUX

Professeur, Université d'Orléans

RAPPORTEURS :

Jean-François AUJOL

Chargé de recherche CNRS, LATP Aix-Marseille

Pierre MARECHAL

Professeur, Université Paul Sabatier

JURY :

Romain ABRAHAM

Professeur, Université d'Orléans

Jean-François AUJOL

Chargé de recherche CNRS, LATP Aix-Marseille

Maitine BERGOUNIOUX

Professeur, Université d'Orléans

Anne-Marie JOLLY

Professeur, Université d'Orléans

Pierre MARECHAL

Professeur, Université Paul Sabatier

Simon MASNOU

Professeur, Université Lyon 1

Remerciements

Mes premiers remerciements sont adressés à ma directrice Maitine Bergounioux, pour avoir accepté d'encadrer cette thèse. Elle m'a permis de travailler dans des conditions optimales en étant d'une extrême disponibilité tout en me laissant une grande autonomie. Toutes les discussions que nous avons pu avoir ont toujours été très enrichissantes, que ce soit à propos des mathématiques ou non. J'espère, dans l'avenir, avoir à nouveau l'occasion de travailler et apprendre à ses côtés, et ainsi profiter de son expérience et de son incroyable intuition mathématique. Je souhaite de plus la remercier sincèrement de m'avoir accordé sa confiance. J'espère en avoir été digne.

Jean-François Aujol et Pierre Maréchal m'ont fait le grand honneur d'accepter la tâche de rapporteur. Je les remercie pour leur relecture minutieuse ainsi que pour tous leur précieux commentaires. Je remercie également Anne-Marie Jolly, Simon Masnou et Romain Abraham de l'intérêt qu'ils ont porté à mon travail en acceptant d'être membres de mon jury de thèse.

J'ai eu la grande chance de pouvoir effectuer ma thèse au MAPMO. Il règne dans ce laboratoire une ambiance extraordinaire rendant les conditions de travail idéales. Les thésards sont écoutés et entièrement intégrés à la vie du labo. Ce fut un véritable plaisir de venir au bureau durant ces trois années.

Je tiens particulièrement à remercier François James qui fut mon tuteur pédagogique dans le cadre de mon monitorat. Il m'a laissé une grande liberté dans mes enseignements, tout en étant très disponible. J'ai toujours pu compter sur son soutien ainsi que sur son aide.

J'aimerais aussi remercier Mounir Haddou pour sa disponibilité de tous les instants pour répondre à mes innombrables questions, pour sa bonne humeur et sa gentillesse. Merci à Emmanuel Trélat, qui m'a permis de me défouler les mercredi soirs au self, même si nous n'avons pas toujours été très assidus. Je me réjouis de pouvoir enseigner à ses côtés cette année. Merci également à mon ami Pierre qui me fait cracher mes poumons toutes les semaines sur les terrains de badminton.

Je n'oublie pas mes amis Nicolas et Romain, qui ont toujours fait preuve d'une extraordinaire patience face à mes lacunes en informatique, ainsi que Christelle, Anne, Marie-France

et Marie Laurence pour leur gentillesse et leur efficacité.

Comment ne pas remercier mes collègues et amis thésards. Nous avons passé ensemble durant ces trois années de très bons moments. Je tiens particulièrement à remercier Cristiana, Chi, Guillaume, Roland, Pierre, Jadeh, Radouen, Bassirou et Sebastien pour leur amitié.

Je profite de l'occasion pour adresser mes plus sincères remerciements à Karim Zayana ainsi qu'à Alain Chabas, tous deux professeurs de mathématiques en CPGE. Ils m'ont donné la chance de pouvoir effectuer des colles durant ces trois années. Ce fut pour moi une expérience très enrichissante.

C'est à ma famille que je dois la possibilité d'écrire aujourd'hui ces quelques lignes. Mes parents ont toujours fait en sorte que mes études soient mon unique préoccupation. Je n'aurais jamais pu arriver où j'en suis sans le soutien inconditionnel qu'ils m'ont apporté.

Merci à Claire de me supporter.

Table des matières

Introduction	1
I Les espaces BV et BV^2	7
1 Théorie de la mesure : définitions et résultats utiles	7
2 Les espaces de Sobolev $W^{1,p}$ et $W^{2,p}$	10
2.1 Définitions et propriétés élémentaires	10
2.2 Approximation par des fonctions régulières	11
3 L'espace BV des fonctions à variation bornée	13
4 L'espace BV^2 des fonctions à hessienne bornée	14
4.1 Définitions et propriétés	14
4.2 Fonctions convexes et BV^2	20
4.3 Résultats d'approximation et d'injection	21
4.3.1 Théorèmes d'injections	21
4.3.2 Théorème d'approximation	22
II Restauration d'images	29
1 Introduction	29
2 Le modèle de Rudin, Osher et Fatemi	31
2.1 Présentation	31
2.2 Discrétisation	31
2.3 Algorithme de projection de Chambolle	32
2.4 Effet "staircasing"	33
3 Modèle d'ordre deux	35
3.1 Introduction	35
3.2 Modèle variationnel	36
3.3 Le problème discrétisé	36
3.3.1 Discrétisation du problème (\mathcal{P}^δ)	37
3.3.2 Calcul explicite de la solution du problème (\mathcal{P}^δ)	38
3.3.3 Cas où $\delta = 0$	39
3.4 Un algorithme de point fixe pour implémenter ∂J_2	40
3.5 Résultats numériques	40
4 Modèle mixte	42
4.1 Introduction	42

4.2	Existence et unicité de la solution	43
4.3	Discrétisation du modèle	44
4.4	Résolution numérique et algorithme	45
4.5	tests numériques	46
III Extraction de texture		49
1	Etat de l'art	49
1.1	Modèle de Meyer	49
1.1.1	L'espace \mathcal{G} des fonctions oscillantes	49
1.1.2	Espace de Meyer discret	50
1.2	Modèle de Vese-Osher	50
1.3	Modèle de Osher-Solé-Vese	51
1.4	Modèle de Aujol-Aubert-Blanc-Feraud-Chambolle	51
1.5	Modèle $TV-L^1$	51
2	Retour sur le modèle d'ordre 2	52
2.1	Introduction	52
3	Amélioration du modèle	53
3.1	Modèle globalement anisotrope	54
3.2	Modèle localement anisotrope	58
3.2.1	Principe	58
3.2.2	Implémentation	61
IV Aspects numériques		65
1	Méthode de Lagrangien Augmenté	65
1.1	Introduction	65
1.2	Approximation de Moreau-Yoshida Généralisée	66
1.3	Système d'optimalité	67
1.4	Méthode de Lagrangien augmenté	69
2	Applications aux modèles $TV - L^1$ et $TV^2 - L^1$	69
2.1	Résolution explicite de l'étape 3	72
2.2	Résolution explicite de l'étape 2	74
2.3	Algorithme	75
2.4	Tests numériques	76
3	Une autre façon de résoudre les modèles $TV-L^1$ et TV^2-L^1	77
4	Modèle d'extraction de texture mixte	78
4.1	Introduction	78
4.2	Discrétisation	78
4.3	Résolution explicite de l'étape 3	80
4.4	Résolution explicite de l'étape 2	80
4.5	Algorithme	81
4.6	Exemple	82

V	Tests numériques	83
1	Algorithme plus rapide pour implémenter Π_{K_1} et Π_{K_2}	83
2	Le modèle ROF2	86
2.1	Influence du nombre d'itérations et critère d'arrêt	86
2.2	Influence de λ	88
2.3	Tests sur des signaux 1D	92
3	Le modèle mixte	95
3.1	Influence de λ	95
3.2	Influence de μ	99
4	Comparaison entre les modèles ROF, ROF2 et mixte	103
4.1	Images moyennement bruitées ($\sigma = 15$)	105
4.2	Images fortement bruitées ($\sigma = 50$)	109
5	Modèle localement anisotrope dans le cadre de l'extraction de texture	113
6	Modèles $TV-L^1$ et TV^2-L^1	119
6.1	Le modèle $TV-L^1$	120
6.2	Le modèle $TV-L^1$ avec terme d'attache aux données L^2	122
6.3	Le modèle TV^2-L^1	124
6.4	Comparaison	126
	Appendices	131
A	Article bfg	133
B	Article sva	145

Introduction

Le traitement d'image connaît de nos jours un essor considérable avec l'apparition d'une technologie de plus en plus pointue. En effet, si d'une part les ordinateurs, téléviseurs, téléphones portables et autres produits de grande consommation, dont les enjeux commerciaux et économiques sont immenses, apportent de nouveaux défis, des domaines plus "fondamentaux" soulèvent eux aussi de nombreux problèmes. En effet, nous pouvons par exemple citer la médecine et l'utilisation des radiographies ou échographies, l'observation de l'infiniment petit, de l'infiniment grand... Dans chacun des cas pré-cités, tout l'enjeu est de pouvoir extraire de l'image le maximum d'information. Cela nécessite généralement un traitement particulier : restauration, décomposition, segmentation...

Le projet ayant conduit à l'élaboration de cette thèse, concerne un sujet particulièrement important dans la région Centre : la conservation des monuments en pierre. Nous nous intéressons plus précisément au tuffeau (figure 1).



FIG. 1 – Château des ducs de Bretagne

Le tuffeau est une roche sédimentaire calcaire à grains fins, contenant une proportion variable de grains détritiques tels que le quartz ou les micas blancs. Comme toute pierre, elle subit inévitablement les ravages du temps. Les causes de dégradation sont multiples : érosion due à l'eau, au vent, à la pollution ou encore la structure même de l'édifice qui lui fait subir différentes contraintes (compression, torsion, flexion...)... Pouvoir préserver ces pierres nécessite de comprendre les mécanismes entraînant leur dégradation. Nous nous intéressons ici plus particulièrement à l'écoulement des fluides à l'intérieur de la pierre, entraînant son érosion et sa modification structurelle par de multiples réactions chimiques, le but final étant de réussir à simuler cet écoulement. Il est alors nécessaire de pouvoir situer les espaces permettant aux fluides de s'infiltrer. Nous travaillons pour cela sur des images obtenues par microtomographie à rayons X (figure 2), dans le but de les segmenter par régions.

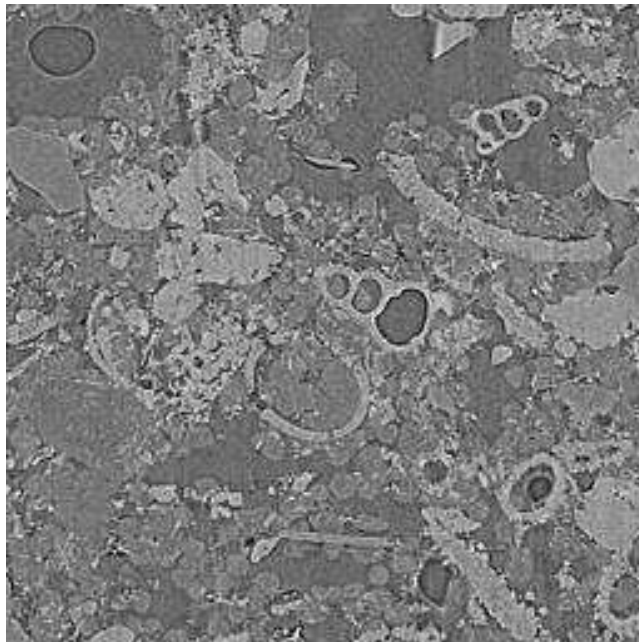


FIG. 2 – Structure du tuffeau par microtomographie à rayons X

Les méthodes actuelles permettent de localiser la plupart des zones de vide (avec plus ou moins de précision). Toutefois, ces zones ne sont pas les seules à permettre un transfert de fluide. En effet, les zones érodées ayant subi des réactions chimiques sont, elles aussi, perméables, et doivent donc être prise en compte dans l'élaboration d'un modèle d'écoulement. Ces zones sont constituées de calcite micritique, et se reconnaissent sur l'image comme étant les zones les plus texturées, zones que nous aimerions situer précisément. Si elles semblent visibles à l'oeil, les techniques actuelles d'extraction de texture ne permettent malheureusement pas d'obtenir de résultats vraiment satisfaisants. En effet, ces images sont extrêmement bruitées et possèdent de nombreux artefacts dus à

l'acquisition des données ainsi qu'aux algorithmes de reconstructions utilisés. Elles sont donc inexploitablement telles quelles et nécessitent un pré-traitement. Le problème est qu'une restauration puissante, si elle fait disparaître le bruit, fait aussi disparaître les zones très texturées qui nous intéressent. Il faut donc trouver un juste équilibre entre élimination du bruit et conservation des détails importants.

En pratique, nous utilisons le très célèbre modèle de Rudin, Osher et Fatemi (ROF), reconnu pour son efficacité en débruitage. Ce modèle engendre toutefois un inconvénient majeur sur lequel nous reviendrons plus en détail dans cette thèse : l'effet *staircasing*. En effet l'image restaurée aura tendance à être constante par morceaux, possédant ainsi des contours artificiels. Ceci est un réel problème si le but final est la segmentation de l'image par régions.

Nous allons, dans ce manuscrit, tenter d'apporter, dans une première partie, une solution au problème de l'effet *staircasing*, ainsi que, dans une seconde partie, une contribution dans le domaine de l'extraction de texture qui nous permettrait de situer précisément les zones très texturées décrites précédemment. En d'autres termes, nous nous intéressons dans ce travail au problème plus général de décomposition d'images. Voici comment s'organise le manuscrit :

Nous consacrons la première partie à l'élaboration d'un modèle de débruitage. Après avoir rappelé quelques notions de théorie de la mesure, ainsi que de nombreux résultats sur l'espace BV des fonctions à variation bornée, nous introduisons l'espace BV^2 des fonctions à hessien borné. Ceci nous permet de proposer un modèle de restauration du second ordre, directement inspiré du modèle ROF consistant moralement à remplacer la variation totale de la fonction par la variation totale de sa dérivée. Après une étude théorique du modèle, nous le discrétisons et proposons de calculer la solution en adaptant des méthodes efficaces introduites pour la résolution du modèle ROF par A. Chambolle. Nous observons que notre modèle fait bel et bien disparaître l'effet *staircasing*, mais engendre en contrepartie un effet de flou, entraînant une légère perte d'information. Afin de remédier à cet inconvénient, nous proposons enfin un modèle de restauration mixte c'est à dire un modèle intermédiaire entre le modèle ROF et notre modèle d'ordre deux nous permettant d'éliminer l'effet *staircasing* tout en préservant un maximum de détails.

La deuxième partie est consacrée à l'extraction de texture. Après un bref récapitulatif des modèles existants les plus efficaces, nous voulons observer l'efficacité du modèle d'ordre deux dans ce contexte particulier. Face à son manque manifeste de performance, nous proposons une amélioration, permettant de rendre le modèle localement anisotrope. Malgré un temps de calcul conséquent au niveau de l'implémentation numérique, nous obtenons des résultats plutôt satisfaisants. Nous nous intéressons ensuite au modèle $TV-L^1$, modèle de décomposition étant reconnu à l'heure actuelle comme l'un des plus performants. Il s'agit simplement du modèle ROF dans lequel la norme L^2 du terme d'attache aux données est remplacée par la norme L^1 . Nous proposons ici une méthode de résolution originale. Ce modèle nous confronte encore une fois au phénomène de *staircasing*, nous conduisant, là encore à proposer un modèle d'ordre deux que nous résolvons de la même manière. Toutefois, s'il permet bien d'éliminer l'effet *staircasing*, la partie texture de la décomposition de l'image est moins satisfaisante. Nous introduisons donc finalement un modèle mixte. Nous ne proposons dans ce travail qu'une étude très rapide de ce dernier modèle qui fera l'objet

d'un travail ultérieur. Dans ce manuscrit, nous ne proposerons des tests numériques que pour des images classiques pour justifier la validité (ou non) des différents modèles, les images de tuffeau étant très particulières et présentant de très nombreuses difficultés.

Voici, en détail, comment se compose ce travail :

Chapitre 1

Ce chapitre est purement théorique. Après avoir rappelé quelques notions de théorie de la mesure ainsi que sur l'espace des fonctions à variation bornée BV , nous définissons l'espace BV^2 des fonctions à hessien borné. Cet espace a été introduit par F. Demengel dans [28]. W. Hinterberger et O. Schertzer proposent également dans [36] une utilisation de cet espace dans le cadre du débruitage. Nous nous attachons ici à adapter à l'espace BV^2 les résultats connus pour l'espace BV .

Chapitre 2

Ce chapitre est entièrement dédié au problème de restauration d'image. Nous commençons par y rappeler ce qu'est le modèle de Rudin, Osher et Fatemi. Etant donnée une image originale u , on suppose qu'elle est dégradée par un bruit additif v (que nous supposons systématiquement Gaussien centré) ainsi que par un opérateur R . Le problème qui se pose alors est de pouvoir reconstruire u à partir de l'image dégradée $u_d = Ru + v$. Nous sommes alors amené à chercher u comme solution du problème

$$\inf_u \|u_d - Ru\|_{L^2}^2. \quad (1)$$

Malheureusement, il s'agit d'un problème mal posé, dont la résolution numérique nécessite l'introduction d'un terme de régularisation. Le problème devient alors

$$\inf_{u \in L^2} \|u_d - Ru\|_{L^2}^2 + L(u), \quad (2)$$

où $L(u)$ est le terme de régularisation. Dans la suite, nous ne considérerons que le cas simplifié où R est l'opérateur identité. Signalons toutefois que, dans le cas général, l'opérateur R n'est pas nécessairement inversible. Nous pouvons par exemple prendre $L(u) = \|\nabla u\|_{L^2}^2$; c'est la très classique régularisation de Tychonov. Toutefois, ce modèle n'est pas vraiment adapté, la minimisation de la norme L^2 du gradient entraînant que l'image reconstruite est beaucoup trop floutée. L'approche de Rudin, Osher et Fatemi (ROF) est beaucoup plus efficace. Elle consiste à considérer la fonction dans l'espace BV et à prendre comme terme de régularisation la variation totale $TV(u)$ de la fonction. Ceci conduit au problème

$$\inf_{u \in BV} \|u_d - u\|_{L^2}^2 + TV(u). \quad (3)$$

Ce modèle est beaucoup plus performant et n'entraîne aucun lissage excessif. Toutefois, il engendre un phénomène rendant l'image reconstruite constante par morceaux : l'effet *staircasing* (voir [29, 44, 45, 46, 48]). Afin de résoudre ce problème, nous introduisons un

modèle d'ordre deux que nous nommerons ROF2. Il consiste à remplacer la variation totale de la fonction par la variation totale de sa dérivée. Nous supposons donc que la fonction est dans l'espace BV^2 et remplaçons la variation totale $TV(u)$ par ce que l'on appellera la variation totale seconde. Ceci conduit au modèle suivant :

$$\inf_{u \in BV^2} \|u_d - u\|_{L^2}^2 + TV^2(u). \quad (4)$$

On observe alors que ce modèle est un modèle de restauration assez efficace, et qu'il permet de répondre significativement au problème de *staircasing*. Toutefois, l'image reconstruite aura tendance à être un peu trop lissée, entraînant une certaine perte d'information au niveau des différents détails de l'image. Nous proposons finalement un modèle de restauration mixte, supposant que l'image reconstruite possède une composante dans BV^2 (la partie lissée) ainsi qu'une composante dans BV (la partie "cartoon"). On introduit ainsi le modèle suivant :

$$\inf_{(u,v) \in BV \times BV^2} \|u_d - u - v\|_{L^2}^2 + TV(u) + TV^2(v). \quad (5)$$

Signalons qu'un modèle équivalent à (5) à déjà été mentionné par A. Chambolle et P-L. Lions dans [23]. Nous proposons, pour résoudre les problèmes mixte et ROF2, une adaptation des méthodes de projections et de descente de gradient semi-implicite introduites par A. Chambolle dans [22].

Chapitre 3

Nous nous intéressons dans ce chapitre au problème de l'extraction de texture. Après un bref récapitulatif des méthodes existantes les plus efficaces, nous proposons d'adapter le modèle ROF2 aux cas de l'extraction de texture. Le modèle n'est a priori pas bon (ce qui se confirme par les tests numériques). En effet, passer à l'ordre deux rend la composante "cartoon" de l'image trop lissée, ce qui implique que la quasi-totalité de l'information géométrique est contenue dans la composante oscillante. Toutefois, nous avons la possibilité de rendre ce modèle anisotrope en modifiant l'opérateur hessien, permettant d'agir sur l'information géométrique de l'image, la faisant en partie disparaître de la composante oscillante.

Chapitre 4

Ce chapitre est consacré aux modèles $TV-L^1$ et TV^2-L^1 , c'est à dire aux modèles ROF et ROF2 dans lesquels la norme L^2 du terme d'attache au données est remplacée par la norme L^1 :

$$\inf_{u \in BV} \|u_d - u\|_{L^1} + TV(u), \quad (6)$$

et

$$\inf_{u \in BV^2} \|u_d - u\|_{L^1} + TV^2(u). \quad (7)$$

Le modèle $TV-L^1$ est reconnu comme étant l'un des modèles d'extraction de texture les plus performants. Toutefois, si la décomposition est satisfaisante, la partie "cartoon" de l'image présente encore une fois cet effet *staircasing*. C'est pourquoi nous introduisons le modèle TV^2-L^1 . Malheureusement, si ce modèle résout bien le problème du *staircasing*, la décomposition de l'image n'est plus satisfaisante. En effet, la composante oscillante comprend, encore une fois, une trop grande partie de l'information géométrique. Comme dans le chapitre consacré au débruitage, nous proposons un modèle mixte :

$$\inf_{(u,v) \in BV \times BV^2} \|u_d - u - v\|_{L^1} + TV(u) + TV^2(v). \quad (8)$$

Nous ne proposons, dans cette thèse, qu'une étude rapide de ce modèle qui fera l'objet de futurs travaux.

Chapitre 5

Nous proposons dans ce chapitre tous les tests numériques, illustrant les différents modèles introduits tout au long de cette thèse.

Chapitre I

Les espaces BV et BV^2

1 Théorie de la mesure : définitions et résultats utiles

Nous rappelons dans cette section des résultats de base de théorie de la mesure, sans démonstration. Le lecteur pourra se référer à [4] pour plus de détails.

Définition I.1. Soit X un ensemble non vide et soit \mathcal{E} un ensemble de parties de X .

1. On dit que \mathcal{E} est une algèbre si $\emptyset \in \mathcal{E}$, $E_1 \cup E_2 \in \mathcal{E}$ et si $X \setminus E_1 \in \mathcal{E}$ pour tous $E_1, E_2 \in \mathcal{E}$.
2. On dit qu'une algèbre \mathcal{E} est une σ -algèbre si, pour toute suite $(E_h) \subset \mathcal{E}$, $\bigcup_h E_h \in \mathcal{E}$.
3. Si (X, τ) est un espace topologique, on note $\mathcal{B}(X)$ la σ -algèbre des boréliens de X , c'est à dire la σ -algèbre engendrée par les ouverts de X .
4. Si \mathcal{E} est une σ -algèbre dans X , on appelle la paire (X, \mathcal{E}) un espace mesuré.

Définition I.2. Soit X un espace vectoriel (sur \mathbb{R} ou \mathbb{C}), \mathcal{E} un sous-ensemble des parties de X et soit $n \in \mathbb{N}^*$.

1. On dit que $\mu : \mathcal{E} \rightarrow \mathbb{R}^n$ est une mesure si $\mu(\emptyset) = 0$ et si pour toute suite $(E_k)_{k \in \mathbb{N}}$ d'éléments deux à deux disjoints de \mathcal{E}

$$\mu \left(\bigcup_{k \in \mathbb{N}} E_k \right) = \sum_{k \in \mathbb{N}} \mu(E_k). \quad (\text{I.1})$$

(X, \mathcal{E}) est alors un espace mesuré. Si $n = 1$, on dit que μ est une mesure réelle. Si $n > 1$, on dit que μ est une mesure vectorielle.

2. Si μ est une mesure, on définit sa variation totale $|\mu|$ pour tout $E \in \mathcal{E}$ par

$$|\mu|(E) := \sup \left\{ \sum_{i=0}^{\infty} |\mu(E_i)| \mid E_i \in \mathcal{E} \text{ deux à deux disjoints, } \bigcup_{i=0}^{\infty} E_i = E \right\}.$$

Contrairement à une mesure réelle positive qui peut, par définition, prendre la valeur $+\infty$, la définition ci-dessus impose la convergence absolue de la série dans (I.1). On en déduit le résultat suivant :

Théorème I.3. *Si μ est une mesure sur (X, \mathcal{E}) , alors $|\mu|$ est une mesure positive finie.*

Définition I.4 (Ensembles μ -négligeables). Soit μ une mesure positive sur l'espace mesuré (X, \mathcal{E}) .

1. On dit que $N \subset X$ est μ -négligeable s'il existe $E \in \mathcal{E}$ tel que $N \subset E$ et $\mu(E) = 0$.
2. Soit \mathcal{E}_μ l'ensemble des parties de X de la forme $F = E \cup N$, avec $E \in \mathcal{E}$ et N μ -négligeable. Alors \mathcal{E}_μ est une σ -algèbre appelée μ -complétion de \mathcal{E} , et on dit que $E \subset X$ est μ -mesurable si $E \in \mathcal{E}_\mu$.

Définition I.5 (Fonctions mesurables). Soit (X, \mathcal{E}) un espace mesuré et (Y, d) un espace métrique.

1. Une fonction $f : X \rightarrow Y$ est dite \mathcal{E} -mesurable si $f^{-1}(A) \in \mathcal{E}$ pour tout ouvert $A \subset Y$.
2. Si μ est une mesure positive sur (X, \mathcal{E}) , la fonction f est dite μ -mesurable si elle est \mathcal{E}_μ -mesurable.

Dans ce qui suit, nous rappelons la définition de l'intégrale ainsi que les notions de fonctions sommables et intégrables pour des fonctions positives μ -mesurables. La définition est ensuite étendue aux fonctions réelles et vectorielles ainsi qu'aux mesures réelles et vectorielles.

Définition I.6 (Intégrales). 1. Une fonction $u : X \rightarrow \mathbb{R}$ est simple si elle ne prend qu'un nombre fini de valeurs, c'est à dire si elle est une combinaison linéaire de fonctions indicatrices.

2. On définit l'intégrale d'une fonction μ -mesurable $u : X \rightarrow [0, \infty]$ en posant :

$$\int_X u \, d\mu := \sup \left\{ \int_X v \, d\mu \mid v \, \mu\text{-mesurable, simple, } v \leq u \right\}.$$

On dit qu'une fonction μ -mesurable $u : X \rightarrow \overline{\mathbb{R}}$ est μ -sommable si

$$\int_X |u| \, d\mu < \infty.$$

3. Soient μ une mesure sur (X, \mathcal{E}) et $u : X \rightarrow \overline{\mathbb{R}}$ une fonction $|\mu|$ -mesurable. On dit que u est μ -sommable si u est $|\mu|$ -sommable. Si de plus μ est une mesure vectorielle \mathbb{R}^n -valuée, on définit

$$\int_X u \, d\mu := \left(\int_X u \, d\mu_1, \dots, \int_X u \, d\mu_n \right).$$

4. Si μ est une mesure réelle et $u = (u_1, \dots, u_k) : X \rightarrow \mathbb{R}^k$ est $|\mu|$ -mesurable, on dit que u est $|\mu|$ -sommable si toutes les composantes sont $|\mu|$ -sommables, et on note

$$\int_X u \, d\mu := \left(\int_X u_1 \, d\mu, \dots, \int_X u_k \, d\mu \right).$$

Définition I.7. Soit μ une mesure positive sur l'espace mesuré (X, \mathcal{E}) , et soit $f \in L^1((X, \mathcal{E}), \mathbb{R}^n)$. On définit la mesure \mathbb{R}^n -valuée

$$(f \llcorner \mu)(B) := \int_B f \, d\mu, \quad \forall B \in \mathcal{E}.$$

La variation totale d'une telle mesure est donnée par la proposition suivante.

Proposition I.1. Soit $f \llcorner \mu$ la mesure définie dans la définition I.7. Alors

$$|f \llcorner \mu|(B) = \int_B |f| \, d\mu, \quad \forall B \in \mathcal{E}.$$

Dans ce qui suit, Ω est un ouvert de \mathbb{R}^n .

Définition I.8. 1. Une mesure positive sur $(\Omega, \mathcal{B}(\Omega))$ est une mesure de Borel. Une mesure de Borel est appelée mesure de Radon positive si elle est finie sur tout compact.

2. Une mesure de $\mathcal{B}(\Omega) \rightarrow \mathbb{R}^n$ est appelée mesure de Radon finie.

Exemple I.1. La mesure de Lebesgue sur $(\Omega, \mathcal{B}(\Omega))$ est une mesure de Radon positive. Nous la noterons \mathcal{L}^n .

Proposition I.2. Soit μ une mesure de Radon \mathbb{R}^n -valuée sur Ω . Alors, pour tout ouvert $A \subset \Omega$, nous avons l'égalité suivante :

$$|\mu|(A) = \sup \left\{ \sum_{i=1}^n \int_X u_i \, d\mu_i \mid u \in \mathcal{C}_c(A), \|u\|_\infty \leq 1 \right\}.$$

Définition I.9. 1. On note $\mathcal{C}_c(\Omega, \mathbb{R}^n)$ l'espace vectoriel réel des fonctions continues à support compact définies sur Ω et à valeurs dans \mathbb{R}^n .

2. Ayant muni l'espace $\mathcal{C}_c(\Omega, \mathbb{R}^n)$ de la norme $\|u\|_\infty = \sup_{x \in \Omega} \|u(x)\|_{\mathbb{R}^n}$, le complété de cet espace pour cette norme est noté $\mathcal{C}_0(\Omega, \mathbb{R}^n)$. Il est constitué des fonctions continues sur Ω qui tendent vers zéro à l'infini.

Rappelons maintenant le théorème de Riesz :

Théorème I.10. (Riesz) Soit L une forme linéaire continue sur $\mathcal{C}_0(\Omega, \mathbb{R}^n)$. Il existe une unique mesure de Radon finie μ \mathbb{R}^n -valuée sur Ω telle que

$$L(u) = \sum_{i=1}^n \int_\Omega u_i \, d\mu_i, \quad \forall u \in \mathcal{C}_0(\Omega, \mathbb{R}^n).$$

De plus,

$$\|L\| = |\mu|(\Omega).$$

Corollaire I.1. Soit L une forme linéaire continue sur $\mathcal{C}_c(\Omega, \mathbb{R}^n)$. Il existe une unique mesure de Radon finie μ \mathbb{R}^n -valuée sur Ω telle que

$$L(u) = \sum_{i=1}^n \int_{\Omega} u_i d\mu_i, \quad \forall u \in \mathcal{C}_c(\Omega, \mathbb{R}^n).$$

Corollaire I.2. Soit $L : \mathcal{C}_c^\infty(\Omega) \rightarrow \mathbb{R}$ une application linéaire positive i.e.

$$L(f) \geq 0, \quad \forall f \in \mathcal{C}_c^2(\Omega), f \geq 0.$$

Alors il existe une mesure de Radon finie μ sur Ω telle que

$$L(f) = \int_{\Omega} f d\mu, \quad \forall f \in \mathcal{C}_c^2(\Omega).$$

Démonstration : Soit un compact $K \subset \Omega$ et soit une fonction régulière ζ telle que ζ soit à support compact, $\zeta \equiv 1$ sur K , $0 \leq \zeta \leq 1$. Ensuite, pour tout $f \in \mathcal{C}_c^\infty(\Omega)$ à support inclus dans Ω , on pose $g \equiv \|f\|_{L^\infty} \zeta - f \geq 0$. Alors, l'inégalité implique

$$0 \leq L(g) = \|f\|_{L^\infty} L(\zeta) - L(f),$$

et donc

$$L(f) \leq C \|f\|_{L^\infty}$$

avec $C \equiv L(\zeta)$. L est donc une forme linéaire continue sur $\mathcal{C}_c^\infty(\Omega)$ que l'on peut prolonger (Hahn-Banach) sur $\mathcal{C}_c(\Omega)$. Les hypothèses du théorème de Riesz sont donc vérifiées. Par suite, il existe une unique mesure de Radon finie sur Ω vérifiant

$$L(f) = \int_{\Omega} f d\mu, \quad f \in \mathcal{C}_c^\infty(\Omega).$$

■

2 Les espaces de Sobolev $W^{1,p}$ et $W^{2,p}$

Nous rappelons ici quelques définitions et résultats élémentaires sur les espaces de Sobolev des fonction définies sur \mathbb{R}^n , dont les dérivées faibles premières et secondes appartiennent à un espace L^p . Pour plus de précisions sur les espaces de Sobolev, nous renvoyons le lecteur à [2], [6], ou encore [35]. Dans cette section, Ω est un ouvert de \mathbb{R}^n .

2.1 Définitions et propriétés élémentaires

Définition I.11. Soit $f \in L^1(\Omega)$, $1 \leq i \leq n$. On dit que $g_i \in L^1(\Omega)$ est la dérivée partielle faible de f par rapport à x_i dans Ω et que $h_{i,j} \in L^1(\Omega)$ est la dérivée partielle faible seconde de f par rapport à (x_i, x_j) dans Ω si,

$$\int_{\Omega} f \frac{\partial \varphi}{\partial x_i} dx = - \int_{\Omega} g_i \varphi dx, \quad \forall \varphi \in \mathcal{C}_c^1(\Omega),$$

et

$$\int_{\Omega} f \frac{\partial^2 \varphi}{\partial x_i \partial x_j} dx = \int_{\Omega} h_{i,j} \varphi dx, \quad \forall \varphi \in \mathcal{C}_c^2(\Omega).$$

Il s'agit ici des dérivées au sens des distributions.

Nous écrivons alors, pour $i = 1, \dots, n$,

$$\frac{\partial f}{\partial x_i} \equiv g_i \text{ et } \frac{\partial^2 f}{\partial x_i \partial x_j} \equiv h_{i,j}.$$

Définition I.12. Soit $1 \leq p \leq \infty$.

1. La fonction f appartient à l'espace

$$W^{1,p}(\Omega)$$

si $f \in L^p(\Omega)$ et si $\frac{\partial f}{\partial x_i}$ existe et appartient à $L^p(\Omega)$ pour $i = 1, \dots, n$.

2. La fonction f appartient à l'espace

$$W^{2,p}(\Omega)$$

si $f \in W^{1,p}(\Omega)$ et si $\frac{\partial^2 f}{\partial x_i \partial x_j}$ existe et appartient à $L^p(\Omega)$ pour $i, j = 1, \dots, n$.

3. La fonction f appartient à l'espace $W_{\text{loc}}^{1,p}(\Omega)$ (respectivement $W_{\text{loc}}^{2,p}(\Omega)$), si $f1_K \in W^{1,p}(\Omega)$ (respectivement $f1_K \in W^{2,p}(\Omega)$) pour tout ensemble compact K inclus dans Ω , où 1_K désigne la fonction indicatrice de K .

Pour $1 \leq p < \infty$, nous munissons $W^{1,p}(\Omega)$ de la norme suivante :

$$\|f\|_{W^{1,p}(\Omega)} \equiv \left(\int_{\Omega} |f|^p + \|\nabla f\|_{\mathbb{R}^n}^p dx \right)^{1/p}.$$

Définition I.13. On dit que

$$f_k \rightarrow f \text{ dans } W^{1,p}(\Omega)$$

si

$$\|f_k - f\|_{W^{1,p}(\Omega)} \rightarrow 0.$$

2.2 Approximation par des fonctions régulières

Définition I.14. 1. Si $\varepsilon > 0$, on note $\Omega_\varepsilon \equiv \{x \in \Omega \mid \text{dist}(x, \partial\Omega) > \varepsilon\}$.

2. On définit la fonction \mathcal{C}^∞ , $\eta : \mathbb{R}^n \rightarrow \mathbb{R}$ de la façon suivante

$$\begin{cases} c \exp\left(\frac{1}{|x|^2-1}\right) & \text{si } |x| < 1 \\ 0 & \text{sinon.} \end{cases}$$

En ajustant la constante pour que

$$\int_{\mathbb{R}^n} \eta(x) dx = 1.$$

On pose alors, pour $\varepsilon > 0$ et $x \in \mathbb{R}^n$,

$$\eta_\varepsilon \equiv \frac{1}{\varepsilon^n} \eta\left(\frac{x}{\varepsilon}\right).$$

η_ε est une approximation de l'unité.

3. Si $f \in L^1_{\text{loc}}(\Omega)$, on définit

$$f^\varepsilon \equiv \eta_\varepsilon * f;$$

c'est à dire, pour $x \in \Omega_\varepsilon$,

$$f^\varepsilon(x) \equiv \int_{\Omega} n_\varepsilon(x-y) f(y) dy.$$

Remarque I.1. Ainsi définie, η_ε est à support compact.

On rappelle le résultat suivant, dont une démonstration est proposée dans [35].

Théorème I.15. 1. Pour tout $\varepsilon > 0$, $f^\varepsilon \in C^\infty(\Omega_\varepsilon)$.

2. Si $f \in \mathcal{C}(\Omega)$, alors

$$f^\varepsilon \rightarrow f$$

uniformément sur tout sous-ensemble compact de Ω .

3. Si $f \in L^p_{\text{loc}}(\Omega)$ pour $1 \leq p < \infty$, alors

$$f^\varepsilon \rightarrow f \text{ dans } L^p_{\text{loc}}(\Omega).$$

4. De plus,

$$f^\varepsilon \rightarrow f \text{ } \mathcal{L}^n \text{ p.p.}$$

5. Si $f \in W^{1,p}_{\text{loc}}(\Omega)$ pour $1 \leq p \leq \infty$, alors

$$\frac{\partial f^\varepsilon}{\partial x_i} = \eta_\varepsilon * \frac{\partial f}{\partial x_i}, \quad i = 1, \dots, n$$

sur Ω_ε .

6. En particulier, si $f \in W^{1,p}_{\text{loc}}(\Omega)$ pour $1 \leq p \leq \infty$, alors

$$f^\varepsilon \rightarrow f \text{ dans } W^{1,p}_{\text{loc}}(\Omega).$$

Nous pouvons remarquer que les résultats 5 et 6 s'adaptent de façon évidente au cas où $f \in W^{2,p}_{\text{loc}}(\Omega)$ pour $1 \leq p \leq \infty$, en considérant les dérivées partielles d'ordre deux de f . Nous donnons maintenant un résultat d'approximation régulière des fonctions de $W^{1,p}(\Omega)$, lui aussi démontré dans [35] :

Théorème I.16. Soit $f \in W^{1,p}(\Omega)$, $1 \leq p < \infty$. Il existe une suite $(f_k)_{k \in \mathbb{N}}$ d'éléments de $W^{1,p}(\Omega) \cap C^\infty(\Omega)$ telle que

$$f_k \rightarrow f \text{ dans } W^{1,p}(\Omega).$$

3 L'espace BV des fonctions à variation bornée

Nous rappelons, dans cette partie la définition de l'espace des fonctions à variation bornée, ainsi que quelques résultats de base. Les ouvrages [4, 35, 6] proposent des études précises de cet espace. Nous y renvoyons le lecteur pour les démonstrations des différents résultats énoncés ci-dessous. Dans ce qui suit, Ω est un ouvert borné de \mathbb{R}^n .

Définition I.17. Soit $u \in L^1(\Omega)$. On dit que u est une fonction à variation bornée dans Ω si la dérivée de u au sens des distributions est une mesure de Radon finie, c'est à dire si

$$\int_{\Omega} u \frac{\partial \varphi}{\partial x_i} dx = - \int_{\Omega} \varphi dD_i u, \quad \forall \varphi \in C_c^\infty(\Omega), \quad i = 1, \dots, n \quad (\text{I.2})$$

pour une mesure \mathbb{R}^n -valuée $Du = (D_1 u, \dots, D_n u)$ dans Ω . L'espace vectoriel de toutes les fonctions à variation bornée est noté $BV(\Omega)$.

Remarque I.2. La formule (I.2) reste valable pour $\varphi \in C_c^1(\Omega)$. Elle peut encore s'écrire

$$\int_{\Omega} u \operatorname{div} \varphi dx = - \sum_{i=1}^n \int_{\Omega} \varphi_i dD_i u, \quad \forall \varphi \in C_c^1(\Omega), \quad \forall \varphi \in C_c^1(\Omega, \mathbb{R}^n). \quad (\text{I.3})$$

Nous pouvons d'ores et déjà remarquer que l'espace de Sobolev $W^{1,1}(\Omega)$ est inclus dans $BV(\Omega)$; en effet, pour tout $u \in W^{1,1}$, la dérivée au sens des distributions est donnée par $-\nabla u \llcorner \mathcal{L}^n$. En particulier, Ω étant borné, on a

$$W^{1,p}(\Omega) \subset BV(\Omega), \quad \forall 1 \leq p \leq +\infty$$

c'est à dire que toute fonction d'un espace de Sobolev est à variation bornée. De plus, d'après la proposition I.1, $|Df|(\Omega) = \|\nabla f\|_{L^1(\Omega, \mathbb{R}^n)}$.

Définition I.18. On dit qu'une fonction u est dans $BV_{\text{loc}}(\Omega)$ si, pour tout compact K inclus dans Ω $u \mathbb{1}_K$ est un élément de $BV(\Omega)$.

Définition I.19. Soit $u \in L^1_{\text{loc}}(\Omega)$. La variation totale de u dans Ω est définie par

$$TV(u, \Omega) := \sup \left\{ \int_{\Omega} u \operatorname{div} \varphi dx \mid \varphi \in C_c^1(\Omega, \mathbb{R}^n), \|\varphi\|_{\infty} \leq 1 \right\}.$$

Le résultat suivant donne la relation entre la variation totale de la fonction u et la variation totale au sens des mesures de la dérivée de u au sens des distributions. Voir [4] pour la démonstration.

Proposition I.3 (Variation des fonctions BV). Une fonction u de $L^1(\Omega)$ est un élément de $BV(\Omega)$ si et seulement si $TV(u, \Omega) < \infty$. De plus $TV(u, \Omega) = |Du|(\Omega)$.

Proposition I.4. L'application

$$\begin{aligned} BV(\Omega) &\rightarrow \mathbb{R}^+ \\ u &\mapsto \|u\|_{BV(\Omega)} = \|u\|_{L^1(\Omega)} + TV(u, \Omega) \end{aligned}$$

est une norme.

On munit désormais l'espace BV de cette norme. Les résultats suivants sont démontrés (par exemple) dans [35].

Théorème I.20 (Semi-continuité inférieure de la variation totale). *L'application $u \mapsto TV(u, \Omega)$ de $BV(\Omega)$ dans \mathbb{R}^+ est semi-continue inférieurement pour la topologie séquentielle de $L^1(\Omega)$. Plus précisément, si $(u_k)_{k \in \mathbb{N}}$ est une suite de fonctions de $BV(\Omega)$ qui converge fortement vers u dans $L^1(\Omega)$, alors*

$$TV(u, \Omega) \leq \liminf_{k \rightarrow +\infty} TV(u_k, \Omega).$$

Théorème I.21 (Approximation régulière). *Pour toute fonction $u \in BV(\Omega)$, il existe une suite de fonctions $(u_k)_{k \in \mathbb{N}}$ de $BV(\Omega) \cap C^\infty(\Omega)$ telle que*

1. $u_k \rightarrow u$ dans $L^1(\Omega)$ et
2. $TV(u_k, \Omega) \rightarrow TV(u, \Omega)$.

Théorème I.22. *L'espace $BV(\Omega)$ muni de la norme*

$$\|u\|_{BV(\Omega)} = \|u\|_{L^1(\Omega)} + TV(u, \Omega)$$

est un espace de Banach.

Terminons par un résultat de compacité démontré dans [4].

Théorème I.23. (Résultats d'injection) *Si $n = 1$,*

$$BV(\Omega) \hookrightarrow L^\infty(\Omega). \tag{I.4}$$

Si $n > 1$,

$$BV(\Omega) \hookrightarrow L^q(\Omega) \quad \text{avec} \quad 1 \leq q \leq \frac{n}{n-1}, \tag{I.5}$$

et l'injection est compacte si $q < \frac{n}{n-1}$.

4 L'espace BV^2 des fonctions à hessienne bornée

4.1 Définitions et propriétés

Dans ce qui suit Ω est un ouvert borné de \mathbb{R}^n .

Définition I.24. Soit $u \in W^{1,1}(\Omega)$. On dit que u est une fonction à hessien borné dans Ω si la dérivée seconde de u au sens des distributions est une mesure de Radon finie, c'est à dire si

$$\int_{\Omega} u \frac{\partial^2 \varphi}{\partial x_i \partial x_j} dx = \int_{\Omega} \varphi dD_{i,j}^2 u, \quad \forall \varphi \in C_c^\infty(\Omega), \quad i, j = 1, \dots, n \tag{I.6}$$

pour une mesure $\mathbb{R}^{n \times n}$ -valuée définie par $D^2 f = \begin{pmatrix} D_{1,1}^2 & D_{1,2}^2 & \cdots & D_{1,n}^2 \\ D_{2,1}^2 & D_{2,2}^2 & \cdots & D_{2,n}^2 \\ \vdots & \vdots & \ddots & \vdots \\ D_{n,1}^2 & D_{n,2}^2 & \cdots & D_{n,n}^2 \end{pmatrix}$, dans Ω .

L'espace vectoriel de toutes les fonctions à hessien borné est noté $BV^2(\Omega)$.

Remarque I.3. La formule (I.6) reste valable pour $\varphi \in \mathcal{C}_c^2(\Omega)$ et elle peut encore s'écrire

$$\int_{\Omega} u H^* \varphi \, dx = \sum_{i,j=1}^n \int_{\Omega} \varphi_{i,j} \, dD_{i,j}^2 u, \quad \forall \varphi \in \mathcal{C}_c^2(\Omega, \mathbb{R}^{n \times n}), \quad (\text{I.7})$$

où $H^* : \varphi \mapsto \sum_{i,j=1}^n \frac{\partial^2 \varphi}{\partial x_i \partial x_j}$ est l'opérateur adjoint de l'opérateur hessien H . Nous pouvons

d'ores et déjà remarquer que l'espace de Sobolev $W^{2,1}(\Omega)$ est inclus dans $BV^2(\Omega)$; en effet, pour tout $u \in W^{2,1}(\Omega)$, la dérivée au sens des distributions est donnée par $Hu \llcorner \mathcal{L}^n$. En particulier, Ω étant borné, on a

$$W^{2,p}(\Omega) \subset BV^2(\Omega), \quad \forall 1 \leq p \leq +\infty.$$

De plus, d'après la proposition I.1, $|Df^2|(\Omega) = \|Hf\|_{L^1(\Omega, \mathbb{R}^{n \times n})}$.

Définition I.25. On dit qu'une fonction u est dans $BV_{\text{loc}}^2(\Omega)$ si, pour tout sous-ensemble compact K de Ω , $u1_K$ est un élément de $BV^2(\Omega)$.

Définition I.26. Soit $u \in W_{\text{loc}}^{1,1}(\Omega)$. La variation seconde de u dans Ω est définie par

$$TV^2(u, \Omega) := \sup \left\{ \int_{\Omega} u H^* \varphi \, dx \mid \varphi \in \mathcal{C}_c^2(\Omega, \mathbb{R}^{n \times n}), \|\varphi\|_{\infty} \leq 1 \right\},$$

$$\text{avec } \|\varphi\|_{\infty} = \sup_{x \in \Omega} \sqrt{\sum_{i,j=1}^n |\varphi_{i,j}(x)|^2}.$$

Proposition I.5 (Variation seconde des fonctions BV^2). Une fonction u de $W^{1,1}(\Omega)$ est un élément de $BV^2(\Omega)$ si et seulement si $TV^2(u, \Omega) < \infty$. De plus, $TV^2(u, \Omega) = |D^2 u|(\Omega)$.

Démonstration : On suppose que $u \in BV^2(\Omega)$. Ainsi, pour tout $\varphi \in \mathcal{C}_c^2(\Omega, \mathbb{R}^{n \times n})$, on a

$$\int_{\Omega} u H^* \varphi \, dx = \sum_{i,j=1}^n \int_{\Omega} \varphi_{i,j} \, dD_{i,j}^2 u.$$

Nous avons, d'après la propriété I.2 et la définition de $TV^2(u, \Omega)$,

$$TV^2(u, \Omega) \leq |D^2 u|(\Omega) < +\infty.$$

Réciproquement, on suppose que $TV^2(u, \Omega) < +\infty$. Un argument d'homogénéité montre que

$$\left| \int_{\Omega} u H^* \varphi \, dx \right| \leq TV^2(u, \Omega) \|\varphi\|_{\infty}, \quad \forall \varphi \in \mathcal{C}_c^2(\Omega, \mathbb{R}^{n \times n}).$$

Comme $\mathcal{C}_c^2(\Omega, \mathbb{R}^n)$ est dense dans $\mathcal{C}_0(\Omega, \mathbb{R}^n)$, il existe une forme linéaire L définie sur $\mathcal{C}_0(\Omega, \mathbb{R}^{n \times n})$, et qui coïncide avec

$$\varphi \mapsto \int_{\Omega} u H^* \varphi \, dx$$

sur $\mathcal{C}_c^2(\Omega, \mathbb{R}^n)$. De plus, on sait que ces formes linéaires ont même norme, et donc $\|L\| \leq TV^2(u, \Omega)$.

D'après le théorème de Riesz, il existe une mesure de Radon, $\mathbb{R}^{n \times n}$ -valuée : $\mu = (\mu_{i,j})_{i,j}$ telle que $\|L\| = |\mu|(\Omega)$, et telle que

$$L(\varphi) = \sum_{i,j=1}^n \int_{\Omega} \varphi_{i,j} \, d\mu_{i,j}, \quad \forall \varphi \in \mathcal{C}_0(\Omega, \mathbb{R}^{n \times n}).$$

D'après (I.7) et l'identité

$$\int_{\Omega} u H^* \varphi \, dx = \sum_{i,j=1}^n \int_{\Omega} \varphi_{i,j} \, d\mu_{i,j}, \quad \forall \varphi \in \mathcal{C}_c^2(\Omega, \mathbb{R}^{n \times n}),$$

on obtient que $u \in BV^2(\Omega)$. Par suite, $D^2u = \mu$, et

$$|D^2u|(\Omega) = |\mu|(\Omega) = \|L\| \leq TV^2(u, \Omega).$$

Ceci implique que $|D^2u|(\Omega) = TV^2(u, \Omega)$.

■

Remarque I.4. Contrairement à l'espace $BV(\Omega)$, une fonction de $BV^2(\Omega)$ est, par définition, très régulière et ne peut donc pas comporter de sauts. Une fonction de $BV^2(\Omega)$ sera donc relativement lisse. Ceci explique l'effet de flou observé lors des tests numériques.

La proposition suivante donne la relation fondamentale entre l'espace $BV(\Omega)$ et $BV^2(\Omega)$:

Proposition I.6. Une fonction f est dans $BV^2(\Omega)$ si et seulement si $f \in W^{1,1}(\Omega)$ et $\frac{\partial f}{\partial x_i} \in BV(\Omega)$ pour $i \in \{1, \dots, n\}$. De plus,

$$TV^2(f, \Omega) \leq \sum_{i=1}^n \left| \frac{\partial f}{\partial x_i} \right|_{BV(\Omega)} \leq n TV^2(f, \Omega). \quad (\text{I.8})$$

Démonstration : Nous remarquons tout d'abord que le fait que $\mathcal{C}_c^\infty(\Omega)$ soit dense dans $\mathcal{C}_c^1(\Omega)$ et dans $\mathcal{C}_c^2(\Omega)$ ainsi que le fait que Du et D^2u soient des mesures de Radon finies impliquent que

$$TV(u, \Omega) = \sup \left\{ \int_{\Omega} u \operatorname{div} \varphi \, dx \mid \varphi \in \mathcal{C}_c^\infty(\Omega, \mathbb{R}^n), \|\varphi\|_\infty \leq 1 \right\},$$

et

$$TV^2(u, \Omega) = \sup \left\{ \int_{\Omega} u H^* \varphi \, dx \mid \varphi \in \mathcal{C}_c^\infty(\Omega, \mathbb{R}^{n \times n}), \|\varphi\|_\infty \leq 1 \right\}.$$

Pour démontrer la proposition, il suffit de démontrer (I.8).

• *Première inégalité :* soit $\varphi \in \mathcal{C}_c^\infty(\Omega, \mathbb{R}^{n \times n})$; on a en intégrant par parties :

$$\begin{aligned} \int_{\Omega} f(x) H^* \varphi(x) \, dx &= \int_{\Omega} f(x) \sum_{i,j=1}^n \frac{\partial^2 \varphi_{ij}}{\partial x_i \partial x_j}(x) \, dx \\ &= - \sum_{i,j=1}^n \int_{\Omega} \frac{\partial f}{\partial x_i} \frac{\partial \varphi_{ij}}{\partial x_j} \, dx \\ &= \sum_{i=1}^n \int_{\Omega} \frac{\partial f}{\partial x_i} \sum_{j=1}^n \frac{\partial(-\varphi_{ij})}{\partial x_j} \, dx \end{aligned}$$

Or $\sum_{j=1}^n \frac{\partial(-\varphi_{ij})}{\partial x_j} = \operatorname{div} L_i(\varphi)$, où $L_i(\varphi)$ est la i -ème ligne de la matrice $-\varphi$. De plus, $L_i(\varphi)$

vérifie clairement la propriété $\|L_i(\varphi)\|_\infty \leq 1$. Par suite,

$$\int_{\Omega} f H^* \varphi \, dx \leq \sum_{i=1}^n \left| \frac{\partial f}{\partial x_i} \right|_{BV(\Omega)}.$$

On obtient alors la première inégalité en passant à la borne supérieure sur φ .

• *Seconde inégalité :* d'après la remarque du début de la démonstration,

$$\left| \frac{\partial f}{\partial x_i} \right|_{BV(\Omega)} = \sup \left\{ \int_{\Omega} \frac{\partial f}{\partial x_i} \operatorname{div} \varphi, \varphi \in \mathcal{C}_c^\infty(\Omega, \mathbb{R}^n) \right\}.$$

Pour toute fonction $\varphi \in \mathcal{C}_c^\infty(\Omega, \mathbb{R}^n)$, on a en intégrant par parties

$$\begin{aligned} \int_{\Omega} \frac{\partial f}{\partial x_i}(x) \operatorname{div} \varphi(x) \, dx &= \int_{\Omega} \frac{\partial f}{\partial x_i}(x) \sum_{j=1}^n \frac{\partial \varphi_j}{\partial x_j}(x) \, dx \\ &= \int_{\Omega} f(x) \sum_{j=1}^n \frac{\partial^2 \varphi_j}{\partial x_j \partial x_i}(x) \, dx. \end{aligned}$$

On définit alors la fonction $\psi^k : \Omega \rightarrow \mathbb{R}^{n \times n}$ de la manière suivante :

$$\psi^k(x) = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \\ \varphi_1(x) & \cdots & \varphi_n(x) \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix},$$

où la ligne non nulle est la k -ième ligne. Ainsi,

$$\sum_{j=1}^n \frac{\partial^2 \varphi_j}{\partial x_k \partial x_j} = \sum_{i,j=1}^n \frac{\partial^2 \psi_{ij}^k}{\partial x_i \partial x_j},$$

d'où, $\forall k \in \{1, \dots, N\}$,

$$\begin{aligned} \int_{\Omega} \frac{\partial f}{\partial x_k}(x) \operatorname{div} \varphi(x) \, dx &= - \int_{\Omega} f(x) \sum_{i,j=1}^n \frac{\partial \psi_{ij}^k}{\partial x_i \partial x_j}(x) \, dx \\ &= - \int_{\Omega} f(x) H^* \psi(x) \, dx \\ &\leq TV^2(f, \Omega). \end{aligned}$$

Ceci étant vrai pour tout $\varphi \in \mathcal{C}_c^\infty(\Omega, \mathbb{R}^n)$, on obtient alors, $\forall i \in \{1, \dots, n\}$,

$$\left| \frac{\partial f}{\partial x_i} \right|_{BV(\Omega)} \leq TV^2(f, \Omega);$$

Par suite,

$$\sum_{i=1}^n \left| \frac{\partial f}{\partial x_i} \right|_{BV(\Omega)} \leq n TV^2(f, \Omega).$$

D'où la conclusion. ■

Remarque I.5. Ceci nous montre que

$$BV^2(\Omega) = \left\{ f \in W^{1,1}(\Omega) \mid \forall i \in \{1, \dots, n\}, \frac{\partial f}{\partial x_i} \in BV(\Omega) \right\}.$$

Théorème I.27 (semi-continuité inférieure). Soit $\{f_k\}_{k \in \mathbb{N}}$ une suite de fonctions de $BV^2(\Omega)$ qui converge vers f fortement dans $W^{1,1}(\Omega)$. Alors

$$TV^2(f, \Omega) \leq \liminf_{k \rightarrow \infty} TV^2(f_k, \Omega).$$

Démonstration : Soit $\varphi \in \mathcal{C}_c^2(\Omega, \mathbb{R}^{n \times n})$. On a

$$\int_{\Omega} f H^* \varphi \, dx = \lim_{k \rightarrow +\infty} \int_{\Omega} f_k H^* \varphi \, dx$$

Or

$$\int_{\Omega} f_k H^* \varphi \, dx \leq TV^2(f_k, \Omega). \quad (\text{I.9})$$

Ainsi,

$$\lim_{k \rightarrow +\infty} \int_{\Omega} f_k H^* \varphi \, dx \leq \liminf_{k \rightarrow +\infty} TV^2(f_k, \Omega).$$

On obtient le résultat en passant à la borne supérieure sur φ . ■

Remarque I.6. *En particulier, si $\liminf_{k \rightarrow \infty} TV^2(f_k, \Omega) < \infty$, alors $f \in BV^2(\Omega)$.*

Théorème I.28. *Muni de la norme*

$$\|f\|_{BV^2(\Omega)} = \|f\|_{BV(\Omega)} + TV^2(f, \Omega),$$

$BV^2(\Omega)$ est un espace de Banach.

Démonstration : Soit $(f_n)_{n \in \mathbb{N}}$ une suite de Cauchy dans $BV^2(\Omega)$. Comme $f_n \in W^{1,1}(\Omega)$, et que donc $TV(f_n, \Omega) + \|f_n\|_{L^1(\Omega)} = \|f_n\|_{W^{1,1}(\Omega)}$, on en déduit que $(f_n)_{n \in \mathbb{N}}$ est une suite de Cauchy dans $W^{1,1}(\Omega)$. Par suite, elle converge vers $f \in W^{1,1}(\Omega)$. D'autre part, elle est bornée dans $BV^2(\Omega)$. Ainsi,

$$\exists M > 0, \quad TV^2(f_n, \Omega) \leq M.$$

D'après le résultat de semi-continuité inférieure, il vient

$$TV^2(f, \Omega) \leq \liminf_{k \rightarrow \infty} TV^2(f_k, \Omega) \leq M < +\infty.$$

Par conséquent, $f \in BV^2(\Omega)$. Soit $\varepsilon > 0$ et $N \in \mathbb{N}$ tels que

$$\forall k, n > N, \quad \|f_n - f_k\|_{BV^2(\Omega)} \leq \varepsilon.$$

Alors

$$TV^2(f_n - f_k) \leq \varepsilon.$$

En fixant n , on obtient ainsi avec la semi-continuité inférieure

$$\forall N > \mathbb{N}, \quad TV^2(f_n - f) \leq \liminf_{k \rightarrow \infty} TV^2(f_n - f_k) \leq \varepsilon.$$

Ceci prouve que $TV^2(f_n - f)$ tend vers 0, d'où le résultat. ■

4.2 Fonctions convexes et BV^2

Théorème I.29. *Soit $f : \Omega \rightarrow \mathbb{R}$ une fonction convexe. Alors, pour tout $(i, j) \in (1, \dots, n)$, il existe une mesure de Radon $\mu_{i,j}$ telle que*

$$\int_{\Omega} f \frac{\partial^2 \varphi}{\partial x_i \partial x_j} = \int_{\Omega} \varphi d\mu_{i,j}, \quad \forall \varphi \in \mathcal{C}_c^2(\Omega).$$

De plus $\mu_{i,j} = \mu_{j,i}$.

Démonstration : Soient ξ un élément de \mathbb{R}^n tel que $|\xi| = 1$ et η_ε une suite régularisante (cf. page 11). On pose $f^\varepsilon = \eta_\varepsilon * f$. Alors f^ε est une fonction convexe régulière, donc

$$D^2 f^\varepsilon \geq 0.$$

Ainsi, pour tout φ dans $\mathcal{C}_c^2(\Omega)$ tel que $\varphi \geq 0$, on a en intégrant par parties

$$\sum_{i,j=1}^n \int_{\Omega} f^\varepsilon \frac{\partial^2 \varphi}{\partial x_i \partial x_j} \xi_i \xi_j dx = \int_{\Omega} \varphi \sum_{i,j=1}^n \frac{\partial^2 f^\varepsilon}{\partial x_i \partial x_j} \xi_i \xi_j dx \geq 0.$$

En effet, f étant convexe,

$$\varphi \sum_{i,j=1}^n \frac{\partial^2 f^\varepsilon}{\partial x_i \partial x_j} \xi_i \xi_j = \varphi \times \langle D^2 f^\varepsilon \xi, \xi \rangle \geq 0.$$

On fait alors tendre ε vers 0, ce qui nous donne

$$L_\xi(\varphi) := \sum_{i,j=1}^n \int_{\Omega} f \frac{\partial^2 \varphi}{\partial x_i \partial x_j} \xi_i \xi_j dx \geq 0.$$

D'après le corollaire I.2 du théorème de Riesz, il existe une mesure de Radon finie μ_ξ telle que

$$L_\xi(\varphi) = \int_{\Omega} \varphi d\mu_\xi, \quad \forall \varphi \in \mathcal{C}_c^2(\Omega).$$

Soit $\mu_{ii} = \mu_{e_i}$ pour $i = 1, \dots, n$ et en notant $(e_i)_i$ la base canonique de \mathbb{R}^n . Alors

$$\int_{\Omega} f \frac{\partial^2 \varphi}{\partial x_i \partial x_i} dx = \int_{\Omega} \varphi d\mu_{i,i}.$$

Si $i \neq j$, on pose $\xi = (e_i + e_j)/\sqrt{2}$. Dans ce cas,

$$\sum_{k,l=1}^n \frac{\partial^2 \varphi}{\partial x_k \partial x_l} \xi_k \xi_l = \frac{1}{2} \left[\frac{\partial^2 \varphi}{\partial x_i \partial x_i} + 2 \frac{\partial^2 \varphi}{\partial x_i \partial x_j} + \frac{\partial^2 \varphi}{\partial x_j \partial x_j} \right].$$

Ainsi,

$$\begin{aligned}
\int_{\Omega} f \frac{\partial^2 \varphi}{\partial x_i \partial x_j} \xi_i \xi_j dx &= \int_{\Omega} f \sum_{k,l=1}^n \frac{\partial^2 \varphi}{\partial x_k \partial x_l} \xi_k \xi_l dx - \frac{1}{2} \left[\int_{\Omega} f \frac{\partial^2 \varphi}{\partial x_i \partial x_i} dx + \int_{\Omega} f \frac{\partial^2 \varphi}{\partial x_j \partial x_j} dx \right] \\
&= \int_{\Omega} \varphi d\mu^{\xi} - \frac{1}{2} \int_{\Omega} \varphi d\mu_{i,i} - \frac{1}{2} \int_{\Omega} \varphi d\mu_{j,j} \\
&= \int_{\Omega} \varphi d\mu_{i,j},
\end{aligned}$$

où $\mu_{i,j} := \mu^{\xi} - \frac{1}{2}\mu_{i,i} - \frac{1}{2}\mu_{j,j}$. Il est alors clair que $\mu_{i,j} = \mu_{j,i}$. ■

Théorème I.30. *Soit $f \in W_{\text{loc}}^{1,1}(\Omega)$ une fonction convexe. Alors*

$$f \in BV_{\text{loc}}^2(\Omega) \text{ pour } i \in \{1, \dots, n\}.$$

Démonstration : Soit K un compact de Ω et soit $\varphi \in \mathcal{C}_c^{\infty}(K, \mathbb{R}^n)$. Pour $k = 1, \dots, n$, on a

$$\begin{aligned}
\int_K \frac{\partial f}{\partial x_k} \operatorname{div} \varphi dx &= - \int_K f \sum_{i=1}^n \frac{\partial^2 \varphi_i}{\partial x_i \partial x_k} dx \\
&= \sum_{i=1}^n \int_K \varphi_i d\mu^{ik} \\
&\leq \sum_{i=1}^n \mu_{i,k}(K) < \infty,
\end{aligned}$$

les $\mu_{i,k}$ étant des mesures de Radon finies. Ainsi, d'après la proposition I.6 $f \in BV_{\text{loc}}^2(\Omega)$. ■

Remarque I.7. *Les théorèmes précédents montrent la relation fondamentale entre les fonctions convexes et l'espace des fonctions à hessien borné.*

4.3 Résultats d'approximation et d'injection

4.3.1 Théorèmes d'injections

Nous citons dans cette partie des résultats d'injection démontrés par F. Demengel dans [28]. Ω est désormais un ouvert borné de \mathbb{R}^n .

Théorème I.31 (Résultats d'injection). *Soit $n > 1$.*

$$BV^2(\Omega) \hookrightarrow W^{1,q}(\Omega) \text{ avec } q \leq \frac{n}{n-1}, \quad (\text{I.10})$$

et l'injection est compacte si $q < \frac{n}{n-1}$. En particulier,

$$BV^2(\Omega) \hookrightarrow L^q(\Omega) \text{ pour } q \leq \frac{n}{n-2} \text{ pour } n > 2, \quad (\text{I.11})$$

$$BV^2(\Omega) \hookrightarrow L^q(\Omega), \quad \forall q \in [1, \infty[, \text{ si } n = 2. \quad (\text{I.12})$$

Remarque I.8. En particulier on a $BV^2(\Omega) \subset L^2(\Omega)$.

Théorème I.32. Si Ω est \mathcal{C}^2 régulier uniforme, alors

$$BV^2(\Omega) \hookrightarrow L^\infty(\Omega) \text{ si } n \leq 2.$$

Théorème I.33. On suppose que Ω est un ouvert borné de \mathbb{R}^n , \mathcal{C}^2 régulier uniforme. Alors

$$BV^2(\Omega) \hookrightarrow \mathcal{C}(\bar{\Omega}).$$

Nous renvoyons le lecteur à [2] page 84 pour la définition d'un ouvert \mathcal{C}^k régulier uniforme.

4.3.2 Théorème d'approximation

Nous donnons ici l'adaptation à l'espace $BV^2(\Omega)$ du résultat d'approximation dans l'espace $BV(\Omega)$ démontré dans [35].

Théorème I.34 (Approximation régulière). Soit $u \in BV^2(\Omega)$. Alors il existe une suite de fonctions $\{u_k\}_{k \in \mathbb{N}} \in BV^2(\Omega) \cap \mathcal{C}^\infty(\Omega)$ telle que

1. $u_k \rightarrow u$ dans $W^{1,1}(\Omega)$ et
2. $TV^2(u_k, \Omega) \rightarrow TV^2(u, \Omega)$.

Démonstration :

1. Soient $\varepsilon > 0$ et $n \in \mathbb{N}^*$. On définit l'ouvert

$$\Omega_k \equiv \left\{ x \in \Omega \mid \text{dist}(x, \partial\Omega) > \frac{1}{m+k} \right\} \cap \mathcal{B}(0, k+m), \quad k \in \mathbb{N}^*,$$

et l'on choisit m suffisamment grand pour que

$$|D^2 f|(\Omega - \Omega_1) < 3\varepsilon. \quad (\star)$$

On pose $\Omega_0 \equiv \emptyset$, et on définit

$$V_k \equiv \Omega_{k+1} - \bar{\Omega}_{k-1}, \quad k \in \mathbb{N}^*.$$

Soit $(\zeta_k)_{k \leq 1}$ une suite de fonctions régulières telles que

$$\begin{cases} \zeta_k \in \mathcal{C}_c^\infty(V_k) & 0 \leq \zeta_k \leq 1 \\ \sum_{k=1}^{\infty} \zeta_k \equiv 1 & \text{sur } \Omega. \end{cases}$$

Soit η l'unité approchée définie page 11. Pour tout k , on choisit $\varepsilon_k > 0$ suffisamment petit pour que

$$\begin{cases} \text{supp}(\eta_{\varepsilon_k} * (f\zeta_k)) \subset V_k \\ \int_{\Omega} |\eta_{\varepsilon_k} * (f\zeta_k) - f\zeta_k| dx < \frac{\varepsilon}{2^k}, \\ \int_{\Omega} |\eta_{\varepsilon_k} * (\nabla f\zeta_k) - \nabla f\zeta_k| dx < \frac{\varepsilon}{2^k}, \\ \int_{\Omega} |\eta_{\varepsilon_k} * (fH\zeta_k) - fH\zeta_k| dx < \frac{\varepsilon}{2^k}. \end{cases} \quad (**)$$

On définit

$$f_{\varepsilon} \equiv \sum_{k=1}^{\infty} \eta_{\varepsilon_k} * (f\zeta_k).$$

Pour tout x dans Ω , la somme ne contient qu'un nombre fini de termes non nuls dans un voisinage suffisamment petit de x . La différentiabilité étant un critère local, on en déduit que

$$f_{\varepsilon} \in C^{\infty}(\Omega).$$

2. Comme

$$f = \sum_{k=1}^{\infty} f\zeta_k,$$

(**) entraîne

$$\|f - f_{\varepsilon}\|_{L^1(\Omega)} \leq \sum_{k=1}^{\infty} \int_{\Omega} |\eta_{\varepsilon_k} * (f\zeta_k) - f\zeta_k| dx < \varepsilon,$$

et

$$\|\nabla f - \nabla f_{\varepsilon}\|_{L^1(\Omega)} \leq \sum_{k=1}^{\infty} \int_{\Omega} |\eta_{\varepsilon_k} * (f\nabla\zeta_k) - f\nabla\zeta_k| dx < \varepsilon.$$

Ainsi, $f_{\varepsilon} \in W^{1,1}(\Omega)$ et

$$f_{\varepsilon} \rightarrow f \text{ dans } W^{1,1}(\Omega).$$

D'autre part, pour $\varphi \in \mathcal{C}_c^2(\Omega, \mathbb{R}^{n \times n})$,

$$\begin{aligned}
\int_{\Omega} f_{\varepsilon} H^* \varphi \, dx &= \sum_{k=1}^{\infty} \int_{\Omega} \eta_{\varepsilon_k} * (f \zeta_k) H^* \varphi \, dx \\
&= \sum_{k=1}^{\infty} \int_{x \in \Omega} \int_{y \in \Omega} \eta_{\varepsilon_k}(x-y) (f \zeta_k)(y) \, dy \, H^* \varphi(x) \, dx \\
&= \sum_{k=1}^{\infty} \int_x \int_y \eta_{\varepsilon_k}(x-y) (f \zeta_k)(y) \, H^* \varphi(x) \, dy \, dx \\
&= \sum_{k=1}^{\infty} \int_y (f(y) \zeta_k(y)) \left[\int_x \eta_{\varepsilon_k}(y-x) \, H^* \varphi(x) \, dx \right] \, dy \\
&= \sum_{k=1}^{\infty} \int_{\Omega} (f(y) \zeta_k(y)) (\eta_{\varepsilon_k} * H^* \varphi)(y) \, dy \\
&= \sum_{k=1}^{\infty} \int_{\Omega} (f \zeta_k) \, H^*(\eta_{\varepsilon_k} * \varphi) \, dy.
\end{aligned}$$

La somme est en réalité finie et avec les inégalités $\|\zeta_k(\eta_{\varepsilon_k} * \varphi)\|_{\infty} \leq 1$ et $0 \leq \zeta_k \leq 1$, on en déduit que $f_{\varepsilon} \in BV^2(\Omega)$.

3. D'après le théorème précédent,

$$TV^2(f, \Omega) \leq \liminf_{\varepsilon \rightarrow 0} TV^2(f_{\varepsilon}). \quad (\star \star \star)$$

4. Soit $\varphi \in \mathcal{C}_c^2(\Omega, \mathbb{R}^{n \times n})$. On a montré précédemment que

$$\int_{\Omega} f_{\varepsilon} H^* \varphi \, dx = \sum_{k=1}^{\infty} \int_{\Omega} (f \zeta_k) \, H^*(\eta_{\varepsilon_k} * \varphi) \, dy.$$

Or

$$H^*(\zeta_k(\eta_{\varepsilon_k} * \varphi)) = \zeta_k H^*(\eta_{\varepsilon_k} * \varphi) + H \zeta_k \cdot (\eta_{\varepsilon_k} * \varphi) + \sum_{i,j=1}^n \frac{\partial \zeta_k}{\partial x_i} \frac{\partial (\eta_{\varepsilon_k} * \varphi)_{ij}}{\partial x_j} + \sum_{i,j=1}^n \frac{\partial \zeta_k}{\partial x_j} \frac{\partial (\eta_{\varepsilon_k} * \varphi)_{ij}}{\partial x_i},$$

i.e.

$$\zeta_k H^*(\eta_{\varepsilon_k} * \varphi) = H^*(\zeta_k(\eta_{\varepsilon_k} * \varphi)) - H \zeta_k \cdot (\eta_{\varepsilon_k} * \varphi) - \sum_{i,j=1}^n \frac{\partial \zeta_k}{\partial x_i} \frac{\partial (\eta_{\varepsilon_k} * \varphi)_{ij}}{\partial x_j} - \sum_{i,j=1}^n \frac{\partial \zeta_k}{\partial x_j} \frac{\partial (\eta_{\varepsilon_k} * \varphi)_{ij}}{\partial x_i};$$

donc

$$\begin{aligned}
\sum_{k=1}^{\infty} \int_{\Omega} (f \zeta_k) \, H^*(\eta_{\varepsilon_k} * \varphi) \, dy &= \sum_{k=1}^{\infty} \int_{\Omega} f \, H^*(\zeta_k(\eta_{\varepsilon_k} * \varphi)) \, dy - \sum_{k=1}^{\infty} \int_{\Omega} f \, H \zeta_k \cdot (\eta_{\varepsilon_k} * \varphi) \, dy \\
&\quad - \sum_{k=1}^{\infty} \int_{\Omega} f \sum_{i,j=1}^n \frac{\partial \zeta_k}{\partial x_i} \frac{\partial (\eta_{\varepsilon_k} * \varphi)_{ij}}{\partial x_j} \, dy \\
&\quad - \sum_{k=1}^{\infty} \int_{\Omega} f \sum_{i,j=1}^n \frac{\partial \zeta_k}{\partial x_j} \frac{\partial (\eta_{\varepsilon_k} * \varphi)_{ij}}{\partial x_i} \, dy.
\end{aligned}$$

En intégrant par parties dans les deux derniers termes, on a

$$- \int_{\Omega} f \sum_{i,j=1}^n \frac{\partial \zeta_k}{\partial x_i} \frac{\partial(\eta_{\varepsilon_k} * \varphi)_{ij}}{\partial x_j} dx = - \int_{\Omega} f \sum_{i,j=1}^n \frac{\partial \zeta_k}{\partial x_j} \frac{\partial(\eta_{\varepsilon_k} * \varphi)_{ij}}{\partial x_i} dx = \int_{\Omega} (f \zeta_k) H^*(\eta_{\varepsilon_k} * \varphi) dx.$$

Ainsi,

$$\begin{aligned} \int_{\Omega} f_{\varepsilon} H^* \varphi dx &= \sum_{k=1}^{\infty} \int_{\Omega} f H^*(\zeta_k(\eta_{\varepsilon_k} * \varphi)) dy - \sum_{k=1}^{\infty} \int_{\Omega} f H \zeta_k \cdot (\eta_{\varepsilon_k} * \varphi) dy \\ &= \sum_{k=1}^{\infty} \int_{\Omega} f H^*(\zeta_k(\eta_{\varepsilon_k} * \varphi)) dy - \sum_{k=1}^{\infty} \int_{\Omega} \varphi \cdot (\eta_{\varepsilon_k} * (f H \zeta_k) - f H \zeta_k) dy \\ &= I_1^{\varepsilon} + I_2^{\varepsilon} \end{aligned}$$

En effet,

$$\begin{aligned} \int_{\Omega} \varphi \cdot (\eta_{\varepsilon_k} * (f H \zeta_k)) dx &= \int_{\Omega} \varphi \cdot [\eta_{\varepsilon_k}(x - y) (f H \zeta_k)(y) dy] dx \\ &= \int_{\Omega} \int_{\Omega} \sum_{i,j=1}^n \varphi_{ij}(x) \eta_{\varepsilon_k}(x - y) f(y) \frac{\partial^2 \zeta_k}{\partial x_i \partial x_j}(y) dy dx \\ &= \sum_{i,j=1}^n \int_{\Omega} f(y) \frac{\partial^2 \zeta_k}{\partial x_i \partial x_j}(y) \left[\int_{\Omega} \varphi_{ij}(x) \eta_{\varepsilon_k}(y - x) dx \right] dy \\ &= \int_{\Omega} \sum_{i,j=1}^n f(y) \frac{\partial^2 \zeta_k}{\partial x_i \partial x_j}(y) \eta_{\varepsilon_k} * \varphi_{ij}(y) dy \\ &= \int_{\Omega} f H \zeta_k \cdot (\eta_{\varepsilon_k} * \varphi) dx. \end{aligned}$$

D'autre part, le lemme [I.1](#) montre que, comme $\sum_{k=1}^{\infty} H \zeta_k = 0$,

$$\sum_{k=1}^{\infty} \int_{\Omega} \varphi \cdot f H \zeta_k dx = 0.$$

Utilisons maintenant le fait que $\|\zeta_k(\eta_{\varepsilon_k} * \varphi)\|_{\infty} \leq 1$ pour tout k dans \mathbb{N}^* , et que tout point de Ω ne peut être que dans au plus trois éléments de $(V_k)_{k \geq 1}$. Ainsi, on obtient

$$\begin{aligned} |I_1^{\varepsilon}| &= \left| \int_{\Omega} f H^*(\zeta_1(\eta_{\varepsilon_1} * \varphi)) dx + \sum_{k=2}^{\infty} \int_{\Omega} f H^*(\zeta_k(\eta_{\varepsilon_k} * \varphi)) dx \right| \\ &\leq |D^2 f|(\Omega) + \sum_{k=2}^{\infty} |D^2 f|(V_k) \\ &\leq |D^2 f|(\Omega) + 3|D^2 f|(\Omega - \Omega_1) \\ &\leq |D^2 f|(\Omega) + 3\varepsilon, \text{ par } (\star). \end{aligned}$$

D'autre part, $(\star\star)$ entraîne

$$|I_2^\varepsilon| \leq \varepsilon.$$

Ainsi,

$$\int_{\Omega} f_\varepsilon H^* \varphi \, dx \leq TV^2(f, \Omega) + 4\varepsilon,$$

puis en passant à la borne supérieure sur φ ,

$$TV^2(f_\varepsilon, \Omega) \leq TV^2(f, \Omega) + 4\varepsilon.$$

Cette inégalité et $(\star\star\star)$ nous permettent de conclure. ■

Lemme I.1. *Dans les conditions du théorème précédent,*

$$\sum_{k=1}^{\infty} H\zeta_k = 0 \implies \sum_{k=1}^{+\infty} \int_{\Omega} \varphi \cdot f H\zeta_k \, dx = 0.$$

Démonstration : Soit $x \in \Omega$. On pose $F_N(x) = \left\langle \varphi(x), f(x) \sum_{k=1}^N H\zeta_k(x) \right\rangle$.

On a

$$\begin{aligned} |F_N(x)| &= \left| \left\langle \varphi(x), f(x) \sum_{k=1}^N H\zeta_k(x) \right\rangle \right| \\ &= |f(x)| \left| \left\langle \varphi(x), \sum_{k=1}^N H\zeta_k(x) \right\rangle \right| \\ &\leq |f(x)| \|\varphi(x)\|_{\mathbb{R}^{n \times n}} \left\| \sum_{k=1}^N H\zeta_k(x) \right\|_{\mathbb{R}^{n \times n}}. \end{aligned}$$

Or

$$\left\| \sum_{k=1}^N H\zeta_k(x) \right\|_{\mathbb{R}^{n \times n}} \leq \left\| \sum_{k=1}^N H\zeta_k \right\|_{\infty} \xrightarrow{N \rightarrow \infty} 0.$$

Ainsi, il existe une constante $K \in \mathbb{R}^+$ telle que

$$\left\| \sum_{k=1}^N H\zeta_k(x) \right\|_{\mathbb{R}^{n \times n}} \leq K.$$

D'autre part,

$$\|\varphi(x)\|_{\mathbb{R}^n} \leq \|\varphi\|_{\infty} \leq 1.$$

On a donc démontré que

$$|F_N(x)| \leq K |f(x)|.$$

f étant dans $L^1(\Omega)$, le théorème de convergence dominée de Lebesgue permet d'invertir intégrale et limite. Par suite,

$$\sum_{k=1}^{+\infty} \int_{\Omega} \varphi \cdot f H\zeta_k \, dx = \int_{\Omega} \varphi \cdot f \sum_{k=1}^{+\infty} H\zeta_k \, dx = 0$$

■

Nous pouvons signaler que, parallèlement, K. Bredies, K. Kunisch et T. Pock ont introduit dans [18] la notion d'espaces de fonctions à variation bornée généralisée d'ordre k . Pour $k = 2$, la définition d'un tel espace est toutefois différentes de celle donnée dans ce chapitre à l'espace BV^2 des fonctions à Hessien borné. Dans ce manuscrit, nous ne reviendrons pas en détail sur ces travaux et renvoyons le lecteur intéressé à [18] pour plus d'information.

Chapitre II

Restauration d'images

1 Introduction

Les modèles variationnels en traitement d'image ont fait l'objet de nombreuses études ces dix dernières années. Il sont utilisés pour la segmentation d'images ainsi que pour le débruitage. Nous nous intéressons ici à ce dernier.

Récemment, beaucoup d'études ont été faites en admettant le principe qu'une image peut être décomposée en de nombreuses composantes, décrivant chacune une propriété particulière de l'image de départ ([7, 8, 51, 52, 53] par exemple). De la même manière, nous supposons ici qu'une image u_d peut être décomposée sous la forme $u_d = u + v$ où u et v sont des fonctions caractérisant différentes parties de u_d (voir [8, 53, 58] par exemple). Le modèle le plus célèbre est certainement le modèle de débruitage de Rudin, Osher et Fatemi (voir [1], [50]). L'image observée est dans ce cas divisée en deux parties $u + v$ où u représente la partie oscillante de l'image (texture et bruit) et v la partie cartoon. Ainsi, nous cherchons une solution de la forme $u + v$ avec $v \in BV(\Omega)$ et $u \in L^2(\Omega)$, où Ω est un ouvert borné ([4, 6]). Nous obtenons ainsi le modèle suivant :

$$\min_{v \in BV(\Omega)} \mu TV(v, \Omega) + \frac{1}{2} \|u_d - v\|_{L^2(\Omega)}^2, \quad (\mathcal{P}_{ROF})$$

où $\mu > 0$. Ce modèle fait l'objet de la première partie de ce chapitre, où nous rappelons le résultat d'existence et d'unicité, ainsi que la discrétisation du modèle et sa résolution numérique par l'algorithme de projection de Chambolle [22]. Ce modèle est particulièrement efficace dans le cadre du débruitage. Toutefois, il engendre des perturbations numériques. En effet, la solution a tendance à être constante par morceaux ce que l'on appelle l'effet "staircasing". Nous mettons en évidence ce phénomène à la fin de la première partie de ce chapitre. Ainsi, bien que le bruit soit éliminé de manière très significative, la solution obtenue n'est pas complètement satisfaisante.

C'est pourquoi nous passons à l'ordre deux en remplaçant dans le modèle (\mathcal{P}_{ROF}) la variation totale $TV(v, \Omega)$ par la variation totale seconde $TV^2(v, \Omega)$. Nous espérons ainsi rendre l'image restaurée plus lisse et faire disparaître cet effet "staircasing". Nous

proposons donc le modèle suivant :

$$\min_{v \in BV^2(\Omega)} \mu TV^2(v, \Omega) + \frac{1}{2} \|u_d - v\|_{L^2(\Omega)}^2, \quad (\mathcal{P}_{ROF2})$$

où $\mu > 0$. L'étude précise de ce modèle fait l'objet de la deuxième partie de ce chapitre. Nous commençons tout d'abord par ajouter à ce modèle un terme de "régularisation" nécessaire à l'obtention d'un résultat d'existence. Nous proposons ensuite une discrétisation du modèle, puis la résolution numérique du problème discrétisé par des méthodes de projections de type Chambolle, directement adaptée de l'étude du modèle (\mathcal{P}_{ROF}) dans [22]. Nous observons alors que le modèle est, lui aussi, efficace dans le cadre de la restauration, est que l'effet "staircasing" disparaît bel et bien. Toutefois, même si les différents contours de l'image sont préservés, un effet de flou apparaît, d'autant plus important que le paramètre μ est grand.

Pour remédier à ce problème, nous proposons, dans la troisième partie de ce chapitre, un nouveau modèle dans lequel nous supposons qu'une image u_d peut être décomposée en trois parties $u_d = u + v + w$. En gardant la même philosophie que pour les modèles ROF et ROF2, nous supposons cette fois que u est la composante cartoon contenant en particulier l'information géométrique, que v est la composante BV^2 lisse et que w contient la partie oscillante de l'image de départ, à savoir le bruit et la texture. Nous faisons ainsi l'hypothèse que $u_d = u + v + w$, avec $u \in BV(\Omega)$, $v \in BV^2(\Omega)$ et $w \in L^2(\Omega)$. Nous obtenons alors un modèle du type

$$\inf_{(u,v) \in BV(\Omega) \times BV^2(\Omega)} \lambda TV(u, \Omega) + \frac{1}{2} \|u_d - u - v\|_{L^2(\Omega)}^2 + \mu TV^2(v, \Omega) + \frac{\delta}{2} \left(\|v\|_{L^2(\Omega)}^2 + \|\nabla v\|_{L^2(\Omega)}^2 \right), \quad (\mathcal{P})$$

où $\lambda, \mu > 0$, et auquel nous ajoutons, comme pour le modèle ROF2, un terme de "régularisation" pour obtenir un résultat d'existence et d'unicité. Nous proposons, comme pour les modèles précédents, la discrétisation et la résolution numérique du problème. Toutefois, les tests numériques mettent en évidence un problème au niveau de la convergence de l'algorithme que nous proposons. En effet, nous remarquons que la vitesse de convergence des algorithmes de projection de type Chambolle est beaucoup trop lente pour obtenir un résultat dans des délais raisonnables. Nous proposons une alternative par l'utilisation d'un algorithme de Nesterov dont P.Weiss montre dans [57] l'efficacité. Nous obtenons ainsi des résultats numériques répondant à nos attentes : v est bien la partie lisse, u contient les contours significatifs de l'image, alors que tout le bruit ainsi que la texture sont contenus dans w . Nous récupérons alors dans $u + v$ une image débruitée sans effet "staircasing" et sans effet de flou aux endroits stratégiques (contours, certains détails de l'image...). Le comportement de ce modèle n'est pas encore totalement expliqué. Ceci fait l'objet d'un travail en cours.

Remarque II.1. *Dans ce chapitre, nous ne parlons volontairement pas des modèles de Meyer [41], Vese-Osher [52], Osher-Solé-Vese [51] où encore Aujol-Aubert [7]. En effet, bien qu'entrant complètement dans le cadre des modèles de décomposition d'une image u_d sous la forme $u_d = u + v$, ces modèles montrent toute leur efficacité dans le cadre de l'extraction de texture. Nous en parlerons donc en détail dans le chapitre suivant.*

2 Le modèle de Rudin, Osher et Fatemi

2.1 Présentation

Le modèle de Rudin, Osher et Fatemi (ROF) est une alternative à la régularisation de Tychonov qui est trop violente. En effet, si le bruit est bien éliminé, le résultat induit en contre-partie un effet de flou d'autant plus important que le débruitage est fort. Ces modèles consistent à diviser l'image en deux parties : $u_d = u + v$ où v est le bruit et u la partie « régulière ». Seul le terme de régularisation change : il s'agit, en général, de la norme H^1 ou de la norme L^2 du gradient pour le modèle de Tychonov et de la variation totale pour ROF. Cette dernière permet de préserver les sauts d'une image, ce qu'une régularisation de type Tychonov interdit. On va donc chercher la solution du problème sous la forme $u + v$ avec $u \in BV(\Omega)$ et $v \in L^2(\Omega)$, minimiser la partie « bruit » et faire porter la régularisation sur la variation totale. Cela conduit à :

$$\min \left\{ TV(u, \Omega) + \frac{1}{2\varepsilon} \|v\|_{L^2(\Omega)}^2 \mid u \in BV(\Omega), v \in L^2(\Omega), u + v = u_d \right\}, \quad (\mathcal{P}_{ROF})$$

avec $\varepsilon > 0$.

Théorème II.1. *Le problème (\mathcal{P}_{ROF}) admet une unique solution.*

2.2 Discrétisation

Afin d'alléger les notations, nous ne considérerons que des images carrées de taille $N \times N$. Tous les résultats énoncés ci-après restent bien sûr vrais pour des images rectangulaires non carrées. On note X l'espace euclidien $\mathbb{R}^{N \times N}$, et $Y = X \times X$. On munit l'espace X du produit scalaire

$$\langle u, v \rangle_X = \sum_{1 \leq i, j \leq N} u_{i,j} v_{i,j} \quad (\text{II.1})$$

Pour définir la variation totale discrète, on introduit d'abord une version discrète de l'opérateur gradient. Si $u \in X$, le gradient ∇u est un élément de Y donné par :

$$(\nabla u)_{i,j} = \left((\nabla u)_{i,j}^1, (\nabla u)_{i,j}^2 \right), \quad (\text{II.2})$$

$$(\nabla u)_{i,j}^1 = \begin{cases} u_{i+1,j} - u_{i,j} & \text{si } i < N \\ 0 & \text{si } i = N, \end{cases} \quad (\text{II.3})$$

et

$$(\nabla u)_{i,j}^2 = \begin{cases} u_{i,j+1} - u_{i,j} & \text{si } j < N \\ 0 & \text{si } j = N. \end{cases} \quad (\text{II.4})$$

La variation totale $TV(u, \Omega)$ de u est alors donnée par :

$$J_1(u) = \sum_{1 \leq i, j \leq N} \left\| (\nabla u)_{i,j} \right\|_{\mathbb{R}^2}, \quad (\text{II.5})$$

où

$$\|(\nabla u)_{i,j}\|_{\mathbb{R}^2} = \left\| \left((\nabla u)_{i,j}^1, (\nabla u)_{i,j}^2 \right) \right\|_{\mathbb{R}^2} = \sqrt{\left((\nabla u)_{i,j}^1 \right)^2 + \left((\nabla u)_{i,j}^2 \right)^2}. \quad (\text{II.6})$$

On introduit également la version discrète de l'opérateur divergence, défini en posant :

$$\text{div} = -\nabla^*, \quad (\text{II.7})$$

où ∇^* est l'opérateur adjoint de ∇ , c'est à dire

$$\forall (p, u) \in Y \times X, \quad \langle -\text{div } p, u \rangle_X = \langle p, \nabla u \rangle_Y.$$

On vérifie que

$$(\text{div } p)_{i,j} = \begin{cases} p_{i,j}^1 - p_{i-1,j}^1 & \text{si } 1 < i < N \\ p_{i,j}^1 & \text{si } i = 1 \\ -p_{i-1,j}^1 & \text{si } i = N \end{cases} + \begin{cases} p_{i,j}^1 - p_{i,j-1}^2 & \text{si } 1 < j < N \\ p_{i,j}^2 & \text{si } j = 1 \\ -p_{i,j-1}^2 & \text{si } j = N. \end{cases} \quad (\text{II.8})$$

2.3 Algorithme de projection de Chambolle

Le problème discrétisé est le suivant :

$$\inf_{u \in X} \left(J_1(u) + \frac{1}{2\lambda} \|f - u\|_X^2 \right). \quad (\mathcal{P}_d^1)$$

Le résultat suivant est montré dans [22] :

Proposition II.1. *La solution de (\mathcal{P}_d^1) est donnée par :*

$$u = f - P_{\lambda K_1}(f), \quad (\text{II.9})$$

où

$$K_1 = \{ \text{div } g \mid g \in Y, \|g_{i,j}\|_{\mathbb{R}^2} \leq 1 \forall i, j \}, \quad (\text{II.10})$$

et où $P_{\lambda K_1}$ est le projecteur orthogonal sur λK_1 .

A. Chambolle donne, dans [22], un algorithme pour calculer $P_{\lambda K_1}(f)$. Le calcul de $P_{\lambda K_1}(f)$ conduit à résoudre le problème

$$\min \{ \|\lambda \text{div } p - f\|_X^2 \mid \|p_{i,j}\|_{\mathbb{R}^2} \leq 1 \forall i, j = 1, \dots, N \}. \quad (\text{II.11})$$

Ce problème peut se résoudre par une méthode de point fixe :

$$p^0 = 0, \quad (\text{II.12a})$$

$$p_{i,j}^{n+1} = \frac{p_{i,j}^n + \tau (\nabla (\text{div } p^n - f/\lambda))_{i,j}}{1 + \tau \left\| (\nabla (\text{div } p^n - f/\lambda))_{i,j} \right\|_{\mathbb{R}^2}} \quad (\text{II.12b})$$

A. Chambolle donne dans [22] une condition suffisante pour que l'algorithme converge :

Théorème II.2. *Supposons que le paramètre τ vérifie $\tau \leq 1/8$. Alors $\lambda \operatorname{div} p^n$ converge vers $P_{\lambda K_1}(f)$ quand n tend vers l'infini.*

Remarque II.2. *Dans [10] J.F. Aujol prouve que l'algorithme modifié*

$$p^0 = 0 \tag{II.13a}$$

$$p_{i,j}^{n+1} = \frac{p_{i,j}^n + \tau(\nabla(\operatorname{div} p^n - u_d/\lambda))_{i,j}}{\max(1, \left\| p_{i,j}^n + \tau(\nabla(\operatorname{div} p^n - u_d/\lambda))_{i,j} \right\|_{\mathbb{R}^2})} \tag{II.13b}$$

converge si $\tau \leq 1/4$.

2.4 Effet “staircasing”

L'effet *staircasing* est un effet bien connu induit par une restauration avec le modèle ROF. Concrètement, cela signifie que se créent sur l'image restaurée, des régions constantes ainsi que des discontinuités. L'étude des discontinuités a été faite dans [21]. Nous reprenons pour définition de l'effet *staircasing*, celle donnée par C. Louchet dans [42] : on définit une région connectée $\mathcal{W} \subset \Omega$, comme un sous-ensemble de Ω ne contenant aucun point isolé. On parlera d'effet *staircasing* dans la version restaurée \bar{u} d'une image bruitée u s'il existe une région connectée $\mathcal{W} \subset \Omega$ contenant au moins deux pixels, telle que

$$\forall x \in \mathcal{W}, \forall x' \in \mathcal{N}_x, x' \in \mathcal{W} \Rightarrow \bar{u}(x) = \bar{u}(x'),$$

où \mathcal{N}_x définit un voisinage du pixel x . Bien évidemment, l'effet *staircasing* est d'autant plus visible que les ensembles \mathcal{W} sont gros et nombreux. Numériquement, cet effet est d'autant plus visible que le paramètre λ dans (\mathcal{P}_d^1) est élevé.

L'effet *staircasing* a d'abord été mis en évidence par Dobson et Santosa ([29]), qui l'ont utilisé pour la reconstruction d'images définies par morceaux. D'autres auteurs ont d'autre part prouvé mathématiquement l'existence de cet artefact : Ring dans [56] et Nikolova dans [44, 45, 46, 48]. Nous n'entrerons pas plus dans le détail sur le sujet, et renvoyons le lecteurs aux articles pré-cités pour plus de précisions. Différents travaux ont été réalisés dans le but de diminuer l'effet de cet artefact. Nous pouvons par exemple citer Blomgren et al. [17], Chan et al. [24, 26], Bouman et Sauer [19], Louchet [42]. Les figures II.3 et II.4 illustrent ce phénomène sur un signal 1D et sur une image. Nous renvoyons le lecteur au chapitre 5 pour plus de tests numériques.

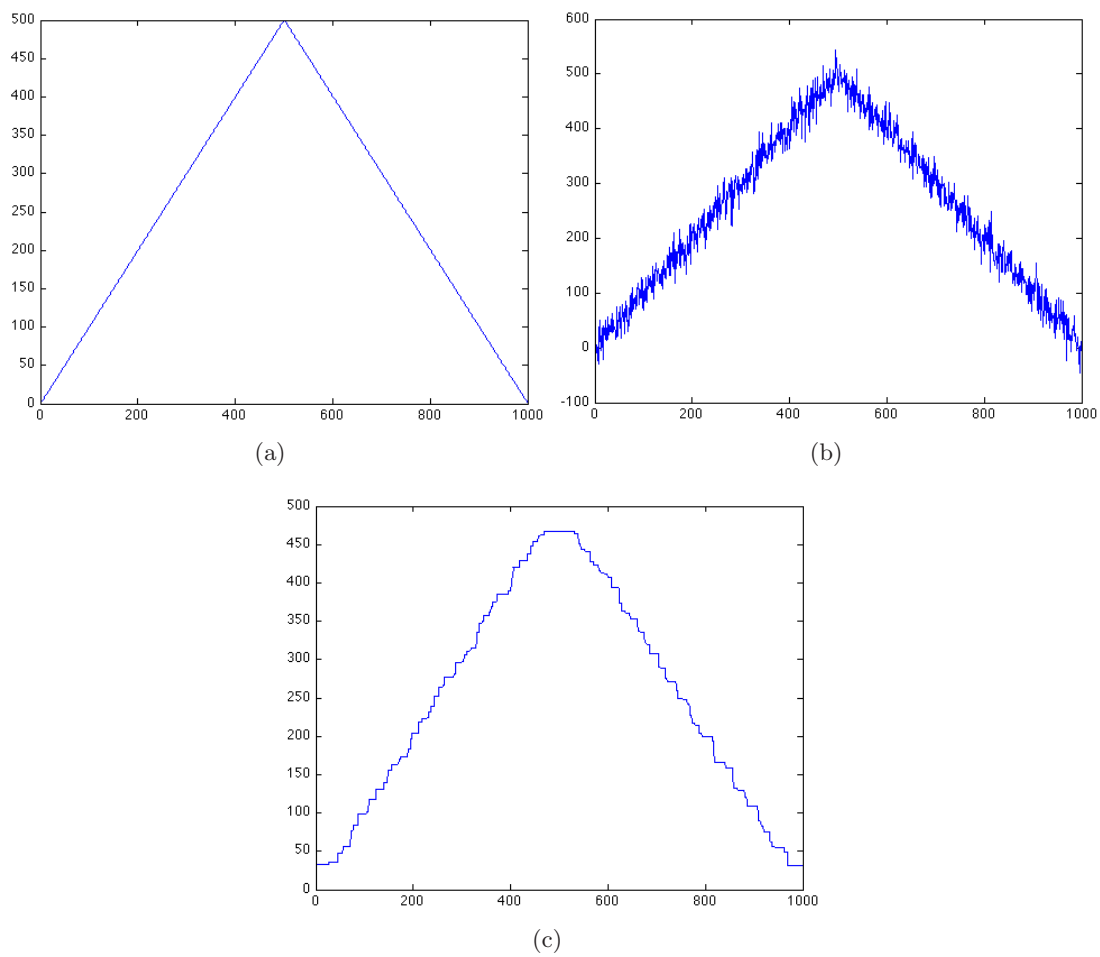


FIG. II.1 – Effet *staircasing* sur un signal de dimension 1. Utilisation du modèle ROF pour restaurer un signal artificiellement bruité ($\sigma = 20$, $\lambda = 100$). On observe que le signal restauré a un aspect "en escalier".

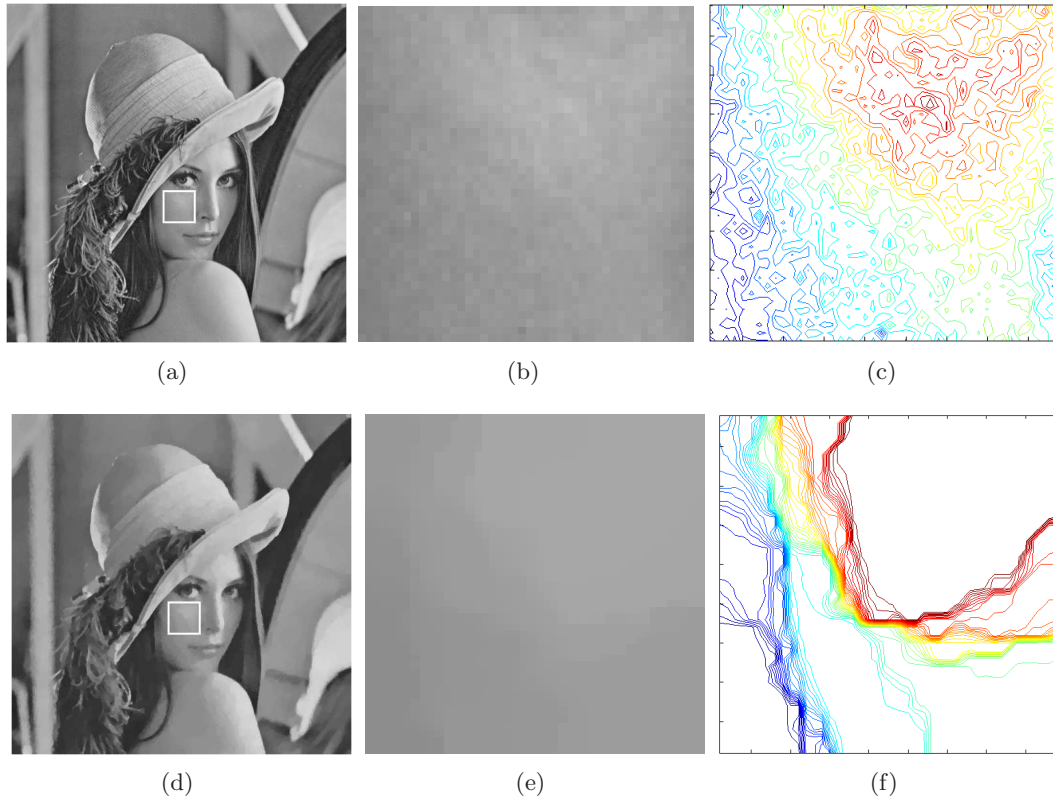


FIG. II.2 – Effet *staircasing* sur une image. (a) est l'image de départ non bruitée, (d) est l'image bruitée ($\sigma = 20$) restaurée par ROF ($\lambda = 25$), (b) et (e) sont les zooms respectivement des parties encadrées de (a) et (d), et (c) et (f) sont les courbes de niveaux respectivement associées à (b) et (e). L'effet *staircasing* est visible sur l'image (e) qui semble constante par morceaux. L'image (f) met en évidence la création de contours fictifs, autours desquels se concentre un grand nombre de courbes de niveaux.

Nous proposons dans la suite un modèle d'ordre deux, directement inspiré du modèle de Rudin, Osher et Fatemi permettant d'éliminer cet artefact.

3 Modèle d'ordre deux

3.1 Introduction

Comme pour le modèle ROF, on suppose que l'image que nous voulons obtenir de l'image de départ u_d peut être décomposée sous la forme $u_d = u + v$, où u et v sont des fonctions caractérisant des parties différentes de u_d (voir [8, 53, 58] par exemple). Les composantes u et v appartiennent à des espaces fonctionnels différents : u est la partie non régulière et appartient à $L^2(\Omega)$ tandis que v est la partie plus régulière de l'image et appartient à $BV^2(\Omega)$ (et non plus $BV(\Omega)$ comme pour le modèle ROF). Nous espérons

obtenir dans le terme $u = u_d - v$ les composants très oscillants de l'image (texture et bruit) ainsi que très probablement les contours. Le fait de passer à l'ordre deux devrait permettre de lisser suffisamment la partie régulière de l'image, afin de faire disparaître significativement l'effet *staircasing*, tout en engendrant un effet de flou acceptable. Nous introduisons donc un modèle du type :

$$\inf_{v \in BV^2(\Omega)} \frac{\|f - v\|_{L^2(\Omega)}^2}{2\lambda} + TV^2(v, \Omega). \quad (\text{II.14})$$

Malheureusement, le passage à l'ordre deux génère des difficultés dans l'obtention d'un résultat d'existence et d'unicité. Nous allons donc ajouter à ce problème un terme permettant d'obtenir ce résultat d'existence de manière aisée. Nous verrons de plus que ce terme régularisant, inutile dans le cas discret, n'a qu'une influence infime et visuellement non détectable sur le résultat final. Tous les résultats énoncés dans cette section ont fait l'objet d'une publication. Nous omettons donc toutes les démonstrations et renvoyons le lecteur à l'annexe B pour tous les détails.

3.2 Modèle variationnel

Suite aux remarques précédentes quant à l'ajout nécessaire d'un nouveau terme, on considère la fonction définie sur $BV^2(\Omega)$ suivante :

$$\mathcal{F}_\delta(v) = \frac{1}{2} \|u_d - v\|_{L^2(\Omega)}^2 + \lambda TV^2(v, \Omega) + \delta \|v\|_{W^{1,1}(\Omega)}, \quad (\text{II.15})$$

où $\lambda, \delta \geq 0$. On cherche alors la solution du problème d'optimisation

$$\inf \{ \mathcal{F}_\delta(v) \mid v \in BV^2(\Omega) \} \quad (\text{II.16})$$

On obtient alors le résultat d'existence et d'unicité suivant :

Théorème II.3. *On suppose $\lambda > 0$ et $\delta > 0$. Alors le problème (\mathcal{P}^δ) a une unique solution v .*

Si l'image est bruitée, le bruit (comme la texture) apparaîtra dans la partie oscillante de l'image $u := u_d - v$. Dans le cadre du débruitage, une telle approche à déjà été utilisée par Hinterberger et Scherzer [36] avec l'espace $BV^2(\Omega)$. Leur algorithme est toutefois différent de celui que nous utilisons ici.

3.3 Le problème discrétisé

Le problème (II.16) peut être réécrit de la façon suivante :

$$\inf_{v \in BV^2(\Omega)} \frac{\|u_d - v\|_{L^2(\Omega)}^2}{2\lambda} + TV^2(v, \Omega) + \delta \|v\|_{W^{1,1}(\Omega)}, \quad (\mathcal{P}^\delta)$$

où $\frac{\delta}{\lambda}$ à été remplacé par δ . Afin de calculer numériquement la solution, nous commençons par détailler le processus de discrétisation.

3.3.1 Discrétisation du problème (\mathcal{P}^δ)

Comme pour le modèle ROF, nous ne considérerons ici que des images carrées de taille $N \times N$. Pour définir une version discrète de la variation totale seconde, nous devons introduire une version discrète de l'opérateur hessien. Si $v \in X$, la hessienne de v , que l'on notera (Hv) est un vecteur de $Z = X^4$ avec $X = \mathbb{R}^{n \times n}$, défini par

$$(Hv)_{i,j} = ((Hv)_{i,j}^{11}, (Hv)_{i,j}^{12}, (Hv)_{i,j}^{21}, (Hv)_{i,j}^{22}),$$

avec, pour $i, j = 1, \dots, N$,

$$\begin{aligned} (Hv)_{i,j}^{11} &= \begin{cases} v_{i+1,j} - 2v_{i,j} + v_{i-1,j} & \text{si } 1 < i < N, \\ v_{i+1,j} - v_{i,j} & \text{si } i = 1, \\ v_{i-1,j} - v_{i,j} & \text{si } i = N, \end{cases} \\ (Hv)_{i,j}^{12} &= \begin{cases} v_{i,j+1} - v_{i,j} - v_{i-1,j+1} + v_{i-1,j} & \text{si } 1 < i \leq N, 1 \leq j < N, \\ 0 & \text{si } i = 1, \\ 0 & \text{si } i = N, \end{cases} \\ (Hv)_{i,j}^{21} &= \begin{cases} v_{i+1,j} - v_{i,j} - v_{i+1,j-1} + v_{i,j-1} & \text{si } 1 \leq i < N, 1 < j \leq N, \\ 0 & \text{si } i = 1, \\ 0 & \text{si } i = N, \end{cases} \\ (Hv)_{i,j}^{22} &= \begin{cases} v_{i,j+1} - 2v_{i,j} + v_{i,j-1} & \text{si } 1 < j < N, \\ v_{i,j+1} - v_{i,j} & \text{si } j = 1, \\ v_{i,j-1} - v_{i,j} & \text{si } j = N. \end{cases} \end{aligned}$$

La variation totale seconde discrète de v est alors définie par

$$J_2(v) = \sum_{1 \leq i,j \leq N} \|(Hv)_{i,j}\|_{\mathbb{R}^4}. \quad (\text{II.17})$$

Nous obtenons le problème discrétisé suivant :

$$\inf_{v \in X} F_\delta^2, \quad (\mathcal{P}_\delta^2)$$

où

$$F_\delta^2 = \frac{\|u_d - v\|_X^2}{2\lambda} + J_2(v) + \frac{\delta}{2} (\|v\|_{l^1} + \|\nabla v\|_{l^1 \times l^1}), \quad (\text{II.18})$$

avec

$$\|v\|_{l^1} := \sum_{1 \leq i,j \leq N} |v_{i,j}|,$$

et

$$\|(v^1, v^2)\|_{l^1 \times l^1} := \sum_{1 \leq i,j \leq N} (|v_{i,j}^1| + |v_{i,j}^2|).$$

Dans le cas de la dimension finie, nous avons un résultat d'existence pour $\delta = 0$.

Théorème II.4. *Le problème (\mathcal{P}_δ^2) admet une unique solution pour $\delta \geq 0$ et $\lambda > 0$.*

Dans la suite, on adapte la méthode utilisée par Chambolle dans le cas de $BV(\Omega)$, à l'espace $BV^2(\Omega)$.

3.3.2 Calcul explicite de la solution du problème (\mathcal{P}^δ)

Les conditions d'optimalité associées au problème (\mathcal{P}^δ) nous amènent à considérer la conjuguée de Legendre-Fenchel de J_2 . Le théorème II.5 en donne une caractérisation. Définissons d'abord l'opérateur adjoint de H (qui est la version discrétisée de l'opérateur "divergence seconde") :

$$\forall p \in X^4, \forall v \in X, \quad \langle H^*p, v \rangle_X = \langle p, Hv \rangle_{X^4}.$$

On vérifie que $H^* : X^4 \rightarrow X$ satisfait, pour tout $p = (p^{11}, p^{12}, p^{21}, p^{22}) \in X^4$

$$(H^*p)_{i,j} = \begin{cases} p_{i-1,j}^{11} - 2p_{i,j}^{11} + p_{i+1,j}^{11} & \text{si } 1 < i < N \\ p_{i+1,j}^{11} - p_{i,j}^{11} & \text{si } i = 1, \\ p_{i-1,j}^{11} - p_{i,j}^{11} & \text{si } i = N, \end{cases} \\ + \begin{cases} p_{i,j-1}^{22} - 2p_{i,j}^{22} + p_{i,j+1}^{22} & \text{si } 1 < j < N, \\ p_{i,j+1}^{22} - p_{i,j}^{22} & \text{si } j = 1, \\ p_{i,j-1}^{22} - p_{i,j}^{22} & \text{si } j = N, \end{cases} \\ + \begin{cases} p_{i,j-1}^{12} - p_{i,j}^{12} - p_{i+1,j-1}^{12} + p_{i+1,j}^{12} & \text{si } 1 < i, j < N, \\ p_{i+1,j}^{12} - p_{i+1,j-1}^{12} & \text{si } i = 1, 1 < j < N, \\ p_{i,j-1}^{12} - p_{i,j}^{12} & \text{si } i = N, 1 < j < N, \\ p_{i+1,j}^{12} - p_{i,j}^{12} & \text{si } 1 < i < N, j = 1, \\ p_{i,j-1}^{12} - p_{i+1,j-1}^{12} & \text{si } 1 < i < N, j = N, \\ p_{i+1,j}^{12} & \text{si } i = 1, j = 1, \\ -p_{i+1,j-1}^{12} & \text{si } i = 1, j = N, \\ -p_{i,j}^{12} & \text{si } i = N, j = 1, \\ p_{i,j-1}^{12} & \text{si } i = N, j = N, \end{cases} \quad (\text{II.19}) \\ + \begin{cases} p_{i-1,j}^{21} - p_{i,j}^{21} - p_{i-1,j+1}^{21} + p_{i,j+1}^{21} & \text{si } 1 < i, j < N, \\ p_{i,j+1}^{21} - p_{i,j}^{21} & \text{si } i = 1, 1 < j < N, \\ p_{i-1,j}^{21} - p_{i-1,j+1}^{21} & \text{si } i = N, 1 < j < N, \\ p_{i,j+1}^{21} - p_{i-1,j+1}^{21} & \text{si } 1 < i < N, j = 1, \\ p_{i-1,j}^{21} - p_{i,j}^{21} & \text{si } 1 < i < N, j = N, \\ p_{i,j+1}^{21} & \text{si } i = 1, j = 1, \\ -p_{i,j}^{21} & \text{si } i = 1, j = N, \\ -p_{i-1,j+1}^{21} & \text{si } i = N, j = 1, \\ p_{i-1,j}^{21} & \text{si } i = N, j = N, \end{cases}$$

On montre alors le résultat suivant :

Théorème II.5. *La conjuguée de Legendre-Fenchel de J_2 est $J_2^* = \mathbf{1}_{K_2}$ où*

$$K_2 := \{H^*p \mid p \in X^4, \|p_{i,j}\|_{\mathbb{R}^4} \leq 1, \forall i, j = 1, \dots, N\} \subset X. \quad (\text{II.20})$$

3.3.3 Cas où $\delta = 0$

On s'intéresse maintenant au cas où $\delta = 0$. En effet, le terme en δ dans la définition [II.15](#) était un outil nécessaire dans le cas de la dimension infinie pour prouver l'existence d'une solution. Cependant, nous avons vu qu'en dimension finie, le problème admet une solution même si $\delta = 0$. Pour $\delta = 0$, nous notons le problème de la façon suivante :

$$\inf_{v \in X} F_d^2(v), \quad (\mathcal{P}_d^2)$$

où

$$F_d^2(v) = \frac{\|u_d - v\|_X^2}{2\lambda} + J_2(v). \quad (\text{II.21})$$

De plus,

Théorème II.6. *Soit $(\delta_n)_n$ une suite de termes de \mathbb{R}_+^* qui tend vers 0 en décroissant, et u_{δ_n} la solution du problème $(\mathcal{P}_{\delta_n}^2)$ pour tout $n > 0$. Alors la suite $(u_{\delta_n})_n$ converge vers la solution u_0 du problème (\mathcal{P}_d^2) .*

Démonstration : Comme u_{δ_n} est solution du problème $(\mathcal{P}_{\delta_n}^2)$, on a

$$F_{\delta_n}^2(u_{\delta_n}) \leq F_{\delta_n}^2(0) = \frac{1}{2\lambda} \|u_d\|_X^2. \quad (\text{II.22})$$

Par suite,

$$\|u_d - u_{\delta_n}\|_X \leq \|u_d\|_X. \quad (\text{II.23})$$

Ainsi, la suite $\|u_{\delta_n}\|_X$ est bornée par une constante indépendante de δ_n . Par suite, il existe $\bar{u} \in X$ tel que (u_{δ_n}) converge (à une sous-suite près) vers \bar{u} . Pour tout élément u de X on a

$$F_{\delta_n}^2(u_{\delta_n}) \leq F_{\delta_n}^2(u). \quad (\text{II.24})$$

Ans, en faisant tendre n vers $+\infty$, il vient

$$F_d^2(\bar{u}) \leq F_d^2(u). \quad (\text{II.25})$$

On en déduit que \bar{u} est solution de (\mathcal{P}_d^2) . Par unicité, il vient $\bar{u} = u_0$. On vient de démontrer que $(u_{\delta_n})_n$ est une suite bornée qui n'admet que u_0 comme valeur d'adhérence. On en conclut que $(u_{\delta_n})_n$ converge vers l'unique solution u_0 du problème (\mathcal{P}_d^2) . ■

Comme dans le cas *BV* (proposition [II.1](#)) nous avons une caractérisation de la solution.

Théorème II.7. *La solution v de (\mathcal{P}_d^2) vérifie*

$$v = u_d - P_{\lambda K_2}(u_d),$$

où $P_{\lambda K_2}$ est la projection orthogonale sur λK_2 .

3.4 Un algorithme de point fixe pour implémenter ∂J_2

Nous proposons ici un algorithme de point fixe de type Chambolle : nous étendons à l'ordre deux les résultats énoncés dans la partie 2.3. Ainsi, nous définissons le problème de point fixe

$$p^0 = 0, \quad (\text{II.26a})$$

$$p_{i,j}^{n+1} = \frac{p_{i,j}^n - \tau (H [H^* p^n - f/\lambda])_{i,j}}{1 + \tau \left\| (H [H^* p^n - f/\lambda])_{i,j} \right\|_{\mathbb{R}^4}}, \quad (\text{II.26b})$$

et obtenons également le critère de convergence suivant démontré dans l'annexe B :

Théorème II.8. *Soit $\tau \leq 1/64$. Alors $\lambda (H^* p^n)_n$ converge vers $P_{\lambda K_2}(f)$.*

3.5 Résultats numériques

Nous ne proposons ici que deux exemples : l'un sur une image, l'autre sur un signal de dimension 1. Ces résultats sont à comparer aux résultats obtenus figures II.3 et II.4. Nous observons ainsi que que le modèle ROF2 apporte une réponse satisfaisante au problème du phénomène de *staircasing*. La contre partie est que l'image restaurée subit une régularisation telle qu'un léger effet de flou apparaît. Nous renvoyons le lecteur à l'annexe B ainsi qu'au chapitre 4 pour plus de tests numériques illustrant ce phénomène.

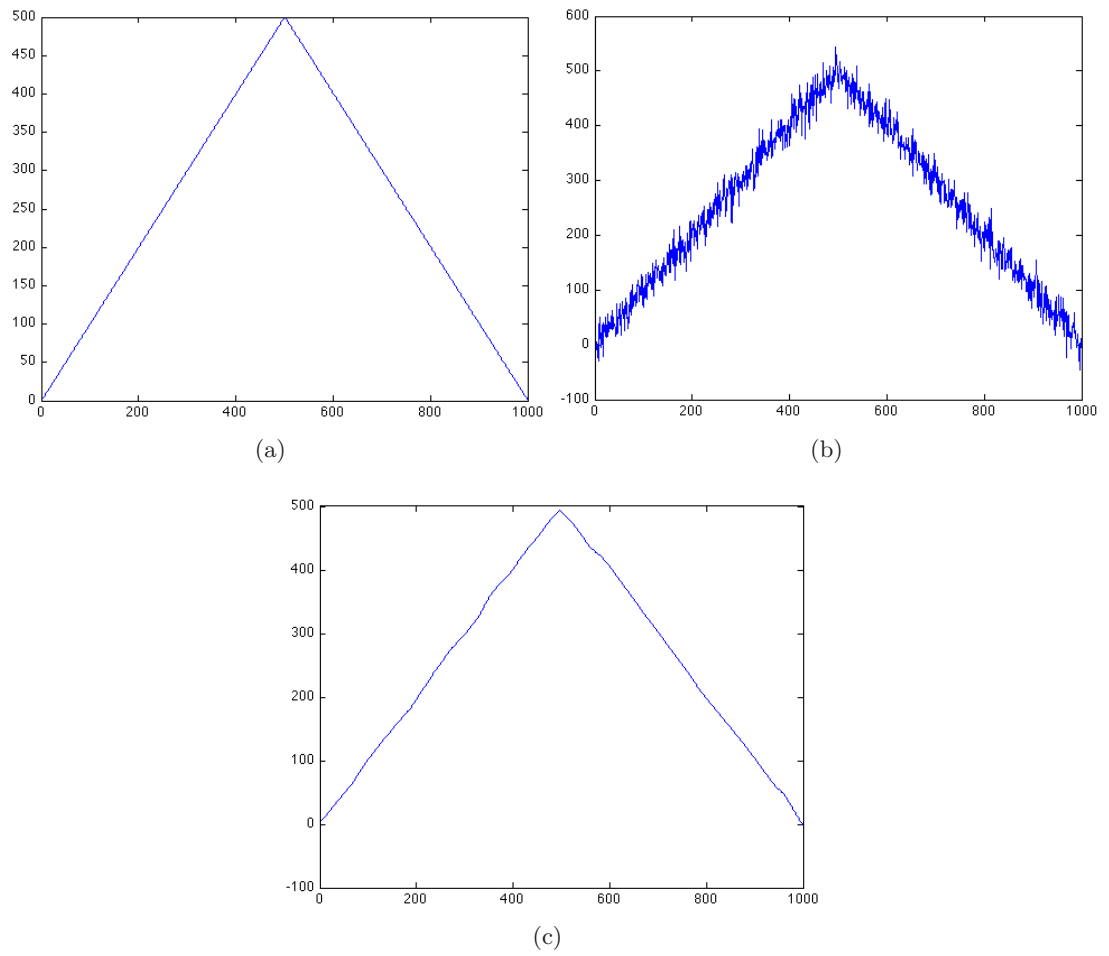


FIG. II.3 – Utilisation du modèle ROF2 pour restaurer un signal artificiellement bruité ($\sigma = 20$, $\lambda = 100$). On observe que le signal restauré est parfaitement lissé : il n'y a plus d'effet *staircasing*.

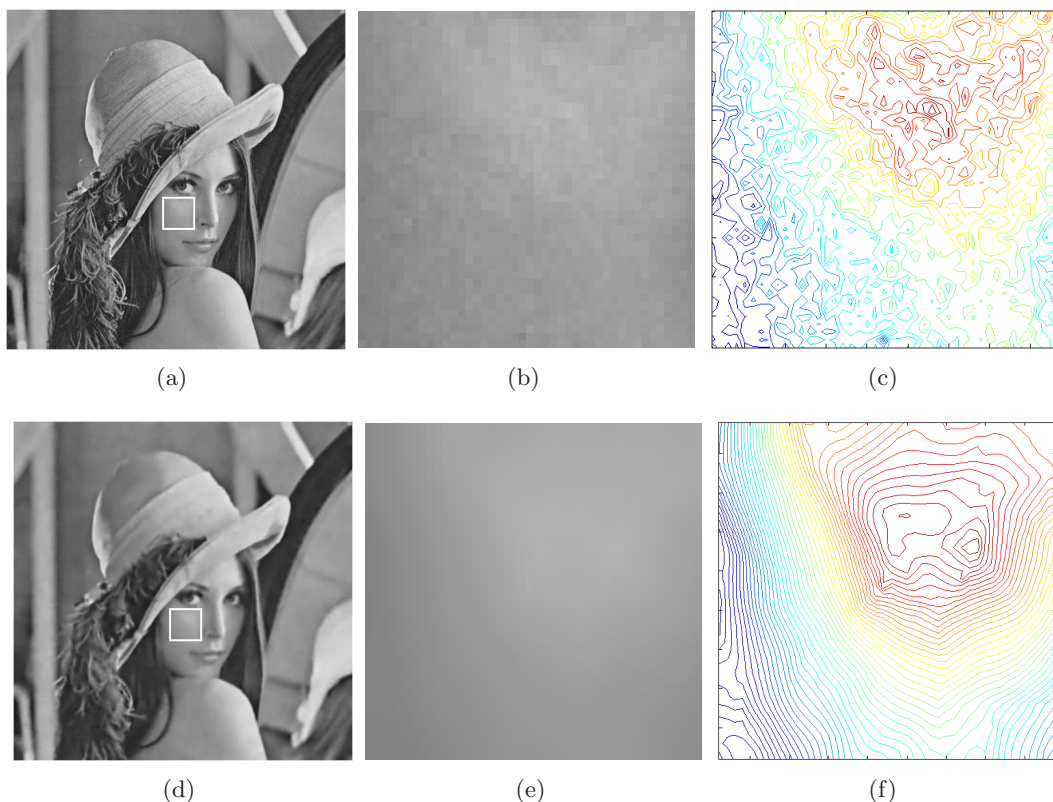


FIG. II.4 – (a) est l'image de départ non bruitée, (d) est l'image bruitée ($\sigma = 20$) restaurée par ROF2 ($\lambda = 25$), (b) et (e) sont les zooms respectivement des parties encadrées de (a) et (d), et (c) et (f) sont les courbes de niveaux respectivement associées à (b) et (e). L'image (e) montre que l'image restaurée est lissée. L'effet *staircasing* a bien disparu. On observe toutefois un léger floutage de l'image restaurée.

4 Modèle mixte

4.1 Introduction

En vue de diminuer l'effet de flou engendré par le modèle ROF2, nous supposons maintenant que l'image de départ se décompose en trois parties : la partie cartoon v , la partie contours u et la partie oscillante de l'image contenant bruit et texture w . De telles décompositions existent déjà [7, 8, 9] utilisant l'espace de Meyer G des fonctions oscillantes [41] (dont nous parlons dans le chapitre 3) à la place de $BV^2(\Omega)$. Toutefois, le modèle que nous proposons ici est très différent : la partie oscillante de l'image est a priori contenue dans le terme $w = u_d - u - v$ alors que v est la composante lisse et que la composante cartoon u contient l'information géométrique de l'image.

Encore une fois, les composantes u , v et w appartiennent à des espaces différents : u

appartient à $BV(\Omega)$, v appartient à $BV^2(\Omega)$ et w appartient à $L^2(\Omega)$. Dans le cadre du débruitage, nous espérons ainsi obtenir, en sommant les composantes u et v , une image restaurée sans l'effet *staircasing* présent dans ROF et sans l'effet de flou engendré par ROF2. Autrement dit, nous espérons par ce modèle, conserver les points forts des modèles ROF et ROF2 tout en évitant leurs effets négatifs.

On considère alors la fonctionnelle suivante, définie sur $BV(\Omega) \times BV^2(\Omega)$:

$$F(u, v) = \frac{1}{2} \|u_d - u - v\|_{L^2(\Omega)}^2 + \lambda TV(u, \Omega) + \mu TV^2(v, \Omega) + \delta \|v\|_{W^{1,1}(\Omega)}, \quad (\text{II.27})$$

avec $\lambda, \mu > 0$, et où le terme en δ est, comme pour le modèle ROF2, un terme régularisant permettant d'obtenir un résultat d'existence. Comme Ω est borné, on a $BV^2(\Omega) \subset H^1(\Omega)$ et nous pouvons ainsi remplacer, dans F , $\|v\|_{W^{1,1}(\Omega)}$ par $\|v\|_{H^1(\Omega)}^2$. C'est ce que nous ferons dans le cas discret afin de faciliter l'implémentation numérique de la solution, le second terme étant différentiable. Le problème que l'on veut résoudre est donc :

$$\inf_{(u,v) \in BV(\Omega) \times BV^2(\Omega)} F(u, v) \quad (\mathcal{P})$$

avec

$$F(u, v) = \frac{1}{2} \|u_d - u - v\|_{L^2(\Omega)}^2 + \lambda TV(u, \Omega) + \mu TV^2(v, \Omega) + \frac{\delta}{2} \|v\|_{H^1(\Omega)}^2 \quad (\text{II.28})$$

4.2 Existence et unicité de la solution

Théorème II.9. *Il existe un unique couple (u, v) réalisant l'infimum de la fonction F définie en II.28.*

Démonstration :

• Existence :

Soit $(u_n, v_n) \in BV(\Omega) \times BV^2(\Omega)$ une suite minimisante,

i.e.

$$\lim_{n \rightarrow +\infty} F(u_n, v_n) = \min_{(u,v) \in K} F(u, v).$$

Alors les suites

$$(TV(u_n, \Omega))_n, (TV^2(v_n, \Omega))_n, \left(\|u_d - u_n - v_n\|_{L^2(\Omega)} \right)_n, \left(\|v_n\|_{L^1(\Omega)} \right)_n \text{ et } \left(\|\nabla v_n\|_{L^1(\Omega)} \right)_n$$

sont bornées.

On rappelle que la norme de $BV^2(\Omega)$ est définie de la manière suivante :

$$\|v_n\|_{BV^2(\Omega)} = \|v_n\|_{L^1(\Omega)} + TV(v_n, \Omega) + TV^2(v_n, \Omega).$$

On sait que $(TV^2(v_n, \Omega))_n$ est bornée ; d'autre part, on a la relation :

$$TV(v_n, \Omega) = \|\nabla v_n\|_{L^1(\Omega)}. \quad (\text{II.29})$$

Ainsi, le fait que $(\nabla v_n)_n$ soit bornée dans $L^1(\Omega)$ signifie que $(TV(v_n, \Omega))_n$ est bornée. On en déduit que $(v_n)_n$ est bornée dans $BV^2(\Omega)$. D'après le théorème I.31, il vient que (v_n) converge fortement (à une sous-suite près) dans $W^{1,1}(\Omega)$ vers un $v^* \in BV^2(\Omega)$. Ainsi, le théorème I.27 nous donne l'inégalité suivante :

$$TV^2(v^*, \Omega) \leq \liminf_{n \rightarrow +\infty} TV^2(v_n, \Omega).$$

On rappelle que la norme de $BV(\Omega)$ est définie de la manière suivante :

$$\|u\|_{BV(\Omega)} = \|u\|_{L^1(\Omega)} + TV(u, \Omega).$$

Le fait que $(v_n)_n$ soit bornée dans $L^1(\Omega)$ implique que $(u_n)_n$ l'est aussi. On déduit alors que $(u_n)_n$ est bornée dans $BV(\Omega)$. Le théorème I.23 nous donne l'existence d'une sous-suite (encore notée u_n) et d'un élément u^* de $BV(\Omega)$ tels que $(u_n)_n$ converge fortement dans $L^1(\Omega)$ vers u^* . On déduit alors du théorème I.20 l'inégalité suivante :

$$TV(u^*, \Omega) \leq \liminf_{n \rightarrow +\infty} TV(u_n, \Omega).$$

Les normes étant continues, on a finalement

$$F(u^*, v^*) \leq \liminf_{n \rightarrow +\infty} F(u_n, v_n) = \min_K F(u, v).$$

Ainsi, (u^*, v^*) minimise le problème \mathcal{P} .

• Unicité :

La fonction est strictement convexe. ■

4.3 Discrétisation du modèle

Nous reprenons les notations des parties précédentes. En remplaçant de plus $\|\cdot\|_{W^{1,1}}$ par $\|\cdot\|_{H^1}^2$, on obtient le problème discrétisé suivant :

$$\inf_{(u,v) \in X^2} F_d(u, v), \quad (\mathcal{P}_d)$$

où

$$F_d(u, v) = \frac{1}{2} \|u_d - u - v\|_X^2 + \lambda J_1(u) + \mu J_2(v) + \frac{\delta}{2} \left(\|v\|_X^2 + \|\nabla v\|_X^2 \right). \quad (\text{II.30})$$

Comme dans le cas des modèles ROF et ROF2, nous avons le résultat d'existence et d'unicité suivant :

Théorème II.10. *Soient $\lambda, \mu, \delta > 0$. Le problème (\mathcal{P}_d) admet une unique solution.*

Remarque II.3. *Contrairement au modèle ROF2 où le terme en δ était inutile en dimension finie, on ne peut s'en passer ici tout en conservant le résultat d'existence précédent. Malgré tout, nous avons observé dans nos tests numériques que de prendre $\delta = 0$ ou δ très petit ne changeait pas visuellement le résultat obtenu. Nous garderons toutefois le terme en δ pour la résolution numérique du problème, quitte à le prendre nul lors des tests proprement dits.*

4.4 Résolution numérique et algorithme

Les outils d'analyse convexe non lisse nous permettent d'obtenir des conditions d'optimalité nécessaires et suffisantes. Plus précisément, (u, v) est solution de (P_d) si et seulement si

$$\begin{cases} 0 \in \partial_u \left(\lambda J_1(u) + \frac{1}{2} \|u_d - u - v\|^2 \right) \\ 0 \in \partial_v \left(\mu J_2(v) + \frac{1}{2} \|u_d - u - v\|_X^2 + \frac{1}{2} \|v\|_X^2 + \frac{1}{2} \|\nabla v\|_X^2 \right), \end{cases} \quad (\text{II.31})$$

En utilisant les propriétés du sous-différentiel, on voit que (II.31) est équivalente à

$$\begin{cases} u = u_d - v - \Pi_{\lambda K_1}(u_d - v) \\ v = u_d - u - \delta v - \delta \nabla v - \Pi_{\mu K_2}(u_d - u - \delta v - \delta \nabla v). \end{cases} \quad (\text{II.32})$$

où K_1 et K_2 sont les convexes fermés définis au chapitre 2 :

$$K_1 = \{\operatorname{div} p \mid p \in X^2, \|p_{i,j}\|_{\mathbb{R}^2} \leq 1 \forall i, j = 1, \dots, N\} \quad (\text{II.33})$$

et

$$K_2 = \{H^* p \mid p \in X^4, \|p_{i,j}\|_{\mathbb{R}^4} \leq 1, \forall i, j = 1, \dots, N\}. \quad (\text{II.34})$$

Nous obtenons ainsi le problème de point fixe

$$(u, v) = G(u, v), \quad (\text{II.35})$$

où

$$\begin{aligned} G : X^2 &\rightarrow X^2 \\ (u, v) &\mapsto \begin{pmatrix} u_d - v - \Pi_{\lambda K_1}(u_d - v) \\ u_d - u - \delta v - \delta \nabla v - \Pi_{\mu K_2}(u_d - u - \delta v - \delta \nabla v) \end{pmatrix}. \end{aligned} \quad (\text{II.36})$$

Afin de résoudre ce problème, nous introduisons la fonction G_α définie de la façon suivante :

$$G_\alpha(u, v) = \begin{pmatrix} u \\ v \end{pmatrix} + \alpha \left(G(u, v) - \begin{pmatrix} u \\ v \end{pmatrix} \right). \quad (\text{II.37})$$

On peut alors proposer l'algorithme suivant :

Algorithme

- **Étape 1** : choix de u_0 et v_0 (par exemple $u_0 = 0$ et $v_0 = u_d$) et $0 < \alpha < 1$.
- **Étape 2** : définition des suites $((u_n, v_n))_n$ de la façon suivante :

$$\begin{cases} u_{n+1} = u_n + \alpha (u_d - v_n - \Pi_{\lambda K_1}(u_d - v_n) - u_n) \\ v_{n+1} = v_n + \alpha (u_d - u_n - \delta v_n - \delta \nabla v_n - \Pi_{\mu K_2}(u_d - u_n - \delta v_n - \delta \nabla v_n) - v_n). \end{cases}$$

- **Étape 3** : si le critère d'arrêt n'est pas satisfait, $k := k + 1$ et retour à 2.

Théorème II.11. *Pour $\alpha \in]0, 1/2[$ et $\delta \leq 0.1$, l'algorithme converge vers le point fixe de G .*

Démonstration :

Il suffit de montrer que la fonction G_α est strictement contractante. On a

$$G_\alpha(u, v) = \left(\begin{array}{c} (1 - \alpha)u + \alpha u_d - \alpha v - \alpha \Pi_{\lambda K_1}(u_d - v) \\ (1 - (1 + \delta)\alpha)v + \alpha u_d - \alpha u - \delta \alpha \nabla v - \alpha \Pi_{\mu K_2}(u_d - u - \delta v - \delta \nabla v) \end{array} \right). \quad (\text{II.38})$$

On munit X de la norme

$$\|u\|_X = \sum_{i,j=1}^n |u_{i,j}|,$$

et X^2 de la norme

$$\|(u, v)\|_{X^2} = \sum_{i,j=1}^n (|u_{i,j}| + |v_{i,j}|) = \|u\|_X + \|v\|_X.$$

Ainsi, pour $(u_1, v_1), (u_2, v_2) \in X^2$, on montre que

$$\|G_\alpha(u_1, v_1) - G_\alpha(u_2, v_2)\|_{X^2} \leq \max(|1 - \alpha|, |1 - (1 + \delta)\alpha|, 8\delta|\alpha|, 2|\alpha|) \|((u_1, v_1) - (u_2, v_2))\|_{X^2}$$

Par suite, nous pouvons d'ores et déjà constater que α doit être strictement compris entre 0 et $\frac{1}{2}$. De plus, pour δ suffisamment petit (inférieur à 0.1 en pratique), nous obtenons l'inégalité suivante :

$$\|G_\alpha(u_1, v_1) - G_\alpha(u_2, v_2)\|_{X^2} \leq \max(|1 - \alpha|, 2|\alpha|) (\|u_1 - u_2\|_X + \|v_1 - v_2\|_X).$$

De plus, il est aisé de voir que si $\alpha \in]0; 1/2[$, alors $\max(|1 - \alpha|, 2|\alpha|) < 1$, ce qui signifie qu'alors G_α est strictement contractante.

Par suite, la suite définie par

$$(u_{n+1}, v_{n+1}) = G_\alpha(u_n, v_n) \quad (\text{II.39})$$

converge vers le point fixe de G_α . Il est de plus aisé de voir que G et G_α ont les mêmes points fixes. On en conclut que l'algorithme converge pour $\alpha \in]0; 1/2[$.

■

Remarque II.4. *Numériquement, on observe que la convergence est plus rapide pour $\alpha = 0.6$.*

4.5 tests numériques

Nous ne proposons ici qu'un seul exemple. Nous renvoyons le lecteur au chapitre 6 pour plus de tests numériques ainsi que des comparaisons entre les différents modèles décrits dans ce chapitre.



FIG. II.5 – Utilisation du modèle mixte pour la restauration d'image. (a) image bruitée par un bruit blanc Gaussien ($\sigma = 25$); (b) composante u de l'a solution; (c) composante u de la solution; (d) image restaurée ($u + v$). On observe que u contient l'information géométrique de l'image et que v est très lissée. La somme des deux nous permet d'obtenir une image restaurée sans effet *staircasing* préservant de façon convenable l'information géométrique de l'image de départ.

Chapitre III

Extraction de texture

1 Etat de l'art

Il existe de nombreux modèles permettant l'extraction de texture. La liste des modèles que nous citons dans cette partie est loin d'être exhaustive. Nous nous contentons de présenter ici les modèles qui sont, à notre connaissance, les plus performants.

1.1 Modèle de Meyer

Yves Meyer a analysé de nombreux modèles de décomposition $u_d = u+v$, et notamment le modèle ROF, mettant en avant le fait que la norme L^2 n'est pas le meilleur choix pour la capture de la partie oscillante d'une image. C'est pourquoi il a introduit dans [41] un espace a priori mieux adapté : l'espace \mathcal{G} des fonctions oscillantes.

1.1.1 L'espace \mathcal{G} des fonctions oscillantes

Définition III.1. $\mathcal{G}(\mathbb{R}^2)$ est l'espace de Banach constitué des fonctions f qui peuvent s'écrire de la façon suivante :

$$f = \frac{\partial g_1}{\partial x_1} + \frac{\partial g_2}{\partial x_2} = \operatorname{div}(g),$$

avec $g_1, g_2 \in L^\infty(\mathbb{R}^2)$, et où $\frac{\partial g_1}{\partial x_1}$ et $\frac{\partial g_2}{\partial x_2}$ sont des dérivées au sens des distributions.

$\mathcal{G}(\mathbb{R}^2)$ est de plus muni de la norme

$$\|v\|_{\mathcal{G}} = \inf \{ \|g\|_{\infty} \mid v = \operatorname{div}(g), g = (g_1, g_2) \in L^\infty(\mathbb{R}^2) \times L^\infty(\mathbb{R}^2) \},$$

où

$$\|g\|_{\infty} = \sup_{x \in \mathbb{R}^2} \|g(x)\|_{\mathbb{R}^2},$$

avec

$$\|g(x)\|_{\mathbb{R}^2} = \sqrt{|g_1|^2 + |g_2|^2}(x).$$

On peut démontrer qu'une suite de fonctions de plus en plus oscillantes voit sa norme \mathcal{G} tendre vers 0. Y. Meyer propose alors dans [41] le modèle suivant :

$$\inf_{(u,v) \in BV \times \mathcal{G} / f=u+v} (TV(u, \Omega) + \alpha \|v\|_{\mathcal{G}}) \quad (\text{III.1})$$

Toutefois, il ne propose pas dans [41] de méthode numérique permettant de résoudre le problème ; c'est pourquoi les modèles suivants ont été proposés. La résolution numérique de ceux-ci nécessite la définition d'une version discrète donnée par J.F. Aujol dans [8].

1.1.2 Espace de Meyer discret

On rappelle que $X = \mathbb{R}^{N \times N}$, où $N \times N$ est la taille de l'image.

Définition III.2. Soit G l'espace vectoriel défini par

$$G = \{v \in X \mid \exists g \in Y \text{ tel que } v = \text{div}(g)\} \quad (\text{III.2})$$

et

$$\|v\|_G = \inf \{ \|g\|_{\infty} \mid v = \text{div}(g), g = (g^1, g^2) \in X^2 \}, \quad (\text{III.3})$$

où

$$\|g\|_{\infty} = \sup_{1 \leq i, j \leq N} \|g_{i,j}\|_{\mathbb{R}^2},$$

avec

$$\|g_{i,j}\|_{\mathbb{R}^2} = \sqrt{|g_{i,j}^1|^2 + |g_{i,j}^2|^2}.$$

J.F. Aujol introduit de plus l'ensemble

$$G_{\mu} = \{v \in G \mid \|v\|_G \leq \mu\}. \quad (\text{III.4})$$

1.2 Modèle de Vese-Osher

L. Vese et S. Osher ont proposé les premiers une méthode pour résoudre le problème de Meyer. Pour cela, ils proposent dans [52] le modèle suivant :

$$\inf_{(u,v) \in BV(\Omega) \times \mathcal{G}(\Omega)} \left(TV(u, \Omega) + \lambda \|u_d - u - v\|_{L^2(\Omega)}^2 + \mu \|v\|_{\mathcal{G}(\Omega)} \right), \quad (\text{III.5})$$

où Ω est un ouvert borné, u la composante cartoon et v la composante oscillante.. Pour calculer la solution, ils remplacent le terme $\|v\|_{\mathcal{G}(\Omega)}$ par $\left\| \sqrt{g_1^2 + g_2^2} \right\|_{L^1(\Omega)}$, puis calculent les équations d'Euler-Lagrange associées. Pour plus de détails, nous renvoyons le lecteur à [52].

1.3 Modèle de Osher-Solé-Vese

Les auteurs modifient dans [51] le modèle (III.1) en remplaçant $\mathcal{G}(\Omega)$ par $W^{-1,2}(\Omega)$. Nous rappelons que pour un ouvert Ω de \mathbb{R}^2 , l'espace $W^{-1,q}(\Omega)$ est défini comme étant le dual de la fermeture de $C_c^1(\Omega)$ dans $W^{1,\frac{q}{q-1}}(\Omega)$. Les auteurs considèrent ainsi le problème suivant :

$$\inf_{u \in BV(\Omega)} \left(TV(u, \Omega) + \lambda \|u_d - u\|_{-1,2}^2 \right), \quad (\text{III.6})$$

qu'ils résolvent en utilisant la théorie des équations aux dérivées partielles. Signalons que J-F. Aujol et A. Chambolle proposent dans [11] un algorithme de type Chambolle pour résoudre ce problème.

1.4 Modèle de Aujol-Aubert-Blanc-Feraud-Chambolle

En réalité, les deux problèmes précédents n'utilisent pas la norme \mathcal{G} mais une approximation de celle-ci. J-F. Aujol, G. Aubert, L. Blanc-Feraud et A. Chambolle proposent dans [8] un modèle utilisant vraiment cette norme. Ils résolvent le problème discrétisé

$$\inf_{(u,v) \in X \times X} J_1(u) + \frac{1}{2\lambda} \|u_d - u - v\|_X^2 + \chi_{G_\mu}(v), \quad (\text{III.7})$$

et proposent un algorithme pour calculer numériquement la solution. Enfin, un lien est établi entre le modèle de Meyer discrétisé et le problème limite associé à (III.7) ($\lambda = 0$), justifiant que le problème (III.7) permet bien de résoudre le problème de Meyer discrétisé. Une étude du modèle est également effectuée dans le cas continu (voir [7]).

1.5 Modèle $TV-L^1$

Il s'agit simplement du modèle ROF dans lequel la norme L^2 est remplacée par la norme L^1 . Le modèle est le suivant :

$$\inf_{u \in BV(\Omega)} \|u_d - u\|_{L^1(\Omega)} + \lambda TV(u, \Omega). \quad (\text{III.8})$$

Il est aisé de montrer que ce problème (III.8) admet au moins une solution (la démonstration est identique à celle donnant l'existence d'une solution pour le modèle ROF). Toutefois, une conséquence de ce changement de norme est que la fonctionnelle à minimiser n'est plus strictement convexe, ce qui implique qu'il n'y a pas nécessairement unicité de la solution (voir par exemple [25]).

Ce modèle a été introduit dans [3] dans le cas de la dimension 1. M. Nikolova a proposé dans [47] d'utiliser ce modèle dans le cadre de la restauration d'image. Même si la composante oscillante d'une image n'a aucune raison d'avoir une norme L^1 faible, il a été montré que le modèle (III.8) est très performant pour la décomposition d'image (voir par exemple [29, 58]).

En montrant que la résolution de ce problème était équivalente à la résolution d'un problème géométrique, V. Duval, J-F. Aujol, et Y. Gousseau ont pu expliquer l'efficacité

de ce modèle dans [30], mettant en évidence que le comportement du modèle dépend de l'échelle des objets : lorsque le paramètre λ varie, un objet disparaît de la composante cartoon (et apparaît donc dans la composante oscillante) lorsque le rapport périmètre/aire est trop petit. Ils proposent de plus dans un algorithme rapide permettant de minimiser (III.8).

Nous revenons en détail sur ce modèle dans le chapitre 4, dans lequel nous proposons une méthode de résolution originale.

2 Retour sur le modèle d'ordre 2

2.1 Introduction

Ce travail fait suite au travail réalisé dans [31] (voir annexe A), dans lequel nous proposons un modèle variationnel du deuxième ordre dans le but d'extraire la texture d'une image, supposant que nous pouvions décomposer l'image u_d en deux parties : une partie cartoon u et une partie v contenant les fortes oscillations (texture + bruit), ainsi que les contours. Le modèle était le suivant :

$$\inf_{(u,v) \in BV(\Omega) \times BV^2(\Omega)} \frac{1}{2} \|u_d - u - v\|_{L^2(\Omega)}^2 + \lambda |u|_{BV(\Omega)} + \mu |v|_{BV^2(\Omega)} + \delta \|v\|_{W^{1,1}(\Omega)}, \quad (\text{III.9})$$

avec $\lambda, \mu, \delta \geq 0$, et où $\|u_d - u - v\|_{L^2(\Omega)}^2$ est le terme d'ajustement aux données et $\delta \|v\|_{W^{1,1}(\Omega)}$ un terme de "pénalisation" permettant d'avoir l'existence et l'unicité d'une solution. Nous ne travaillerons pas ici avec ce même modèle, mais avec le modèle de type Tychonov proposé dans [15] dans le cadre du débruitage, en supposant que l'image étudiée est dans l'ensemble $BV^2(\Omega)$ des fonctions à hessien borné, ce qui revient à poser $u = 0$ dans le modèle ci-dessus. Nous utiliserons donc le modèle suivant :

$$\inf_{v \in BV^2(\Omega)} \frac{1}{2} \|u_d - v\|_{L^2(\Omega)}^2 + \mu |v|_{BV^2(\Omega)} + \delta \|v\|_{W^{1,1}(\Omega)}, \quad (\text{III.10})$$

avec $\mu, \delta \geq 0$.

Nous préférons ici étudier le modèle réduit (III.10) plutôt que (III.9) pour des raisons numériques. En effet, le temps de calcul lors de l'implémentation est très nettement inférieur.

Nous rappelons que v représente la partie cartoon et $w := u_d - v$ la partie contenant texture/bruit et une partie des contours. Nous pouvons considérer un modèle de décomposition d'image cartoon/texture efficace, lorsque toute la partie oscillante de l'image est contenue dans la composante w , et que la composante v contient la partie régulière de l'image ainsi que les contours. Il s'avère qu'alors le modèle (III.10), s'il est un bon modèle de restaurartion, n'est pas efficace pour l'extraction de texture. En effet, c'est ici la partie oscillante w qui contient la majorité des contours de l'image de départ. Le but de ce travail est donc de proposer un moyen de faire disparaître les contours de la composante w . Pour cela, nous conservons l'idée introduite dans [31] de rendre le modèle globalement anisotrope. Ceci permet d'améliorer la décomposition mais ne permet pas de la rendre réellement satisfaisante. Nous proposons ici d'améliorer cette idée en remplaçant

l'anisotropie globale par une anisotropie locale, afin que chaque contour de l'image ait un traitement particulier.

Nous mettrons, dans un premier temps, en évidence les limites des modèles original et globalement anisotrope, puis introduirons le modèle localement anisotrope. Dans cette partie, nous ne proposerons des tests numériques que pour deux images (figure III.1). Nous renvoyons le lecteur au chapitre 6 pour plus de tests numériques.

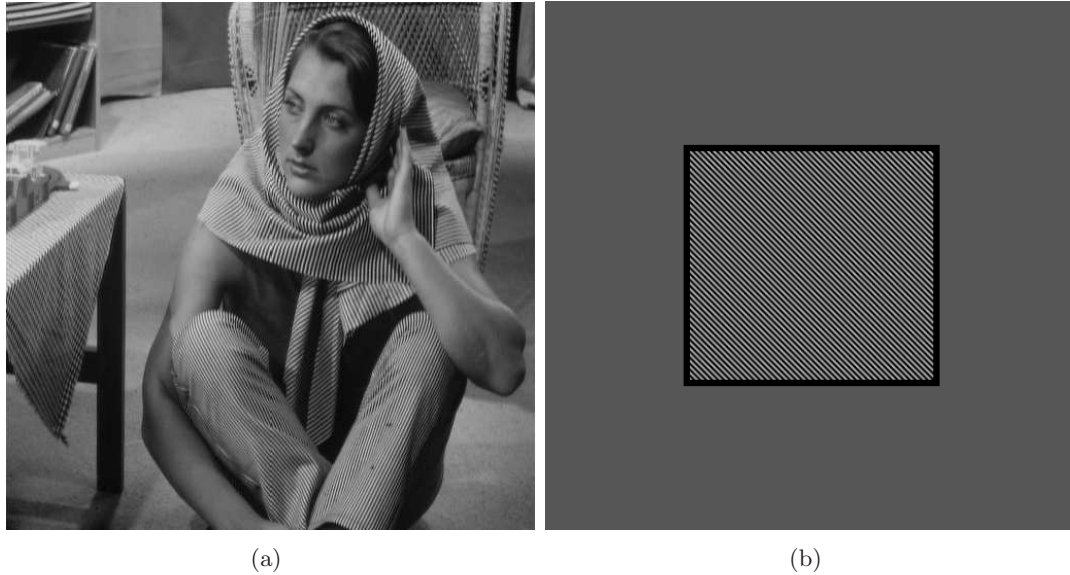


FIG. III.1 – images tests

3 Amélioration du modèle

Nous avons observé dans [15] que le modèle du second ordre engendre un phénomène de flou sur la partie régulière. Ceci signifie que les contours sont en quelque sorte trop “extraits” de la composante BV^2 , ce qui implique qu'ils apparaissent dans leur grande majorité dans la composante “texture”, ce qui rend le modèle de décomposition peu efficace dans le cadre de l'extraction de texture. Ce problème est illustré par la figure III.2.

Nous pouvons voir que, si la composante oscillante contient toute la texture, elle contient beaucoup trop d'information géométrique. En effet, la figure III.2 montre que les différents contours (bras, pieds de table pour l'image III.1 (a), contour du carré pour l'image III.1 (b)...) sont bien présents dans la partie très oscillante. Nous avons proposé dans [31] une réponse très partielle à ce problème : rendre l'opérateur hessien globalement anisotropique afin de privilégier des directions.

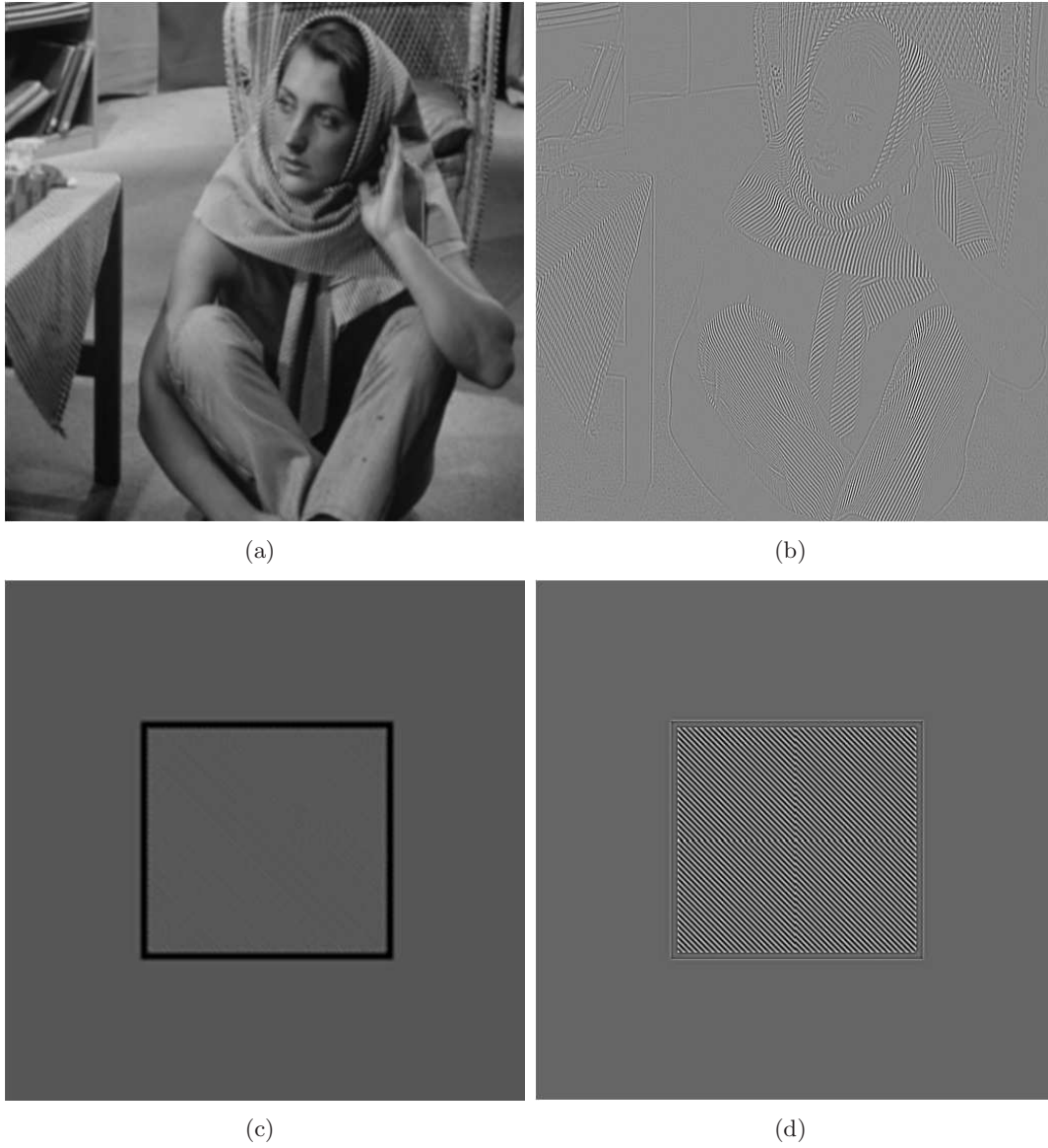


FIG. III.2 – (a) et (b) sont les composantes BV^2 des images de départ. (c) et (d) sont les composantes oscillantes recalibrée.

3.1 Modèle globalement anisotrope

Nous pouvons rendre le modèle anisotrope en modifiant l'opérateur hessien. Nous rappelons que l'opérateur hessien discret est défini par

$$(Hv)_{i,j} = ((Hv)_{i,j}^{11}, (Hv)_{i,j}^{12}, (Hv)_{i,j}^{21}, (Hv)_{i,j}^{22}), \quad v \in X,$$

avec, pour $i, j = 1, \dots, N$,

$$\begin{aligned}
(Hv)_{i,j}^{11} &= \begin{cases} v_{i+1,j} - 2v_{i,j} + v_{i-1,j} & \text{si } 1 < i < N, \\ v_{i+1,j} - v_{i,j} & \text{si } i = 1, \\ v_{i-1,j} - v_{i,j} & \text{si } i = N, \end{cases} \\
(Hv)_{i,j}^{12} &= \begin{cases} v_{i,j+1} - v_{i,j} - v_{i-1,j+1} + v_{i-1,j} & \text{si } 1 < i \leq N, 1 \leq j < N, \\ 0 & \text{si } i = 1, \\ 0 & \text{si } i = N, \end{cases} \\
(Hv)_{i,j}^{21} &= \begin{cases} v_{i+1,j} - v_{i,j} - v_{i+1,j-1} + v_{i,j-1} & \text{si } 1 \leq i < N, 1 < j \leq N, \\ 0 & \text{si } i = 1, \\ 0 & \text{si } i = N, \end{cases} \\
(Hv)_{i,j}^{22} &= \begin{cases} v_{i,j+1} - 2v_{i,j} + v_{i,j-1} & \text{si } 1 < j < N, \\ v_{i,j+1} - v_{i,j} & \text{si } j = 1, \\ v_{i,j-1} - v_{i,j} & \text{si } j = N. \end{cases}
\end{aligned}$$

Il suffit d'annuler l'une des composante pour qu'elle n'ait plus d'effet lors du traitement de l'image. Concrètement, nous voyons que si nous posons $(Hv)^{1,1} = 0$, les sauts horizontaux ne sont plus détectés, ce qui signifie que les contours verticaux de l'image ne sont plus lissés. Ainsi, après traitement, nous retrouvons ces contours dans la composante cartoon alors qu'ils ont complètement disparu de la composante oscillante.

De la même manière, nous pouvons décider d'agir sur les contours horizontaux ($(Hv)^{2,2} = 0$), ou sur les contours à la fois verticaux et horizontaux ($(Hv)^{1,1} = (Hv)^{2,2} = 0$).

C'est cette dernière possibilité que nous illustrons avec la figure III.3 en comparant les résultats avec et sans anisotropie. Nous observons que les contours horizontaux et verticaux ont bien disparu de la composante oscillante w alors que celle-ci contient la grande majorité de l'information oscillante de l'image de départ.

Malheureusement, bien que nous puissions agir sur des contours obliques en faisant effectuer à l'image une rotation, nous ne pouvons seulement traiter que deux directions simultanément, en imposant de plus qu'elles soient perpendiculaires entre elles. De plus, ajouter de l'anisotropie afin de perdre les contours dans une certaine direction implique que nous perdrons aussi une très grande partie de l'information sur toute texture orientée dans la même direction. La figure III.4 met en évidence tous ces problèmes : nous testons sur une image sans texture, mais proposant des contours dans toutes les directions, ainsi que sur une image très proche de l'image III.1(b), mais dont la texture est orientée, cette fois-ci, horizontalement, c'est à dire dans le sens de contours que nous souhaitons voir disparaître de la composante oscillante. On observe alors que, dans ces conditions, le modèle globalement anisotrope se révèle être inefficace (figure III.4).

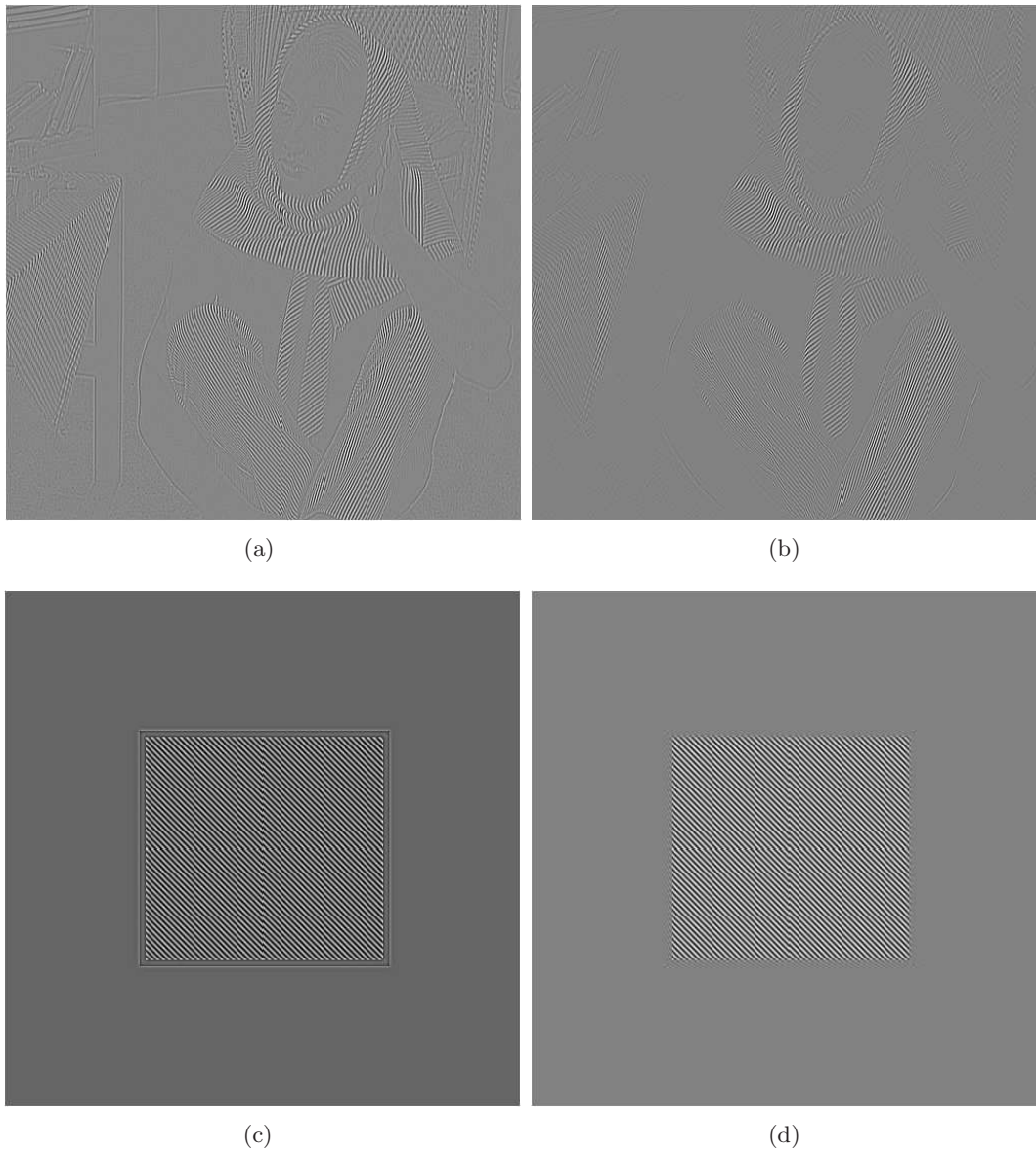


FIG. III.3 – (a) et (c) sont les composantes oscillantes des images obtenues à l'aide du modèle original, et (b) et (d) sont celles obtenues à l'aide du modèle globalement anisotrope. On observe que les contours verticaux et horizontaux disparaissent. Toutefois, les contours obliques subsistent (cf les jambes et bras de barbara). De plus, on remarque une perte d'information au niveau de la texture orientée verticalement ou horizontalement (cf les bandes horizontales du foulard de barbara).

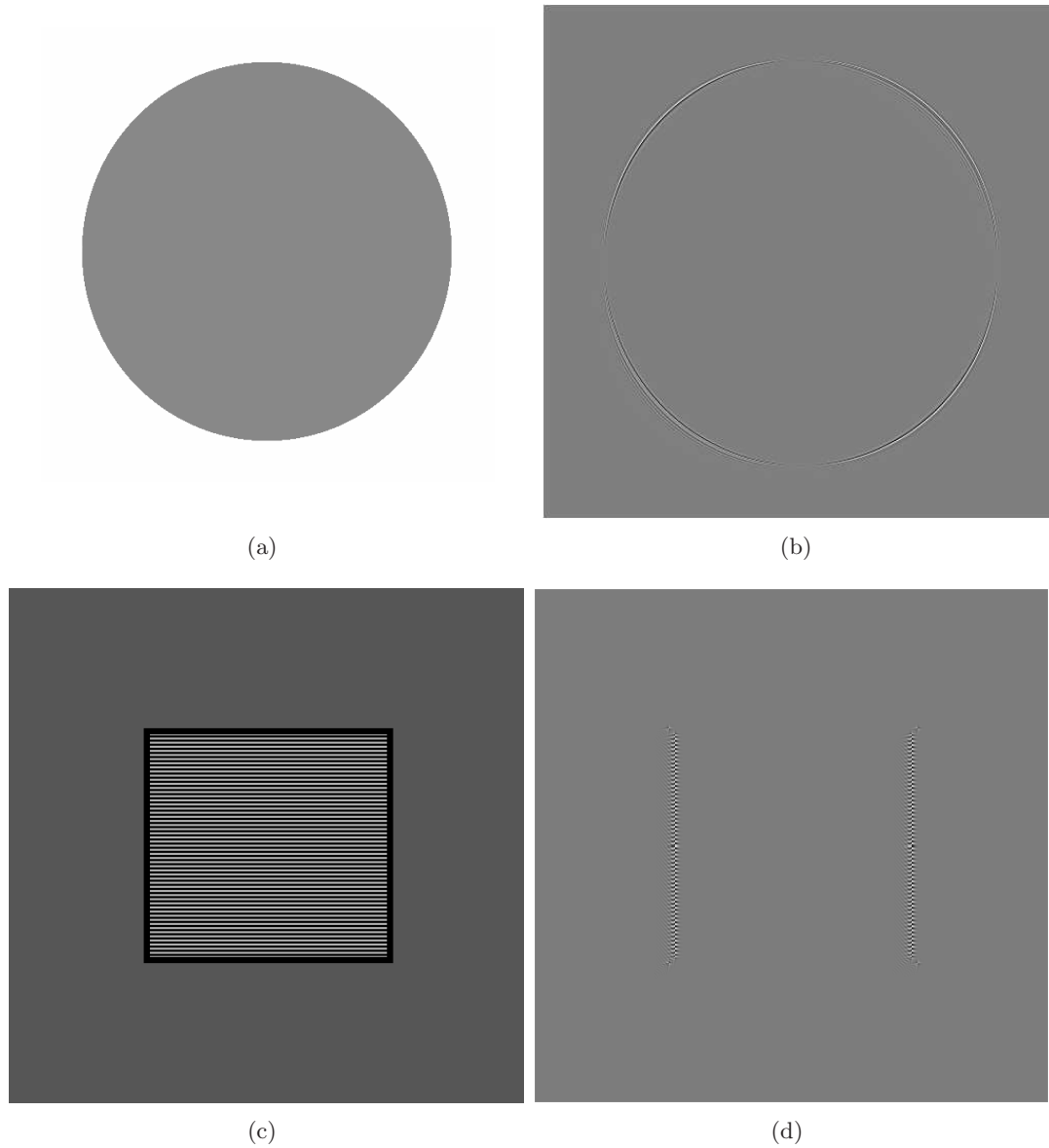


FIG. III.4 – (a) et (c) sont les images originales. (b) et (d) représentent la texture extraite à l'aide du modèle globalement anisotrope. On observe l'inefficacité du modèle : les contours du cercle sont quasiment intacts tandis que la texture horizontale est totalement perdue.

Pour remédier à ces difficultés, nous n'allons plus rendre le modèle anisotrope globalement, mais localement en chacun des points composant les contours de l'image.

3.2 Modèle localement anisotrope

3.2.1 Principe

Nous avons mis en évidence les limites du modèle anisotrope en mettant en avant le fait que nous ne pouvons l'orienter que dans deux directions simultanément, ce qui ne permet de traiter qu'une partie des contours, et le fait que nous perdons nécessairement l'information oscillante de l'image dans ces deux directions. Nous proposons alors de traiter les contours localement. Voici la démarche :

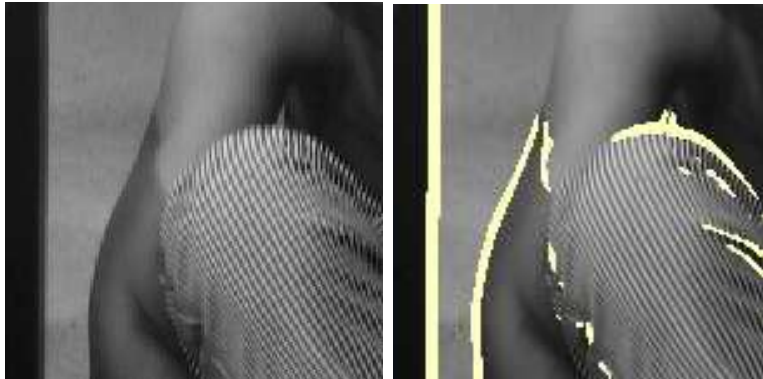
- **Etape 1 :** Nous détectons dans un premier temps les points qui posent problème (figure III.5), à savoir les contours qui apparaissent avec la texture et dont on veut se débarrasser. Chacun des pixels restants est traité avec l'algorithme original.
- **Etape 2 :** Ensuite, nous calculons en ces points le vecteur gradient, dont la valeur des composantes nous indique l'angle α que forme localement le contour avec l'horizontal.
- **Etape 3 :** Nous extrayons une matrice de taille $p \times p$ centrée en chacun de ces points et effectuons une rotation locale d'angle α calculé précédemment autour de chaque pixel d'intérêt. Il ne reste qu'à appliquer à la nouvelle matrice l'algorithme original. Il faut prendre une matrice suffisamment grande (par exemple $p = 5$) pour éviter les effets de bords. Ceci est illustré figure III.6, pour un pixel détecté à la première étape.

Remarque III.1. *Signalons que le procédé peut-être systématisé en introduisant des matrices de rotations locales, que nous appliquons à l'opérateur hessien discret (voir [16])*

Nous espérons ainsi faire disparaître significativement les contours tout en préservant la quasi totalité de la texture. La figure III.7 illustre le résultat lorsque nous traitons chacun des points détectés. On observe bien la disparition des contours sur la composante oscillante alors que la texture est elle presque totalement conservée. La contre-partie est que nous obtenons une composante cartoon qui contient toujours une partie de la texture, notamment au voisinage des pixels composant les contours.



(a)



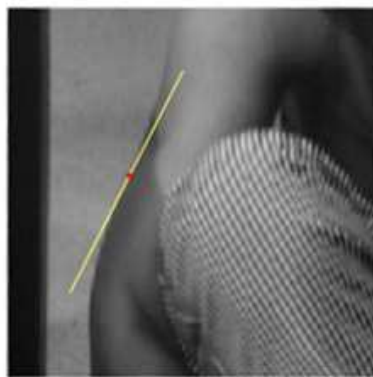
(b)

(c)

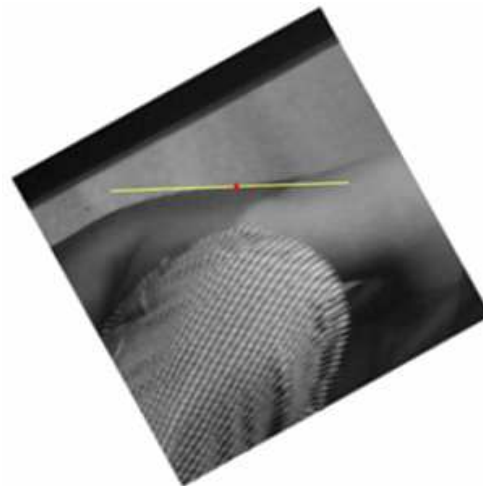
FIG. III.5 – (c) mise en évidence des pixels en lesquels nous appliquerons le modèle anisotrope. Les autres points seront traités avec le modèle ROF2 brut.



(a)



(b)



(c)

FIG. III.6 – (a) choix d'un des pixels mis en évidence figure III.5 (c). Après avoir déterminer l'angle formé entre l'horizontale et la tangente en ce point, nous faisons effectuer à l'image une rotation afin que cette tangente devienne horizontale. Il ne reste plus qu'à appliquer sur un voisinage de ce pixel le modèle globalement anisotrope.

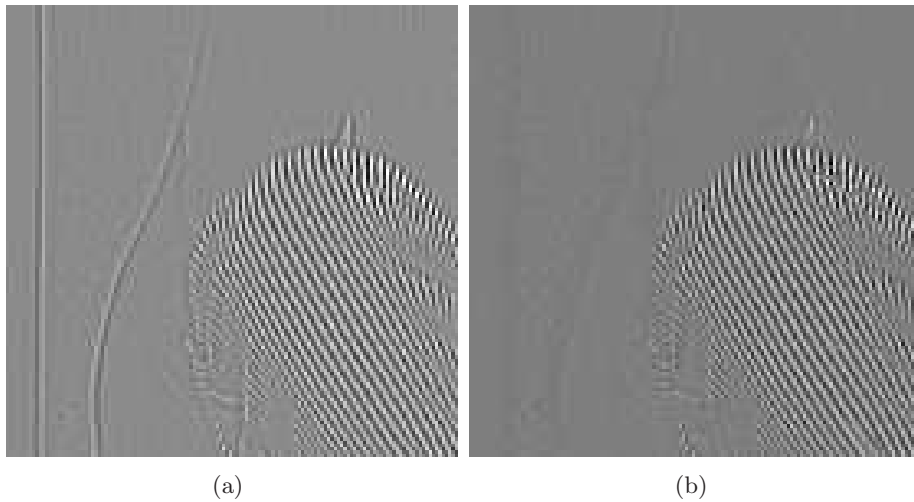


FIG. III.7 – A gauche : texture extraite à l’aide du modèle brut. A droite : texture extraite à l’aide du modèle localement anisotrope. On observe que la texture est conservée dans sa quasi totalité alors que les contours disparaissent, quelles que soient leurs directions.

On observe bien que la texture est récupérée alors que les contours indésirables ont presque disparu. En pratique, nous sommes confrontés à quelques difficultés que nous explicitons dans la partie suivante.

3.2.2 Implémentation

La première grande difficulté est de détecter précisément les points qui posent problème. Il faut en détecter autant que possible, tout en évitant d’y inclure des points qui composent en réalité ce que l’on définit dans l’image comme texture, car se serait autant d’information importante perdue. Un détecteur de contour classique (type gradient) sur l’image de départ est donc à proscrire, car il ne permet pas de distinguer précisément les contours de la texture (le gradient est élevé dans les deux cas).

D’autre part, un détecteur de contours plus précis, serait également à proscrire, pour des raisons liées à la nature même de notre modèle. En effet, nous avons pu observer dans [15] que, même si le modèle préserve bien les contours, il va avoir tendance à les diffuser. Cette diffusion est d’autant plus importante que le nombre d’itérations ainsi que la valeur du paramètre λ sont élevés. Cela signifie en particulier qu’un point de l’image qui n’appartient pas à un contour mais qui en est suffisamment proche en subira les effets lors du traitement. Ainsi, un détecteur de contour, aussi performant soit-il, ne pourra détecter qu’une partie des points problématiques. Ce phénomène est illustré en figure (III.8), sur une image qui ne contient volontairement pas de texture. On peut voir que, malgré l’anisotropie, des traits verticaux sont toujours apparents.

Afin de remédier à ce problème, nous faisons subir à l’image un premier traitement avec l’algorithme original. Nous choisissons le paramètre λ en fonction de l’image, assez grand afin de la lisser suffisamment pour que la texture disparaisse. Les contours sont alors

diffusés (phénomène décrit au chapitre 2), et un gradient sur cette image lissée permet ainsi de détecter tous les points influencés par les contours, qu'un détecteur de contours classique quel qu'il soit ne pouvait détecter sur l'image de départ. La figure III.8 met en évidence les pixels qui subissent l'influence des contours. Nous pouvons voir, qu'alors, les contours disparaissent presque totalement de la composante oscillante.

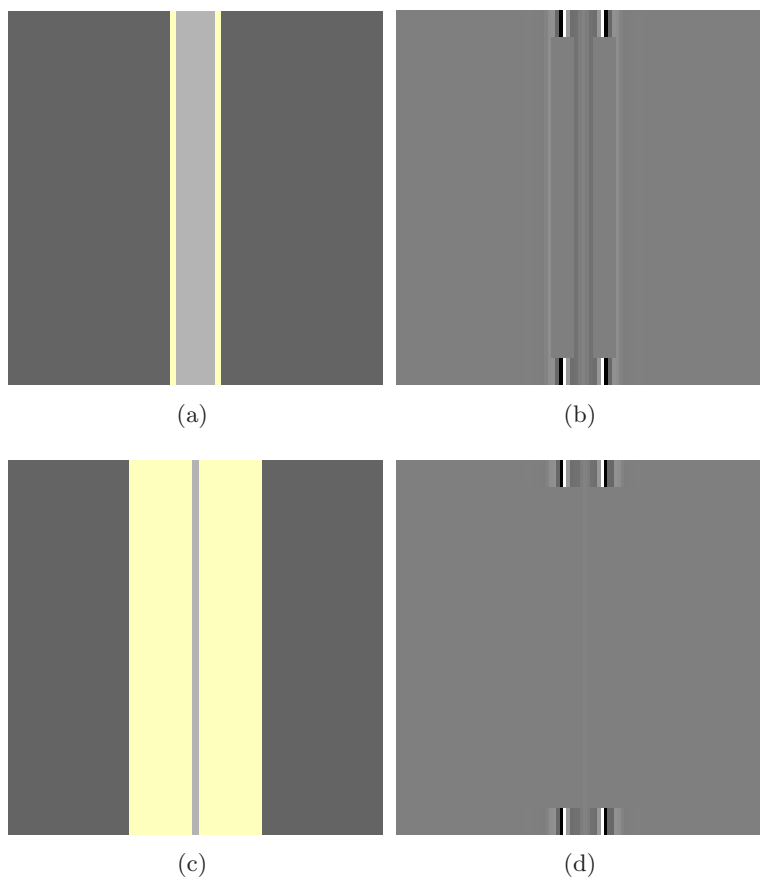


FIG. III.8 – Détection des points à traiter "localement" sur une image non texturée. Sur les images de gauches sont mis en évidences les points subissant un traitement local anisotrope. Les images de droite sont les résultats d'extraction de texture associés. Nous détectons sur l'image (a) les contours à l'aide d'un gradient et faisons subir à ces points un traitement anisotrope ; on observe en (b) qu'une partie des contours subsistent dans la partie oscillante. Enfin, la figure (e) met en évidence les points détectés à l'aide d'un gradient sur la partie cartoon de l'image ayant déjà subi l'algorithme ROF2 brut. On observe que les points détectés sont beaucoup plus nombreux. La composante oscillante de la solution associée est alors débarassées de la quasi totalité des contours. Nous n'avons pas effectué de traitement anisotrope sur les bords de l'image afin de pouvoir comparer facilement les résultats avec et sans anisotropie globale ou locale

De plus, un pré-traitement de l'image faisant disparaître une très grande partie de la texture, les points relevés ne correspondent, dans leur grande majorité, qu'à des contours. On observe sur la figure III.9 le résultat obtenu sur l'image complète III.1 (a).



FIG. III.9 – Résultat du modèle localement anisotrope sur l'image entière. On observe à droite que nous conservons dans la composante oscillante une grande partie de la texture malgré une légère perte d'information en certains endroits. La composante lisse, quant à elle, conserve, en ces endroits, une partie de la texture.

Si la perte d'information utile dans la composante oscillante est très acceptable comparativement à l'information géométrique que nous éliminons, la composante lisse n'est pas réellement satisfaisante. En effet, trop d'information oscillante y est conservée. Il semblerait donc cohérent de parler, non pas d'un modèle de décomposition, mais plutôt d'un modèle d'extraction et de localisation de texture.

Chapitre IV

Aspects numériques

1 Méthode de Lagrangien Augmenté

Nous décrivons dans cette partie les résultats du livre de K. Ito et K. Kunisch [38].

1.1 Introduction

Soient

- \mathcal{X} et \mathcal{H} sont des espaces de Hilbert,
- \mathcal{C} un sous ensemble convexe fermé de \mathcal{X} ,
- Λ est un élément de $\mathcal{L}(\mathcal{X}, \mathcal{H})$.
- $f : \mathcal{X} \rightarrow \mathbb{R}$ est une fonction convexe \mathcal{C}^1 ,
- $\varphi : \mathcal{H} \rightarrow (-\infty, \infty]$ est une fonction propre, convexe mais pas nécessairement \mathcal{C}^1 .

Nous voulons ici résoudre des problèmes d'optimisation du type

$$\min_{x \in \mathcal{C}} f(x) + \varphi(\Lambda x). \quad (\text{IV.1})$$

Une théorie du type multiplicateurs de Lagrange est développée pour traiter la non différentiabilité de φ : Soient $x, \gamma \in \mathcal{H}$ et soit $c > 0$. On définit la famille d'approximations de Moreau-Yoshida généralisées $\varphi_c(x, \gamma)$ par

$$\varphi_c(x, \gamma) = \inf_{u \in \mathcal{H}} \left\{ \varphi(\Lambda x - u) + \langle \gamma, u \rangle_{\mathcal{H}} + \frac{c}{2} \|u\|_H^2 \right\}. \quad (\text{IV.2})$$

Ceci est équivalent à une approche de Lagrangien augmenté. En effet, on remarque que (IV.1) est équivalent à

$$\begin{cases} \min_{x \in \mathcal{C}} f(x) + \varphi(x - u) \\ \text{avec } x \in \mathcal{C} \text{ et } u = 0 \text{ dans } \mathcal{H}. \end{cases} \quad (\text{IV.3})$$

En traitant la contrainte d'égalité $u = 0$ dans (IV.3) par la méthode de Lagrangien augmenté, on obtient le problème de minimisation

$$\min_{x \in \mathcal{C}, u \in \mathcal{H}} f(x) + \varphi(\Lambda x - u) + \langle \gamma, u \rangle_{\mathcal{H}} + \frac{c}{2} \|u\|_{\mathcal{H}}^2, \quad (\text{IV.4})$$

où $\gamma \in \mathcal{H}$ est un multiplicateur de Lagrange et c un paramètre pénalisant strictement positif. De manière équivalente, (IV.4) peut s'écrire

$$\min_{x \in \mathcal{C}} F_c(x, \gamma) := f(x) + \varphi_c(\Lambda x, \gamma). \quad (\text{IV.5})$$

1.2 Approximation de Moreau-Yoshida Généralisée

Les résultats énoncés dans cette partie proviennent de [38]. Il existe, bien entendu, d'autres références sur le sujet ([20, 14]).

Définition IV.1 (Opérateur Monotone). Soit A une multi-application de \mathcal{X} dans \mathcal{X}^* . On note $D(A)$ le domaine de A .

1. A est monotone si, pour tout $x_1, x_2 \in D(A)$ et $y_1 \in Ax_1, y_2 \in Ax_2$,

$$\langle y_1 - y_2, x_1 - x_2 \rangle \geq 0. \quad (\text{IV.6})$$

2. A est maximal monotone si toute extension monotone de A coïncide avec A .

Soit φ une fonction convexe s.c.i. sur X . Pour $x_1^* \in \partial\varphi(x_1)$ et $x_2^* \in \partial\varphi(x_2)$, nous avons

$$\begin{cases} \varphi(x_1) - \varphi(x_2) \leq \langle x_2^*, x_1 - x_2 \rangle, \\ \varphi(x_2) - \varphi(x_1) \leq \langle x_1^*, x_2 - x_1 \rangle. \end{cases}$$

Par suite, $\langle x_1^* - x_2^*, x_1 - x_2 \rangle \geq 0$. Ainsi, $\partial\varphi$ est un opérateur monotone sur $\mathcal{X} \times \mathcal{X}^*$.

Théorème IV.2 (Rockafellar). Soit \mathcal{X} un espace de Banach réel. Si φ est une fonction convexe, propre et s.c.i. sur \mathcal{X} , alors $\partial\varphi$ est un opérateur maximal de \mathcal{X} dans \mathcal{X}^* .

Démonstration : cf [38] page 105. ■

Soit \mathcal{H} un espace de Hilbert réel, identifié à son dual \mathcal{H}^* .

Définition IV.3. Soit A un opérateur maximal monotone sur \mathcal{H} . Pour $r > 0$, on appelle résolvante de A l'opérateur $T_r^A = (I + rA)^{-1}$. L'approximation de Yoshida A_r de A est définie par $A_r(x) = \frac{1}{r}(x - T_r^A(x))$.

L'opérateur A_r est uni-valué, monotone, lipschitzien (de constante r), et $A_r x \in AT_r x$ pour tout $x \in \mathcal{H}$.

Par la suite, nous nous intéresserons au cas où $A = \partial\varphi$. La proposition qui suit nous permet d'obtenir une expression concrète de l'approximation de Yoshida de cet opérateur.

Proposition IV.1. Soit $\phi : \mathcal{H} \rightarrow]-\infty, +\infty]$ une fonction convexe, propre, sci sur \mathcal{H} . Alors

$$T_r^{\partial\phi}(x) = \arg \min_y \left\{ \frac{1}{2} \|y - x\|_{\mathcal{H}}^2 + r\phi(y) \right\}. \quad (\text{IV.7})$$

$\varphi_c(x, \gamma)$ est une approximation de $\varphi(x)$ au sens suivant :

Théorème IV.4. Pour $x, \gamma \in \mathcal{H}$, l'infimum de (IV.2) est atteint en l'unique point

$$u_c(x, \gamma) = x - T_{\frac{1}{c}} \left(x + \frac{\gamma}{c} \right).$$

$\varphi_c(x, \gamma)$ est convexe, Lipschitz continûment Fréchet différentiable en x , et

$$\varphi'_c(x, \gamma) = A_{\frac{1}{c}} \left(x + \frac{\gamma}{c} \right).$$

De plus, $\lim_{c \rightarrow \infty} \varphi_c(x, \gamma) = \varphi(x)$, et, pour tout $x, \gamma \in \mathcal{H}$,

$$\varphi \left(T_{\frac{1}{c}} \left(x + \frac{\gamma}{c} \right) \right) - \frac{1}{2c} \|\gamma\|_{\mathcal{H}}^2 \leq \varphi_c(x, \gamma) \leq \varphi(x).$$

Le théorème suivant donne une caractérisation équivalente de l'inclusion $\gamma \in \partial\varphi(x)$.

Théorème IV.5.

1. Si $\gamma \in \partial\varphi(x)$ pour $x, \gamma \in \mathcal{H}$, alors $\gamma = \varphi'_c(x, \gamma)$ pour tout $c > 0$.
2. Réciproquement, s'il existe $c > 0$ tel que $\gamma = \varphi'_c(x, \gamma)$, alors $\gamma \in \partial\varphi(x)$.

1.3 Système d'optimalité

Nous rappelons que nous supposons que $f : \mathcal{X} \rightarrow \mathbb{R}$ et $\varphi : \mathcal{H} \rightarrow (-\infty, \infty]$ sont sci, convexes, que f est continûment différentiable, et que φ est propre. De plus, $\Lambda \in \mathcal{L}(\mathcal{X}, \mathcal{H})$ et \mathcal{C} est un sous-ensemble convexe fermé de \mathcal{X} . Nous ajoutons les hypothèses suivantes :

(H1) f et φ sont positives.

(H2) Il existe $\sigma > 0$ tel que, pour tous $x_1, x_2 \in \mathcal{C}$,

$$\langle f'(x_1) - f'(x_2), x_1 - x_2 \rangle_{X^*, X} \geq \sigma \|x_1 - x_2\|_X^2,$$

et

(H3) $\varphi \circ \Lambda$ est continue sur \mathcal{C} .

La condition (H2) entraîne que la fonction $f + \varphi$ est coercive. Etant de plus convexe et propre on déduit le résultat suivant :

Théorème IV.6. Le problème (IV.1) admet une unique solution.

Nous introduisons maintenant le problème de minimisation régularisé suivant, dépendant de γ :

$$\min_{x \in \mathcal{C}} F_c(x, \gamma), \quad (\mathcal{P}_c^\gamma)$$

où

$$F_c(x, \gamma) = f(x) + \varphi_c(\Lambda x, \gamma),$$

avec φ_c défini par (IV.2).

Par ailleurs, le théorème IV.4. montre que $x \mapsto \varphi_c(\Lambda x, \gamma)$ est minorée par $-\frac{1}{2c} \|\gamma\|_{\mathcal{H}}^2$. Ainsi,

Théorème IV.7. *Il existe un unique couple (u_c, v_c) réalisant l'infimum de F_c sur \mathcal{C} .*

Théorème IV.8 (Multiplicateurs de Lagrange). *Soit \hat{x} un minimiseur de (IV.1). Alors il existe $\hat{\gamma} \in \mathcal{H}$ vérifiant*

$$\begin{cases} \langle f'(\hat{x}), x - \hat{x} \rangle_{\mathcal{X}^*, \mathcal{X}} + \langle \hat{\gamma}, \Lambda(x - \hat{x}) \rangle_{\mathcal{H}} = 0 \quad \forall x \in \mathcal{C} \\ \hat{\gamma} = \varphi'_c(\Lambda \hat{x}, \hat{\gamma}) \end{cases} \quad (\text{IV.8})$$

Démonstration : \hat{x} est un minimiseur de (IV.1). Par suite, pour $x \in \mathcal{C}$ et pour $0 < t < 1$, nous avons $\hat{x} + t(x - \hat{x}) \in \mathcal{C}$ et

$$f(\hat{x} + t(x - \hat{x})) + \varphi(\Lambda(\hat{x} + t(x - \hat{x}))) \geq f(\hat{x}) + \varphi(\Lambda \hat{x}).$$

Comme

$$\varphi(\Lambda((1-t)\hat{x} + tx)) - \varphi(\Lambda \hat{x}) \leq t(\varphi(\Lambda x) - \varphi(\Lambda \hat{x})),$$

il vient

$$\frac{1}{t}(f(\hat{x} + t(x - \hat{x})) - f(\hat{x})) + (\varphi(\Lambda x) - \varphi(\Lambda \hat{x})) \geq 0.$$

En passant à la limite quand t tend vers 0^+ , on obtient le résultat suivant :

$$\forall x \in \mathcal{C}, \quad \langle f'(\hat{x}), x - \hat{x} \rangle_{\mathcal{X}^*, \mathcal{X}} + \varphi(\Lambda x) - \varphi(\Lambda \hat{x}) \geq 0. \quad (\text{IV.9})$$

L'inégalité (IV.9) peut se récrire

$$\varphi(\Lambda x) \geq \varphi(\Lambda \hat{x}) + \langle -f'(\hat{x}), x - \hat{x} \rangle_{\mathcal{X}^*, \mathcal{X}}, \quad (\text{IV.10})$$

ce qui signifie que $-f'(\hat{x}) \in \partial(\varphi \circ \Lambda)(\hat{x})$, c'est à dire $-f'(\hat{x}) \in \Lambda^* \partial \varphi(\Lambda \hat{x})$. Ainsi, il existe $\hat{\gamma} \in \partial \varphi(\Lambda \hat{x})$ tel que $-f'(\hat{x}) = \Lambda^* \hat{\gamma}$. Ceci implique, d'après le théorème IV.5, que

$$\hat{\gamma} = \varphi'_c(\Lambda \hat{x}, \hat{\gamma}).$$

Ainsi,

$$\begin{aligned} \langle f'(\hat{x}), x - \hat{x} \rangle_{\mathcal{X}^*, \mathcal{X}} + \langle \hat{\gamma}, \Lambda(x - \hat{x}) \rangle_{\mathcal{H}} &= \langle f'(\hat{x}), x - \hat{x} \rangle_{\mathcal{X}^*, \mathcal{X}} + \langle \Lambda^* \hat{\gamma}, x - \hat{x} \rangle_{\mathcal{H}} \\ &= \langle f'(\hat{x}), x - \hat{x} \rangle_{\mathcal{X}^*, \mathcal{X}} + \langle -f'(\hat{x}), x - \hat{x} \rangle_{\mathcal{H}} \\ &= 0, \end{aligned}$$

d'où la conclusion. ■

1.4 Méthode de Lagrangien augmenté

Nous donnons dans cette partie un algorithme dont les résultats précédents permettent de montrer qu'il converge vers la solution \hat{x} de (IV.1).

Algorithme

- **Etape 1** : On choisit une valeur de départ $\gamma_1 \in \mathcal{H}$, un réel $c > 0$ et on pose $k = 1$.
- **Etape 2** : Soit $\gamma_k \in \mathcal{H}$ fixé. On cherche $x_k \in \mathcal{C}$ qui vérifie

$$f(x_k) + \varphi_c(\Lambda x_k, \gamma_k) = \min_{x \in \mathcal{C}} f(x) + \varphi_c(\Lambda x, \gamma_k).$$

- **Etape 3** : $\gamma_{k+1} = \varphi'_c(\Lambda x_k, \gamma_k)$.
- **Etape 4** : Si le critère d'arrêt n'est pas satisfait, on pose $k := k + 1$ et on retourne à 2.

Théorème IV.9. *La suite (x_k, γ_k) est bien définie et satisfait*

$$\|x_k - \hat{x}\|_{\mathcal{H}} \rightarrow 0, \quad \text{quand } k \rightarrow \infty.$$

Démonstration : cf [38] page 117. ■

2 Applications aux modèles $TV - L^1$ et $TV^2 - L^1$

Nous allons utiliser l'algorithme décrit précédemment afin de résoudre le modèle $TV - L^1$:

$$\inf \mu \|u_d - u\|_{L^1(\Omega)} + \lambda TV(u, \Omega), \quad (TVL^1)$$

où $\lambda > 0$. Nous résoudrons de la même manière le modèle $TV^2 - L^1$:

$$\inf_{u \in BV^2(\Omega)} \mu \|u_d - u\|_{L^1(\Omega)} + \lambda TV^2(u, \Omega) + \xi \|u\|_{W^{1,1}(\Omega)}, \quad (TV^2L^1)$$

où $\lambda, \mu, \xi > 0$, et où le terme en ξ est, comme pour le modèle ROF2, un terme permettant d'obtenir le résultat d'existence et d'unicité, et qui a vocation à être extrêmement proche de 0. Nous avons donc le résultat suivant :

Théorème IV.10. *Les problèmes (TVL^1) et (TV^2L^1) admettent une solution.*

Démonstration : La démonstration est identique aux démonstrations de l'existence des théorèmes II.1 et II.3. ■

Remarque IV.1. *Contrairement aux problèmes (\mathcal{P}_{ROF}) et (\mathcal{P}_{ROF2}) , il n'y a pas nécessairement ici unicité de la solution. Ceci vient du fait que la norme L^1 n'est pas strictement convexe.*

En utilisant les notations ainsi que les méthodes de discrétisation décrites au chapitre 2, nous nous intéressons aux deux problèmes discrets, dont nous allons implémenter numériquement la solution en utilisant l'algorithme de Lagrangien augmenté décrit dans la partie 1.4. Pour les deux précédents modèles, le problème discret a la forme suivante :

$$\inf_{u \in X} \mu \|u_d - u\|_{l^1} + \lambda J_i(u) + \frac{\xi_i}{2} (\|u\|_{l^1} + \|\nabla u\|_{l^1 \times l^1}). \quad (\text{IV.11})$$

où $\mu, \lambda, \xi_2 > 0, \xi_1 = 0$, et où

$$\begin{aligned} \|u\|_{l^1} &= \sum_{i,j} |u_{i,j}|, \\ \|(u^1, u^2)\|_{l^1 \times l^1} &= \sum_{i,j} (|u_{i,j}^1| + |u_{i,j}^2|), \\ J_1(u) &= \sum_{i,j} \|(\nabla u)_{i,j}\|_{\mathbb{R}^2}, \end{aligned}$$

et

$$J_2(v) = \sum_{i,j} \|(Hv)_{i,j}\|_{\mathbb{R}^4}.$$

Nous avons de plus le résultat d'existence suivant :

Théorème IV.11. *Le problème (IV.11) admet une solution pour $\lambda, \mu > 0$ et $\xi_2 \geq 0$.*

Démonstration : La fonctionnelle à minimiser est convexe et coercive. ■

Remarque IV.2. *Comme pour le cas continu, il n'y a pas nécessairement unicité de la solution du problème (IV.11).*

Ainsi, comme dans le cas du modèle d'ordre deux décrit au chapitre II, nous poserons systématiquement $\xi_2 = 0$. En effet, la différence entre les résultats obtenus avec $\xi_2 = 0$ et $\xi_2 > 0$ très petit ne sont, encore une fois, pas visibles. Nous nous intéressons donc au problème suivant :

$$\min_{u \in X} \mu \|u_d - u\|_{l^1} + \lambda J_i(u), \quad (\mathcal{P}_d^i)$$

où $\mu, \lambda > 0$ et $i \in \{1, 2\}$.

Toutefois, la méthode de Lagrangien augmenté décrite dans la partie précédente ne peut fonctionner que si la fonctionnelle à minimiser possède une partie différentiable au sens de Fréchet, qui vérifie de plus la condition (H2). En effet, la démonstration du théorème IV.9 montre que la convergence de l'algorithme défini page 69 ne converge que

grâce à cette partie différentiable. Nous ajouterons donc, aux deux modèles ci-dessus, un terme Fréchet-différentiable, assurant la convergence de l'algorithme.

Ainsi, nous obtenons le modèle

$$\min_{u \in X} F_\delta^i(u), \quad (\mathcal{P}_\delta^i)$$

où

$$F_\delta^i(u) := \mu \|u_d - u\|_{l^1} + \lambda J_i(u) + \frac{\delta}{2} \left(\|u\|_X^2 + \|u_d - u\|_X^2 \right),$$

et où $X = \mathbb{R}^{N \times N}$ avec $N \times N$ la taille de l'image, $\delta > 0$ et $i \in \{1, 2\}$. Nous rappelons que

$$\|u\|_X = \left(\sum_{i,j} u_{i,j}^2 \right)^{1/2}. \quad (\text{IV.12})$$

Nous avons le résultat d'existence et d'unicité suivant :

Théorème IV.12. *Soient $\lambda, \mu, \delta > 0$. Les problèmes (\mathcal{P}_δ^1) et (\mathcal{P}_δ^2) admettent une unique solution.*

Démonstration : Le résultat vient du fait que les deux fonctionnelles à minimiser sont coercives et strictement convexes. ■

Remarque IV.3. *Lors des tests numériques, nous choisirons $\delta > 0$ le plus proche possible de zéros afin de limiter au maximum l'effet du terme différentiable.*

Nous écrivons maintenant le problème différemment afin de coller le plus possible aux notations de la première section de ce chapitre. Le problème devient :

$$\min_{(u,v) \in \mathcal{C}} F_\delta^i(u, v), \quad (\tilde{\mathcal{P}}_\delta^i)$$

où

$$F_\delta^i(u, v) := \mu \|v\|_{l^1} + \lambda J_i(u) + \frac{\delta}{2} \left(\|u\|_X^2 + \|v\|_X^2 \right),$$

et

$$\mathcal{C} = \{(u, v) \in X \times X \mid v = u_d - u\}.$$

Lors des tests numériques, $\delta > 0$ sera toujours pris le plus proche possible de zéro. De ce fait, le terme en delta n'aura qu'une influence infime sur les résultats numériques tout en permettant d'assurer la convergence de l'algorithme vers la solution du problème. Nous avons de plus le résultat suivant :

Proposition IV.2. *Soit $(\delta_n)_n$ une suite d'éléments de \mathbb{R}_+^* tendant vers 0 en décroissant. Pour tout n , on note $(u_{\delta_n}, v_{\delta_n})$ la solution du problème $(\tilde{\mathcal{P}}_{\delta_n}^i)$. Alors toute valeur d'adhérence de la suite $((u_{\delta_n}, v_{\delta_n}))_n$ est une solution de $(\tilde{\mathcal{P}}_d^i)$.*

Démonstration : Soit (u_0, v_0) une valeur d'adhérence de la suite $((u_{\delta_n}, v_{\delta_n}))_n$ et soit $\left((u_{\delta_{n_k}}, v_{\delta_{n_k}})\right)_k$ la sous-suite associée. Comme $(u_{\delta_{n_k}}, v_{\delta_{n_k}})$ est solution du problème $(\mathcal{P}_{\delta_{n_k}})$, on a, pour tout $(u, v) \in \mathcal{C}$,

$$F_{\delta_{n_k}}^i \left((u_{\delta_{n_k}}, v_{\delta_{n_k}}) \right) \leq F_{\delta_{n_k}}^i (u, v). \quad (\text{IV.13})$$

Ainsi, en faisant tendre k vers $+\infty$, il vient

$$F^i(u_0, v_0) \leq F^i(u, v). \quad (\text{IV.14})$$

Ceci étant vrai pour tout $(u, v) \in \mathcal{C}$, on en déduit que (u_0, v_0) est un minimiseur de F^i . ■

Remarque IV.4. *Il est facile de montrer que si la solution (\hat{u}, \hat{v}) du problème $(\tilde{\mathcal{P}}_d^i)$ est unique, alors toute suite $((u_{\delta_n}, v_{\delta_n}))_n$ de solutions des problèmes $(\tilde{\mathcal{P}}_{\delta_n}^i)$ converge vers (\hat{u}, \hat{v}) . En effet, une telle suite est nécessairement bornée, et la démonstration de la proposition IV.2 montre qu'elle n'admet qu'une unique valeur d'adhérence.*

Afin d'alléger les notations, et par analogie avec la première section de ce chapitre, nous noterons dans la suite :

$$\mathcal{X} = \mathcal{H} = X \times X,$$

$$\mathcal{Q}(v) := \|v\|_{L^1},$$

et

$$F_\delta(u, v) = f(u, v) + \varphi(u, v),$$

où

$$\begin{cases} f : (u, v) \mapsto \frac{\delta}{2} \left(\|u\|_X^2 + \|v\|_X^2 \right), \\ \varphi : (u, v) \mapsto \mu \mathcal{Q}(v) + \lambda J_i(u), \\ \Lambda = Id. \end{cases} \quad (\text{IV.15})$$

Il est alors évident que f et φ sont s.c.i., convexes, que f est continûment différentiable et vérifie (H2), et que φ est propre. Nous nous trouvons donc bien dans le cadre théorique de la partie précédente.

2.1 Résolution explicite de l'étape 3

Proposition IV.3. *Soient $c > 0$, $\gamma = (\gamma^1, \gamma^2) \in X \times X$ et $x = (u, v) \in X \times X$. Alors*

$$\varphi'_c(x, \gamma) = \begin{pmatrix} \Pi_{\lambda K_i} (cu + \gamma^1) \\ \Pi_{\mu \bar{\mathcal{B}}_\infty} (cv + \gamma^2) \end{pmatrix},$$

où $i \in \{1, 2\}$, avec K_1 et K_2 définis en (II.10) et (II.20), et où

$$\bar{\mathcal{B}}_\infty = \left\{ u \in X \mid \max_{i,j} |u_{i,j}| \leq 1 \right\}.$$

Démonstration : On a $\varphi : X \times X \rightarrow]-\infty, +\infty]$. Par suite, pour $c > 0$ et $x = (u, v) \in X \times X$,

$$T_{\frac{1}{c}}^{\partial\varphi}(x) := T_{\frac{1}{c}}(x) = \arg \min_{y \in X \times X} \left\{ \frac{1}{2} \|y - x\|_{X \times X}^2 + \frac{1}{c} \varphi(y) \right\}.$$

Ceci équivaut à

$$x - T_{\frac{1}{c}}(x) \in \frac{1}{c} \partial\varphi(T_{\frac{1}{c}}(x)).$$

En notant $T_{\frac{1}{c}}(x) = (u_c, v_c)$, ceci équivaut encore à

$$\begin{cases} u - u_c & \in \frac{\lambda}{c} \partial J_i(u_c) \\ v - v_c & \in \frac{\mu}{c} \partial Q(v_c), \end{cases}$$

puis à,

$$\begin{cases} u_c & \in \partial_u J_i^* \left(\frac{c}{\lambda} (u - u_c) \right) \\ v_c & \in \partial_v Q^* \left(\frac{c}{\mu} (v - v_c) \right). \end{cases}$$

Or $Q : v \mapsto \|v\|_{l^1}$ et $l^1(X)$ est un espace de Banach, donc nous savons que $Q^* = \chi_{\bar{B}_\infty}$, où $\bar{B}_\infty = \{u \in X \mid \max_{i,j} |u_{i,j}| \leq 1\}$ (voir [55]).

Ainsi, comme J_i^* et Q^* sont les fonctions indicatrices de deux ensembles convexes fermés K_i et \bar{B}_∞ , nous savons que pour tous $r_1, r_2 > 0$,

$$\begin{cases} u_c = r_1 \left[\frac{c}{\lambda} (u - u_c) + \frac{u_c}{r_1} - \Pi_{K_i} \left(\frac{c}{\lambda} (u - u_c) + \frac{u_c}{r_1} \right) \right] \\ v_c = r_2 \left[\frac{c}{\mu} (v - v_c) + \frac{v_c}{r_2} - \Pi_{\bar{B}_\infty} \left(\frac{c}{\mu} (v - v_c) + \frac{v_c}{r_2} \right) \right]. \end{cases}$$

En posant $r_1 = \frac{\lambda}{c}$ et $r_2 = \frac{\mu}{c}$, il vient

$$\begin{cases} u_c = u - \frac{\lambda}{c} \Pi_{K_i} \left(\frac{c}{\lambda} u \right) \\ v_c = v - \frac{\mu}{c} \Pi_{\bar{B}_\infty} \left(\frac{c}{\mu} v \right). \end{cases}$$

Nous venons de montrer que

$$T_{\frac{1}{c}}(u, v) = \begin{pmatrix} u - \frac{\lambda}{c} \Pi_{K_i} \left(\frac{c}{\lambda} u \right) \\ v - \frac{\mu}{c} \Pi_{\bar{B}_\infty} \left(\frac{c}{\mu} v \right) \end{pmatrix}.$$

D'autre part nous savons d'après le théorème IV.4 et les premières définitions que

$$\varphi'_c(x, \gamma) = c \left(x + \frac{\gamma}{c} - T_{\frac{1}{c}} \left(x + \frac{\gamma}{c} \right) \right).$$

Le calcul est simple et nous donne bien le résultat énoncé. ■

2.2 Résolution explicite de l'étape 2

Pour traiter l'étape 2, nous devons résoudre le problème suivant : Pour $\gamma \in \mathcal{H}$ fixé, on cherche $\bar{x} \in \mathcal{C}$ qui vérifie

$$F_c(\bar{x}, \gamma) = \min_{x \in \mathcal{C}} F_c(x, \gamma), \quad (\text{IV.16})$$

où

$$F_c(x, \gamma) = f(x) + \varphi_c(\Lambda x, \gamma). \quad (\text{IV.17})$$

Définissons le lagrangien associé :

$$\mathcal{L}_\gamma(u, v, \beta) = F_c(u, v, \gamma) + \langle \beta, u_d - u - v \rangle_X \quad (\text{IV.18})$$

Comme l'application $(u, v) \mapsto u_d - u - v$ est linéaire et surjective, on sait qu'il existe un multiplicateur de Lagrange $\bar{\beta} \in \mathcal{H}$ qui vérifie

$$\begin{cases} \mathcal{L}'_\gamma(\bar{u}, \bar{v}, \bar{\beta}) = 0, \\ u_d - \bar{u} - \bar{v} = 0. \end{cases} \quad (\text{IV.19})$$

Posons $g(u, v) = \langle u_d - u - v, \beta \rangle$.

Pour tout $(h, k) \in \mathcal{H}$, on a

$$\mathcal{L}'_\gamma(\bar{u}, \bar{v}, \bar{\beta}) = f'(\bar{u}, \bar{v}) + \varphi'_c((\bar{u}, \bar{v}), \gamma) + g'(u, v, w)$$

Or,

$$\begin{aligned} f'(u, v) &= (\delta u, \delta v), \\ g'(u, v) &= (\beta, \beta). \end{aligned}$$

et la proposition [IV.3](#) nous donne la valeur de $\varphi'_c((\bar{u}, \bar{v}), \gamma)$.

Ainsi, [\(IV.44\)](#) équivaut à

$$\begin{cases} \delta \bar{u} + \Pi_{\lambda K_1} (c\bar{u} + \gamma^1) + \bar{\beta} = 0 \\ \delta \bar{v} + \Pi_{\mu \bar{B}_\infty} (c\bar{v} + \gamma^2) + \bar{\beta} = 0 \\ \bar{v} = u_d - \bar{u}, \end{cases} \quad (\text{IV.20})$$

ce qui équivaut encore à

$$\delta \bar{u} + \Pi_{\lambda K_1} (c\bar{u} + \gamma^1) - \delta(u_d - \bar{u}) - \Pi_{\mu \bar{B}_\infty} (c(u_d - \bar{u}) + \gamma^2) = 0. \quad (\text{IV.21})$$

Nous obtenons donc le problème de point fixe suivant :

$$\Psi(\bar{u}) = \bar{u}, \quad (\text{IV.22})$$

avec

$$\Psi(u) = (1 + \delta)u + \Pi_{\lambda K_1} (cu + \gamma^1) - \delta(u_d - u) - \Pi_{\mu \bar{B}_\infty} (c(u_d - u) + \gamma^2). \quad (\text{IV.23})$$

Afin de résoudre ce problème à l'aide d'un algorithme de point fixe, nous relaxons la fonction Ψ . On introduit ainsi la fonction Ψ_α définie par

$$\Psi_\alpha(u) = u + \alpha (\Psi(u) - u). \quad (\text{IV.24})$$

Il est alors aisé de vérifier (démonstration analogue à celle du théorème [II.11](#)) que la fonction Ψ_α est strictement contractante pour $\alpha \in]-1; 0[$, ce qui implique que l'algorithme de point fixe associé converge. De plus Ψ_α et Ψ possèdent les mêmes points fixes.

2.3 Algorithme

On implémente alors l'algorithme suivant :

- **Etape 1** : On choisit une valeur de départ $\gamma_1 \in \mathcal{H}$, un réel $c > 0$ et on pose $k = 1$.
- **Etape 2** : Soit $\gamma_k \in \mathcal{H}$ fixé. On cherche $x_k = (u_k, v_k) \in \mathcal{H}$ qui vérifie

$$x_{k+1} = \Psi_\alpha(x_k)$$

- **Etape 3** : On pose

$$\gamma_{k+1} = \begin{pmatrix} \Pi_{\lambda K_i}(cu_k + \gamma_k^1) \\ \Pi_{\mu \bar{\mathcal{B}}_\infty}(cv_k + \gamma_k^2) \end{pmatrix}.$$

- **Etape 4** : Si le critère d'arrêt n'est pas satisfait, on pose $k := k + 1$ et on retourne à 2.

2.4 Tests numériques

Nous ne proposons ici qu'un seul test.



FIG. IV.1 – Modèle $TV-L^1$, $\lambda = 30$, $\mu = 15$. La composante u est à gauche, la composante v à droite.



FIG. IV.2 – Modèle TV^2-L^1 , $\lambda = 30$, $\mu = 40$. La composante u est à gauche, la composante v à droite.

3 Une autre façon de résoudre les modèles $TV-L^1$ et TV^2-L^1

Afin de résoudre le modèle $TV-L^1$, J-F. Aujol *et al.* introduisent dans [12] le modèle discret suivant :

$$\min_{u \in X} \frac{1}{2} \|u_d - u - v\|_X^2 + \mu \|v\|_{l^1} + \lambda J_1(u). \quad (\text{IV.25})$$

En d'autres termes, la condition d'attache aux données est cette fois-ci incluse directement dans la fonctionnelle à minimiser. La différence avec le modèle décrit dans la partie précédente est que $u + v$ n'est plus exactement égal au terme de départ, c'est à dire que le terme $w := u_d - u - v$ contient de l'information. En utilisant encore une fois les méthodes de Lagrangien augmenté, nous allons proposer un algorithme afin de résoudre le modèle suivant :

$$\min_{u \in X} \frac{1}{2} \|u_d - u - v\|_X^2 + \mu \|v\|_{l^1} + \lambda J_i(u), \quad (\text{IV.26})$$

où $\mu, \lambda > 0$ et $i \in \{1, 2\}$. Le terme de régularisation permettant de se situer dans le cadre de la théorie et de démontrer la convergence de l'algorithme vers une solution, on s'intéresse au problème

$$\min_{u \in X} \frac{1}{2} \|u_d - u - v\|_X^2 + \mu \|v\|_{l^1} + \lambda J_i(u) + \frac{\delta}{2} \left(\|u\|_X^2 + \|v\|_X^2 \right), \quad (\text{IV.27})$$

où $\mu, \lambda > 0$ et $i \in \{1, 2\}$. En appliquant rigoureusement la même étape que dans la partie précédente, on obtient l'algorithme suivant :

Algorithme

- **Etape 1** : On choisit une valeur de départ $\gamma_1 \in \mathcal{H}$, un réel $c > 0$ et on pose $k = 1$.
- **Etape 2** : Soit $\gamma_k \in \mathcal{H}$ fixé. On cherche $x_k = (u_k, v_k) \in \mathcal{H}$ qui vérifie

$$x_{k+1} = \Psi_\alpha(x_k)$$

- **Etape 3** : On pose

$$\gamma_{k+1} = \begin{pmatrix} \Pi_{\lambda K_i} (cu_k + \gamma_k^1) \\ \Pi_{\mu \bar{B}_\infty} (cv_k + \gamma_k^2) \end{pmatrix}.$$

- **Etape 4** : Si le critère d'arrêt n'est pas satisfait, on pose $k := k + 1$ et on retourne à 2.
où

$$\Psi_\alpha(u, v) = (u, v) + \alpha (\Psi(u, v) - (u, v)), \quad (\text{IV.28})$$

avec

$$\Psi(u, v) = \begin{pmatrix} (2 + \delta)u + v - u_d + \Pi_{\lambda K_i} (cu_k + \gamma_k^1) \\ u + (2 + \delta)v - u_d + \Pi_{\mu \bar{B}_\infty} (cv_k + \gamma_k^2) \end{pmatrix}. \quad (\text{IV.29})$$

4 Modèle d'extraction de texture mixte

4.1 Introduction

On s'intéresse désormais au modèle suivant :

$$\inf_{(u,v) \in BV(\Omega) \times BV^2(\Omega)} \mathcal{F}(u, v), \quad (\text{IV.30})$$

où

$$\mathcal{F}(u, v) = \nu \|u_d - u - v\|_{L^1(\Omega)} + \lambda TV(u, \Omega) + \mu TV^2(v, \Omega) + \xi \|v\|_{W^{1,1}(\Omega)}. \quad (\text{IV.31})$$

Nous avons le résultat d'existence suivant :

Théorème IV.13. *Il existe un unique couple $(u, v) \in BV(\Omega) \times BV^2(\Omega)$ réalisant l'infimum de \mathcal{F} .*

Démonstration : La démonstration est identique à la démonstration du théorème II.9. ■

4.2 Discrétisation

Comme pour le modèle mixte (voir page 42) et contrairement aux modèles ROF2 et TV^2-L^1 , un terme de "régularisation" est encore indispensable au modèle discret pour obtenir l'existence d'une solution. Nous allons donc conserver le terme $\xi \|v\|_{W^{1,1}(\Omega)}$ dans le cas discret et lors de la résolution numérique. Ainsi, comme pour le modèle mixte, nous remplaçons le terme $\|v\|_{W^{1,1}(\Omega)}$ par $\|v\|_{H^1(\Omega)}^2$, ce dernier terme étant différentiable, et donc très facile à implémenter numériquement.

Nous nous intéressons donc finalement au problème suivant :

$$\inf_{(u,v) \in BV(\Omega) \times BV^2(\Omega)} \mathcal{F}(u, v), \quad (\text{IV.32})$$

où

$$\mathcal{F}(u, v) = \nu \|u_d - u - v\|_{L^1(\Omega)} + \lambda TV(u, \Omega) + \mu TV^2(v, \Omega) + \frac{\xi}{2} \|v\|_{H^1(\Omega)}^2. \quad (\text{IV.33})$$

Remarque IV.5. *Lors des tests numériques, $\xi > 0$ sera pris extrêmement petit.*

En utilisant les résultats de discrétisation obtenus dans le chapitre II, on obtient le problème discrétisé suivant :

$$\inf_{(u,v) \in X^2} F(u, v), \quad (\text{IV.34})$$

où

$$F(u, v) = \nu \|u_d - u - v\|_{l^1} + \lambda J_1(u) + \mu J_2(v) + \frac{\xi}{2} \left(\|v\|_X^2 + \|\nabla v\|_{X^2}^2 \right). \quad (\text{IV.35})$$

Là encore, nous avons un résultat d'existence et d'unicité.

Théorème IV.14. *Il existe un unique couple $(u, v) \in X^2$ réalisant l'infimum de F .*

Démonstration : La fonction à minimiser est coercive et strictement convexe. ■

Le terme en ξ est important pour montrer l'existence dans le cadre continu. Toutefois, dans le cas discret, le terme $\|\nabla v\|_{X^2}^2$ n'est plus indispensable. Nous étudierons donc le problème discret suivant :

$$\inf_{(u,v) \in X^2} F(u, v), \quad (\text{IV.36})$$

où

$$F(u, v) = \nu \|u_d - u - v\|_{L^1} + \lambda J_1(u) + \mu J_2(v) + \frac{\xi}{2} \|v\|_X^2. \quad (\text{IV.37})$$

Encore une fois, nous rappelons que l'utilisation de la méthode de Lagrangien augmenté pour l'implémentation numérique impose à la fonctionnelle F de posséder une partie Fréchet-différentiable. C'est le cas ici, mais le terme différentiable ne vérifie pas la condition (H2). Ainsi, comme pour les modèles $TV - L^1$ et $TV^2 - L^1$, nous ajoutons à F une partie différentiable, de sorte que (H2) soit vérifiée. On obtient alors le problème discrétisé suivant :

$$\inf_{(u,v) \in \mathcal{C}} F_\delta(u, v), \quad (\mathcal{P}_\delta)$$

où

$$F_\delta(u, v, w) = \nu \|w\|_{L^1} + \lambda J_1(u) + \mu J_2(v) + \frac{\delta}{2} \|(u, w)\|_X^2 + \frac{\xi}{2} \|v\|_X^2, \quad (\text{IV.38})$$

et

$$\mathcal{C} = \{(u, v, w) \in X^3 \mid u_d = u + v + w\}. \quad (\text{IV.39})$$

Afin d'alléger les notations, nous posons $\delta = \xi$. De même, et par analogie avec la partie 1.4, nous notons dans la suite :

$$\mathcal{X} = \mathcal{H} = X \times X \times X,$$

$$\mathcal{Q}(w) := \|w\|_{L^1},$$

et

$$F_\delta(u, v, w) = f(u, v, w) + \varphi(u, v, w),$$

où

$$\begin{cases} f : (u, v, w) \mapsto \frac{\delta}{2} \left(\|u\|_X^2 + \|w\|_X^2 \right) + \frac{\xi}{2} \left(\|v\|_X^2 + \|\nabla v\|_{X^2}^2 \right), \\ \varphi : (u, v, w) \mapsto \lambda J_1(u) + \mu J_2(v) + \nu \mathcal{Q}(w), \\ \Lambda = Id. \end{cases} \quad (\text{IV.40})$$

Encore une fois, nous voyons que f et φ sont s.c.i., convexes, que f est continûment différentiable et vérifie (H2), et que φ est propre. Nous pouvons donc utiliser les résultats de la première section de ce chapitre.

4.3 Résolution explicite de l'étape 3

Une démonstration analogue à celle de la proposition IV.3 nous donne le résultat suivant :

Proposition IV.4. *Soient $c > 0$, $\gamma = (\gamma^1, \gamma^2, \gamma^3) \in X^3$ et $x = (u, v, w) \in X^3$. Alors*

$$\varphi'_c(x, \gamma) = \begin{pmatrix} \Pi_{\lambda K_1}(cu + \gamma^1) \\ \Pi_{\mu K_2}(cv + \gamma^2) \\ \Pi_{\nu \bar{B}_\infty}(cw + \gamma^3) \end{pmatrix},$$

avec K_1 et K_2 définis en (II.10) et (II.20), et où

$$\bar{B}_\infty = \left\{ u \in X \mid \max_{i,j} |u_{i,j}| \leq 1 \right\}.$$

4.4 Résolution explicite de l'étape 2

Pour traiter l'étape 2, nous devons résoudre le problème suivant : Pour $\gamma \in \mathcal{H}$ fixé, on cherche $\bar{x} \in \mathcal{C}$ qui vérifie

$$F_c(\bar{x}, \gamma) = \min_{x \in \mathcal{C}} F_c(x, \gamma), \quad (\text{IV.41})$$

où

$$F_c(x, \gamma) = f(x) + \varphi_c(\Lambda x, \gamma). \quad (\text{IV.42})$$

Définissons le lagrangien associé :

$$\mathcal{L}_\gamma(u, v, w, \beta) = F_c(u, v, w, \gamma) + \langle \beta, u_d - u - v - w \rangle_X \quad (\text{IV.43})$$

Comme l'application $(u, v, w) \mapsto u_d - u - v - w$ est linéaire et surjective, on sait qu'il existe un multiplicateur de Lagrange $\bar{\beta} \in \mathcal{H}$ qui vérifie

$$\begin{cases} \mathcal{L}'_\gamma(\bar{u}, \bar{v}, \bar{w}, \bar{\beta}) = 0, \\ u_d - \bar{u} - \bar{v} - \bar{w} = 0. \end{cases} \quad (\text{IV.44})$$

Posons $g(u, v, w) = \langle u_d - u - v - w, \beta \rangle$.

On a

$$\mathcal{L}'_\gamma(\bar{u}, \bar{v}, \bar{w}, \bar{\beta}) = f'(\bar{u}, \bar{v}, \bar{w}) + \varphi'_c((\bar{u}, \bar{v}, \bar{w}), \gamma) + g'(\bar{u}, \bar{v}, \bar{w})$$

Or,

$$\begin{aligned} f'(u, v, w) &= (\delta u, \xi(v - \operatorname{div}(\nabla v)), \delta w), \\ g'(u, v, w) &= (\beta, \beta, \beta). \end{aligned}$$

et la proposition IV.3 nous donne la valeur de $\varphi'_c((\bar{u}, \bar{v}), \gamma)$.

Ainsi, (IV.44) équivaut à

$$\begin{cases} \delta \bar{u} + \Pi_{\lambda K_1}(c\bar{u} + \gamma^1) + \bar{\beta} = 0 \\ \xi \bar{v} - \xi \operatorname{div}(\nabla \bar{v}) + \Pi_{\mu K_2}(c\bar{v} + \gamma^2) + \bar{\beta} = 0 \\ \delta \bar{w} + \Pi_{\nu \bar{B}_\infty}(c\bar{w} + \gamma^3) + \bar{\beta} = 0 \\ u_d - \bar{u} - \bar{v} - \bar{w} = 0, \end{cases} \quad (\text{IV.45})$$

ce qui équivaut encore à

$$\begin{cases} \delta\bar{u} - \delta\bar{w} + \Pi_{\lambda K_1}(c\bar{u} + \gamma^1) - \Pi_{\nu\bar{B}_\infty}(c\bar{w} + \gamma^3) = 0 \\ \xi\bar{v} - \xi\operatorname{div}(\nabla\bar{v}) - \delta\bar{w} + \Pi_{\mu K_2}(c\bar{v} + \gamma^2) - \Pi_{\nu\bar{B}_\infty}(c\bar{w} + \gamma^3) = 0 \\ u_d - \bar{u} - \bar{v} - \bar{w} = 0. \end{cases} \quad (\text{IV.46})$$

Nous avons donc à résoudre le problème de point fixe suivant :

$$(\bar{u}, \bar{v}, \bar{w}) = \Psi(\bar{u}, \bar{v}, \bar{w}), \quad (\text{IV.47})$$

où

$$\Psi(\bar{u}, \bar{v}, \bar{w}) = \begin{pmatrix} (1 + \delta)\bar{u} - \delta\bar{w} + \Pi_{\lambda K_1}(c\bar{u} + \gamma^1) - \Pi_{\nu\bar{B}_\infty}(c\bar{w} + \gamma^3) \\ (1 + \xi)\bar{v} - \xi\operatorname{div}(\nabla\bar{v}) - \delta\bar{w} + \Pi_{\mu K_2}(c\bar{v} + \gamma^2) - \Pi_{\nu\bar{B}_\infty}(c\bar{w} + \gamma^3) \\ u_d - \bar{u} - \bar{v} \end{pmatrix}. \quad (\text{IV.48})$$

Pour résoudre ce problème de point fixe, nous nous intéressons au problème relaxé

$$(\bar{u}, \bar{v}, \bar{w}) = \Psi_\alpha(\bar{u}, \bar{v}, \bar{w}), \quad (\text{IV.49})$$

où

$$\Psi_\alpha(u, v, w) = (u, v, w) + \alpha(\Psi(u, v, w) - (u, v, w)). \quad (\text{IV.50})$$

4.5 Algorithme

On implémente alors l'algorithme suivant :

- **Etape 1** : On choisit une valeur de départ $\gamma_1 \in \mathcal{H}$, un réel $c > 0$ et on pose $k = 1$.
- **Etape 2** : Soit $\gamma_k \in \mathcal{H}$ fixé. On cherche $x_k = (u_k, v_k, w_k) \in \mathcal{H}$ qui vérifie

$$x_{k+1} = \Psi_\alpha(x_k)$$

- **Etape 3** : On pose

$$\gamma_{k+1} = \begin{pmatrix} \Pi_{\lambda K_1}(cu_k + \gamma_k^1) \\ \Pi_{\mu K_2}(cv_k + \gamma_k^2) \\ \Pi_{\nu\bar{B}_\infty}(cw_k + \gamma_k^3) \end{pmatrix}.$$

- **Etape 4** : Si le critère d'arrêt n'est pas satisfait, on pose $k := k + 1$ et on retourne à 2.

4.6 Exemple



FIG. IV.3 – Modèle TV^2-L^1 , $\lambda = 30$, $\mu = 40$. La composante u est à gauche, la composante v à droite.

On observe sur cet exemple que si la composante “cartoon” contient trop de texture, la composante “texture” elle semble très convaincante. En effet, seule l’information très oscillante est présente : les détails du visage sont très peu visibles, de même pour les contours (bras, pieds de table, livres en arrière plan). Ce qui est intéressant dans cette décomposition est que nous perdons quasiment tous les détails “perturbateurs” alors que nous préservons toute la texture. Nous n’irons pas plus loin dans l’analyse de ce modèle car très peu de tests numériques ont été effectués à l’heure actuelle. L’étude précise de ce modèle fera l’objet d’un futur travail.

Chapitre V

Tests numériques

Les tests numériques ont été effectués sur différentes machines, ne nous permettant pas retranscrire précisément la durée de calcul. Les différents algorithmes ont été implémentés avec le logiciel MATLAB ©. Nous donnons également la définition du rapport signal sur bruit (SNR) que nous utiliserons pour les différents tests :

$$SNR(v) = 20 \log_{10} \left(\frac{\|u\|_{L^2}}{\|u - v\|_{L^2}} \right),$$

où u est l'image de départ non bruitée et v l'image restaurée.

1 Algorithme plus rapide pour implémenter Π_{K_1} et Π_{K_2}

Lors de la résolution numérique du modèle mixte (cf page 42), nous avons été confrontés à un problème de convergence de l'algorithme proposé page 45. En effet, alors que la théorie prouve que le modèle discret admet une unique solution, les tests numériques nous ont permis d'obtenir des solutions bien différentes, dépendant notamment des conditions initiales de l'algorithme page 45. Ceci nous a fait remettre en cause la vitesse de convergence des algorithmes de type Chambolle. Ceux-ci concernent la résolution des deux problèmes de minimisation suivants, que le calcul des projections Π_{K_1} et Π_{K_2} nécessitent :

$$\min\{\|\lambda \operatorname{div} p - f\|_X^2 \mid \|p_{i,j}\|_{\mathbb{R}^2} \leq 1 \forall i, j = 1, \dots, N\}, \quad (\text{V.1})$$

et

$$\min\{\|\lambda H^* p - f\|_X^2 \mid \|p_{i,j}\|_{\mathbb{R}^4} \leq 1 \forall i, j = 1, \dots, N\}, \quad (\text{V.2})$$

où $X = \mathbb{R}^{N \times N}$.

Nous proposons ici l'utilisation d'un algorithme différent introduit par Nesterov dans [43] afin de calculer la solution d'un problème du type

$$\min_{x \in X} f(x) + \Psi(x), \quad (\text{V.3})$$

où X est un espace vectoriel de dimension finie, f une fonction L -Lipschitz différentiable et Ψ une fonction convexe. Nous n'utiliserons pas directement l'algorithme proposé dans

[43] mais celui proposé par Weiss *et al.* dans [57] dans le cas où la fonction Ψ est la fonction indicatrice d'un ensemble convexe K , ce qui est le cas de (V.1) et (V.2).

Voici l'algorithme :

Algorithme

Entrée : nombre d'itérations N , $x^0 \in K$.

Sortie : x^N (une estimée de la solution x^*)

$k = -1$

$g^{(-1)} = 0$

$x^0 \in K$

$L =$ la constante de Lipschitz de ∇f

pour k allant de 0 à $N - 1$

$\eta^k = \nabla f(x^k)$
$y^k = \Pi_K \left(x^k - \frac{\eta^k}{L} \right)$
$g^k = g^{k-1} + \frac{k+1}{2} \eta^k$
$\nu^k = \Pi_K \left(x^0 - \frac{g^k}{L} \right)$
$x^{k+1} = \frac{2}{k+3} \nu^k + \frac{k+1}{k+3} y^k$

fin

Afin de comparer les vitesses de convergence de cet algorithme et de ceux de type Chambolle introduits dans le chapitre 2, nous proposons un test numérique utilisant le modèle ROF. Nous effectuons un test sur l'image bruitée suivante :



FIG. V.1 – Image bruitée par un bruit blanc gaussien ($\sigma = 25$).

Nous utilisons, pour débruiter l'image [V.1](#), le modèle ROF une fois avec l'algorithme de Chambolle, une fois avec l'algorithme décrit ci-dessus sur un très grand nombre d'itérations (20000 dans les deux cas). Nous obtenons ainsi deux résultats. Afin de déterminer celui qui est le plus proche de la solution réelle, nous appliquons aux deux résultats la fonctionnelle de départ à minimiser, et décrétons que la meilleure approximation est celle qui renvoie la valeur la plus petite. Nous regardons ensuite à quelle vitesse les deux algorithmes permettent de s'approcher de la solution. On obtient la figure [V.2](#).

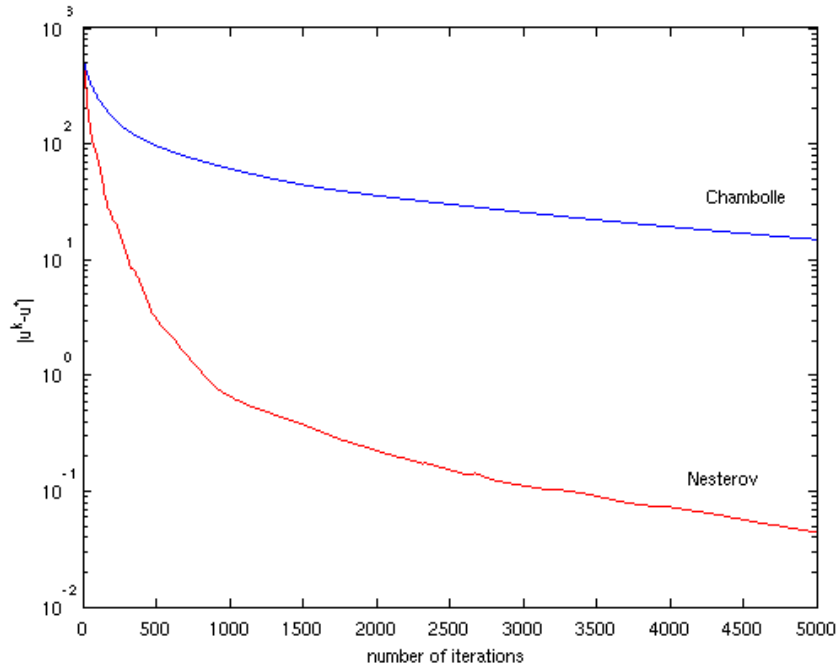


FIG. V.2 – Norme de la différence entre la composante cartoon et la solution u^* pour les algorithmes de Chambolle et Nesterov, en fonction du nombre d'itérations.

Les résultats sont identiques pour l'implémentation numérique de Π_{K_2} . On observe ainsi que l'on obtient au bout d'environ 400 itérations avec l'algorithme de Nesterov un résultat proche de celui obtenu avec l'algorithme de Chambolle au bout de 5000 itérations, avec un temps de calcul entre itérations à peine plus grand (avec l'ordinateur que nous avons utilisé, le temps de calcul passe de plus d'une heure à une dizaine de minute pour l'image [V.1](#)). De plus, on observe que, malgré ce nombre très important d'itérations, les algorithmes de type Chambolle ne permettent pas d'obtenir une bonne approximation de la solution réelle. Ceci explique les incohérences numériques mentionnées précédemment qui disparaissent bien lorsque l'algorithme de Nesterov est utilisé. En effet, c'est comme cela que nous avons pu obtenir les résultats numériques de la figure [II.5](#).

Nous utiliserons donc, dans tous les tests numériques de ce chapitre, l'algorithme de Nesterov afin de calculer Π_{K_1} et Π_{K_2} .

2 Le modèle ROF2

$$\inf_{v \in X} \frac{\|u_d - v\|_X^2}{2\lambda} + J_2(v). \quad (\mathcal{P}_d^2)$$

2.1 Influence du nombre d'itérations et critère d'arrêt

Nous voulions, dans un premier temps, utiliser le SNR comme critère d'arrêt, par exemple en stoppant les calculs au moment où sa valeur diminuait. Ceci n'est pas une méthode fiable. Prenons l'exemple de l'image texturée de la figure V.3. Alors que l'image de droite est clairement moins bruitée que l'image de gauche, son SNR est beaucoup plus faible. Ceci s'explique par le fait que la texture est lissée au même titre que le bruit lors de la restauration. Ainsi, une image bruitée non restaurée sera plus "proche" de l'image idéale qu'une image sans bruit mais sans texture.

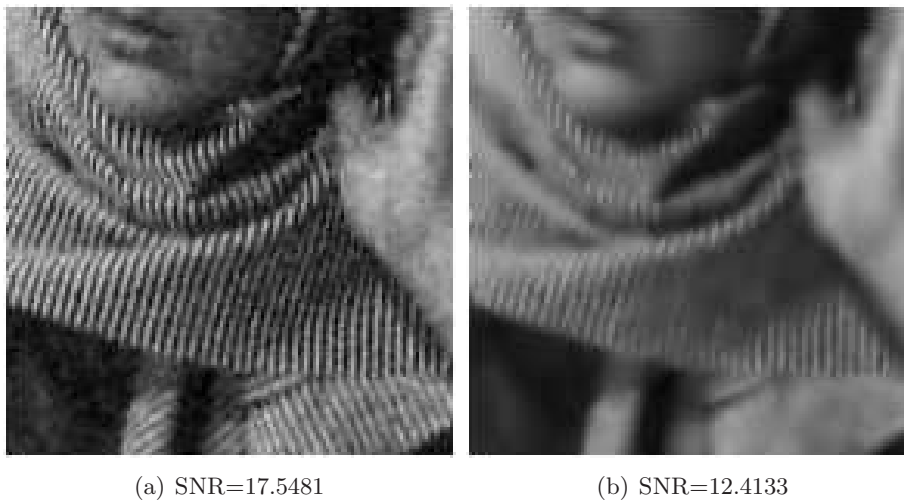


FIG. V.3 – (a) est l'image bruitée par un bruit blanc gaussien ($\sigma = 25$). (b) est l'image ayant subi un débruitage par le modèle mixte ($\lambda = 30$, $\mu = 20$). Le SNR de (a) est plus élevé même si l'image (b) ne contient plus de bruit. Ceci vient du fait qu'une grande partie de la texture de (b) a également disparue.

Nous allons donc utiliser comme critère d'arrêt un nombre maximal d'itération arbitrairement défini. La figure V.4 montre que ce choix n'est pas si mauvais. En effet, on observe que la convergence est plutôt rapide, et surtout que l'évolution du résultat observé n'est plus visible à l'oeil à partir de l'itération 300. Ceci pouvant bien évidemment varier en fonction des différentes images, nous ferons systématiquement tourner les algorithmes de Nesterov permettant de calculer Π_{K_1} et Π_{K_2} exactement 500 itérations afin de garder une certaine marge.

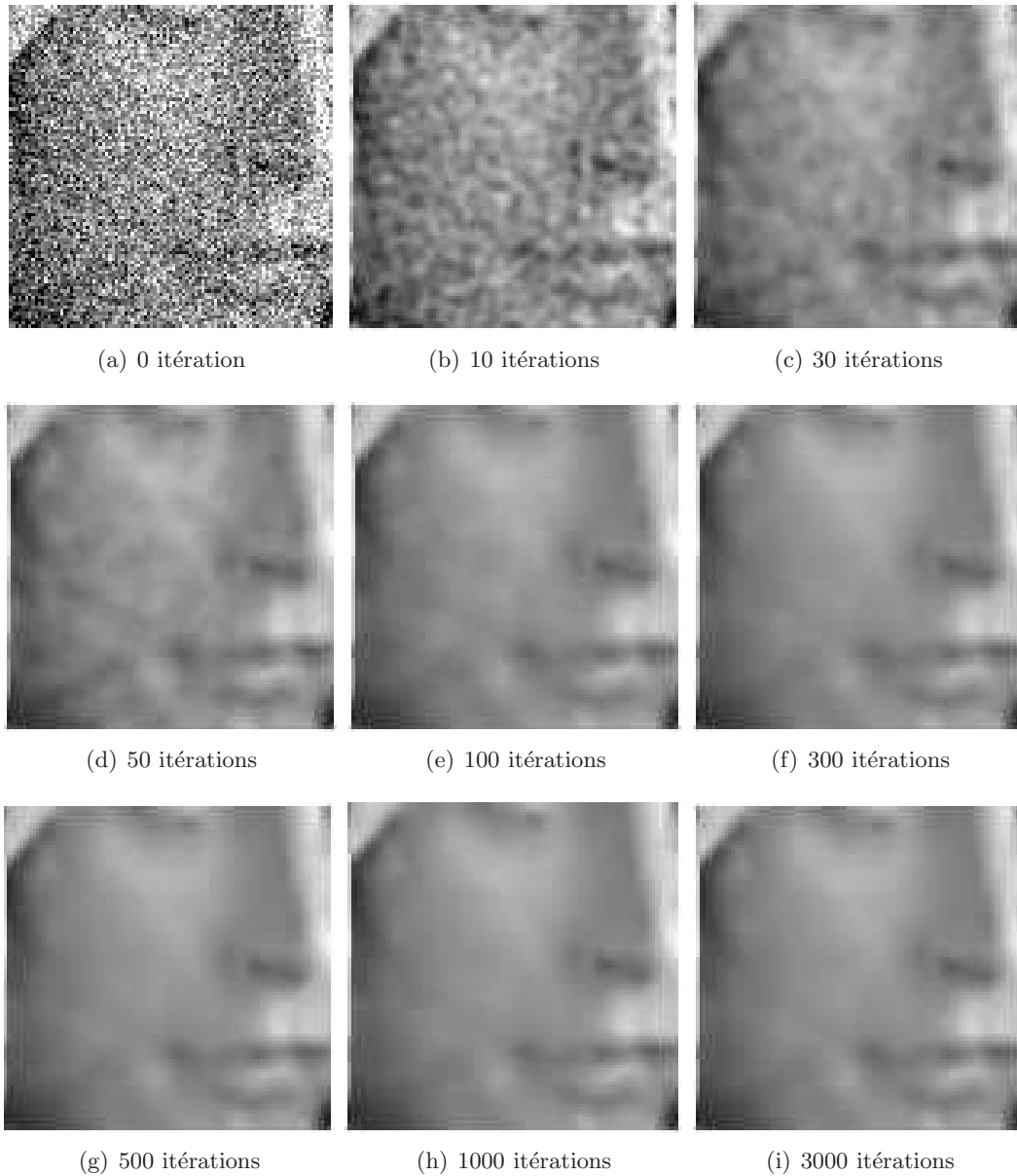


FIG. V.4 – Influence du nombre d'itérations sur une image fortement bruitée par un bruit blanc gaussien ($\sigma = 50$).

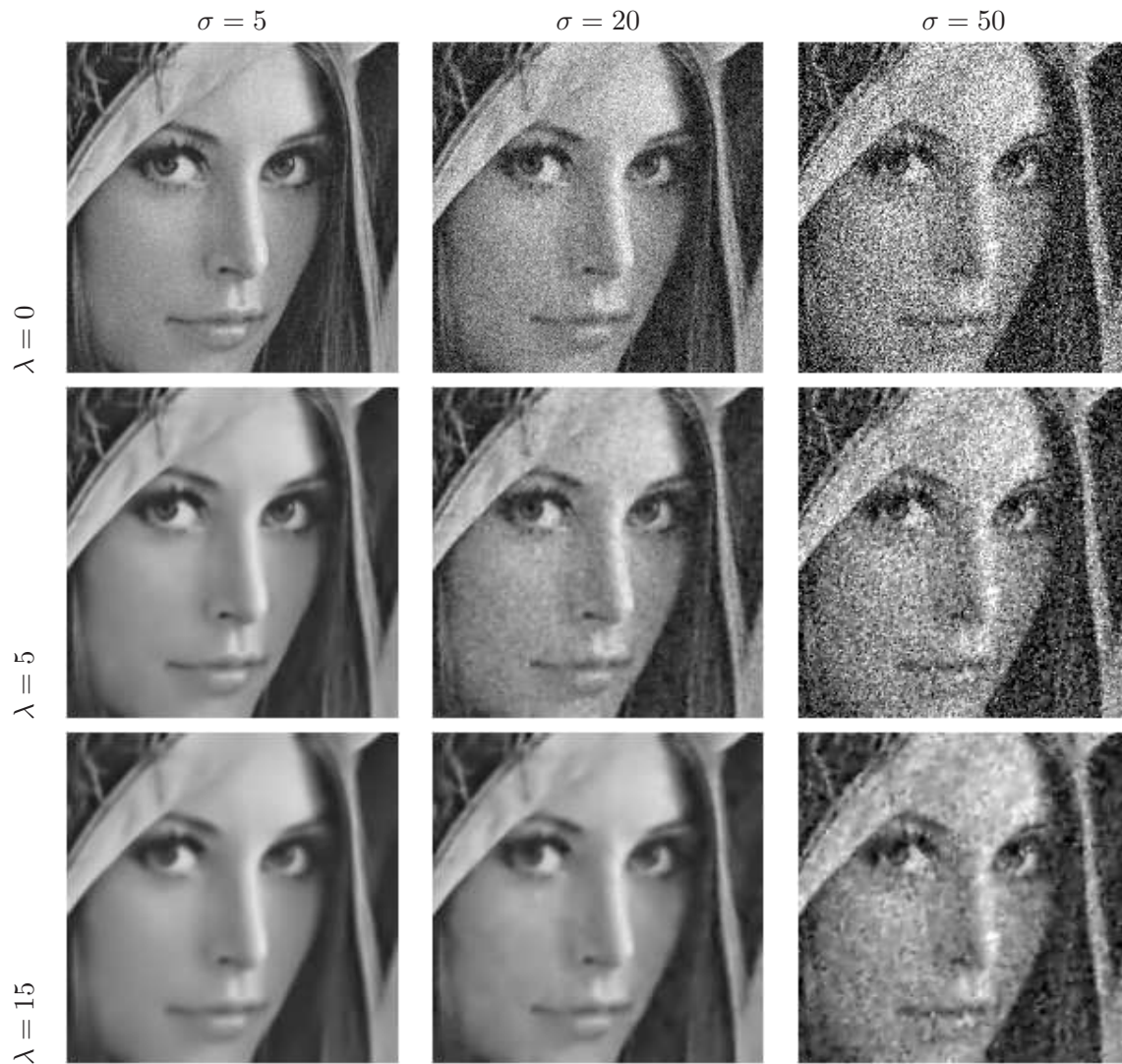
2.2 Influence de λ 

FIG. V.5 – Influence de λ (A-1). On observe que, quelque soit l'intensité du bruit, celui-ci diminue au fur et à mesure que la valeur de λ augmente.



FIG. V.6 – Influence de λ (A-2). Une fois le bruit éliminé, l'image devient de plus en plus floue et les détails (tels que ceux du visage) commencent à disparaître.

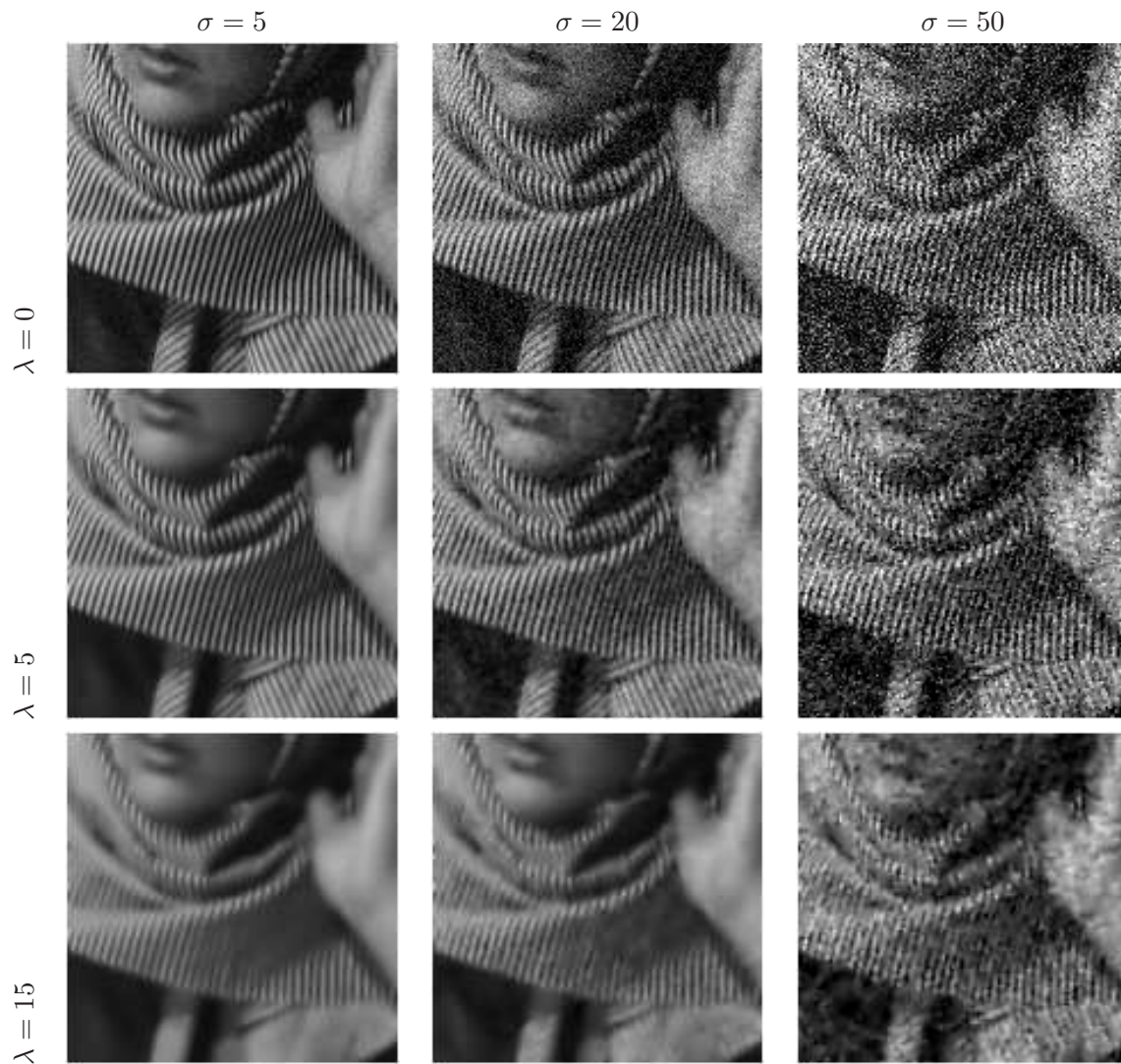


FIG. V.7 – Influence de λ (B-1). Les observations sont identiques à celles de la figure V.5. Nous pouvons de plus remarquer que, en augmentant, λ fait peu à peu disparaître la texture.



FIG. V.8 – Influence de λ (B-2). Les observations sont identiques à celles de la figure V.6. On observe de plus que la texture disparaît totalement à partir de $\lambda = 50$.

	$\sigma = 5$	$\sigma = 20$	$\sigma = 50$
$\lambda = 0$	28.3505	16.3112	8.3177
$\lambda = 5$	28.2685	23.5266	11.9536
$\lambda = 15$	24.5297	24.0058	18.0281
$\lambda = 25$	22.8855	22.6980	20.2467
$\lambda = 50$	20.8984	20.8394	20.0870
$\lambda = 100$	19.1367	19.0836	18.8376

FIG. V.9 – SNR des images des figures V.5 et V.6. On remarque que les SNR les plus élevés correspondent bien aux résultats semblant visuellement être les meilleurs.

	$\sigma = 5$	$\sigma = 20$	$\sigma = 50$
$\lambda = 0$	26.0215	14.1034	6.1484
$\lambda = 5$	19.8996	17.5491	9.3543
$\lambda = 15$	13.1383	13.2460	12.1783
$\lambda = 25$	11.4814	11.5236	11.5104
$\lambda = 50$	11.1039	11.0807	10.9680
$\lambda = 100$	10.7688	10.7502	10.7200

FIG. V.10 – SNR des images des figures V.7 et V.8. Cette fois-ci, nous observons que les SNR les plus élevés correspondent à des images contenant encore une part importante de bruit, et qui ne sont visuellement pas les meilleures. Ceci s'explique par le fait que l'image de départ est très texturée, et que l'échelle de cette texture est relativement petite. Ceci entraîne que celle-ci disparaît, avec la restauration, en même temps que le bruit, expliquant les valeurs plus basses du SNR avant même que l'image ne semble devenir vraiment floue.

2.3 Tests sur des signaux 1D

Nous proposons ici d'effectuer quelques tests numériques sur des signaux artificiels en de dimension 1. Nous comparerons les résultats obtenus à l'aide du modèle ROF2 avec ceux obtenus par le modèle ROF, mettant en évidence la disparition de l'effet *staircasing*.

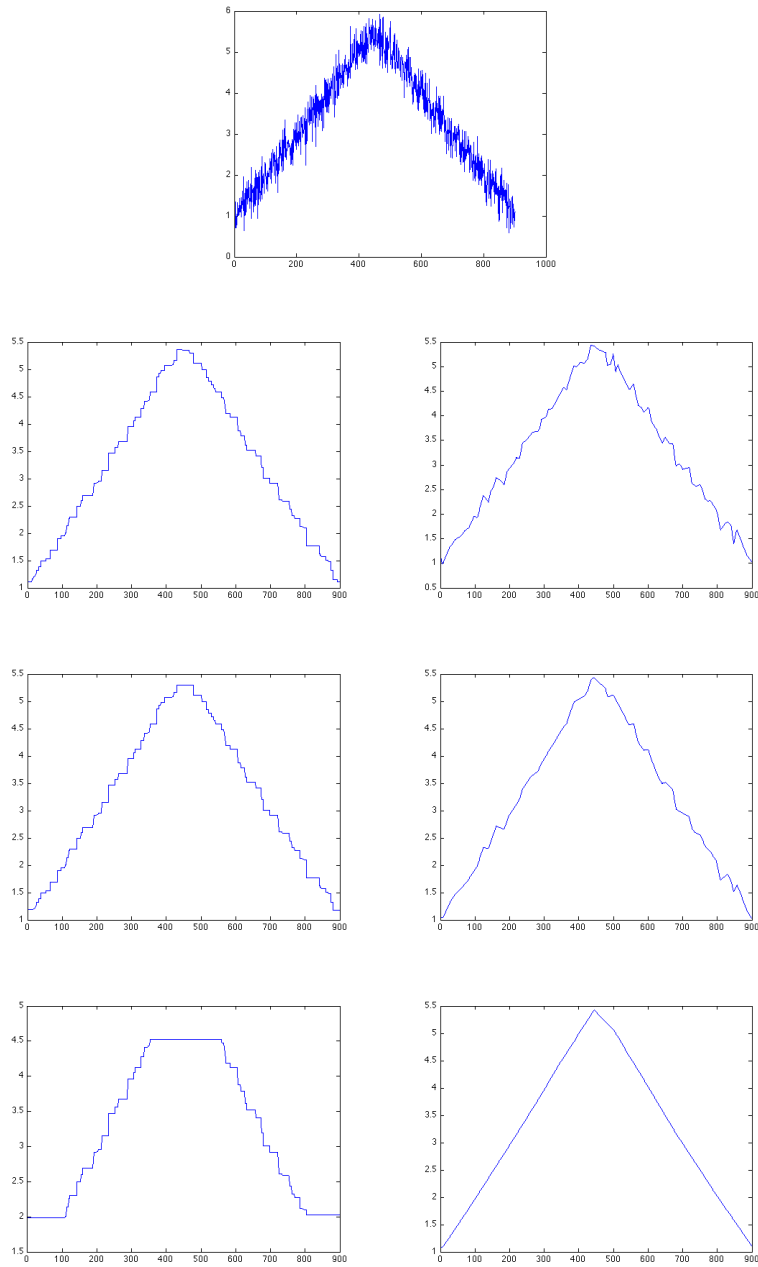


FIG. V.11 – Test sur un signal bruité. Nous comparons ici les effets des modèles ROF (colonne de gauche, de haut en bas : $\lambda = 1$, $\lambda = 2$, $\lambda = 50$) et ROF2 (colonne de droite, de haut en bas : $\lambda = 1$, $\lambda = 2$, $\lambda = 50$). On observe dans les deux cas une diminution du bruit, avec l'apparition d'un effet *staircasing* dès les premières itérations pour ROF

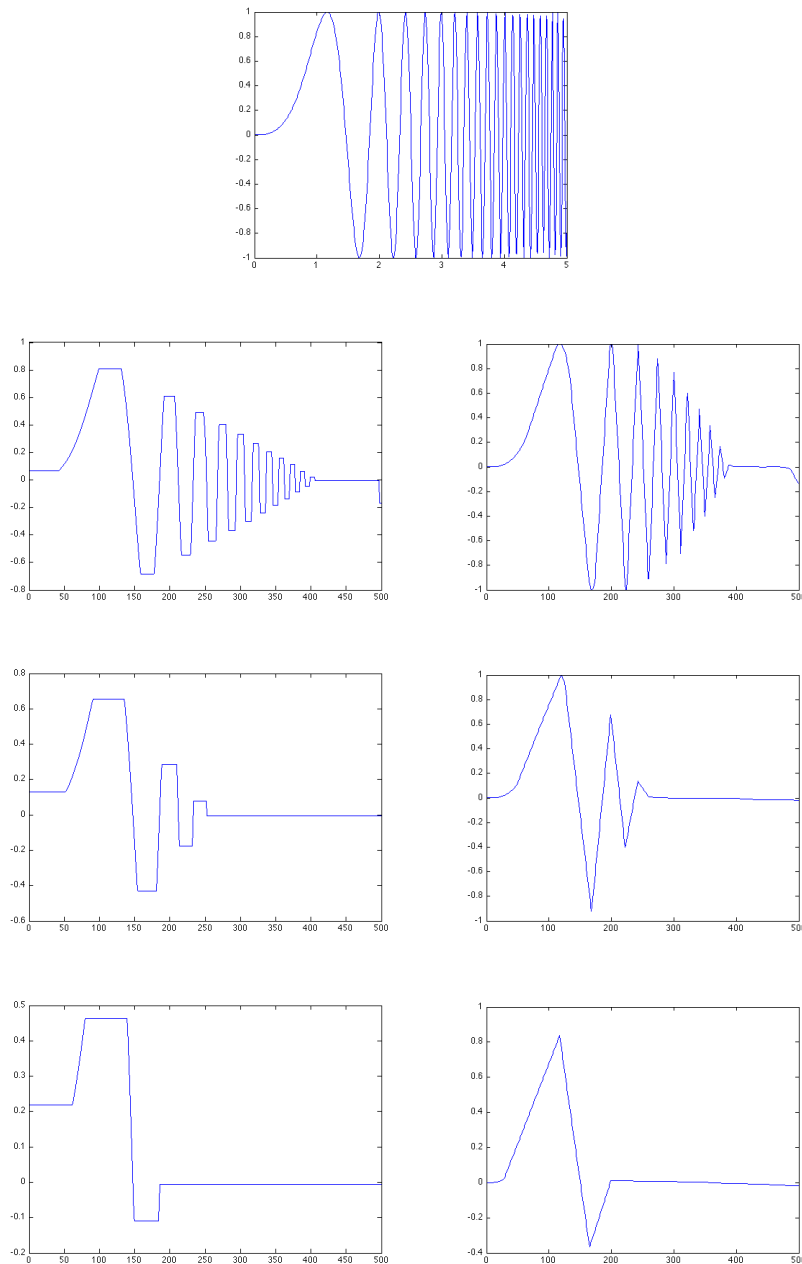


FIG. V.12 – Test sur un signal artificiel ($\sin(x^3)$). L’utilisation du modèle ROF (colonne de gauche, de haut en bas : $\lambda = 2, \lambda = 5, \lambda = 10$) entraîne une diminution de l’amplitude du signal lorsque λ augmente. Un effet de *staircasing* est présent à chaque étape, et augmente lorsque λ grandit. Le débruitage ROF2 (colonne de droite, de haut en bas : $\lambda = 5, \lambda = 30, \lambda = 50$) entraîne également une diminution de l’amplitude, mais sans effet *staircasing*. Les valeurs du paramètre λ sont choisies de telle sorte que les deux signaux présentent “à l’œil” des niveaux de débruitage équivalents.

3 Le modèle mixte

$$\inf_{(u,v) \in X^2} \frac{1}{2} \|u_d - u - v\|_X^2 + \lambda J_1(u) + \mu J_2(v) + \frac{\delta}{2} (\|v\|_X^2 + \|\nabla v\|_X^2). \quad (\mathcal{P}_d)$$

3.1 Influence de λ

Dans les tests suivants, nous utilisons des images bruitées avec un bruit blanc gaussien vérifiant $\sigma = 20$. Nous fixons de plus $\mu = 30$.

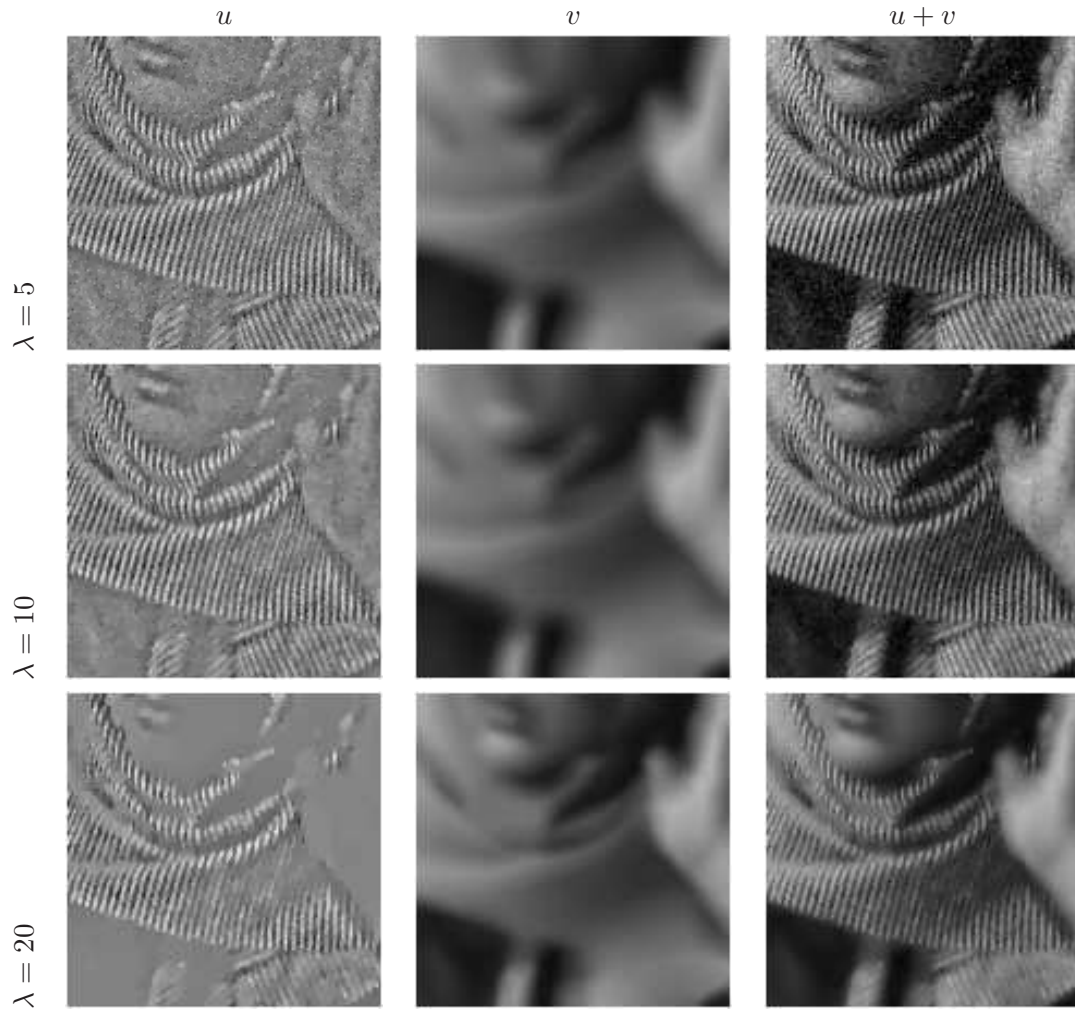


FIG. V.13 – Influence de λ (A-1). $\sigma = 20$ et $\mu = 30$. On observe que plus λ augmente par rapport à μ , plus la composante u perd de l'information et plus v en récupère. On observe pour $\lambda = 20$ que l'image restaurée ne possède plus de bruit alors que la texture est très bien conservée.

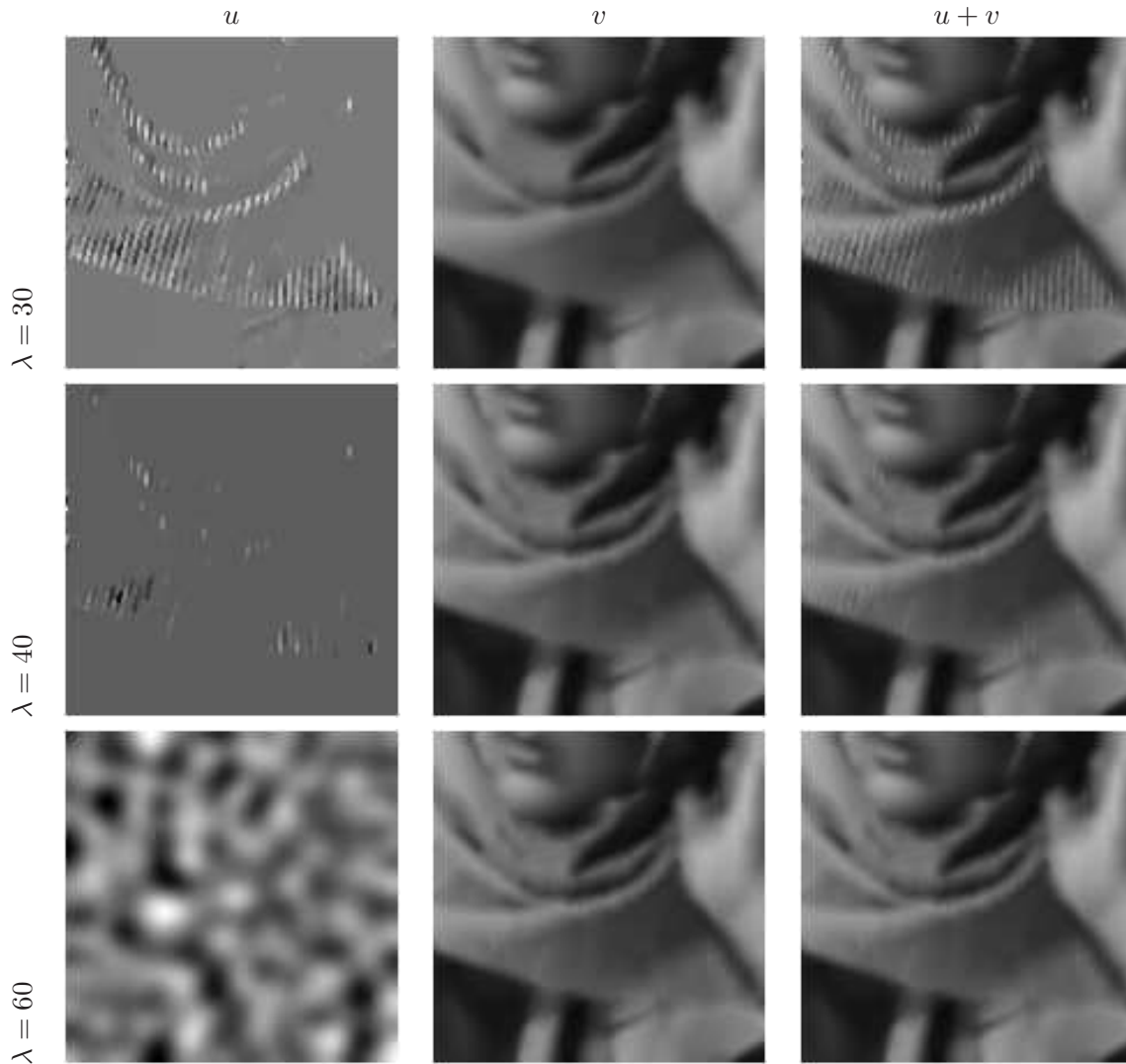


FIG. V.14 – Influence de λ (A-2). $\sigma = 20$ et $\mu = 30$. Lorsque λ augmente, l'information va jusqu'à disparaître totalement de u . De son côté, la composante v devient de moins en moins floue. Il est intéressant de remarquer que le bruit disparaissant peu à peu ne se retrouve pas dans v (il se retrouve en fait dans la composante $u_d - u - v$).

λ	$\lambda = 5$	$\lambda = 10$	$\lambda = 20$	$\lambda = 30$	$\lambda = 40$	$\lambda = 60$
SNR	16.4241	16.9026	14.3784	12.3017	11.4036	11.3308

FIG. V.15 – Images des figures V.13 et V.14.

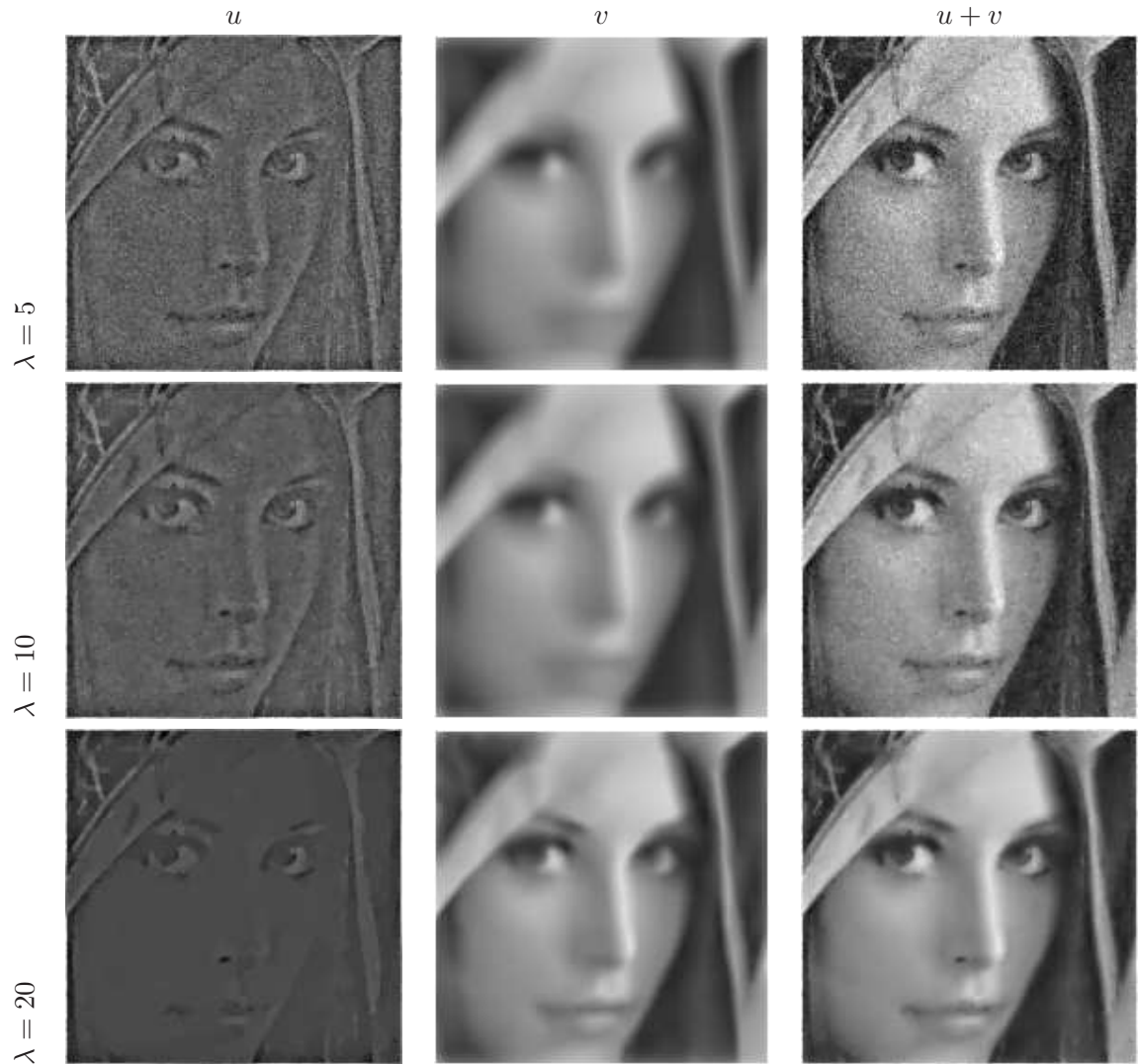


FIG. V.16 – Influence de λ (B-1). $\sigma = 20$ et $\mu = 30$. Les observations sont identiques à celles de la figure V.13 : la composante u perd de plus en plus d'information au fur et à mesure que λ augmente, d'abord le bruit, puis de plus en plus de détail. On observe pour $\lambda = 20$ que l'image restaurée $u+v$ ne possède plus de bruit alors que les détails (visage et plumes du chapeau) sont très bien conservés.

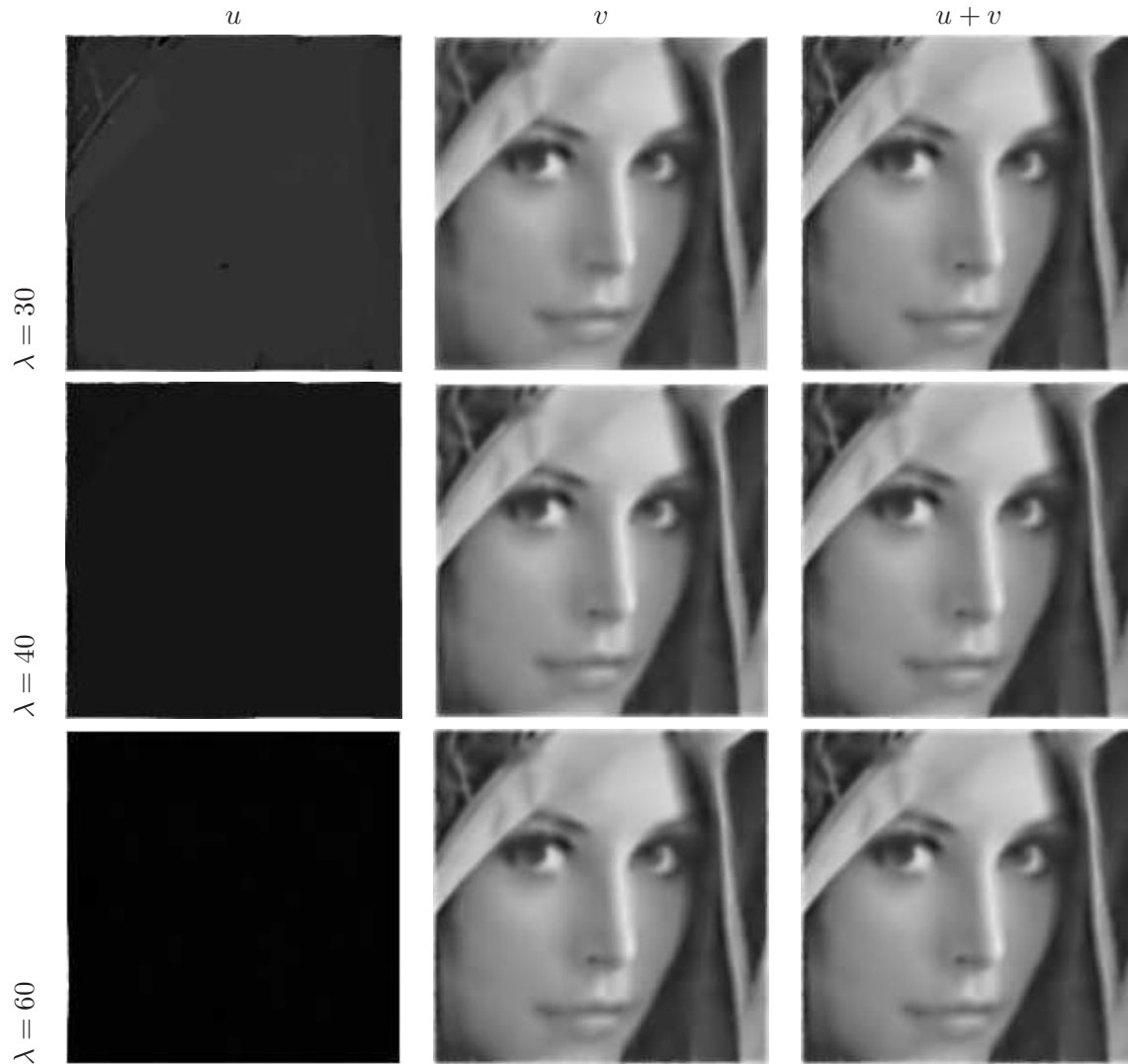


FIG. V.17 – Influence de λ (B-2). $\sigma = 20$ et $\mu = 30$. On observe pour cette image que l'information disparaît totalement de u un peu plus rapidement. Encore une fois, on remarque que la valeur de σ est suffisamment élevée pour que la composante v ne contienne jamais de bruit.

λ	$\lambda = 5$	$\lambda = 10$	$\lambda = 20$	$\lambda = 30$	$\lambda = 40$	$\lambda = 60$
SNR	20.1769	23.0344	22.9052	21.2345	20.2334	19.1735

FIG. V.18 – Images des figures V.16 et V.17.

3.2 Influence de μ

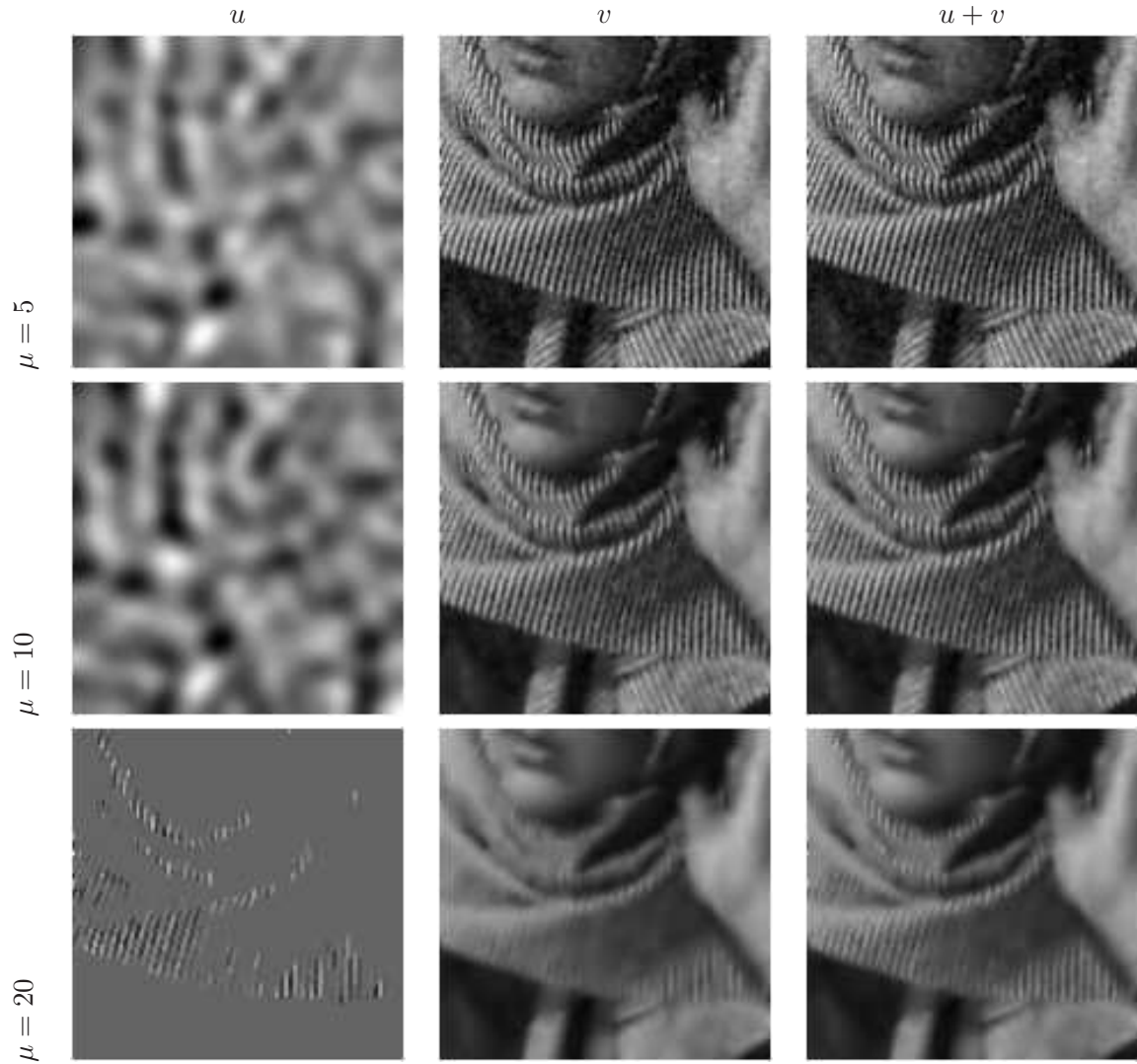


FIG. V.19 – Influence de μ (A-1). $\sigma = 20$ et $\lambda = 30$. Cette fois ci, nous avons fixé λ suffisamment élevé pour que le bruit n'apparaisse pas dans la composante u . L'augmentation de μ régularise la composante v , faisant disparaître assez rapidement le bruit. On remarque de plus que lorsque la texture commence à son tour à disparaître de la composante v , elle commence à apparaître dans la composante u .



FIG. V.20 – Influence de μ (A-2). $\sigma = 20$ et $\lambda = 30$. On observe que la composante v floute de plus en plus alors que la composante u récupère de plus en plus d'information (texture et géométrie). Pour $\mu = 60$ la composante u semble constante par morceaux. L'image restaurée $u + v$ semble ne plus trop évoluer pour μ suffisamment grand.

μ	$\mu = 5$	$\mu = 10$	$\mu = 20$	$\mu = 30$	$\mu = 40$	$\mu = 60$
SNR	17.5481	15.2500	12.4133	12.3017	12.3169	12.3641

FIG. V.21 – Images des figures V.19 et V.20.



FIG. V.22 – Influence de μ (B-1). $\sigma = 20$ et $\lambda = 30$. Les observations sont identiques à celles de la figure V.19 à ceci près que l'information commence à apparaître dans la composante u pour des valeurs de μ plus élevées.



FIG. V.23 – Influence de μ (B-2). $\sigma = 20$ et $\lambda = 30$. Comparativement à la figure V.20, nous sommes obligés de prendre des valeurs de μ beaucoup plus importantes pour obtenir des résultats équivalents. De plus, pour $\mu = 120$, il semblerait que l'image restaurée $u+v$ commence à devenir constante par morceaux.

μ	$\mu = 10$	$\mu = 30$	$\mu = 50$	$\mu = 70$	$\mu = 100$	$\mu = 120$
SNR	23.4240	21.2345	21.2170	21.2315	21.2889	21.3366

FIG. V.24 – Images des figures V.22 et V.23.

4 Comparaison entre les modèles ROF, ROF2 et mixte

Les trois modèles sont très différents, surtout le modèle mixte qui fait apparaître deux paramètres. Comparer leurs résultats de façon rigoureuse semble donc d'une extrême difficulté. Nous avons donc procédé empiriquement, en choisissant les paramètres de sorte que l'image semble subir des restaurations équivalentes. Nous proposons des tests sur deux images, bruitées à différentes intensités.

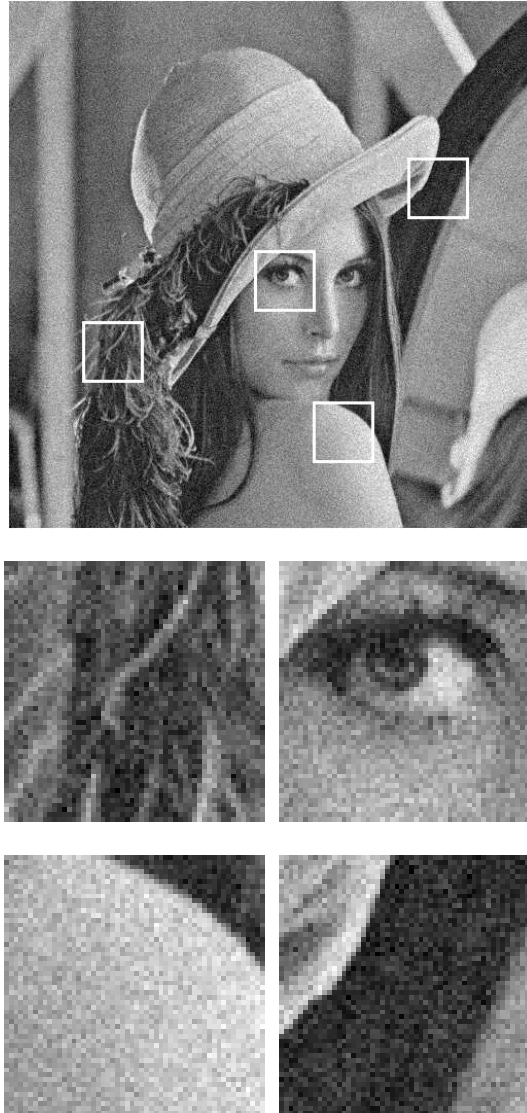


FIG. V.25 – Image test 1



FIG. V.26 – Image test 2

4.1 Images moyennement bruitées ($\sigma = 15$)

FIG. V.27 – Les effets *staircasing* du modèle ROF ainsi que l'effet de flou engendré par le modèle ROF2 sont bien visibles. Il est intéressant de remarquer que le modèle mixte semble se comporter comme le modèle ROF sur les zones possédant des détails (tête, herbe autour de la patte), et comme le modèle ROF2 ailleurs.



FIG. V.28 – Tests sur l'image complète. De haut en bas : ROF (SNR=21.6218), ROF2 (SNR=19.9328), mixte (21.4838).



FIG. V.29 – Encore une fois, nous remarquons que le modèle mixte semble se comporter comme le modèle ROF là où il y a des détails (plumes du chapeau, oeil) et comme le modèle ROF2 ailleurs.



FIG. V.30 – Tests sur l'image complète. De haut en bas et de gauche à droite : ROF (SNR=25.4147), ROF2 (SNR=24.9550), mixte (SNR=25.0423).

4.2 Images fortement bruitées ($\sigma = 50$)

FIG. V.31 – L'augmentation de l'intensité du bruit permet de mettre en valeur les différences des trois modèles. Le modèle mixte est clairement flou sur les zones homogènes comme le modèle ROF2 mais possède beaucoup plus d'information au niveau des détails, où un léger effet de *staircasing* semble apparaître.

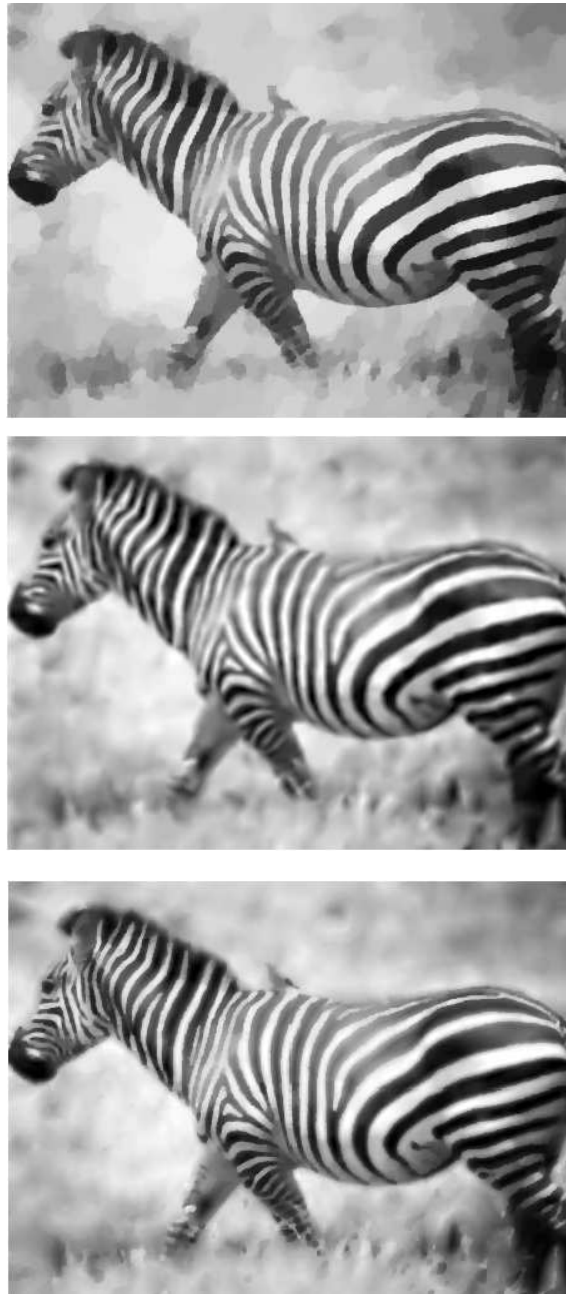


FIG. V.32 – Tests sur l'image complète. De haut en bas : ROF (SNR=16.4469), ROF2 (SNR=16.2236), mixte (SNR=17.1342).



FIG. V.33 – Comme pour l'image précédente, le modèle mixte semble être un bon compromis entre le modèle ROF et le modèle ROF2 : zones homogènes bien lissées et bonne conservation des détails.



FIG. V.34 – Tests sur l'image complète. De haut en bas et de gauche à droite : ROF (SNR=20.4493), ROF2 (SNR=20.5524), mixte (SNR=21.2467).

5 Modèle localement anisotrope dans le cadre de l'extraction de texture

Le principe ayant précisément été exposé au chapitre 3, nous nous contentons ici de proposer des résultats obtenus sur des images assez variées. Nous n'avons pas jugé utile de proposer une étude du comportement du modèle par rapport aux différents paramètres car il s'agit ici du modèle ROF2, rendu anisotrope en certains points (là où l'on situe les contours). La détection des contours est une étape primordiale. Nous rappelons que pour la faire, nous appliquons le modèle ROF2 classique afin de faire disparaître la majeure partie de la texture, puis nous calculons le gradient de l'image. Les points où celui-ci dépasse un certain seuil sont considérés comme "contours" faisant l'objet d'un traitement anisotrope. Il s'agit de la valeur de niveau de gris à partir duquel un pixel du module du gradient de l'image de départ est considéré comme composant un contour. Le choix de ce seuil se fait empiriquement : il faut le choisir suffisamment grand pour ne pas détecter les résidus de textures potentiellement présents et suffisamment petit pour détecter le plus de contours possibles. Pour illustrer nos propos nous appliquons le modèle localement anisotrope à l'image de la figure V.35. Les figures V.36 et V.37 présentent deux résultats différents selon le choix du seuil.



FIG. V.35 –



FIG. V.36 – En haut le seuil est égal à 5, en bas, il est égal à 2. On observe que, plus le seuil est élevé, plus la composante cartoon de l'image contient de l'information.

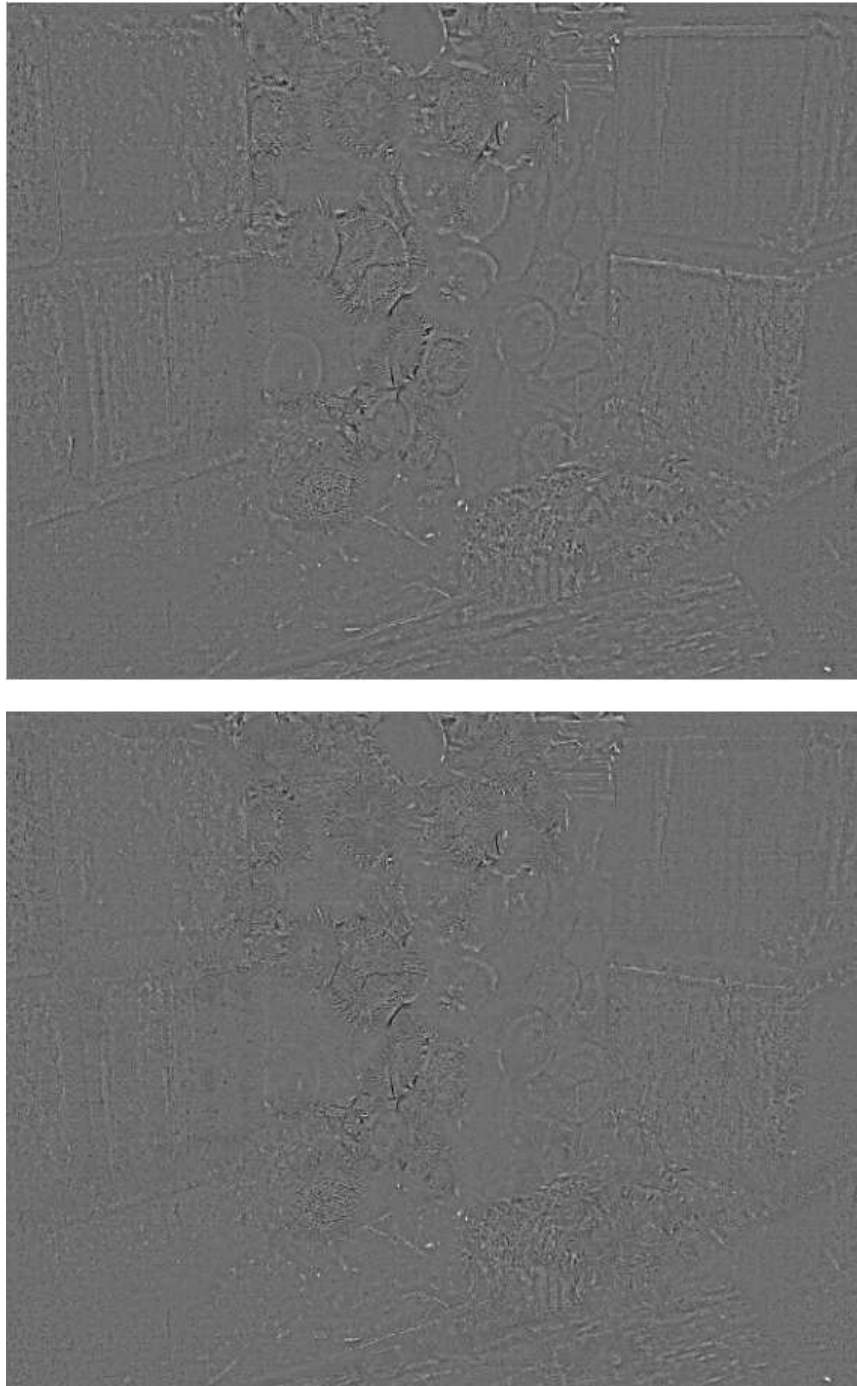


FIG. V.37 – En haut le seuil est égal à 5, en bas, il est égal à 2. On observe que, plus le seuil est élevé, moins la composante oscillante de l'image contient de contours et d'information.



FIG. V.38 – A gauche : composante cartoon et composante texture obtenue avec le modèle ROF2 brut. A droite : composante cartoon et composante texture obtenue avec le modèle ROF2 anisotrope (seuil égal à 2). Dans les deux cas $\lambda = 100$. On observe que l'ajout d'anisotropie permet de transférer la très grande partie de l'information géométrique de l'image de la partie texture à la partie cartoon. Ce phénomène est particulièrement visible sur la silhouette du petit garçon.

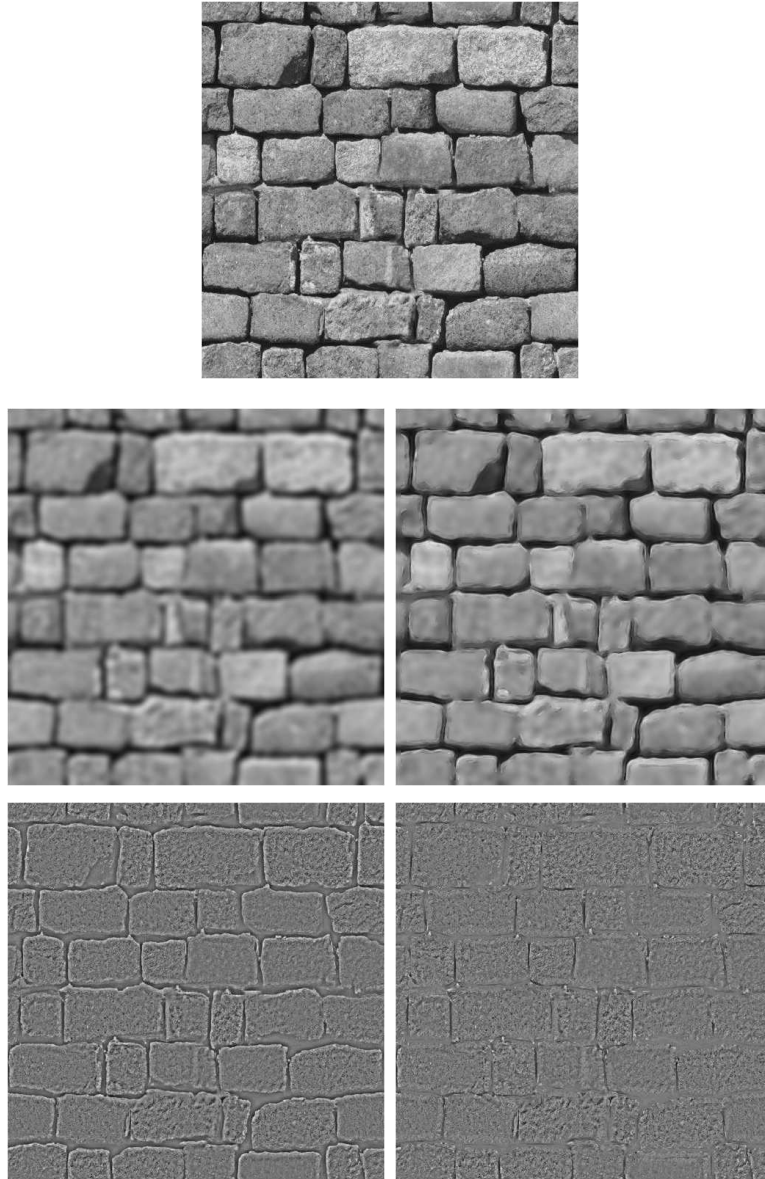


FIG. V.39 – A gauche : composante cartoon et composante texture obtenue avec le modèle ROF2 brut. A droite : composante cartoon et composante texture obtenue avec le modèle ROF2 anisotrope (seuil égal à 2). Dans les deux cas $\lambda = 100$. On observe que l'anisotropie locale permet d'éliminer de la texture les lignes formées par les joints entre chaque pierre. Celles-ci réapparaissent dans la composante cartoon.

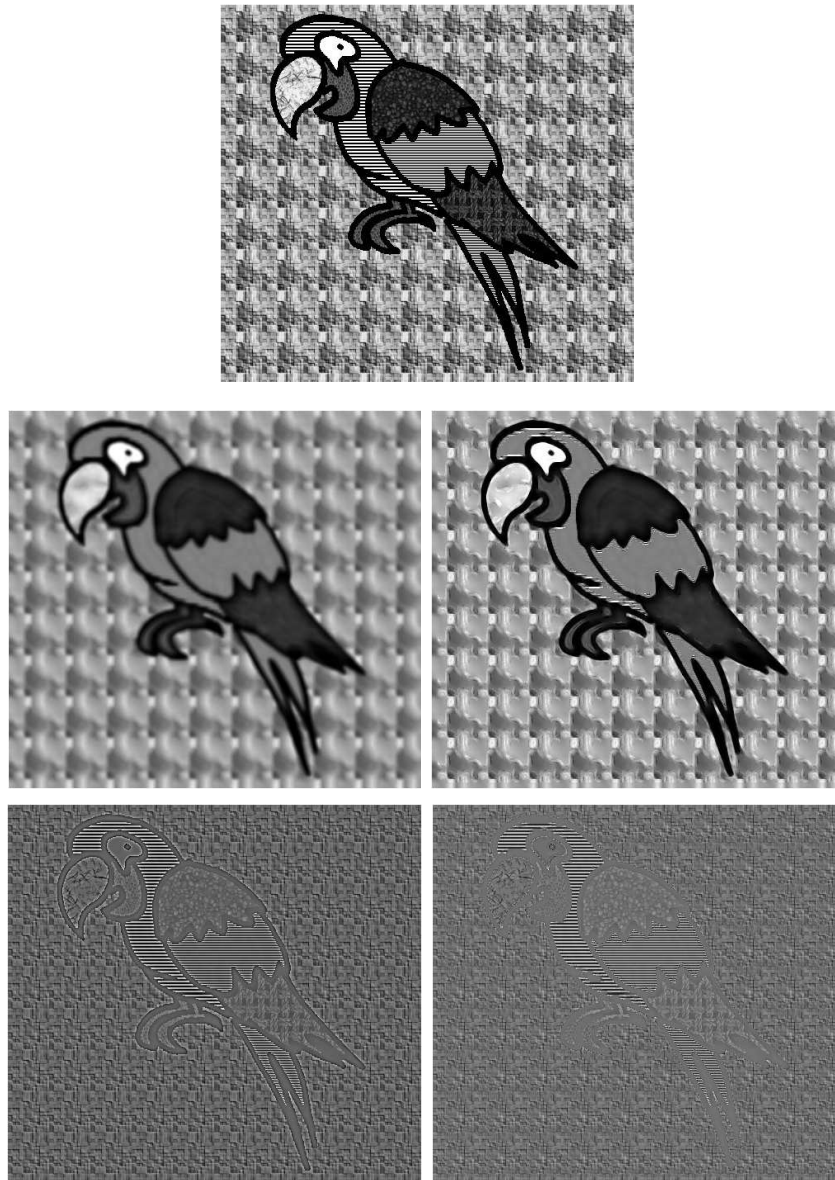


FIG. V.40 – Cette image est à la base un dessin en noir et blanc auquel nous avons ajouté manuellement différentes textures. A gauche : composante cartoon et composante texture obtenue avec le modèle ROF2 brut. A droite : composante cartoon et composante texture obtenue avec le modèle ROF2 anisotrope (seuil égal à 2). Dans les deux cas $\lambda = 100$. On observe encore une fois que la texture est très bien conservée avec le modèle anisotrope, débarrassée de la plupart des éléments géométriques de l'image.

En résumé, on observe dans tous les cas que le modèle localement anisotrope permet d'obtenir une décomposition satisfaisante dans le sens où la composante oscillante est débarrassée d'une grande partie de l'information géométrique mais conserve malgré tout une très grande partie de la texture. La composante cartoon est quant à elle, moins floue. Toutefois, on peut remarquer que certains contours artificiels y apparaissent : il s'agit de tous les pixels détectés comme des contours lors du pré-traitement de l'image par l'algorithme brut. Nous pouvons atténuer ce phénomène en augmentant la valeur du seuil ; mais nous détecterons alors moins de contours. Nous devons de plus signaler que le temps de calcul de ce modèle est très conséquent, à cause du fait que tous les pixels détectés comme contours lors du pré-traitement sont traités un par un.

6 Modèles $TV-L^1$ et TV^2-L^1

Nous rappelons la version du modèle discret :

$$\min_{u \in X} F_\delta^i(u), \quad (\mathcal{P}_\delta^i)$$

où

$$F_\delta^i(u) := \mu \|u_d - u\|_{L^1} + \lambda J_i(u) + \frac{\delta}{2} \left(\|u\|_X^2 + \|u_d - u\|_X^2 \right),$$

avec $i \in \{1, 2\}$.

Pour les tests numériques, nous décidons de ne faire varier que le paramètre μ . En effet, seule est importante la valeur du rapport λ/μ . Dans les tests qui suivent, nous posons $\lambda = 30$.

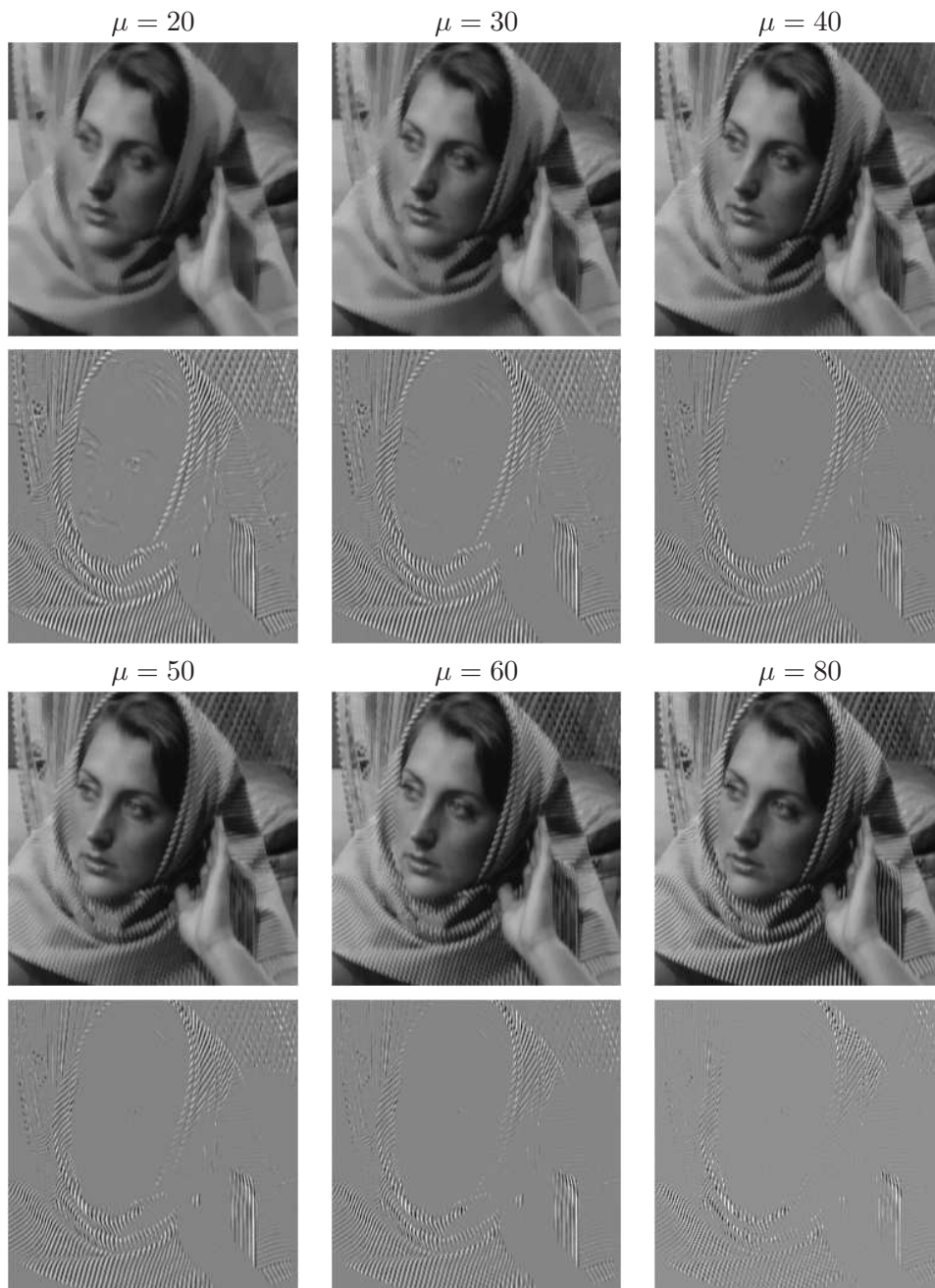
6.1 Le modèle $TV-L^1$ 

FIG. V.41 – $TV-L^1$ (A). $\lambda = 30$. On observe, comme attendu, que l'information oscillante se déplace de la partie texture vers la partie cartoon au fur et à mesure que μ augmente.

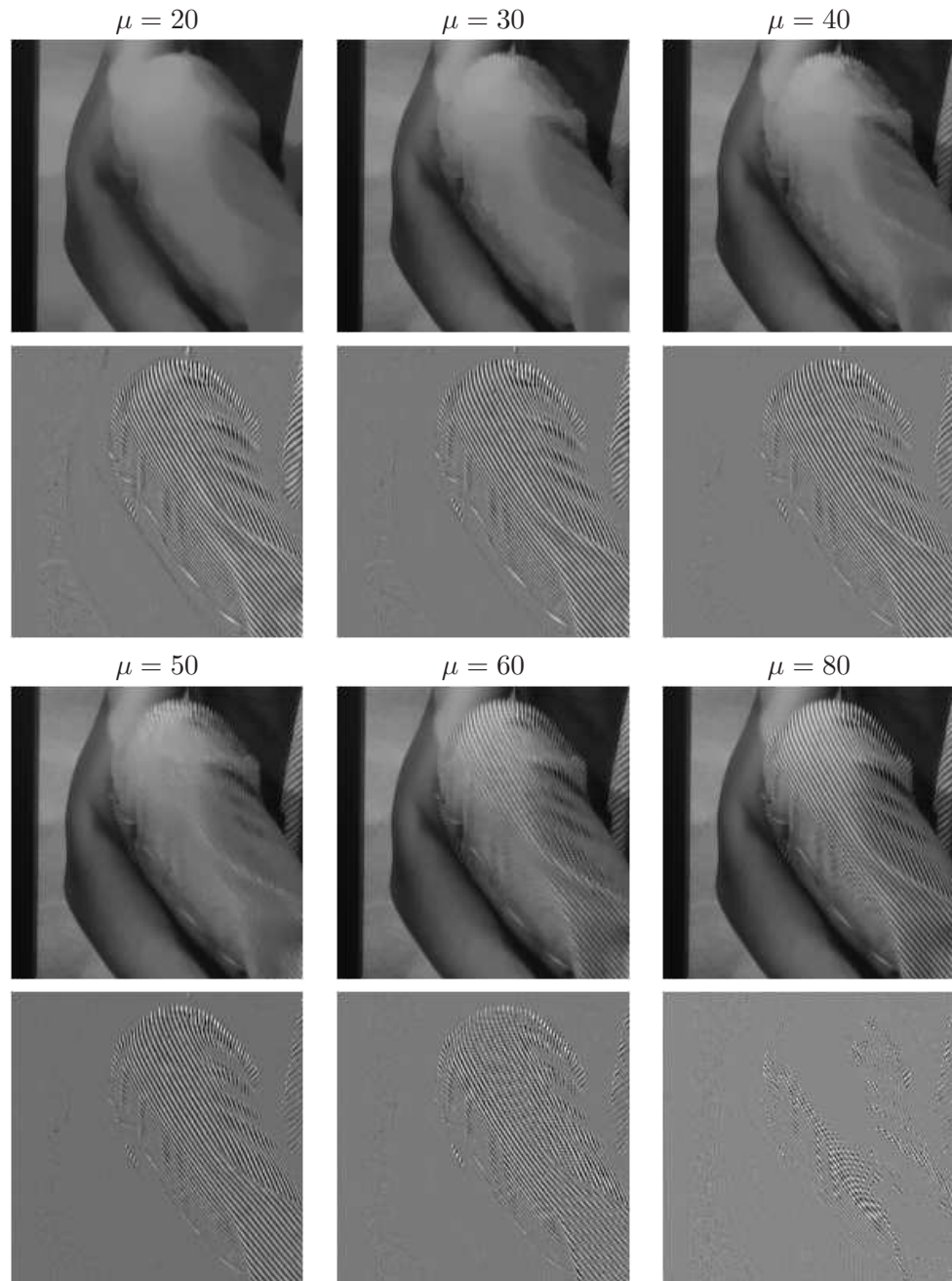


FIG. V.42 – $TV-L^1$ (B). $\lambda = 30$. Les remarques sont identiques à celles de la figure V.41. On peut observer au passage que l'effet *staircasing* est encore présent sur la partie cartoon.

6.2 Le modèle $TV-L^1$ avec terme d'attache aux données L^2

Nous proposons ici de comparer les modèles $TV-L^1$ introduit page 69 et celui possédant un terme d'attache aux données L^2 introduit page 77. Alors que dans le premier cas la décomposition vérifie obligatoirement $u + v = u_d$, ce n'est pas vrai dans le second : en effet, le terme L^2 peut contenir de l'information. La figure V.44 illustre ce propos. La figure V.43 montre de façon plus claire le fait que la somme des deux composantes n'est plus nécessairement égale à l'image de départ. Ceci permet à la fois d'obtenir une composante "cartoon" sans la moindre texture et une composante "texture" nétoyée des informations parasites. Nous ne proposons ici que des tests utilisant la variation totale. Nous signalons toutefois que le comportement est identique lorsque l'on utilise la variation totale seconde.

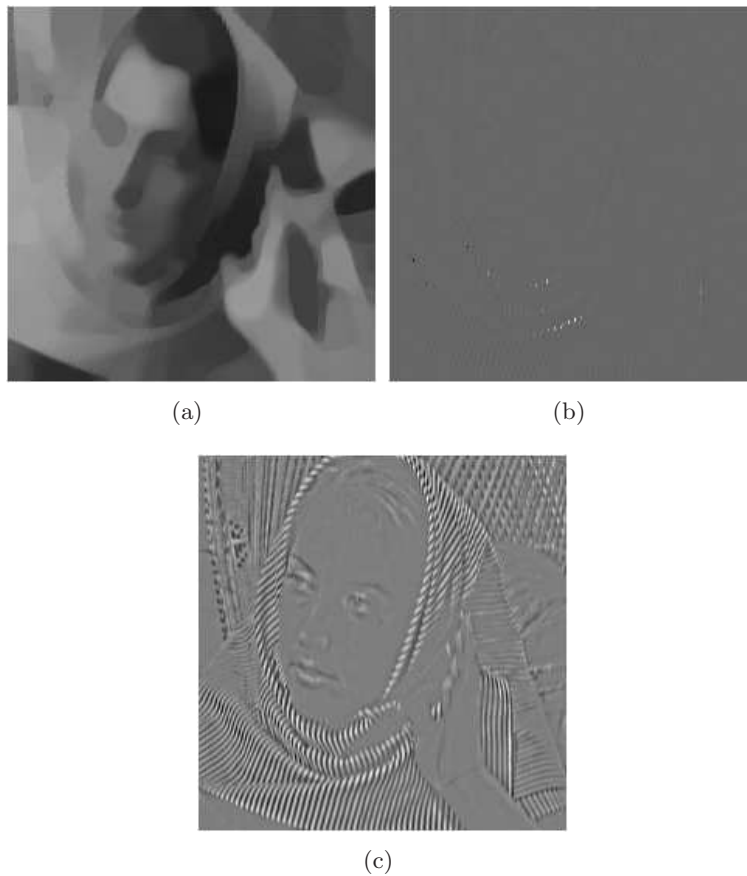


FIG. V.43 – Modèle $TV-L^1$ avec terme d'attache aux données L^2 . $\lambda = \mu = 90$. (a) est la composante u , (b) la composante v et (c) le reste ($u_d - u - v$). Il est évident sur cette décomposition que la somme $u + v$ n'est pas égale à l'image d'origine.



FIG. V.44 – A gauche : Le modèle $TV-L^1$. A droite : Le modèle $TV-L^1$ avec terme d’attache aux données L^2 . $\lambda = 30$, $\mu = 15$. On observe que pour deux composantes “cartoon” équivalentes, la partie “texture” de l’image obtenue avec le modèle possédant le terme d’attache aux données L^2 contient moins de détails “parasites” (pied de table, visage, bras).

Dans la suite nous n’utiliserons que les modèle définis dans la partie 2.

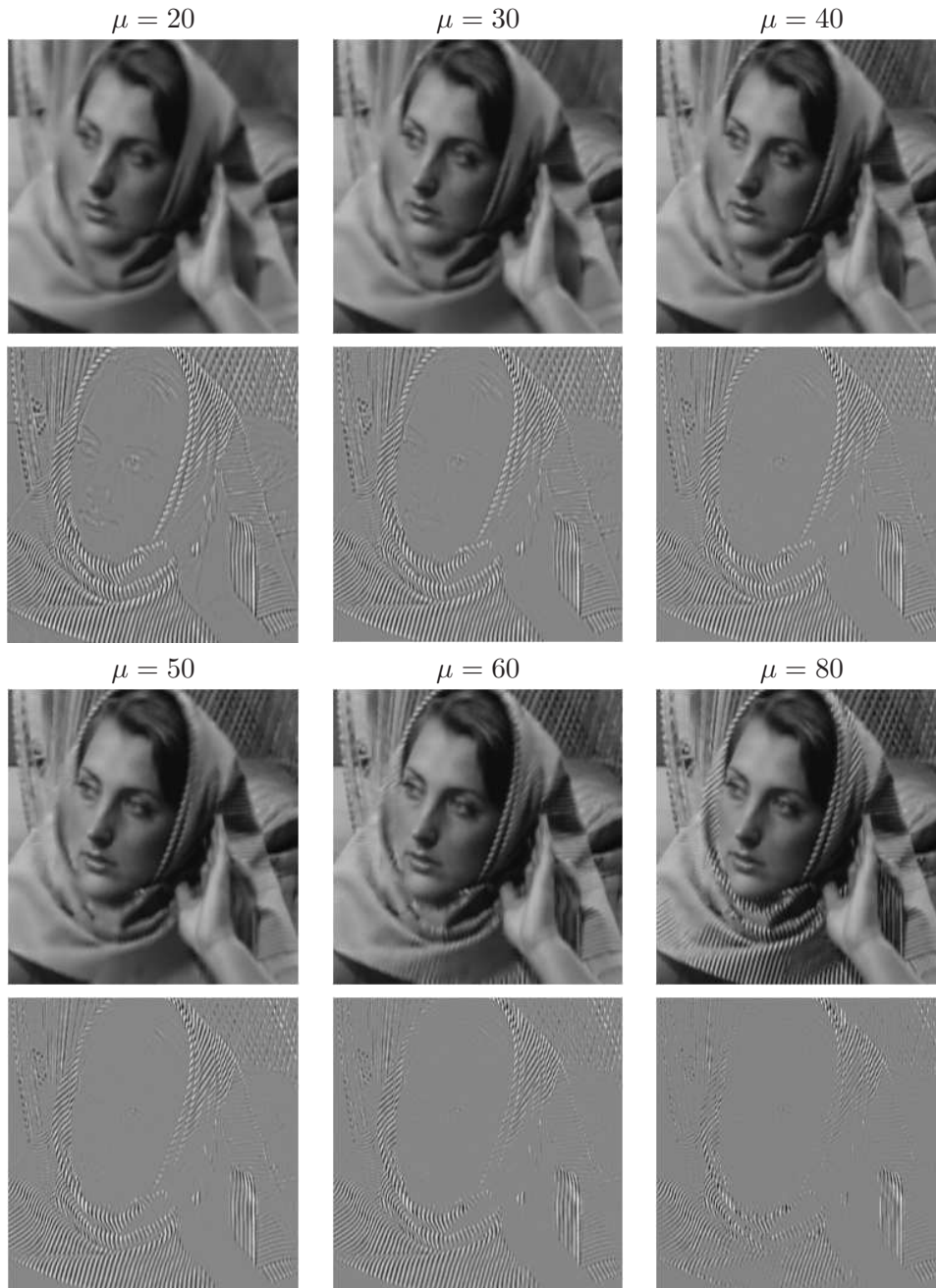
6.3 Le modèle TV^2-L^1 

FIG. V.45 – TV^2-L^1 (A). $\lambda = 30$. Les observations sont identiques à celles faites pour le modèle $TV-L^1$. On observe que le résultat semble particulièrement bon pour $\mu = 30$. On peut de plus remarquer l'absence d'un quelconque effet *staircasing*.

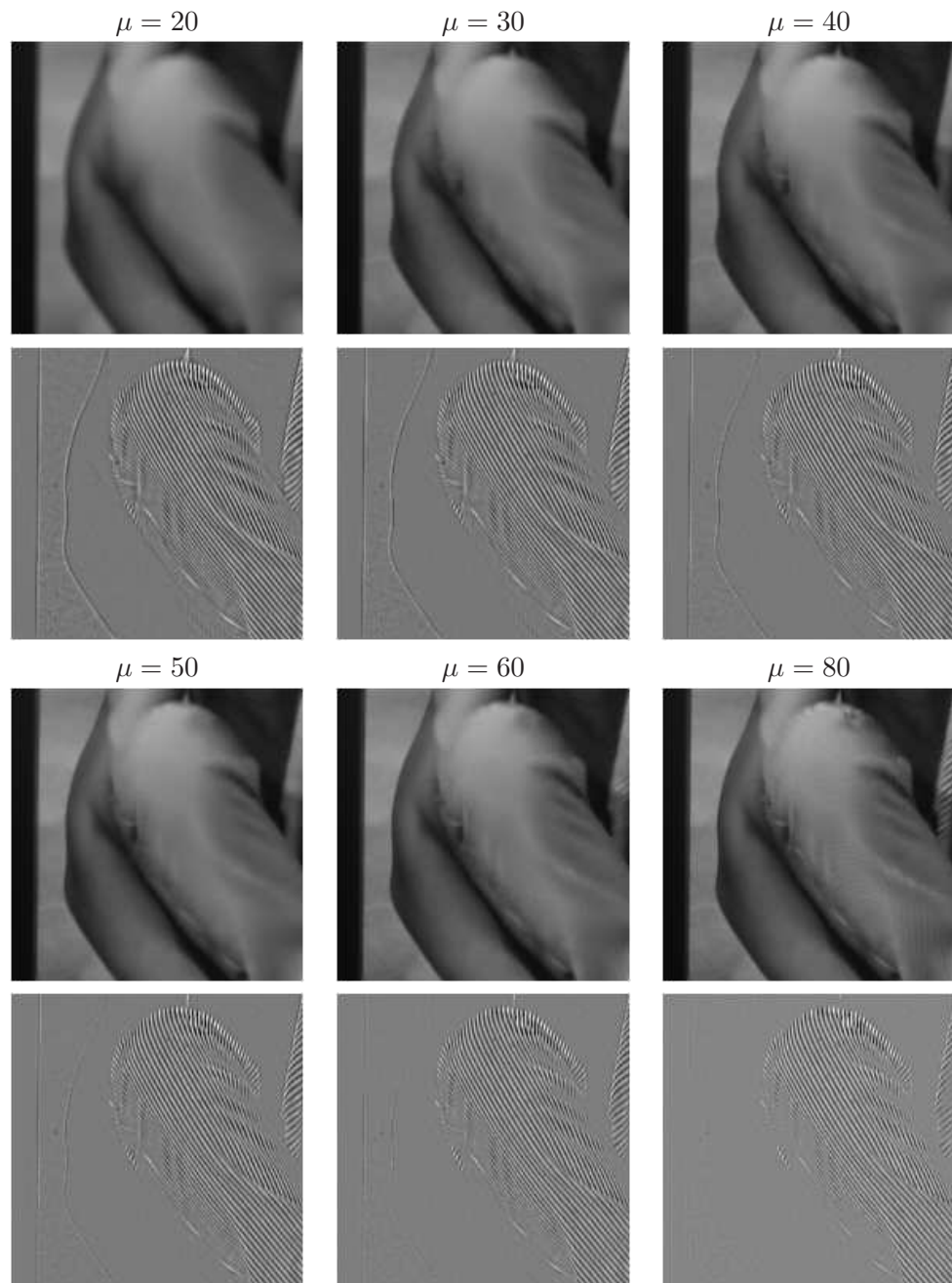


FIG. V.46 – TV^2-L^1 (B). $\lambda = 30$. Ce jeu de tests fait apparaître le défaut majeur de ce modèle. En effet, il est visible que la partie “texture” contient beaucoup trop d’information géométrique.

6.4 Comparaison

Nous allons ici comparer les modèles $TV-L^1$ et TV^2-L^1 . Nous comparons indépendamment les composantes “cartoon” et “texture” dans le sens où nous comparons les meilleures composantes “cartoon” des deux modèles ainsi que les meilleures composantes “texture”. Ces composantes n’appartiennent pas obligatoirement à la même décomposition. Nous nous permettons cette liberté car les tests ont montré que l’utilisation des modèles avec terme d’attache aux données L^2 nous permet d’obtenir de tels résultats (en effet, dans ce cas la somme des deux composante n’est plus nécessairement égale à l’image d’origine).

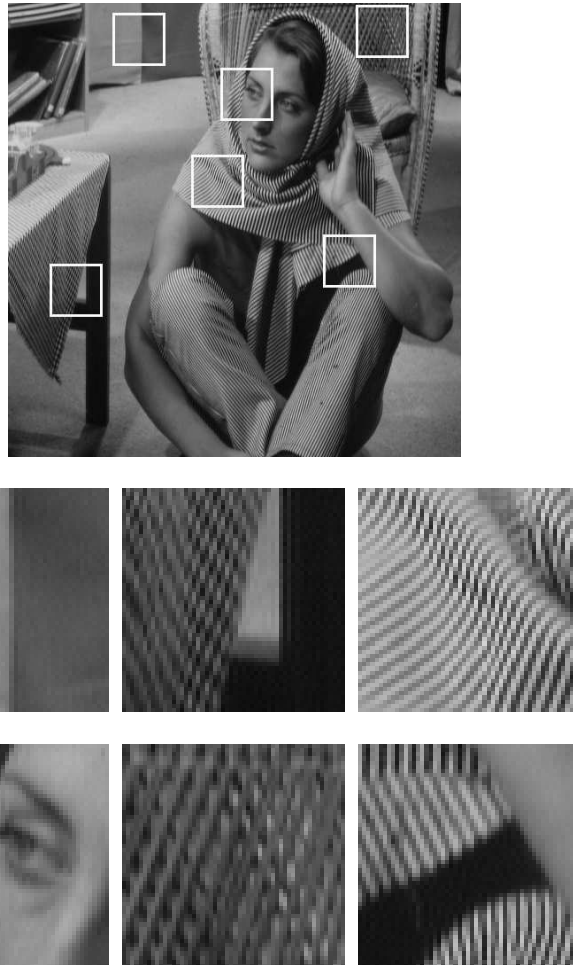


FIG. V.47 – Image test.

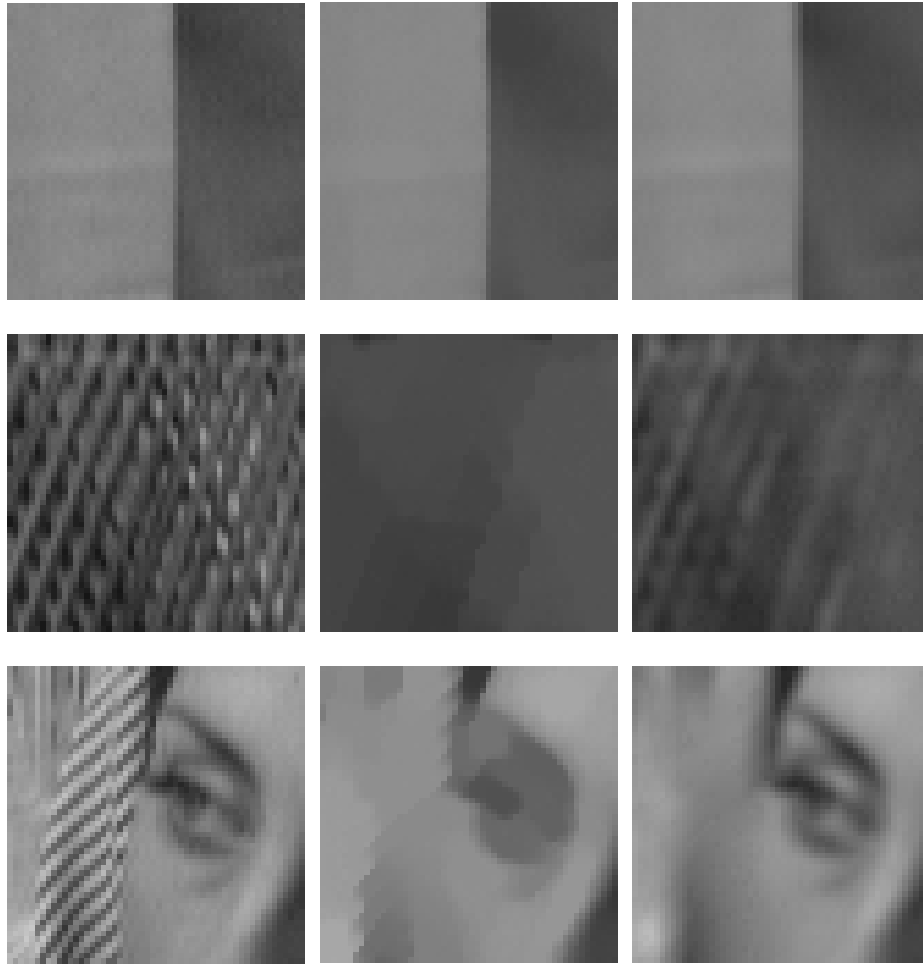


FIG. V.48 – Comparaison de la composante “cartoon” (A). A gauche : Image test. Au centre : Modèle $TV-L^1$. A droite : Modèle TV^2-L^1 . Les paramètres sont pris de sorte à obtenir des composantes comparables. On observe que l’effet *staircasing* présent au centre est absent à droite.



FIG. V.49 – Comparaison de la composante “cartoon” (B). A gauche : Image test. Au centre : Modèle $TV-L^1$. A droite : Modèle TV^2-L^1 . Les observations sont identiques à celles de la figure précédente.

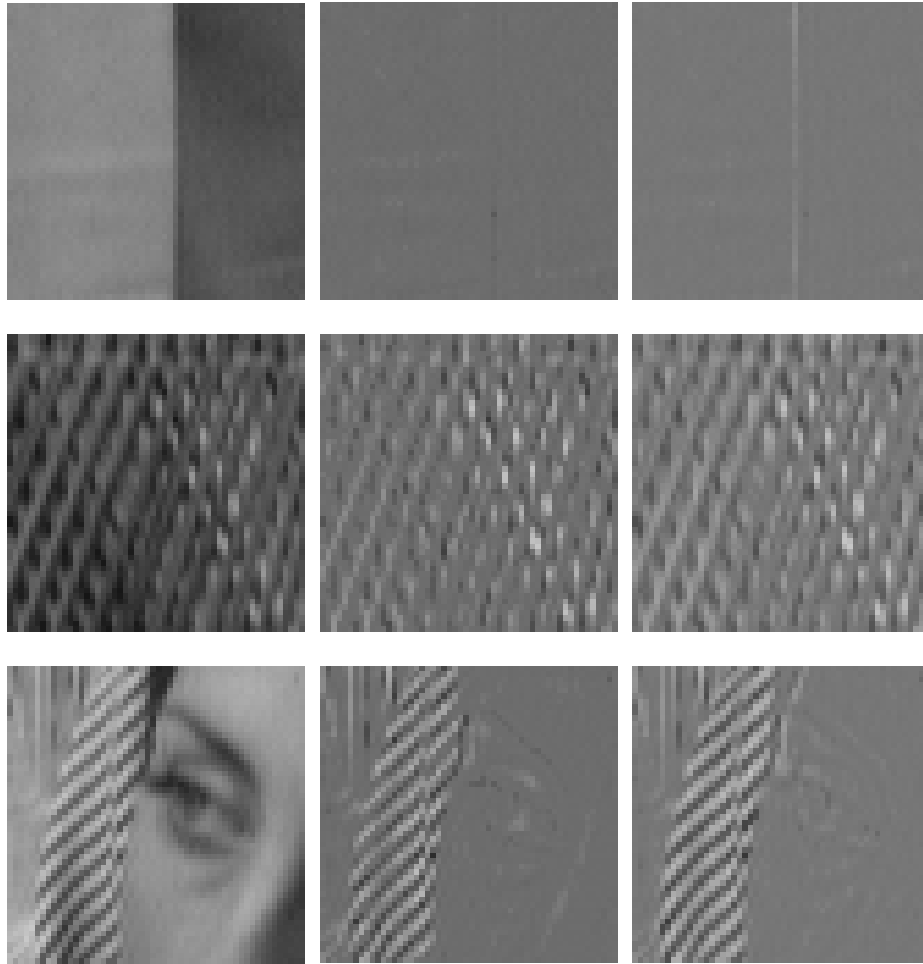


FIG. V.50 – Comparaison de la composante “texture” (A). A gauche : Image test. Au centre : Modèle $TV-L^1$. A droite : Modèle TV^2-L^1 . Les paramètres sont pris de sorte à obtenir des composantes comparables. La première ligne met en évidence les limites du modèle TV^2-L^1 : la composante contient trop d’information géométrique.

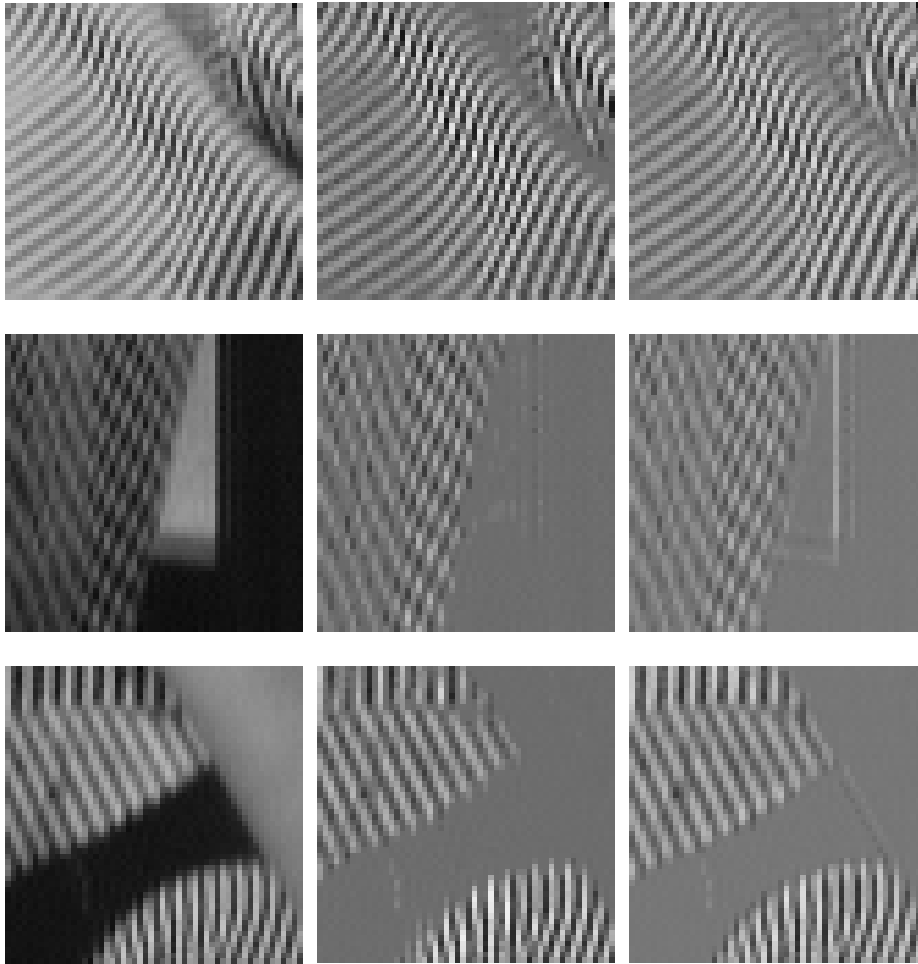


FIG. V.51 – Comparaison de la composante “texture” (B). A gauche : Image test. Au centre : Modèle $TV-L^1$. A droite : Modèle TV^2-L^1 . Ce sont ici les lignes 2 et 3 qui mettent en évidence le fait que la composante “texture” contient trop d’information géométrique.

Les différents tests tendent à montrer que, sur des images possédant peu d’information géométrique, le modèle TV^2-L^1 se révèle aussi efficace voire meilleur que le modèle $TV-L^1$ (absence d’effet *staircasing* sur la composante “cartoon”). En revanche, sur des images contenant une certaine quantité d’information géométrique, la composante “texture” du modèle TV^2-L^1 n’est pas aussi satisfaisante que celle obtenue avec le modèle $TV-L^1$: elle contient trop d’information géométrique. C’est dans le but de résoudre ce problème que nous avons introduit un modèle mixte d’extraction de texture. Ce modèle fera l’objet d’un futur travail.

Appendices

Annexe A

Article bfg

A variational model for image texture identification

R. Echegut and L. Piffet

Université d'Orléans, Département de Mathématiques (MAPMO), UFR Sciences,
Bâtiment de mathématiques, Route de Chartres B.P. 6759, 45067 Orléans cedex 2,
FRANCE
`echegut.romain@wanadoo.fr`, `Loic.Piffet@univ-orleans.fr`

Variational models in image processings have been extensively studied during the past decade. There are used for segmentation processes (geodesic or geometric contours) and restoration and textures extraction purpose as well. Roughly speaking image restoration problems are severely ill posed and a Tychonov- like regularization process is needed. The general form of such models consists in the minimization of a “energy” functional :

$$\mathcal{F}(u) = \|u - u_d\|_X + \mathcal{R}(u) , u \in Y \subset X ,$$

where X, Y are (real) Banach spaces, \mathcal{R} is a regularization operator, u_d is the observed (or measured) image and u is the image to recover. The first term is the fitting data term and the second one permits to get a problem which is no longer ill posed. The most famous model is the Rudin- Osher-Fatemi denoising model (see [11]) which leads to the following problem:

$$\min_{u \in BV(\Omega)} \lambda |u|_{BV(\Omega)} + \frac{1}{2} \|u_d - u\|_{L^2(\Omega)}^2. \quad (\mathcal{P}_{ROF})$$

Here, we are interested in textures extraction and/or image denoising. Recently people considered that an image can be decomposed in many components, each component describing a particular property . of the image (see [6, 12, 14] for example). We agree with this idea and assume that the image f we want to recover from the data u_d can be decomposed as $f = u + v$ where u and v are functions that belong to different functional spaces: u is the “texture” part which involves (periodic or not) details (and noise as well) while v is a more regular part (usually called the “cartoon” component).

In a first section, we present the functional framework, introducing the BV^2 space, and the general variational model we consider. In section 2, we focus on numerical implementation and present the discretization process. Numerical tests are reported in the last section.

1 Functional framework and model

1.1 The $BV^2(\Omega)$ space

Let Ω be an open bounded subset of \mathbb{R}^n , $n \geq 2$ (practically $n = 2$). Following Demengel [9], we define the space of bounded hessian functions that we call $BV^2(\Omega)$. We recall that the space $BV(\Omega)$ of bounded variation functions (see [2, 4, 3]) is defined as

$$BV(\Omega) = \{u \in L^1(\Omega) \mid \Phi(u) < +\infty\},$$

where

$$\Phi(u) = \sup \left\{ \int_{\Omega} u(x) \operatorname{div} \xi(x) dx \mid \xi \in \mathcal{C}_c^1(\Omega), \|\xi\|_{\infty} \leq 1 \right\}. \quad (1)$$

The space $BV(\Omega)$, endowed with the norm $\|u\|_{BV(\Omega)} = \|u\|_{L^1} + \Phi(u)$, is a Banach space. The derivative in the sense of the distributions of every $u \in BV(\Omega)$ is a bounded Radon measure, denoted Du , and $\Phi(u) = \int_{\Omega} |Du|$ is the total variation of Du . We extend this definition to the second derivative (in the distributional sense). Recall that the Sobolev space

$$W^{1,1}(\Omega) = \{u \in L^1(\Omega) \mid \nabla u \in L^1(\Omega)\}$$

where ∇u stands for the first order derivative of u (in the sense of distributions).

Definition 1. *A function $u \in W^{1,1}(\Omega)$ is Hessian bounded if*

$$|u|_{BV^2(\Omega)} := \sup \left\{ \int_{\Omega} \langle \nabla u, \operatorname{div}(\phi) \rangle_{\mathbb{R}^n} \mid \phi \in \mathcal{C}_c^2(\Omega, \mathbb{R}^{n \times n}), \|\phi\|_{\infty} \leq 1 \right\} < \infty,$$

where

$$\operatorname{div}(\phi) = (\operatorname{div}(\phi_1), \operatorname{div}(\phi_2), \dots, \operatorname{div}(\phi_n)),$$

with

$$\forall i, \phi_i = (\phi_i^1, \phi_i^2, \dots, \phi_i^n) \in \mathbb{R}^n \quad \text{and} \quad \operatorname{div}(\phi_i) = \sum_{k=1}^n \frac{\partial \phi_i^k}{\partial x_k}.$$

For more information on the $BV^2(\Omega)$ space, see [9, 13].

1.2 The variational model

We now assume that the image we want to recover from the data u_d can be written as $f = u + v$ where u is in $BV(\Omega)$ and v is in $BV^2(\Omega)$. Such decompositions have already been performed [5, 6, 4] using the ‘‘Meyer’’ space of oscillating function [10] instead of $BV^2(\Omega)$. So far, the modelization we propose is not the same: the oscillating component will be included in the

non regular part u while v involves the cartoon and the contours. We consider the following function defined on $BV(\Omega) \times BV^2(\Omega)$:

$$F(u, v) = \frac{1}{2} \|u_d - u - v\|_{L^2(\Omega)}^2 + \lambda |u|_{BV(\Omega)} + \mu |v|_{BV^2(\Omega)} + \delta \|\nabla v\|_{W^{1,1}(\Omega)}, \quad (2)$$

where $\lambda, \mu, \delta \geq 0$. We are looking for a solution to the optimisation problem

$$\inf_{(u,v) \in BV(\Omega) \times BV^2(\Omega)} F(u, v) \quad (\mathcal{P})$$

The first term $\|u_d - u - v\|_{L^2(\Omega)}^2$ of F is the fitting data term. Other terms are Tychonov-like regularization terms. Note that the δ -term is not useful from the modelling point of view. It is only a tool that allows to prove existence of solutions. We shall choose $\delta = 0$ for numerical tests.

If the image is noisy, the noise is considered as a texture and will be involved in u : more precisely v will be the part of the image without the oscillating component, that is the denoised part. In [7] we focus on the denoising process taking only v into account (and assuming that $u = 0$ so that $u_d - v$ is the noise). We now give an existence and uniqueness result for the general problem (\mathcal{P}) (see [7] for the proof).

Theorem 1. *Assume that $\lambda > 0, \mu > 0$ and $\delta > 0$. Problem (\mathcal{P}) has a unique solution (u, v) .*

2 Numerical implementation

2.1 Discretization of the problem

We assume for simplicity that the image is squared with size $N \times N$. We note $X := \mathbb{R}^{N \times N} \simeq \mathbb{R}^{N^2}$ endowed with the usual inner product and the associated euclidean norm

$$\langle u, v \rangle_X := \sum_{1 \leq i, j \leq N} u_{i,j} v_{i,j}, \quad \|u\|_X := \sqrt{\sum_{1 \leq i, j \leq N} u_{i,j}^2}. \quad (3)$$

It is classical to define the discrete total variation as following (see for example [4]) : the discrete gradient of the numerical image $u \in X$ is $\nabla u \in X^2$ defined by

$$(\nabla u)_{i,j} = \left((\nabla u)_{i,j}^1, (\nabla u)_{i,j}^2 \right), \quad (4)$$

where

$$(\nabla u)_{i,j}^1 = \begin{cases} u_{i+1,j} - u_{i,j} & \text{if } i < N \\ 0 & \text{if } i = N, \end{cases} \quad \text{and} \quad (\nabla u)_{i,j}^2 = \begin{cases} u_{i,j+1} - u_{i,j} & \text{if } j < N \\ 0 & \text{if } j = N. \end{cases}$$

The (discrete) total variation $|u|_{BV(\Omega)}$ is given by

$$J_1(u) = \sum_{1 \leq i, j \leq N} \left\| (\nabla u)_{i,j} \right\|_{\mathbb{R}^2}, \quad (5)$$

where

$$\left\| (\nabla u)_{i,j} \right\|_{\mathbb{R}^2} = \left\| \left((\nabla u)_{i,j}^1, (\nabla u)_{i,j}^2 \right) \right\|_{\mathbb{R}^2} = \sqrt{\left((\nabla u)_{i,j}^1 \right)^2 + \left((\nabla u)_{i,j}^2 \right)^2}.$$

The discrete divergence operator div is the adjoint operator of the gradient operator ∇ :

$$\forall (p, u) \in X^2 \times X, \quad \langle -\operatorname{div} p, u \rangle_X = \langle p, \nabla u \rangle_Y,$$

so that

$$(\operatorname{div} p)_{i,j} = \begin{cases} p_{i,j}^1 - p_{i-1,j}^1 & \text{if } 1 < i < N \\ p_{i,j}^1 & \text{if } i = 1 \\ -p_{i-1,j}^1 & \text{if } i = N \end{cases} + \begin{cases} p_{i,j}^1 - p_{i,j-1}^2 & \text{if } 1 < j < N \\ p_{i,j}^2 & \text{if } j = 1 \\ -p_{i,j-1}^2 & \text{if } j = N. \end{cases} \quad (6)$$

To define a discrete version of the second order total variation we have to introduce the discrete Hessian operator. As for the gradient operator, we define it by finite differences. So, for any $v \in X$, the Hessian matrix of v , denoted Hv is identified to a X^4 vector:

$$(Hv)_{i,j} = \left((Hv)_{i,j}^{11}, (Hv)_{i,j}^{12}, (Hv)_{i,j}^{21}, (Hv)_{i,j}^{22} \right).$$

The discrete second order total variation $|v|_{BV^2(\Omega)}$ of v is defined as

$$J_2(v) = \sum_{1 \leq i, j \leq N} \left\| (Hv)_{i,j} \right\|_{\mathbb{R}^4}. \quad (7)$$

As in the BV case, we may compute the adjoint operator of H (which is the discretized “second divergence” operator) :

$$\forall p \in X^4, \quad \forall v \in X \quad \langle H^* p, v \rangle_X = \langle p, Hv \rangle_{X^4}. \quad (8)$$

and we deduce an numerical expression for H^* from the equality (8). The discretized problem stands

$$\inf_{(u,v) \in X^2} \frac{1}{2} \|u_d - u - v\|_X^2 + \lambda J_1(u) + \mu J_2(v) + \delta(|v| + J_1(v)), \quad (\mathcal{P}_d)$$

where

$$|v| := \sum_{1 \leq i, j \leq N} |v_{i,j}|.$$

In the finite dimensional case we have an existence result with $\delta = 0$.

Theorem 2. *Problem \mathcal{P}_d has a unique solution for every $\delta \geq 0$ and $\lambda > 0$.*

For numerical purpose we shall set $\delta = 0$. In fact, we have performed tests with $\delta = 0$ and very small $\delta \neq 0$ (as required by the theory to get a solution to problem $\tilde{\mathcal{P}}$) and results were identical. So, to simplify numerical implementation, we consider the following discretized problem :

$$\inf_{(u,v) \in X^2} \frac{1}{2} \|u_d - u - v\|_X^2 + \lambda J_1(u) + \mu J_2(v). \quad (\tilde{\mathcal{P}}_d)$$

2.2 Algorithm

Using non smooth analysis tools (for convex functions) it is easy to derive (necessary and sufficient) optimality conditions. More precisely (u, v) is solution of (\mathcal{P}_d) if and only if

$$\begin{cases} 0 \in \partial \left(\lambda J_1(u) + \frac{1}{2} \|u_d - u - v\|^2 \right) \\ 0 \in \partial \left(\mu J_2(v) + \frac{1}{2} \|u_d - u - v\|^2 \right), \end{cases} \quad (9)$$

where ∂J is the classical subdifferential of J . Using subdifferential properties, we see that (9) is equivalent to

$$\begin{cases} u = u_d - v - \Pi_{\lambda K_1}(u_d - v) \\ v = u_d - u - \Pi_{\mu K_2}(u_d - u). \end{cases} \quad (10)$$

where K_1 and K_2 are closed convex sets. Chambolle [8] proved that

$$K_1 = \{ \operatorname{div} p \mid p \in Y, \|p_{i,j}\|_{\mathbb{R}^2} \leq 1 \ \forall i, j = 1, \dots, N \} \quad (11)$$

in the $BV(\Omega)$ setting and we may prove similarly that

$$K_2 = \{ H^* p \mid p \in Z, \|p_{i,j}\|_{\mathbb{R}^4} \leq 1, \ \forall i, j = 1, \dots, N \}, \quad (12)$$

(see [7]). Moreover, Chambolle [8] proposed a fixed point algorithm to compute $\Pi_{\lambda K_1}$ and we are able to extend this result to the second order case. Therefore, we set the following fixed point algorithm:

$$p^0 = 0 \quad (13a)$$

$$p_{i,j}^{n+1} = \frac{p_{i,j}^n - \tau(H[H^* p^n - u_d/\lambda])_{i,j}}{1 + \tau\|(H[H^* p^n - u_d/\lambda])_{i,j}\|_{\mathbb{R}^4}}. \quad (13b)$$

which convergence is proved in [7] :

Theorem 3. *Let be $\tau \leq 1/64$. Then $\lambda(H^* p^n)_n$ converges to $P_{\lambda K_2}(u_d)$.*

So, we propose the following algorithm :

- **Step 1** : We choose u_0 et v_0 (for example, $u_0 = 0$ et $v_0 = u_d$) and $0 < \alpha < 1$.
- **Step 2** : define the sequences $((u_n, v_n))_n$ as follows:

$$\begin{cases} u_{n+1} = u_n + \alpha (u_d - v_n - \Pi_{\lambda K_1}(u_d - v_n) - u_n) \\ v_{n+1} = v_n + \alpha (u_d - u_n - \Pi_{\mu K_2}(u_d - u_n) - v_n). \end{cases}$$

- **Step 3** : if a stopping criterion is not satisfied, set $k := k + 1$ and go back to 2.

We are not able to prove the convergence of this algorithm, but we observed that, numerically, it always converges for appropriate values of α . Nevertheless, we propose in [13] another algorithm which convergence is rigorously proved .

3 Numerical tests and comments

We test the model on two images: The first one is a synthetic image wher texture has been artificially added, and the second one is the well known “Barbara” benchmark, often used in texture extraction.

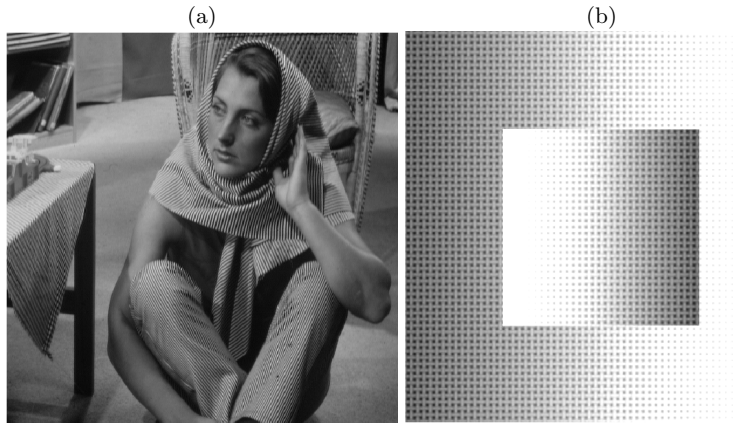


Fig. 1. Original images.

We perform many tests with respect to the different parameters. We only present here the most significant : $\alpha = 0.6$, $\lambda = 0.5$ and $\mu = 100$. Let us first report on the iterations number effect with image (a).

If we are only interested in the texture part, we can observe in fig 2 that we get back all the texture. Unfortunately, most of geometrical information (that we don’t want to get back here) is also kept, and we observe the involved geometric part is getting more important as the iteration number is growing.



Fig. 2. Number of iterations: first line: 60; second line: 200; third line: 2000.

We see in fig 3 that we can choose a large number of iterations for the image (b) texture extraction because of the inner structure of image (b). On the other hand, we have to limit this number for the image (a). We give an example image (b) with a very large iterations number:

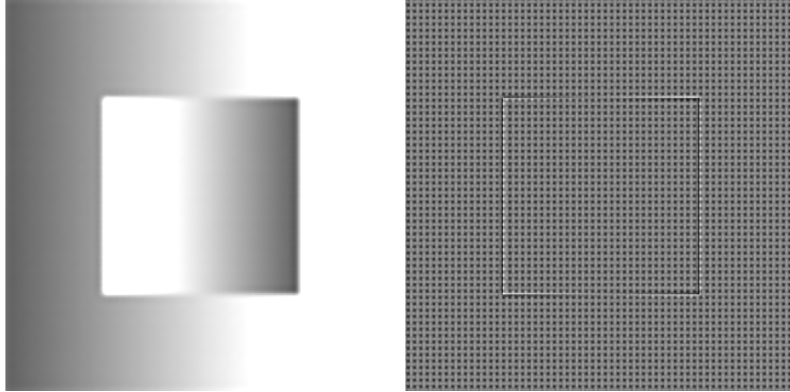


Fig. 3. Number of iterations: 2000.

In addition, we see that the too many geometrical information remains together with the texture in the oscillating part: this is a bad point. Nevertheless, our main goal is to locate the texture and we don't need to work with the cartoon part anymore once it has been identified. We do not need to recover all the texture but only a significant part to identify it. In that case, we propose a method that permits to improve the results significantly: we modify the hessian operator to make it anisotropic. More precisely, we reinforce chosen directions. As texture is made of oscillating information, we hope that we shall keep most of it while many contour lines disappear. We specially, act on the vertical and horizontal components of the hessian operator. To deal with non vertical and horizontal lines, we just have to let the image rotate. In the following test, we have replaced the hessian operator by the operator H' defined for all $v \in X^4$ by :

$$\forall (i, j) \in \{1, \dots, N\}^2, (H'v)_{i,j} = (0, (Hv)_{i,j}^{12}, (Hv)_{i,j}^{21}, 0).$$

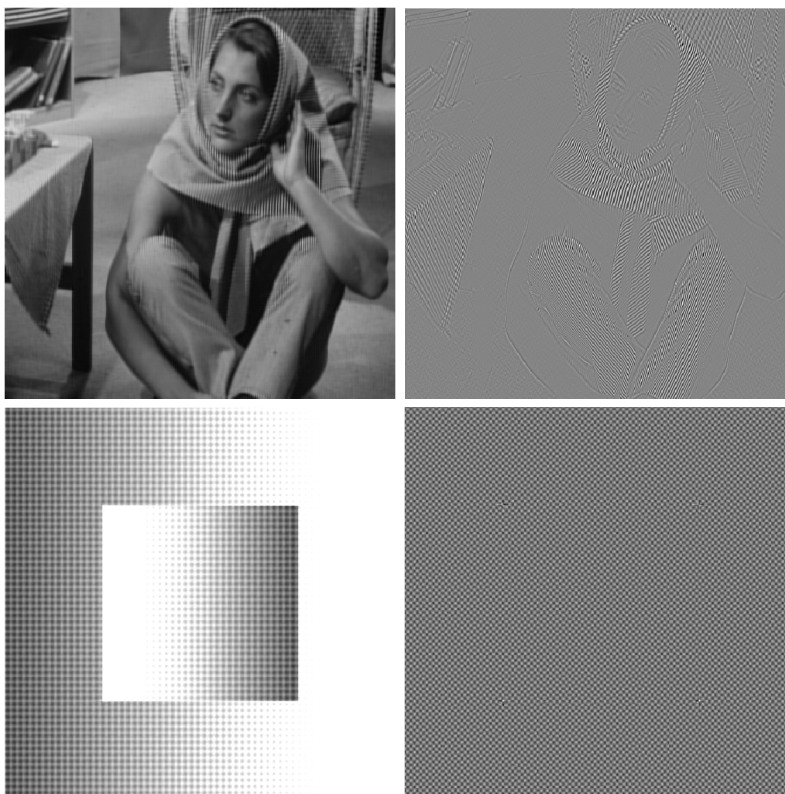


Fig. 4. Test with the anisotropic operator H' . Number of iterations: first line: 60; second line: 2000.

We can see on fig 4 that we keep most of the texture without geometrical information. Of course, this method is only efficient on contour lines which are beelines.

4 Conclusion

The model permits to extract texture from an image, but the texture part still contains too much geometric information. Thus, to recover what we are interested in, we have to use the algorithm with a limited number of iterations. Moreover, we have noticed that we recover too many contour lines as well. The asset of this model is that we can make it anisotropic, modifying the hessian operator in an appropriate way. Therefore we get rid of geometrical

information, but we lose part of the texture as well. Nevertheless, if our goal is just to locate texture on an image, this loss remains acceptable.

References

1. Acar R, Vogel C R (1994) Analysis of bounded variation penalty methods for ill-posed problems. *Inverse Problems*, 10: 1217–1229
2. Ambrosio L, Fusco N, Pallara D (2000) *Functions of bounded variation and free discontinuity problems*. Oxford mathematical monographs, Oxford University Press.
3. Attouch H, Buttazzo, Michaille G (2006) *Variational analysis in Sobolev and BV spaces : applications to PDEs and optimization*. MPS-SIAM series on optimization
4. Aubert G, Kornprobst P (2006) *Mathematical Problems in Image Processing, Partial Differential Equations and the Calculus of Variations*. Applied Mathematical Sciences 147, Springer Verlag
5. Aubert G, Aujol J-F (2005) Modeling very oscillating signals, application to image processing. *Applied Mathematics and Optimization*, 51: 163-182
6. Aubert G, Aujol J-F, Blanc-Feraud L, Chambolle A (2005) Image decomposition into a bounded variation component and an oscillating component. *Journal of Mathematical Imaging and Vision*, 22: 71–88
7. Bergounioux M, Piffet L (2010) A $BV^2(\Omega)$ model for image denoising and/or texture extraction , submitted, *Set Valued Analysis*
8. Chambolle A (2004) An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*, 20: 89–97
9. Demengel F (1984) Fonctions à hessien borné. *Annales de l’institut Fourier*, 34: 155–190
10. Meyer Y (2002) *Oscillating patterns in image processing and nonlinear evolution equations*. Vol. 22 of University Lecture Series, AMS
11. Osher S, Fatemi E, Rudin L (1992) Nonlinear total variation based noise removal algorithms. *Physica D* 60:259–268
12. Osher S, Vese L, (2004) Image denoising and decomposition with total variation minimization and oscillatory functions. Special issue on mathematics and image analysis. *Journal of Mathematical Imaging and Vision*, 20: 7–18
13. Piffet L (2010) *Modèles variationnels pour l’extraction de textures 2D*, PhD Thesis, Université d’Orléans
14. Yin W, Goldfarb D, Osher S (2007) A comparison of three total variation based texture extraction models. *Journal of Visual Communication and Image Representation*, 18: 240–252

Annexe B

Article sva

A Second-Order Model for Image Denoising

Maitine Bergounioux · Loic Piffet

Received: 7 August 2010 / Accepted: 9 August 2010
© Springer Science+Business Media B.V. 2010

Abstract We present a variational model for image denoising and/or texture identification. Noise and textures may be modelled as oscillating components of images. The model involves a L^2 -data fitting term and a Tychonov-like regularization term. We choose the BV^2 norm instead of the classical BV norm. Here BV^2 is the bounded hessian function space that we define and describe. The main improvement is that we do not observe staircasing effects any longer, during denoising process. Moreover, texture extraction can be performed with the same method. We give existence results and present a discretized problem. An algorithm close to the one set by Chambolle (J Math Imaging Vis 20:89–97, 2004) is used: we prove convergence and present numerical tests.

Keywords Second order total variation · Image reconstruction · Denoising · Texture · Variational method

Mathematics Subject Classifications (2010) 65D18 · 68U10 · 65K10

1 Introduction

Variational models in image processing have been extensively studied during the past decade. There are used for segmentation processes (geodesic or geometric contours) and restoration purpose as well. We are mainly interested in the last item which

M. Bergounioux (✉) · L. Piffet
UFR Sciences, Math., Labo. MAPMO, UMR 6628, Université d'Orléans,
Route de Chartres, BP 6759, 45067 Orléans cedex 2, France
e-mail: maitine.bergounioux@univ-orleans.fr

L. Piffet
e-mail: loic.piffet@univ-orleans.fr

involves denoising or deblurring methods and textures extraction as well. Roughly speaking image restoration problems are severely ill posed and a Tychonov-like regularization process is needed. The general form of such models consists in the minimization of an “energy” functional:

$$\mathcal{F}(u) = \|u - u_d\|_X + \mathcal{R}(u), \quad u \in Y \subset X,$$

where X, Y are (real) Banach spaces, \mathcal{R} is a regularization operator, u_d is the observed (or measured) image and u is the image to recover or to denoise. The first term is the fitting data term and the second one permits to get a problem which is no longer ill posed via a regularization process. The most famous model is the Rudin–Osher–Fatemi denoising model (see [1, 16]). This model involves a regularization term that preserves discontinuities, what a classical H^1 -Tychonov regularization method does not. The observed image to recover is split in two parts $u_d = u + v$ where u represents the oscillating component (noise or texture) and v is the smooth part (often called the *cartoon* component). So we look for the solution as $u + v$ with $v \in BV(\Omega)$ and $u \in L^2(\Omega)$, where $BV(\Omega)$ is the functions of bounded variation space defined on an open set Ω [2, 4]. The regularization term involves only the cartoon component v , while the remainder term $u_d - v$ represents the noise to be minimized. We get

$$\min_{v \in BV(\Omega)} \mu |v|_{BV(\Omega)} + \frac{1}{2} \|u_d - v\|_{L^2(\Omega)}^2, \quad (\mathcal{P}_{\text{ROF}})$$

where $\mu > 0$. This problem has a unique solution in $BV(\Omega)$. This functional space is the good one to deal with discontinuous functions that imply that the derivative (in the distribution sense) may be a measure (for example a Dirac measure).

This model is used for denoising purpose. However, the use of the BV norm implies numerical perturbations. The computed solution turns to be piecewise constant which is called the “staircasing effect”. Therefore, though noise can be successfully removed, the solution is not satisfactory. This variational model has been improved using different functional spaces, for the data fitting term or the regularizing term.

Recently people considered that an image can be decomposed into many components, each component describing a particular property of the image ([5, 6, 17–19] and references therein for example). We shall assume as well that the image we want to recover from the data u_d can be decomposed as $f = u + v$ where u and v are functions that characterize different parts of f (see [6, 19, 22] for example).

Components u and v belong to different functional spaces: u is the non regular part and belongs to $L^2(\Omega)$ while v is a more regular part and belongs to $BV^2(\Omega)$ (that we define in the sequel). The remainder term $u = f - v$ involves the oscillating component (as noise and/or texture) and possibly contours. Such decompositions have been already performed [5–7] using the so called Meyer-space of oscillating functions G [15] instead of $BV^2(\Omega)$. So far, the modelling we propose is not the same: the oscillating component will be a priori included in the non regular remainder term part $u := f - v$ while v involves the cartoon. Our philosophy is different than the one used in Meyer approach. Here we perform a second order analysis to get sharper result so that we look for the smooth (cartoon) part. The oscillating part, texture

and/or noise and possible contours will be part of the remainder term and are not modelled a priori.

The paper is organized as follows: in the next section, we present the functional framework and the space $BV^2(\Omega)$ with useful properties. Section 3 is devoted to the variational model. In Section 4, we focus on the discretization process. We present a Chambolle [9] like algorithm in Section 5 and numerical tests are reported in the last section.

2 The Space $BV^2(\Omega)$

Let Ω be an open bounded subset of \mathbb{R}^n , $n \geq 2$ (practically $n = 2$) smooth enough (with the cone property and C^1 for example). Following Demengel [10], we define the space of bounded hessian functions that we call $BV^2(\Omega)$. We first recall the definition and the main properties of the space $BV(\Omega)$ of functions of bounded variation (see [2, 4, 7] for example), defined by

$$BV(\Omega) = \{u \in L^1(\Omega) \mid \Phi_1(u) < +\infty\},$$

where

$$\Phi_1(u) := \sup \left\{ \int_{\Omega} u(x) \operatorname{div} \xi(x) \, dx \mid \xi \in C_c^1(\Omega), \|\xi\|_{\infty} \leq 1 \right\}. \quad (1)$$

The space $BV(\Omega)$, endowed with the norm $\|u\|_{BV(\Omega)} = \|u\|_{L^1} + \Phi_1(u)$, is a Banach space. The derivative in the sense of distributions of every $u \in BV(\Omega)$ is a bounded Radon measure, denoted Du , and $\Phi_1(u) = \int_{\Omega} |Du|$ is the total variation of u . We next recall standard properties of functions of bounded variation [2, 4].

Proposition 1 *Let Ω be an open subset of \mathbb{R}^n with Lipschitz boundary.*

1. *For every $u \in BV(\Omega)$, the Radon measure Du can be decomposed into $Du = \nabla u \, dx + D^s u$, where $\nabla u \, dx$ is the absolutely continuous part of Du with respect of the Lebesgue measure and $D^s u$ is the singular part.*
2. *The mapping $u \mapsto \Phi_1(u)$ is lower semi-continuous (denoted in short lsc) from $BV(\Omega)$ to \mathbb{R}^+ for the $L^1(\Omega)$ topology.*
3. *$BV(\Omega) \subset L^2(\Omega)$ with continuous embedding, if $n = 2$.*
4. *$BV(\Omega) \subset L^p(\Omega)$ with compact embedding, for every $p \in [1, 2)$, if $n = 2$.*

Now we extend this definition to the second derivative (in the distributional sense). Recall that the Sobolev space is defined as

$$W^{1,1}(\Omega) = \{u \in L^1(\Omega) \mid \nabla u \in L^1(\Omega)\}$$

where ∇u stands for the first order derivative of u (in the sense of distributions).

Definition 1 A function $u \in W^{1,1}(\Omega)$ is Hessian bounded if

$$\Phi_2(u) := \sup \left\{ \int_{\Omega} \langle \nabla u, \operatorname{div}(\xi) \rangle_{\mathbb{R}^n} \mid \xi \in \mathcal{C}_c^2(\Omega, \mathbb{R}^{n \times n}), \|\xi\|_{\infty} \leq 1 \right\} < \infty,$$

where

$$\operatorname{div}(\xi) = (\operatorname{div}(\xi_1), \operatorname{div}(\xi_2), \dots, \operatorname{div}(\xi_n)), \quad (2)$$

with

$$\forall i, \xi_i = (\xi_i^1, \xi_i^2, \dots, \xi_i^n) \in \mathbb{R}^n \quad \text{and} \quad \operatorname{div}(\xi_i) = \sum_{k=1}^n \frac{\partial \xi_i^k}{\partial x_k}.$$

$BV^2(\Omega)$ is defined as the space of $W^{1,1}(\Omega)$ functions such that $\Phi_2(u) < +\infty$.

Remark 1 If $V = \mathbb{R}^{n \times n}$, $\|\xi\|_{\infty} = \sup_{x \in \Omega} \sqrt{\sum_{i,j=1}^n |\xi_i^j(x)|^2}$.

We give thereafter many useful properties of $BV^2(\Omega)$ (proofs can be found in [10, 20]).

Theorem 1 The space $BV^2(\Omega)$ endowed with the following norm

$$\|f\|_{BV^2(\Omega)} := \|f\|_{W^{1,1}(\Omega)} + \Phi_2(f) = \|f\|_{L^1} + \|\nabla f\|_{L^1} + \Phi_2(f), \quad (3)$$

where Φ_2 is given by Eq. 2, is a Banach space.

Proposition 2 A function u belongs to $BV^2(\Omega)$ if and only if $u \in W^{1,1}(\Omega)$ and $\frac{\partial u}{\partial x_i} \in BV(\Omega)$ for $i \in \{1, \dots, n\}$. In particular

$$\Phi_2(u) \leq \sum_{i=1}^n \Phi_1\left(\frac{\partial u}{\partial x_i}\right) \leq n \Phi_2(u).$$

Remark 2 The previous result shows that

$$BV^2(\Omega) = \left\{ u \in W^{1,1}(\Omega) \mid \forall i \in \{1, \dots, n\}, \frac{\partial u}{\partial x_i} \in BV(\Omega) \right\}.$$

We get a lower semi-continuity result for the semi-norm Φ_2 as well.

Theorem 2 The operator Φ_2 is lower semi-continuous from $BV^2(\Omega)$ endowed with the strong topology of $W^{1,1}(\Omega)$ to \mathbb{R} . More precisely, if $\{u_k\}_{k \in \mathbb{N}}$ is a sequence of $BV^2(\Omega)$ that strongly converges to u in $W^{1,1}(\Omega)$ then

$$\Phi_2(u) \leq \liminf_{k \rightarrow \infty} \Phi_2(u_k).$$

Remark 3 In particular, if $\liminf_{k \rightarrow \infty} \Phi_2(u_k) < \infty$, then $u \in BV^2(\Omega)$.

We have embedding results as well:

Theorem 3 (Demengel [10]) *Assume $n \geq 2$. Then*

$$BV^2(\Omega) \hookrightarrow W^{1,q}(\Omega) \text{ with } q \leq \frac{n}{n-1}, \tag{4}$$

with continuous embedding. Moreover the embedding is compact if $q < \frac{n}{n-1}$. In particular

$$BV^2(\Omega) \hookrightarrow L^q(\Omega) \text{ for } q \leq \frac{n}{n-2} \text{ if } n > 2 \tag{5}$$

$$BV^2(\Omega) \hookrightarrow L^q(\Omega), \quad \forall q \in [1, \infty[, \text{ if } n = 2. \tag{6}$$

In the sequel, we set $n = 2$ and Ω is a subset of \mathbb{R}^2 , so that $BV^2(\Omega) \subset H^1(\Omega)$ with continuous embedding and $BV^2(\Omega) \subset W^{1,1}(\Omega)$ with compact embedding.

3 The Variational Model

We now assume that the image we want to recover from the data u_d can be decomposed as $f = u + v$ where u and v are functions that characterize different parts of f (see [6, 19, 22] for example). Components u and v belong to different functional spaces: v is the (smooth) second order part and belongs to $BV^2(\Omega)$ while u is the remainder term that should involve noise and/or texture.

We consider the following cost functional defined on $BV^2(\Omega)$:

$$F(v) = \frac{1}{2} \|u_d - v\|_{L^2(\Omega)}^2 + \lambda \Phi_2(v) + \delta \|v\|_{W^{1,1}(\Omega)}, \tag{7}$$

where $\lambda, \delta \geq 0$. We are looking for a solution to the optimization problem

$$\inf\{ F(v) \mid v \in BV^2(\Omega) \} \tag{P}$$

The first term $\|u_d - v\|_{L^2(\Omega)}^2$ of F is the fitting data term. Here we have chosen the L^2 -norm for simplicity but any L^p norm can be used ($p \in [2, +\infty)$). We shall investigate in a future work the very case where $p = 1$; indeed we need to develop an approximation process to deal with this additional non differentiable term. Other terms are Tychonov-like regularization terms. The term $\lambda \Phi_2(v) + \delta \|v\|_{W^{1,1}(\Omega)}$ is nothing else that the $BV^2(\Omega)$ norm of v . However, we have split it because the δ -part is not useful from the modelling point of view. It is only necessary to prove existence of solutions. We shall choose $\delta = 0$ for numerical tests.

If the image is noisy, the noise is considered as a texture and will be involved in the remainder term $u := u_d - v$: more precisely v will be the part of the image without the oscillating component, that is the denoised part and u is expected to involve noise, contours (and part of texture). In the sequel we shall focus on the

denoising process. Such an approach has already been used by Hinterberger and Scherzer [14] with the $BV^2(\Omega)$ space. Their algorithm is different from the one we use here. Note that if we decide to split the function we look for in more than two components ($f = u + v + w$, where u is the texture part, v the cartoon and w the noise for example, with appropriate models) then the minimization problem is a structured optimization problem to which standard decomposition methods can be applied (for example alternating minimization (Gauss Seidel) or parallel methods). Therefore, the study of (\mathcal{P}) is quite significant from this point of view. For parallel methods one may consult [3].

First we give a general existence and uniqueness result for problem (\mathcal{P}) .

Theorem 4 *Assume that $\lambda > 0$ and $\delta > 0$. Problem (\mathcal{P}) has a unique solution v .*

Proof We first prove existence by using the direct method in calculus of variations. Let $v_n \in BV^2(\Omega)$ be a minimizing sequence, i.e.

$$\lim_{n \rightarrow +\infty} F(v_n) = \inf\{ F(v) \mid v \in BV^2(\Omega) \} < +\infty.$$

The sequence $(v_n)_{n \in \mathbb{N}}$ is bounded in $BV^2(\Omega)$. With the compactness result of Theorem 3, this yields that $(v_n)_{n \in \mathbb{N}}$ strongly converges (up to a subsequence) in $W^{1,1}(\Omega)$ to $v^* \in BV^2(\Omega)$. Theorem 2 gives the following:

$$\Phi_2(v^*) \leq \liminf_{n \rightarrow +\infty} \Phi_2(v_n).$$

So

$$F(v^*) \leq \liminf_{n \rightarrow +\infty} F(v_n) = \min_{v \in BV^2(\Omega)} F(v),$$

and v^* is a solution to (\mathcal{P}) . Uniqueness is straightforward with the strict convexity of F due to the term $\|u_d - v\|_{L^2(\Omega)}^2$. \square

Remark 4 It is an open question to know if we really need $\delta > 0$. We could expect a Poincaré-Wirtinger inequality in the $BV^2(\Omega)$ -space which is not very difficult to prove using appropriate “density” results. However, the boundary conditions we have to add are not clear (we have to deal with “second order” Neumann boundary conditions). It is not straightforward to get the suitable boundary conditions that allow to assert that the semi-norm Φ_2 is equivalent to the $BV^2(\Omega)$ -norm. If we assume $\delta = 0$ we are not sure to prove existence of a solution without additional assumptions and we lose uniqueness.

4 The Discretized Problem

Problem (\mathcal{P}) can be written equivalently as

$$\inf_{v \in BV^2(\Omega)} \frac{\|u_d - v\|_{L^2(\Omega)}^2}{2\lambda} + \Phi_2(v) + \delta \|v\|_{W^{1,1}(\Omega)}, \quad (\tilde{\mathcal{P}})$$

where δ has been replaced by $\frac{\delta}{\lambda}$. We are going to compute the solution numerically. We first present the discretization process.

4.1 Discretization of Problem $\tilde{\mathcal{P}}$

We assume for simplicity that the image is squared with size $N \times N$. We note $X := \mathbb{R}^{N \times N} \simeq \mathbb{R}^{N^2}$ endowed with the usual inner product and the associated Euclidean norm

$$\langle u, v \rangle_X := \sum_{1 \leq i, j \leq N} u_{i,j} v_{i,j}, \quad \|u\|_X := \sqrt{\sum_{1 \leq i, j \leq N} u_{i,j}^2}. \tag{8}$$

We set $Y = X \times X$. It is classical to define the discrete total variation as following (see for example [7]): the discrete gradient of the numerical image $u \in X$ is $\nabla u \in Y$:

$$(\nabla u)_{i,j} = \left((\nabla u)_{i,j}^1, (\nabla u)_{i,j}^2 \right), \tag{9}$$

where $(\nabla u)_{i,j}^1 = \begin{cases} u_{i+1,j} - u_{i,j} & \text{if } i < N \\ 0 & \text{if } i = N, \end{cases}$ and $(\nabla u)_{i,j}^2 = \begin{cases} u_{i,j+1} - u_{i,j} & \text{if } j < N \\ 0 & \text{if } j = N. \end{cases}$ The (discrete) total variation corresponding to $\Phi_1(u)$ is given by

$$J_1(u) = \sum_{1 \leq i, j \leq N} \|(\nabla u)_{i,j}\|_{\mathbb{R}^2}, \tag{10}$$

where $\|(\nabla u)_{i,j}\|_{\mathbb{R}^2} = \left\| \left((\nabla u)_{i,j}^1, (\nabla u)_{i,j}^2 \right) \right\|_{\mathbb{R}^2} = \sqrt{\left((\nabla u)_{i,j}^1 \right)^2 + \left((\nabla u)_{i,j}^2 \right)^2}$.

The discrete divergence operator div is the opposite of the adjoint operator of the gradient operator ∇ :

$$\forall (p, u) \in Y \times X, \quad \langle -\text{div } p, u \rangle_X = \langle p, \nabla u \rangle_Y,$$

so that

$$(\text{div } p)_{i,j} = \begin{cases} p_{i,j}^1 - p_{i-1,j}^1 & \text{if } 1 < i < N \\ p_{i,j}^1 & \text{if } i = 1 \\ -p_{i-1,j}^1 & \text{if } i = N \end{cases} + \begin{cases} p_{i,j}^2 - p_{i,j-1}^2 & \text{if } 1 < j < N \\ p_{i,j}^2 & \text{if } j = 1 \\ -p_{i,j-1}^2 & \text{if } j = N. \end{cases} \tag{11}$$

To define a discrete version of the second order total variation Φ_2 we have to introduce the discrete Hessian operator. For any $v \in X$, the Hessian matrix of v , denoted

Hv is identified to a X^4 vector: $(Hv)_{i,j} = \left((Hv)_{i,j}^{11}, (Hv)_{i,j}^{12}, (Hv)_{i,j}^{21}, (Hv)_{i,j}^{22} \right)$, with, for every $i, j = 1, \dots, N$,

$$(Hv)_{i,j}^{11} = \begin{cases} v_{i+1,j} - 2v_{i,j} + v_{i-1,j} & \text{if } 1 < i < N, \\ v_{i+1,j} - v_{i,j} & \text{if } i = 1, \\ v_{i-1,j} - v_{i,j} & \text{if } i = N, \end{cases}$$

$$(Hv)_{i,j}^{12} = \begin{cases} v_{i,j+1} - v_{i,j} - v_{i-1,j+1} + v_{i-1,j} & \text{if } 1 < i \leq N, 1 \leq j < N, \\ 0 & \text{if } i = 1, \\ 0 & \text{if } i = N, \end{cases}$$

$$(Hv)_{i,j}^{21} = \begin{cases} v_{i+1,j} - v_{i,j} - v_{i+1,j-1} + v_{i,j-1} & \text{if } 1 \leq i < N, 1 < j \leq N, \\ 0 & \text{if } i = 1, \\ 0 & \text{if } i = N, \end{cases}$$

$$(Hv)_{i,j}^{22} = \begin{cases} v_{i,j+1} - 2v_{i,j} + v_{i,j-1} & \text{if } 1 < j < N, \\ v_{i,j+1} - v_{i,j} & \text{if } j = 1, \\ v_{i,j-1} - v_{i,j} & \text{if } j = N. \end{cases}$$

The discrete second order total variation corresponding to $\Phi_2(v)$ is defined as

$$J_2(v) = \sum_{1 \leq i, j \leq N} \|(Hv)_{i,j}\|_{\mathbb{R}^4}. \quad (12)$$

The discretized problem stands

$$\inf_{v \in X} \frac{\|u_d - v\|_X^2}{2\lambda} + J_2(v) + \delta(|v| + J_1(v)), \quad (\mathcal{P}_d)$$

where

$$|v| := \sum_{1 \leq i, j \leq N} |v_{i,j}|.$$

In the finite dimensional case we have an existence result with $\delta = 0$.

Theorem 5 *Problem (\mathcal{P}_d) has a unique solution for every $\delta \geq 0$ and $\lambda > 0$.*

Proof The cost functional

$$F_\delta := \frac{\|u_d - v\|_X^2}{2\lambda} + J_2(v) + \delta(|v| + J_1(v)),$$

is continuous and coercive because of the term $\|u_d - v\|_X^2$. In addition it is strictly convex so that we get the result. \square

For numerical purpose we shall set $\delta = 0$. In fact, we have performed tests with $\delta = 0$ and very small $\delta \neq 0$ (as required by the theory to get a solution to problem $\tilde{\mathcal{P}}$): results were identical.

In the sequel, we adapt the method by Chambolle in the $BV(\Omega)$ -case, to the second order framework. We briefly recall the original result of [9]. Consider the finite dimensional optimization problem derived from the discretization of the ROF model:

$$\inf_{u \in X} \left(J_1(u) + \frac{1}{2\lambda} \|u_d - u\|_X^2 \right). \tag{P_d^1}$$

The following result holds:

Proposition 3 *The solution to (P_d^1) is given by*

$$u = u_d - P_{\lambda K_1}(u_d), \tag{13}$$

where

$$K_1 = \{ \operatorname{div} g \mid g \in Y, \|g_{i,j}\|_{\mathbb{R}^2} \leq 1 \ \forall i, j = 1, \dots, N \},$$

and $P_{\lambda K_1}$ is the orthogonal projector operator on λK_1 .

We have a very similar result that we present in next sections. We first describe the conjugate function of J_2 .

4.2 The J_2 Legendre–Fenchel Conjugate Function

As the function J_2 is positively homogeneous, the Legendre–Fenchel conjugate

$$J_2^*(v) = \sup_u \langle u, v \rangle_X - J_2(u),$$

is the characteristic function of a closed, convex set K

$$J_2^*(v) = \mathbf{1}_K(v) = \begin{cases} 0 & \text{if } v \in K \\ +\infty & \text{else.} \end{cases}$$

As $J_2^{**} = J_2$, we get $J_2(v) = \sup_{u \in K} \langle v, u \rangle_X$. We use the inner scalar product of X^4 :

$$\langle p, q \rangle_{X^4} = \sum_{1 \leq i, j \leq N} \left(p_{i,j}^1 q_{i,j}^1 + p_{i,j}^2 q_{i,j}^2 + p_{i,j}^3 q_{i,j}^3 + p_{i,j}^4 q_{i,j}^4 \right),$$

for every $p = (p^1, p^2, p^3, p^4), q = (q^1, q^2, q^3, q^4) \in X^4$. So, for every $v \in X$,

$$J_2(v) = \sup_{p \in \mathcal{C}} \langle p, Hv \rangle_{X^4}, \tag{14}$$

where the feasible set is

$$\mathcal{C} := \{ p \in X^4 \mid \|p_{i,j}\|_{\mathbb{R}^4} \leq 1, \ \forall 1 \leq i, j \leq N \}.$$

Let us compute the adjoint operator of H (which is the discretized “second divergence” operator) :

$$\forall p \in X^4, \forall v \in X \quad \langle H^* p, v \rangle_X = \langle p, Hv \rangle_{X^4}.$$

We verify that $H^* : X^4 \rightarrow X$ satisfies for every $p = (p^{11}, p^{12}, p^{21}, p^{22}) \in X^4$

$$\begin{aligned} (H^* p)_{i,j} = & \begin{cases} p_{i-1,j}^{11} - 2p_{i,j}^{11} + p_{i+1,j}^{11} & \text{if } 1 < i < N \\ p_{i+1,j}^{11} - p_{i,j}^{11} & \text{if } i = 1, \\ p_{i-1,j}^{11} - p_{i,j}^{11} & \text{if } i = N, \end{cases} \\ & + \begin{cases} p_{i,j-1}^{22} - 2p_{i,j}^{22} + p_{i,j+1}^{22} & \text{if } 1 < j < N, \\ p_{i,j+1}^{22} - p_{i,j}^{22} & \text{if } j = 1, \\ p_{i,j-1}^{22} - p_{i,j}^{22} & \text{if } j = N, \end{cases} \\ & + \begin{cases} p_{i,j-1}^{12} - p_{i,j}^{12} - p_{i+1,j-1}^{12} + p_{i+1,j}^{12} & \text{if } 1 < i, j < N, \\ p_{i+1,j}^{12} - p_{i+1,j-1}^{12} & \text{if } i = 1, 1 < j < N, \\ p_{i,j-1}^{12} - p_{i,j}^{12} & \text{if } i = N, 1 < j < N, \\ p_{i+1,j}^{12} - p_{i,j}^{12} & \text{if } 1 < i < N, j = 1, \\ p_{i,j-1}^{12} - p_{i+1,j-1}^{12} & \text{if } 1 < i < N, j = N, \\ p_{i+1,j}^{12} & \text{if } i = 1, j = 1, \\ -p_{i+1,j-1}^{12} & \text{if } i = 1, j = N, \\ -p_{i,j}^{12} & \text{if } i = N, j = 1, \\ p_{i,j-1}^{12} & \text{if } i = N, j = N, \end{cases} \quad (15) \\ & + \begin{cases} p_{i-1,j}^{21} - p_{i,j}^{21} - p_{i-1,j+1}^{21} + p_{i,j+1}^{21} & \text{if } 1 < i, j < N, \\ p_{i,j+1}^{21} - p_{i,j}^{21} & \text{if } i = 1, 1 < j < N, \\ p_{i-1,j}^{21} - p_{i-1,j+1}^{21} & \text{if } i = N, 1 < j < N, \\ p_{i,j+1}^{21} - p_{i-1,j+1}^{21} & \text{if } 1 < i < N, j = 1, \\ p_{i-1,j}^{21} - p_{i,j}^{21} & \text{if } 1 < i < N, j = N, \\ p_{i,j+1}^{21} & \text{if } i = 1, j = 1, \\ -p_{i,j}^{21} & \text{if } i = 1, j = N, \\ -p_{i-1,j+1}^{21} & \text{if } i = N, j = 1, \\ p_{i-1,j}^{21} & \text{if } i = N, j = N, \end{cases} \end{aligned}$$

Theorem 6 *The Legendre–Fenchel conjugate of J_2 is $J_2^* = \mathbf{1}_{K_2}$ where*

$$K_2 := \{ H^* p \mid p \in X^4, \|p_{i,j}\|_{\mathbb{R}^4} \leq 1, \forall i, j = 1, \dots, N \} \subset X. \quad (16)$$

Proof We have already mentioned that $J_2^* = \mathbf{1}_K$ where K is of a non empty, closed, convex set subset of X . We first show that $K_2 \subset K$. Let be $v \in K_2$. The discretized version of the definition of the second order total variation gives

$$J_2(v) = \sup_{\xi \in K_2} \langle \xi, v \rangle_X,$$

and $\langle \xi, v \rangle - J_2(v) \leq 0$ for every $\xi \in K_2$ and $v \in X$. This gives for every $v^* \in K_2$

$$J_2^*(v^*) = \sup_{v \in K_2} \langle v^*, v \rangle - J_2(v) \leq 0. \tag{17}$$

As J_2^* takes only one finite value, we get $J_2^*(v^*) = 0$, which yields that $v^* \in K$. Therefore, $K_2 \subset K$ and as K is closed we finally obtain

$$\bar{K}_2 \subset K.$$

In particular

$$J_2(v) = \sup_{\xi \in K_2} \langle v, \xi \rangle \leq \sup_{\xi \in \bar{K}_2} \langle v, \xi \rangle \leq \sup_{\xi \in K} \langle v, \xi \rangle = \sup_{\xi \in K} \langle v, \xi \rangle - J_2^*(\xi) = J_2^{**}(v).$$

As $J_2^{**} = J_2$, we have

$$\sup_{\xi \in K_2} \langle v, \xi \rangle = \sup_{\xi \in \bar{K}_2} \langle v, \xi \rangle = \sup_{\xi \in K} \langle v, \xi \rangle. \tag{18}$$

Now, let us assume there exists $v^* \in K$ such that $v^* \notin \bar{K}_2$. One may strictly separate v^* and the closed convex set \bar{K}_2 . There exists $\alpha \in \mathbb{R}$ and v_0 such that

$$\langle v_0, v^* \rangle > \alpha \geq \sup_{u \in \bar{K}_2} \langle v_0, u \rangle.$$

We obtain

$$\sup_{\xi \in K} \langle v_0, \xi \rangle \geq \langle v_0, v^* \rangle > \alpha \geq \sup_{u \in \bar{K}_2} \langle v_0, u \rangle = \sup_{u \in K} \langle v_0, u \rangle,$$

that gives a contradiction. We have proved that $K = \bar{K}_2$. As K_2 is closed we get $K = K_2$. □

Remark 5 We proved the previous result for the convenience of the reader. Indeed, one may note that J_2 is the support function of K_2 which is the conjugate function of the indicator function 1_{K_2} of K_2 (see [12] p. 19). Therefore, as K_2 is closed and convex and J_2 is continuous we get $J_2^* = 1_{K_2}^{**} = 1_{K_2}$.

4.3 Case Where $\delta = 0$

Now we focus on the case where $\delta = 0$. Indeed, the δ -term in Definition 7 was needed as a tool in the infinite dimensional framework to prove existence of solutions to (\mathcal{P}) .

However, we have seen that the finite dimensional problem has a solution even if $\delta = 0$, which is the most interesting case. The problem we have to solve turns to be

$$\min_{v \in X} \frac{\|u_d - v\|_X^2}{2\lambda} + J_2(v). \tag{P_d^2}$$

As in the BV-case (Proposition 3) we have a characterization of the solution.

Theorem 7 *The solution v of (P_d^2) verifies*

$$v = u_d - P_{\lambda K_2}(u_d),$$

where $P_{\lambda K_2}$ is the orthogonal projector operator on λK_2 .

Proof The proof is similar to the one by Chambolle [9] but we give it for convenience. The solution v to (P_d^2) is characterized by

$$0 \in \partial \left(J_2(v) + \frac{1}{2\lambda} \|v - u_d\|_2^2 \right) = \frac{v - u_d}{\lambda} + \partial J_2(v),$$

that is $\frac{u_d - v}{\lambda} \in \partial J_2(v)$. As J_2 is proper, convex and continuous, then

$$v^* \in \partial J_2(v) \iff v \in \partial J_2^*(v^*).$$

So

$$v \in \partial J_2^* \left(\frac{u_d - v}{\lambda} \right) \iff 0 \in -v + \partial J_2^* \left(\frac{u_d - v}{\lambda} \right) \iff 0 \in \frac{u_d - v}{\lambda} - \frac{u_d}{\lambda} + \frac{1}{\lambda} \partial J_2^* \left(\frac{u_d - v}{\lambda} \right).$$

This means that $w = \frac{u_d - v}{\lambda}$ is the solution to

$$\min_w \frac{1}{2} \left\| w - \frac{u_d}{\lambda} \right\|_X^2 + \frac{1}{\lambda} \partial J_2^*(w).$$

As $J_2^* = \mathbf{1}_{K_2}$ this implies that $\frac{u_d - v}{\lambda}$ is the orthogonal projection of $\frac{u_d}{\lambda}$ on K_2 :

$$\frac{u_d - v}{\lambda} = P_{K_2} \left(\frac{u_d}{\lambda} \right).$$

As $P_{K_2}(\frac{u_d}{\lambda}) = \frac{1}{\lambda} P_{\lambda K_2}(u_d)$, this ends the proof. □

5 A Fixed-Point Algorithm to Compute ∂J_2

In [9] Chambolle proposed a fixed-point algorithm to compute $P_{\lambda K_1}(f)$ (and $\partial J_1(f)$). Let us briefly recall the main idea that we use again in the BV^2 context. To compute $P_{\lambda K_1}(f)$ we have to solve

$$\min \{ \|\lambda \operatorname{div} p - u_d\|_X^2 \mid \|p_{i,j}\|_{\mathbb{R}^2} \leq 1 \ \forall i, j = 1, \dots, N \}. \tag{19}$$

that can be solved with a fixed-point method:

$$p^0 = 0 \tag{20a}$$

$$p_{i,j}^{n+1} = \frac{p_{i,j}^n + \tau (\nabla (\operatorname{div} p^n - u_d/\lambda))_{i,j}}{1 + \tau \|(\nabla (\operatorname{div} p^n - u_d/\lambda))_{i,j}\|_{\mathbb{R}^2}} \tag{20b}$$

In addition, a convergence result is provided:

Theorem 8 ([9], Theorem 3.1) *Assume that τ satisfies $\tau \leq 1/8$. Then $\lambda \operatorname{div} p^n$ converges to $P_{\lambda K_1}(u_d)$ as $n \rightarrow +\infty$.*

Remark 6 In [8] J.F. Aujol proves that the modified algorithm:

$$p^0 = 0 \tag{21a}$$

$$p_{i,j}^{n+1} = \frac{p_{i,j}^n + \tau (\nabla (\operatorname{div} p^n - u_d/\lambda))_{i,j}}{\max\left(1, \|p_{i,j}^n + \tau (\nabla (\operatorname{div} p^n - u_d/\lambda))_{i,j}\|_{\mathbb{R}^2}\right)} \tag{21b}$$

converges if $\tau \leq 1/4$.

We extend this result to the second-order case. To compute $P_{\lambda K_2}(u_d)$ we have to solve

$$\min \left\{ \|\lambda H^* p - u_d\|_X^2 \mid p \in X^4, \|p_{i,j}\|_{\mathbb{R}^4}^2 - 1 \leq 0, i, j = 1, \dots, N \right\}. \tag{P'}$$

Let us denote $F(p) = \|\lambda H^* p - u_d\|_X^2$ and

$$g_{i,j}(p) = \|p_{i,j}\|_{\mathbb{R}^4}^2 - 1 = (p_{i,j}^{11})^2 + (p_{i,j}^{12})^2 + (p_{i,j}^{21})^2 + (p_{i,j}^{22})^2 - 1.$$

First order optimality conditions give the existence of Lagrange multipliers $\alpha_{i,j}$, $(i, j) \in \{1, \dots, N\}^2$, such that

$$\nabla F(p) + \sum_{i,j=1}^N \alpha_{i,j} \nabla g_{i,j}(p) = 0, \tag{22a}$$

$$\alpha_{i,j} \geq 0 \text{ and } \alpha_{i,j} g_{i,j}(p) = 0, (i, j) \in \{1, \dots, N\}^2. \tag{22b}$$

It is easy to see that $\nabla F(p) = 2\lambda H [\lambda H^* p - u_d]$ and that

$$\sum_{i,j=1}^N \alpha_{i,j} \nabla g_{i,j}(p) = 2\alpha_{i,j} \left((p_{i,j}^{11}, p_{i,j}^{22}, p_{i,j}^{12}, p_{i,j}^{21}) \right)_{1 \leq i, j \leq N}.$$

Therefore relations (22) are equivalent to

$$\forall 1 \leq i, j \leq N \quad (H[\lambda H^* p - u_d])_{i,j} + \alpha_{i,j} p_{i,j} = 0, \tag{23a}$$

$$\forall 1 \leq i, j \leq N \quad \alpha_{i,j} \geq 0 \text{ and } \alpha_{i,j} g_{i,j}(p) = 0. \tag{23b}$$

Let us compute the multipliers $\alpha_{i,j}$ more precisely:

- If $\alpha_{i,j} > 0$ then $\|p_{i,j}\|_{\mathbb{R}^4} = 1$.
- If $\alpha_{i,j} = 0$ then $(H[\lambda H^* p - u_d])_{i,j} = 0$.

In both cases we get

$$\forall 1 \leq i, j \leq N \quad \alpha_{i,j} = \left\| (H[\lambda H^* p - u_d])_{i,j} \right\|_{\mathbb{R}^4}$$

and we finally obtain the following equality: $\forall (i, j) \in \{1, \dots, N\}^2$,

$$(H[\lambda H^* p - u_d])_{i,j} + \left\| (H[\lambda H^* p - u_d])_{i,j} \right\|_{\mathbb{R}^4} p_{i,j} = 0. \tag{24}$$

We use a semi-implicit gradient method to solve these equations, namely:

Choose $\tau > 0$, and

1. Set $p^0 = 0, n = 0$
2. p^n is known. For $(i, j) \in \{1, \dots, N\}^2$, compute $p_{i,j}^{n+1}$ with

$$p_{i,j}^n = p_{i,j}^{n+1} + \tau \left((H[H^* p^n - u_d/\lambda])_{i,j} + \left\| (H[H^* p^n - u_d/\lambda])_{i,j} \right\|_{\mathbb{R}^4} p_{i,j}^{n+1} \right).$$

This is equivalent to

$$p_{i,j}^{n+1} = \frac{p_{i,j}^n - \tau (H[H^* p^n - u_d/\lambda])_{i,j}}{1 + \tau \left\| (H[H^* p^n - u_d/\lambda])_{i,j} \right\|_{\mathbb{R}^4}}. \tag{25}$$

The algorithm step τ is related to the adjoint operator H^* norm that we call κ in the sequel and we first give a κ estimate:

Lemma 1 *The adjoint operator H^* norm, κ satisfies $\kappa \leq 8$.*

Proof The definition of κ reads $\kappa = \sup_{\|p\|_{X^4} \leq 1} \|H^* p\|$.

As $\|H^* p\|_X = \sup_{q \in \mathcal{B}_X(0,1)} \langle H^* p, q \rangle_X$ and $\forall q \in X^4 \langle H^* p, q \rangle_X = \langle p, Hq \rangle_{X^4} \leq \|Hq\|_{X^4} \|p\|_{X^4}$, we get

$$\|H^* p\|_X \leq \|H\| \|p\|_{X^4}. \tag{26}$$

For any $q \in X^4$

$$\begin{aligned} \|Hq\|_{X^4}^2 &= \sum_{1 \leq i, j \leq N} \left[(q_{i+1, j} - 2q_{i, j} + q_{i-1, j})^2 + (q_{i, j+1} - q_{i, j} - q_{i-1, j+1} + q_{i-1, j})^2 \right. \\ &\quad \left. + (q_{i+1, j} - q_{i, j} - q_{i+1, j-1} + q_{i, j-1})^2 + (q_{i, j+1} - 2q_{i, j} + q_{i, j-1})^2 \right] \\ &\leq 4 \sum_{1 \leq i, j \leq N} \left[q_{i+1, j}^2 + q_{i, j}^2 + q_{i, j}^2 + q_{i-1, j}^2 + q_{i, j+1}^2 + q_{i, j}^2 + q_{i-1, j+1}^2 + q_{i-1, j}^2 \right. \\ &\quad \left. + q_{i+1, j}^2 + q_{i, j}^2 + q_{i+1, j-1}^2 + q_{i, j-1}^2 + q_{i, j+1}^2 + q_{i, j}^2 + q_{i, j}^2 + q_{i, j-1}^2 \right] \\ &\leq 4 \times 16 \|q\|_X^2. \end{aligned}$$

Finally $\|H\| \leq 8$, and with relation (26), $\|H^* p\|_X \leq 8 \|p\|_{X^4}$. We deduce that $\kappa \leq 8$ (and $\kappa^2 \leq 64$). \square

Theorem 9 *Let be $\tau \leq 1/64$. Then $\lambda (H^* p^n)_n$ converges to $P_{\lambda K_2}(u_d)$ as $n \rightarrow +\infty$.*

Proof It is easy to check (by induction) that for every $n \geq 0$ and i, j , $\|p_{i, j}^n\|_{\mathbb{R}^4} \leq 1$. Set $n > 0$ and let be $\eta^n = (p^n - p^{n+1})/\tau$. Denote by $(U_n)_n$ the sequence defined by:

$$U_n = \|H^* p^n - u_d/\lambda\|_X^2, \quad \forall n \geq 0.$$

The outline of the proof is the following: we first prove that $(U_n)_n$ is a (nonnegative) non increasing sequence and therefore it is convergent. The proof provides an equality (29) that allows to show that the sequences $(p^n)_n$ and $(p^{n+1})_n$ have the same cluster points which are the unique solution to a projection problem. Therefore the whole sequence $(p^n)_n$ converges.

- We first prove that $(U_n)_n$ is a non increasing sequence if $\tau \leq 1/\kappa^2$ where κ is the norm of H^* . We get

$$\begin{aligned} U_{n+1} &= \|H^*(p^n - \tau \eta^n) - u_d/\lambda\|_X^2 \\ &= \langle -\tau H^* \eta^n + (H^* p^n - u_d/\lambda), -\tau H^* \eta^n + (H^* p^n - u_d/\lambda) \rangle_X \\ &= U_n - 2\tau \langle H^* \eta^n, H^* p^n - u_d/\lambda \rangle_X + \tau^2 \|H^* \eta^n\|_X^2 \\ &= U_n - 2\tau \langle \eta^n, H [H^* p^n - u_d/\lambda] \rangle_{X^4} + \tau^2 \|H^* \eta^n\|_X^2 \\ &\leq U_n - \tau \left[2 \langle \eta^n, H [H^* p^n - u_d/\lambda] \rangle_{X^4} - \kappa^2 \tau \|\eta^n\|_{X^4}^2 \right] \\ &\leq U_n - \tau \left[\sum_{i, j=1}^N \langle 2\eta_{i, j}^n, (H [H^* p^n - u_d/\lambda])_{i, j} \rangle_{\mathbb{R}^4} - \kappa^2 \tau \|\eta_{i, j}^n\|_{\mathbb{R}^4}^2 \right] \end{aligned}$$

As

$$\eta_{i,j}^n = (H [H^* p^n - u_d/\lambda])_{i,j} + \rho_{i,j}^n,$$

with

$$\rho_{i,j}^n = \left\| (H [\lambda H^* p^n - u_d/\lambda])_{i,j} \right\|_{\mathbb{R}^4} p_{i,j}^{n+1};$$

we obtain

$$\begin{aligned} & \left\langle 2\eta_{i,j}^n, (H [H^* p^n - u_d/\lambda])_{i,j} \right\rangle_{\mathbb{R}^4} - \kappa^2 \tau \left\| \eta_{i,j}^n \right\|_{\mathbb{R}^4}^2 \\ &= \left\langle \eta_{i,j}^n, (H [H^* p^n - u_d/\lambda])_{i,j} \right\rangle_{\mathbb{R}^4} + \left\langle \eta_{i,j}^n, (H [H^* p^n - u_d/\lambda])_{i,j} \right\rangle_{\mathbb{R}^4} \\ &\quad - \kappa^2 \tau \left\| \eta_{i,j}^n \right\|_{\mathbb{R}^4}^2 = \left\langle \eta_{i,j}^n, \eta_{i,j}^n - \rho_{i,j}^n \right\rangle_{\mathbb{R}^4} \\ &\quad + \left\langle (H [H^* p^n - u_d/\lambda])_{i,j} + \rho_{i,j}^n, (H [H^* p^n - u_d/\lambda])_{i,j} \right\rangle_{\mathbb{R}^4} - \kappa^2 \tau \left\| \eta_{i,j}^n \right\|_{\mathbb{R}^4}^2 \\ &= \left\| \eta_{i,j}^n \right\|_{\mathbb{R}^4}^2 - \left\langle \eta_{i,j}^n, \rho_{i,j}^n \right\rangle_{\mathbb{R}^4} + \left\| (H [H^* p^n - u_d/\lambda])_{i,j} \right\|_{\mathbb{R}^4}^2 \\ &\quad + \left\langle \rho_{i,j}^n, (H [H^* p^n - u_d/\lambda])_{i,j} \right\rangle_{\mathbb{R}^4} - \kappa^2 \tau \left\| \eta_{i,j}^n \right\|_{\mathbb{R}^4}^2 \\ &= \left\| \eta_{i,j}^n \right\|_{\mathbb{R}^4}^2 - \left\langle \eta_{i,j}^n, \rho_{i,j}^n \right\rangle_{\mathbb{R}^4} + \left\| (H [H^* p^n - u_d/\lambda])_{i,j} \right\|_{\mathbb{R}^4}^2 \\ &\quad + \left\langle \rho_{i,j}^n, \eta_{i,j}^n - \rho_{i,j}^n \right\rangle_{\mathbb{R}^4} - \kappa^2 \tau \left\| \eta_{i,j}^n \right\|_{\mathbb{R}^4}^2 \\ &= (1 - \kappa^2 \tau) \left\| \eta_{i,j}^n \right\|_{\mathbb{R}^4}^2 + \left(\left\| (H [H^* p^n - u_d/\lambda])_{i,j} \right\|_{\mathbb{R}^4}^2 - \left\| \rho_{i,j}^n \right\|_{\mathbb{R}^4}^2 \right). \end{aligned}$$

Finally

$$U_{n+1} \leq U_n - \tau \left[\sum_{i,j=1}^N (1 - \kappa^2 \tau) \left\| \eta_{i,j}^n \right\|_{\mathbb{R}^4}^2 + \left(\left\| (H [H^* p^n - u_d/\lambda])_{i,j} \right\|_{\mathbb{R}^4}^2 - \left\| \rho_{i,j}^n \right\|_{\mathbb{R}^4}^2 \right) \right] \quad (27)$$

As $\left\| p_{i,j}^{n+1} \right\| \leq 1$, $\left\| \rho_{i,j}^n \right\|_{\mathbb{R}^4} \leq \left\| (H [\lambda H^* p^n - u_d/\lambda])_{i,j} \right\|_{\mathbb{R}^4}$ this yields that if $\tau \leq 1/\kappa^2$, then the sequence (U_n) is non increasing. Thus, the sequence $(U_n)_n$ is convergent to some m . Moreover, relation (27) gives

$$0 \leq \sum_{i,j=1}^N (1 - \kappa^2 \tau) \left\| \eta_{i,j}^n \right\|_{\mathbb{R}^4}^2 + \left(\left\| (H [H^* p^n - u_d/\lambda])_{i,j} \right\|_{\mathbb{R}^4}^2 - \left\| \rho_{i,j}^n \right\|_{\mathbb{R}^4}^2 \right) \leq \frac{U_n - U_{n+1}}{\tau}. \quad (28)$$

Passing to the limit in relation (28) gives

$$\lim_{n \rightarrow +\infty} \sum_{i,j=1}^N \left[(1 - \kappa^2 \tau) \|\eta_{i,j}^n\|_{\mathbb{R}^4}^2 + \left(\|(H[H^* p^n - u_d/\lambda])_{i,j}\|_{\mathbb{R}^4}^2 - \|\rho_{i,j}^n\|_{\mathbb{R}^4}^2 \right) \right] = 0. \quad (29)$$

- Let us prove that we may extract a subsequence $(p^{n_k})_k$ of $(p^n)_n$ such that $(p^{n_k})_k$ and $(p^{n_{k+1}})_k$ have the same limit. As $(p^n)_n$ is bounded, we call \bar{p} one cluster point and $(p^{n_k})_k$ the corresponding subsequence of $(p^n)_n$ such that $p^{n_k} \rightarrow \bar{p}$, as $k \rightarrow +\infty$. Let us call \bar{p}' the limit (up to a subsequence) of $p^{n_{k+1}}$. We obtain with relation (25)

$$\bar{p}'_{i,j} = \frac{\bar{p}_{i,j} - \tau (H[H^* \bar{p} - u_d/\lambda])_{i,j}}{1 + \tau \|(H[H^* \bar{p} - u_d/\lambda])_{i,j}\|_{\mathbb{R}^4}}. \quad (30)$$

We note $\bar{\rho}$ et $\bar{\eta}$ the respective limits of ρ^{n_k} et η^{n_k} when k tends to $+\infty$. With Eq. 29, we obtain

$$\sum_{i,j=1}^N \left[(1 - \kappa^2 \tau) \|\bar{\eta}_{i,j}\|_{\mathbb{R}^4}^2 + \left(\|(H[H^* \bar{p} - u_d/\lambda])_{i,j}\|_{\mathbb{R}^4}^2 - \|\bar{\rho}_{i,j}\|_{\mathbb{R}^4}^2 \right) \right] = 0.$$

As the terms of the sum are nonnegative, we get

$$\forall i, j \quad (1 - \kappa^2 \tau) \|\bar{\eta}_{i,j}\|_{\mathbb{R}^4}^2 = 0 \quad \text{and} \quad \|(H[H^* \bar{p} - u_d/\lambda])_{i,j}\|_{\mathbb{R}^4}^2 - \|\bar{\rho}_{i,j}\|_{\mathbb{R}^4}^2 = 0.$$

- If $\kappa^2 \tau < 1$, then $\bar{\eta}_{i,j} = 0$ for all i, j , and so $\bar{p}' = \bar{p}$.
- If $\kappa^2 \tau = 1$, then, for all i, j , $\|\bar{\rho}_{i,j}\|_{\mathbb{R}^4} = \|(H[H^* \bar{p} - u_d/\lambda])_{i,j}\|_{\mathbb{R}^4}$, that is to say

$$\|(H[H^* \bar{p} - u_d/\lambda])_{i,j}\|_{\mathbb{R}^4} \|\bar{p}'_{i,j}\|_{\mathbb{R}^4} = \|(H[H^* \bar{p} - u_d/\lambda])_{i,j}\|_{\mathbb{R}^4}.$$

This implies

$$\|\bar{p}'_{i,j}\|_{\mathbb{R}^4} = 1 \quad \text{or} \quad \|(H[H^* \bar{p} - u_d/\lambda])_{i,j}\|_{\mathbb{R}^4} = 0.$$

- If $\|(H[H^* \bar{p} - u_d/\lambda])_{i,j}\|_{\mathbb{R}^4} = 0$, then relation (30) gives $\bar{p}'_{i,j} = \bar{p}_{i,j}$.
- If $\|\bar{p}'_{i,j}\|_{\mathbb{R}^4} = 1$, then

$$\begin{aligned} 1 &= \frac{\|\bar{p}_{i,j} - \tau (H[H^* \bar{p} - u_d/\lambda])_{i,j}\|_{\mathbb{R}^4}}{1 + \tau \|(H[H^* \bar{p} - u_d/\lambda])_{i,j}\|_{\mathbb{R}^4}} \\ &\leq \frac{\|\bar{p}_{i,j}\|_{\mathbb{R}^4} + \tau \|(H[H^* \bar{p} - u_d/\lambda])_{i,j}\|_{\mathbb{R}^4}}{1 + \tau \|(H[H^* \bar{p} - u_d/\lambda])_{i,j}\|_{\mathbb{R}^4}}, \end{aligned}$$

Fig. 1 Test images (“Shapes” (size: 740×740) and “Lena” (size: 512×512))



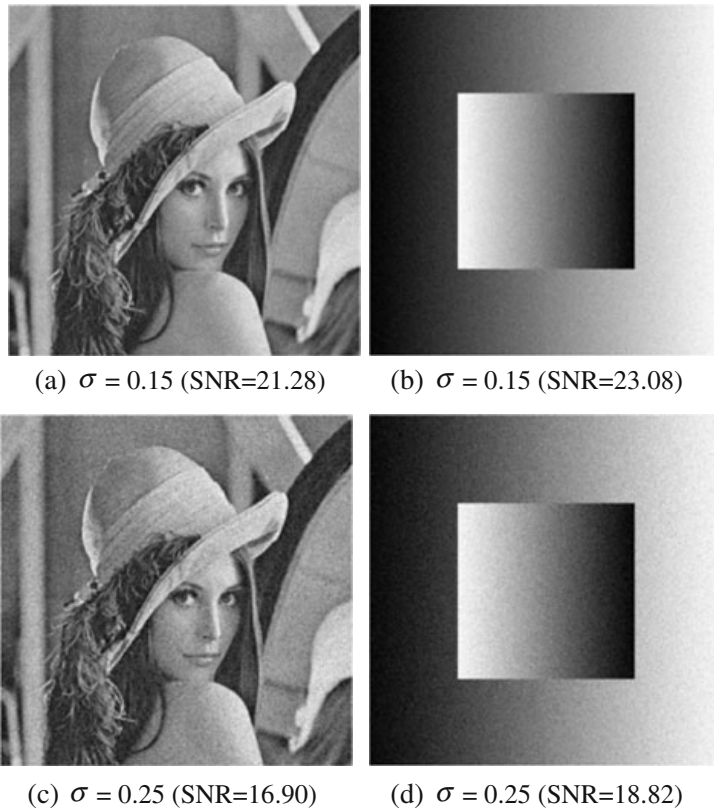
therefore $\|\bar{p}_{i,j}\|_{\mathbb{R}^4} \geq 1$ and (together with $\|\bar{p}_{i,j}\|_{\mathbb{R}^4} \leq 1$) $\|\bar{p}_{i,j}\|_{\mathbb{R}^4} = 1$. We deduce that

$$\|\bar{p}_{i,j} - \tau(H[H^*\bar{p} - u_d/\lambda])_{i,j}\|_{\mathbb{R}^4} = \|\bar{p}_{i,j}\|_{\mathbb{R}^4} + \|\tau(H[H^*\bar{p} - u_d/\lambda])_{i,j}\|_{\mathbb{R}^4}.$$

Since triangular inequality turns to be an equality, there exists $\beta \in \mathbb{R}^*$ so that

$$\tau(H[H^*\bar{p} - u_d/\lambda])_{i,j} = \beta \bar{p}_{i,j}.$$

Fig. 2 Noisy images—white Gaussian noise with standard deviation σ



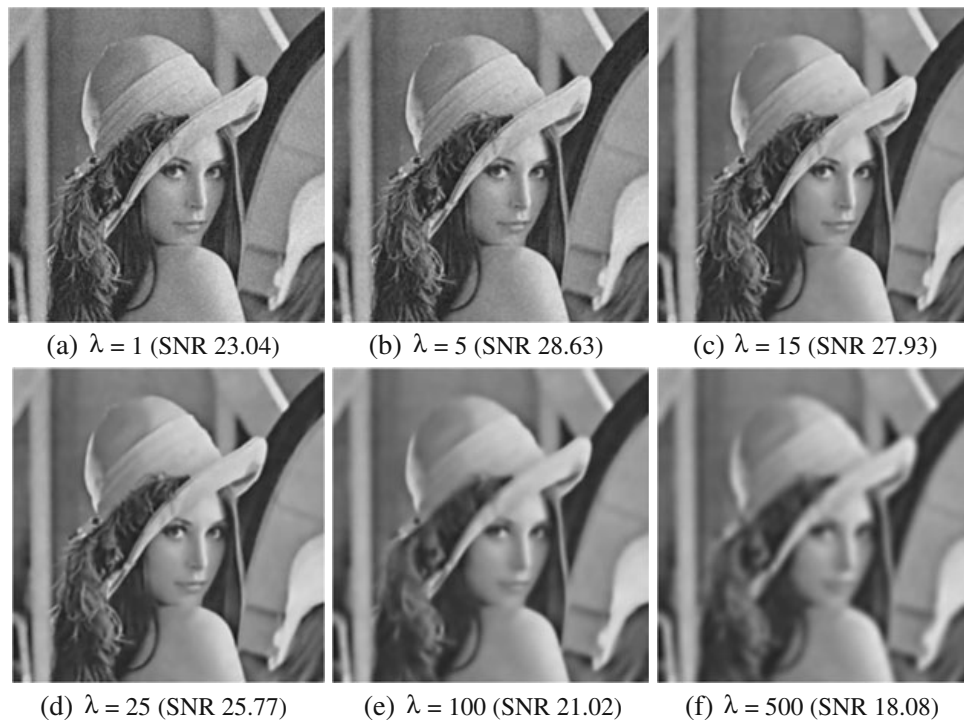


Fig. 3 Solution v —standard deviation $\sigma = 0.15 - 5,000$ iterations

As $|\bar{p}'_{i,j}| = 1$, relation (30) implies $\bar{p}'_{i,j} = \frac{1 - \beta}{1 + |\beta|} \bar{p}_{i,j}$, so that $\left| \frac{1 - \beta}{1 + |\beta|} \right| = 1$; this yields $\beta \leq 0$ and $\bar{p}'_{i,j} = \bar{p}_{i,j}$.

Fig. 4 Solution v —standard deviation $\sigma = 0.25$ and $\lambda = 25 - 5,000$ iterations

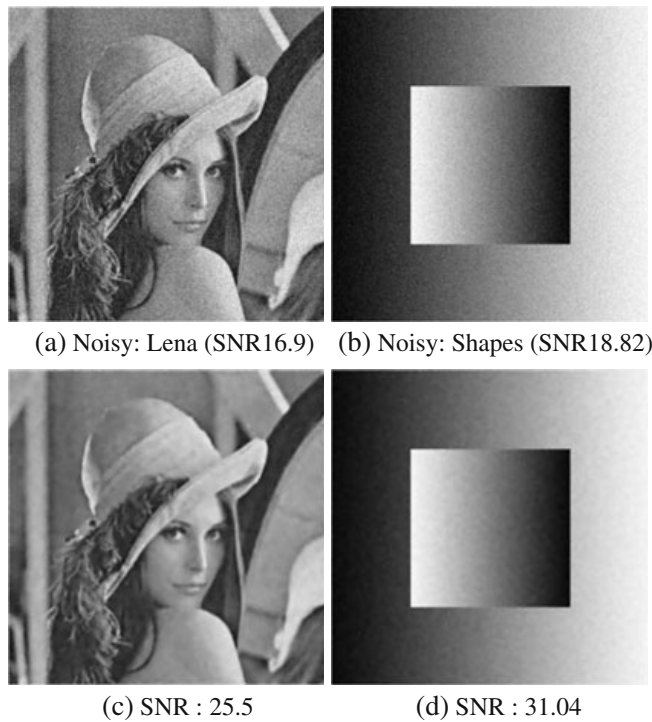
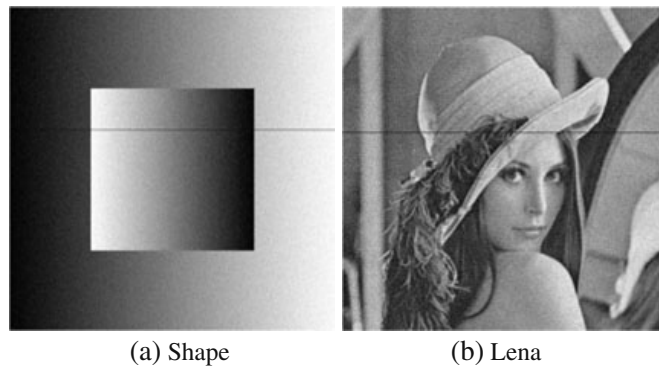


Fig. 5 Noisy images with line— $\sigma = 0.15$



Finally, $\bar{p} = \bar{p}'$, and

$$\forall i, j, \quad (H[\lambda H^* \bar{p} - u_d])_{i,j} + \left\| (H[\lambda H^* \bar{p} - u_d])_{i,j} \right\|_{\mathbb{R}^4} \bar{p}_{i,j} = 0.$$

This is the Euler equation for (\mathcal{P}') . Therefore \bar{p} is a solution to (\mathcal{P}') . With uniqueness of the projection, we deduce that the sequence $(\lambda H^* p^n)_n$ converges to $P_{\lambda \kappa_2}(u_d)$. We conclude with Lemma 1 since $\tau \leq 1/64$ then $\tau \leq 1/\kappa^2$. \square

6 Numerical Results

In this section we briefly present numerical tests for the denoising process. A full comparison with existing methods will be performed in a forthcoming paper.

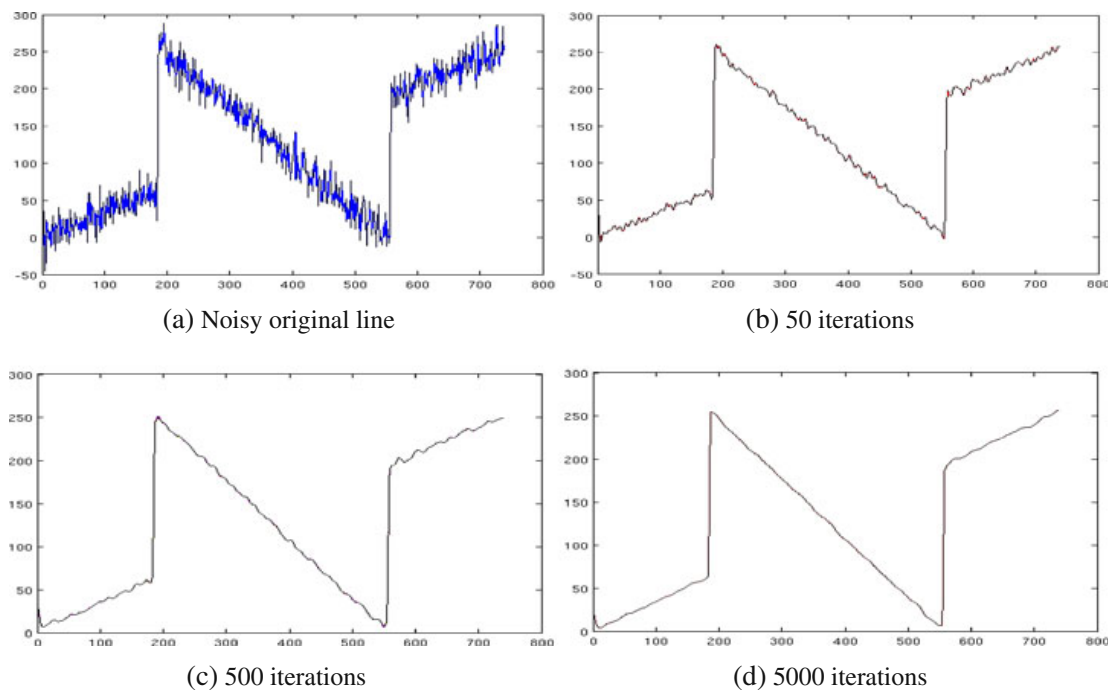


Fig. 6 Sensitivity with respect to the number of iterations— $\sigma = 0.15$, $\lambda = 50$ —slice of “Shapes” image

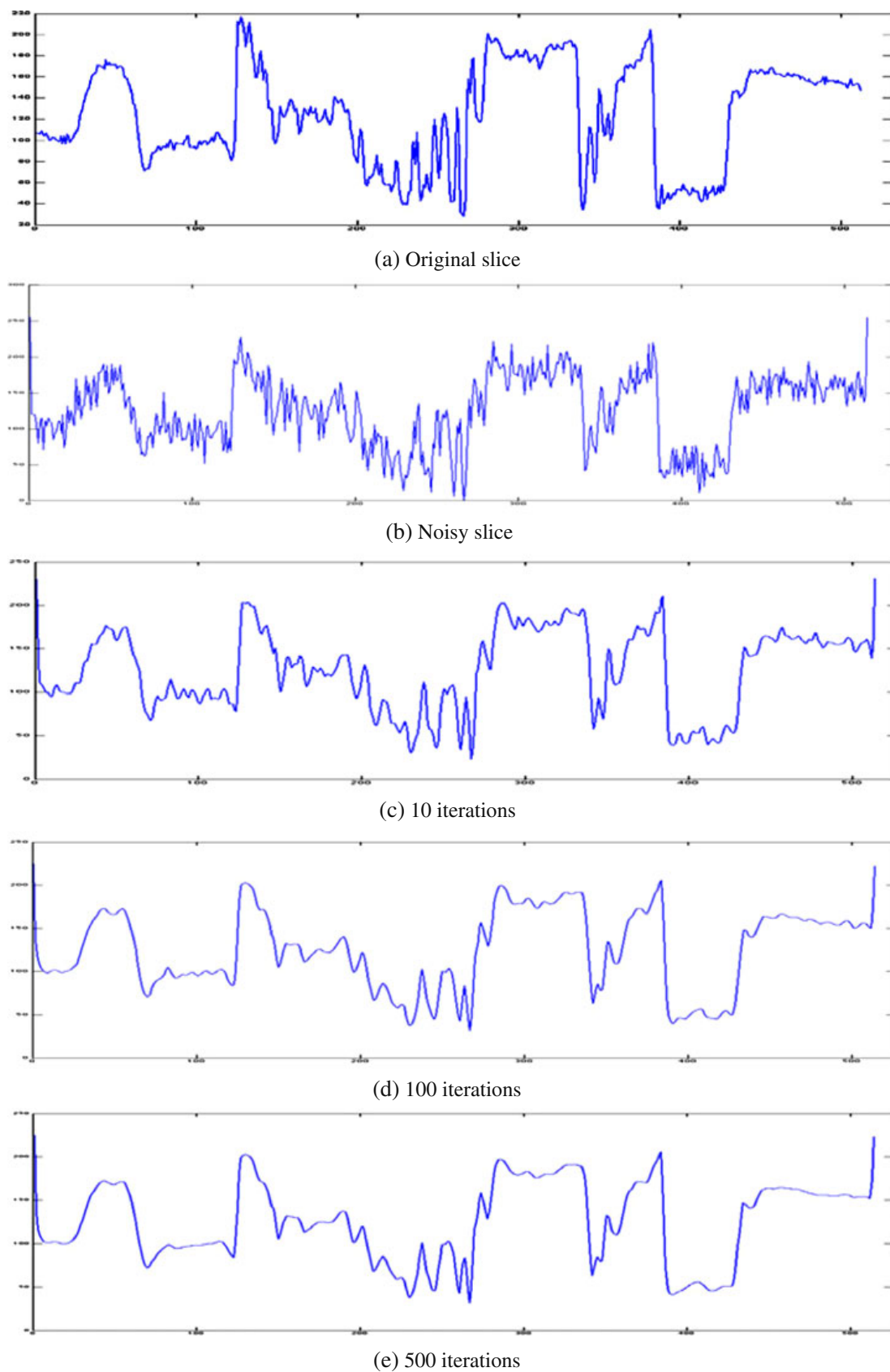


Fig. 7 Sensitivity with respect to the number of iterations— $\sigma = 0.15, \lambda = 15$ —slice of “Lena” image

6.1 Examples

Throughout this section, we consider the following images (Fig. 1) that are degraded with a white Gaussian noise with standard deviation σ (Fig. 2).

We perform numerical tests with different values of σ . In any case, we observe that the model is quite efficient for image restoration. Moreover, we note that we lose details information when parameter λ increases, what was expected. However,

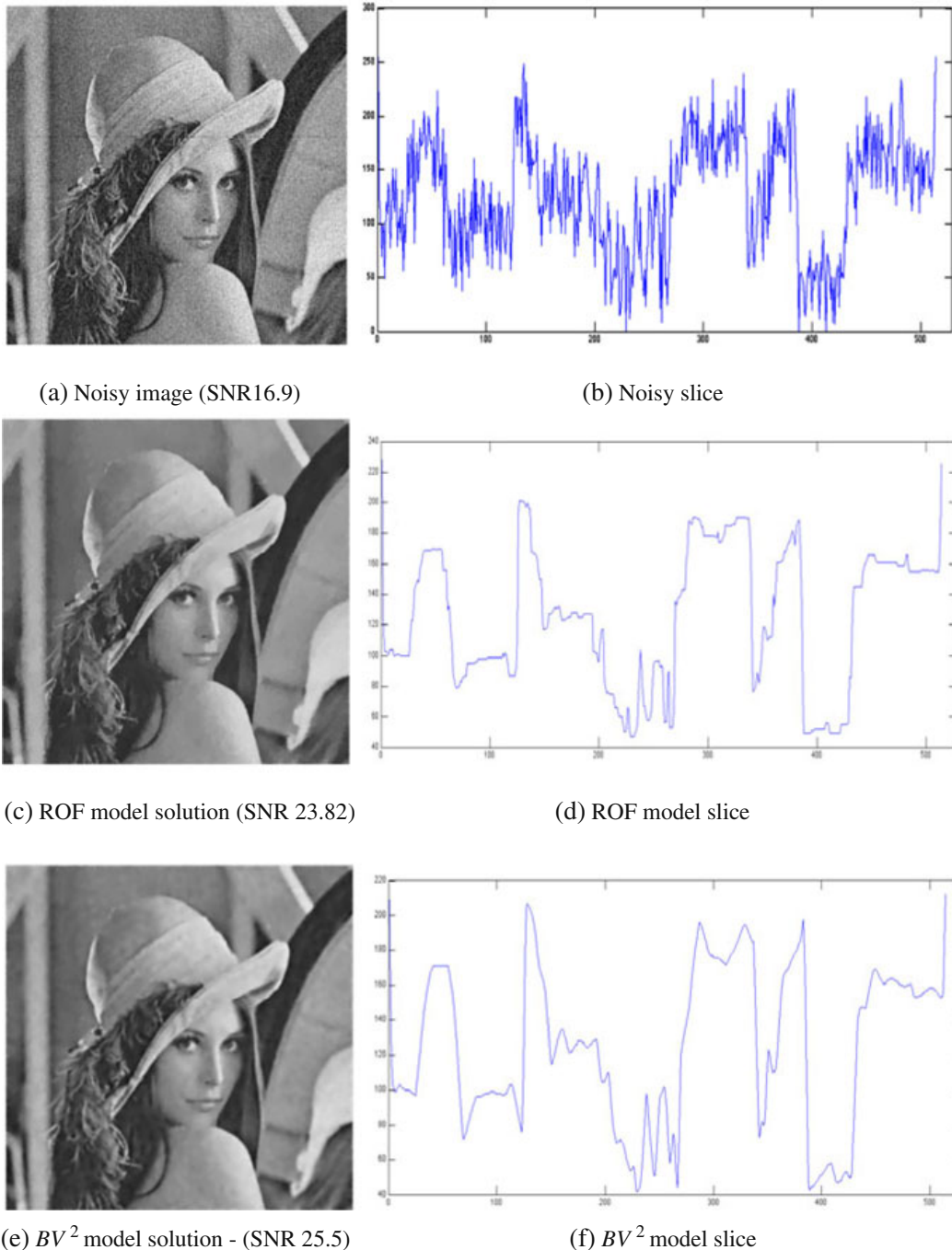


Fig. 8 Comparison between ROF and BV^2 models— $\sigma = 0.25$, $\lambda = 25$, 5,000 iterations

especially when the “observed” image is very noisy, we have a blurriness (subjective) feeling, that we do not have when restoration is performed with the standard ROF model. Checking what happens precisely on slices (lines) of the image (Fig. 8 for example), we remark that the BV^2 -model keeps contour information pretty well, anyway better than expected watching the image.

Numerical tests have been performed on different machines so that we cannot report precisely on the CPU time. However, the result is quite satisfactory after few

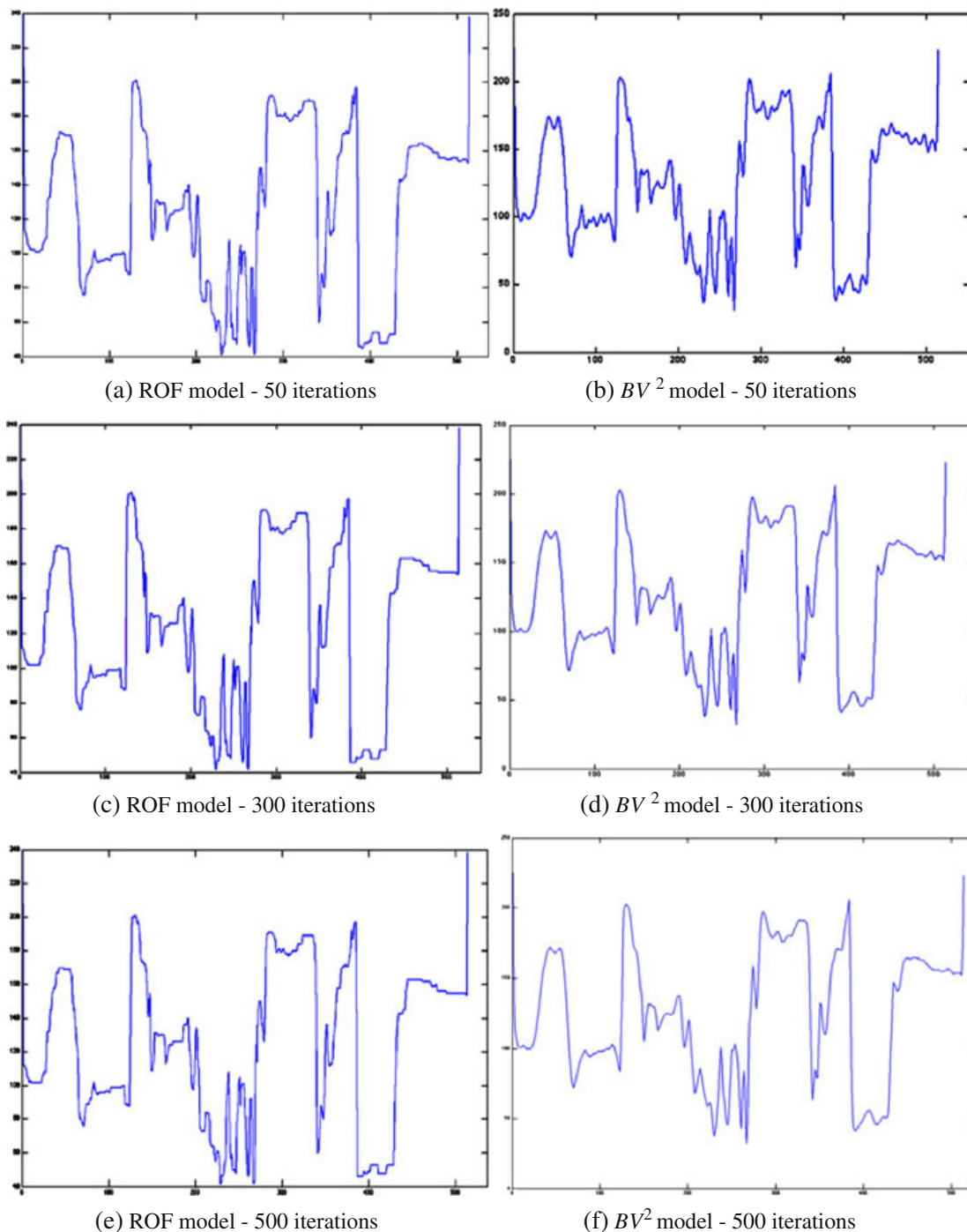


Fig. 9 Comparison between ROF and BV^2 models— $\sigma = 0.15$, $\lambda = 15$ (Lena slice)

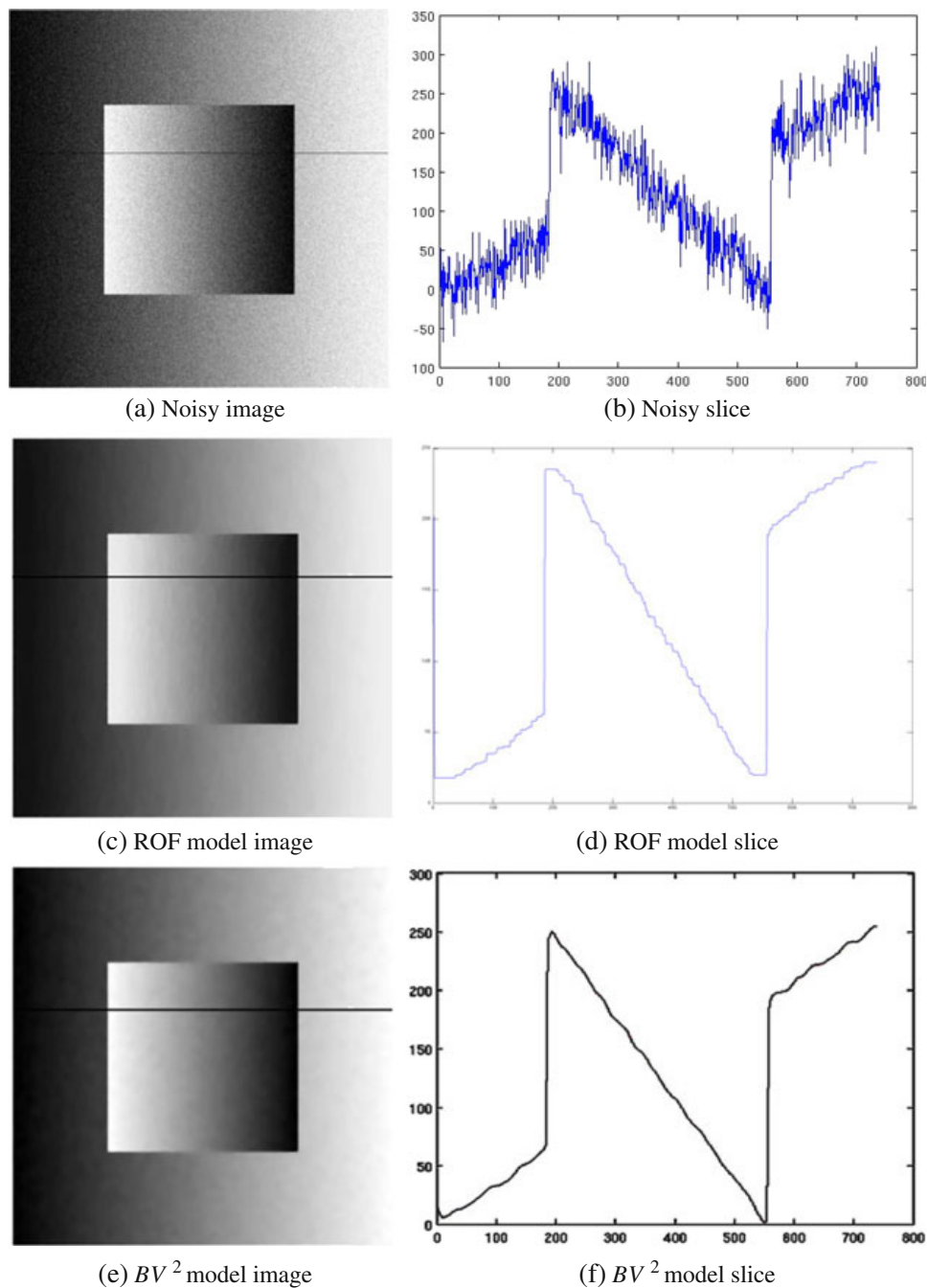
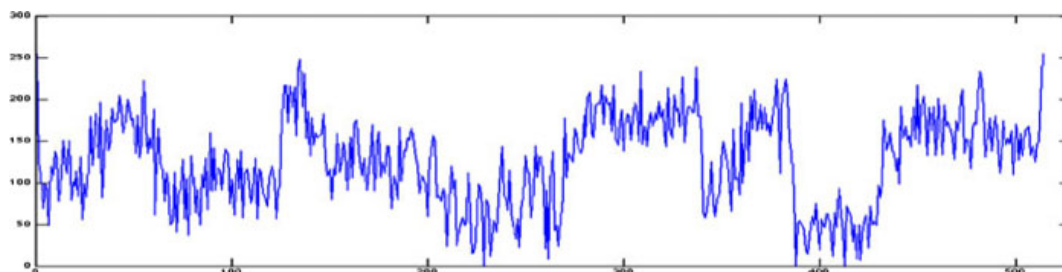


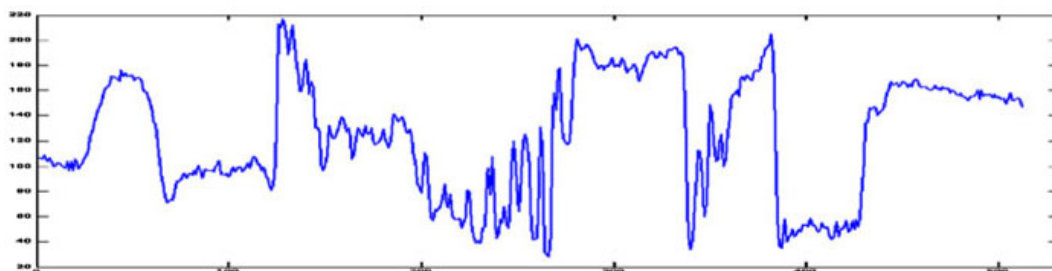
Fig. 10 Comparison between ROF and BV^2 models— $\sigma = 0.25$, $\lambda = 50$, 10,000 iterations

iterations so that the process is not too slow. In what follows, the stopping criterion has been set to a maximal number of iterations it_{\max} that can be chosen arbitrary large. The algorithms have been implemented with MATLAB[®] software. We give also the Signal to Noise Ratio (SNR):¹ $SNR(v) = 20 \log_{10} \left(\frac{\|u\|_{L^2}}{\|u-v\|_{L^2}} \right)$, where u is the

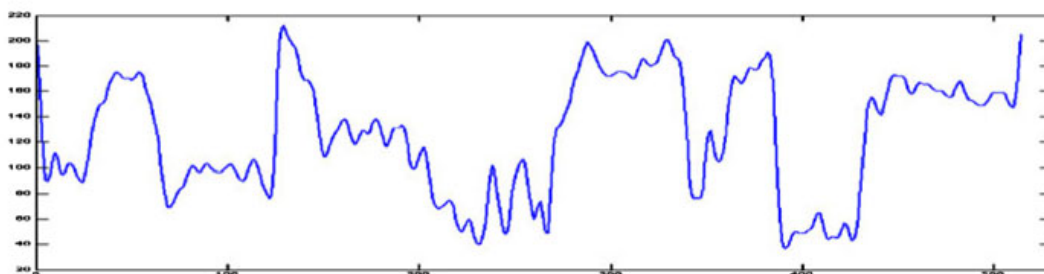
¹There are different ways to compute this quantity with MATLAB[®]. We used the following syntax: $SNR = 20 * \log_{10}(\text{norm}(u(:)) / \text{norm}(u(:) - v(:)))$.



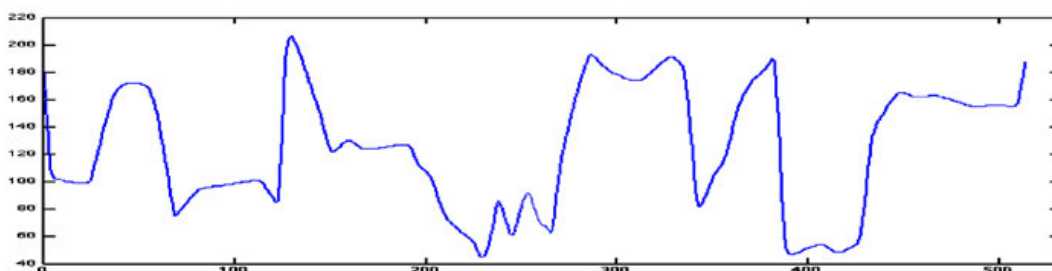
(a) Noisy slice



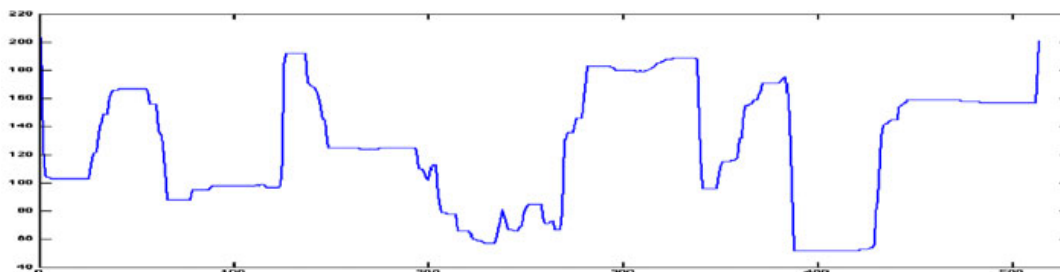
(b) Original slice



(c) BV^2 model - 50 iterations.



(d) BV^2 model - 5000 iterations.



(e) ROF model - 5000 iterations.

Fig. 11 Zoom on “Lena” slices— $\sigma = 0.25$, $\lambda = 50$

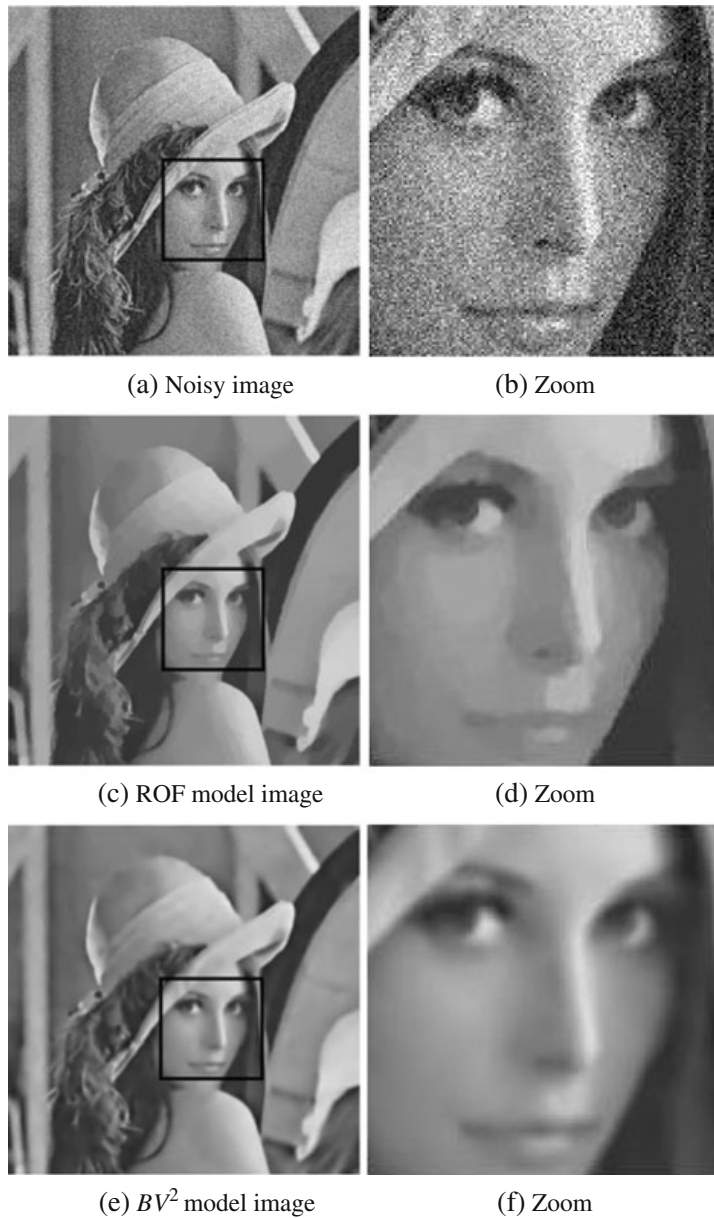
expected image and v is the restored one. $\text{SNR}(u_d)$ gives the observed SNR (with the noisy input image).

We have performed tests for two σ values. In the first example $\sigma = 0.15$ and we stopped after $\text{it}_{\max} = 5,000$ iterations (Fig. 3) and in the second case $\sigma = 0.25$ (the noise is more important). The noisy images are shown in Fig. 2.

6.2 Sensibility with Respect to λ Parameter

As expected, we see on Fig. 3 that the smoothing process is more efficient when λ is large. For both images, the result is satisfactory for $\lambda \simeq 25$ (Fig. 4).

Fig. 12 Staircasing effect— $\sigma = 0.25$, $\lambda = 50$, 10,000 iterations



6.3 Sensitivity with Respect to Iterations Number it_{max}

We fix $\lambda = 15$ and $\sigma = 0.15$ (Fig. 5). Figures 6 and 7 give the behavior of a slice (line) during iterations (we can see more easily how noise is removed). The algorithm converges well: the quality of restoration is improved as the number of iterations grows. Noise is removed and contours are preserved.

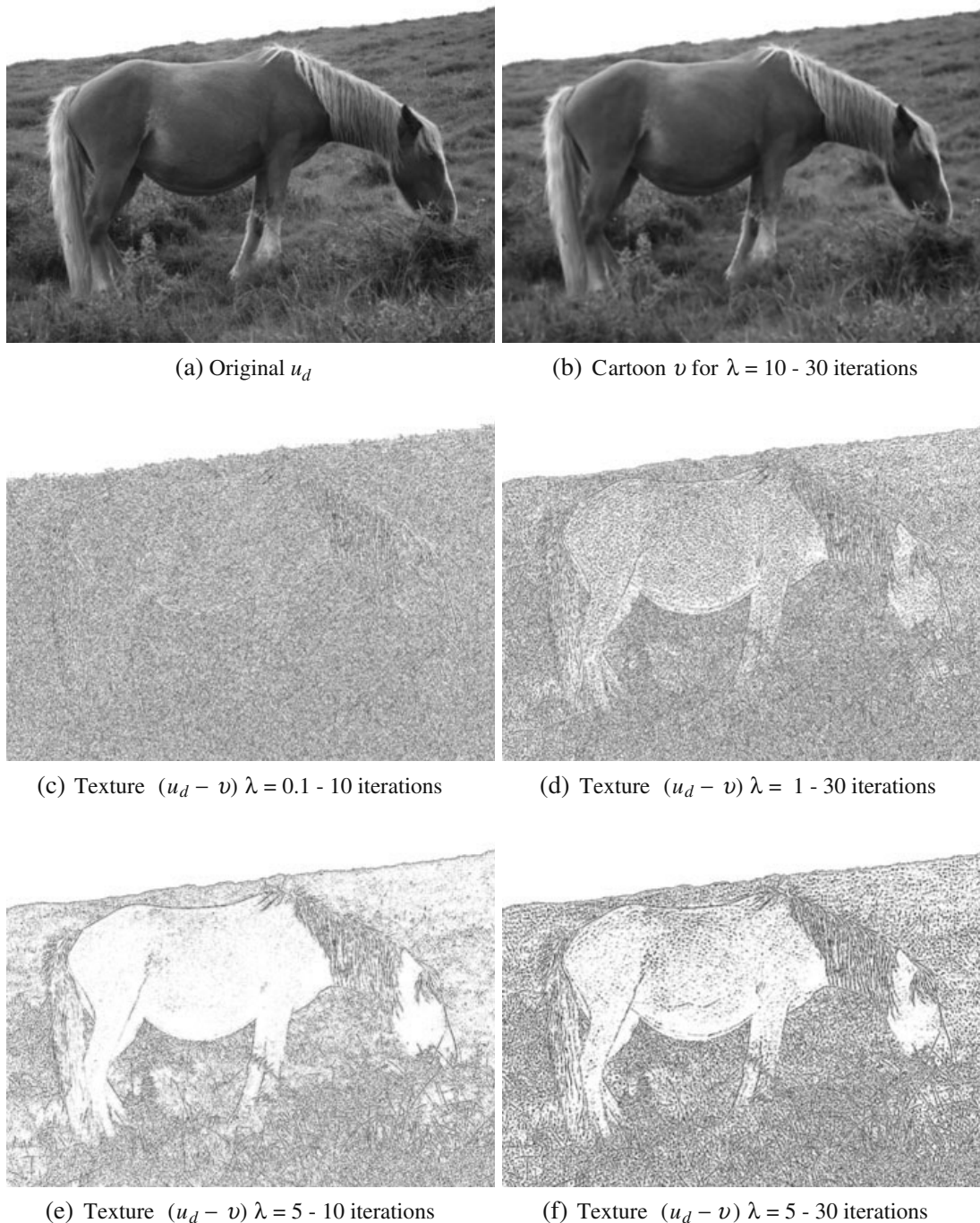


Fig. 13 Texture extraction—rescaled texture image $u_d - v$ at 10 and 15 iterations

6.4 Comparison with the Classical Rudin–Osher–Fatemi Model

It is well known that the ROF model makes staircasing effect appear, since the resulting image is piecewise constant on large areas (Fig. 8). We first compare the two models on test images that are not very noisy. In Fig. 9 we see that piecewise constant areas appear with ROF, which is not the case with the BV^2 model. To focus on this fact, we have selected a line of the image that meets contours (Fig. 10).

Figures 8 and 11 are obtained for $\lambda = 25$ and $\lambda = 50$ respectively and 5,000 iterations. Figures 10 and 12 are obtained for 10,000 iterations and $\lambda = 50$ and show precisely what happens: the image restored with ROF is clearly piecewise constant, and the BV^2 restored one seems to be blurred (Fig. 12). However, this is an optical effect: considering a slice shows that the BV^2 model removes noise significantly and contours are better preserved: the amplitude of high peaks that correspond to contours is not changed, which is not the case in ROF-model (Fig. 11).

6.5 Texture Extraction

We do not report much on texture extraction process. The parameter tuning is slightly different but the algorithm behaves similarly (see [11]). Many variational

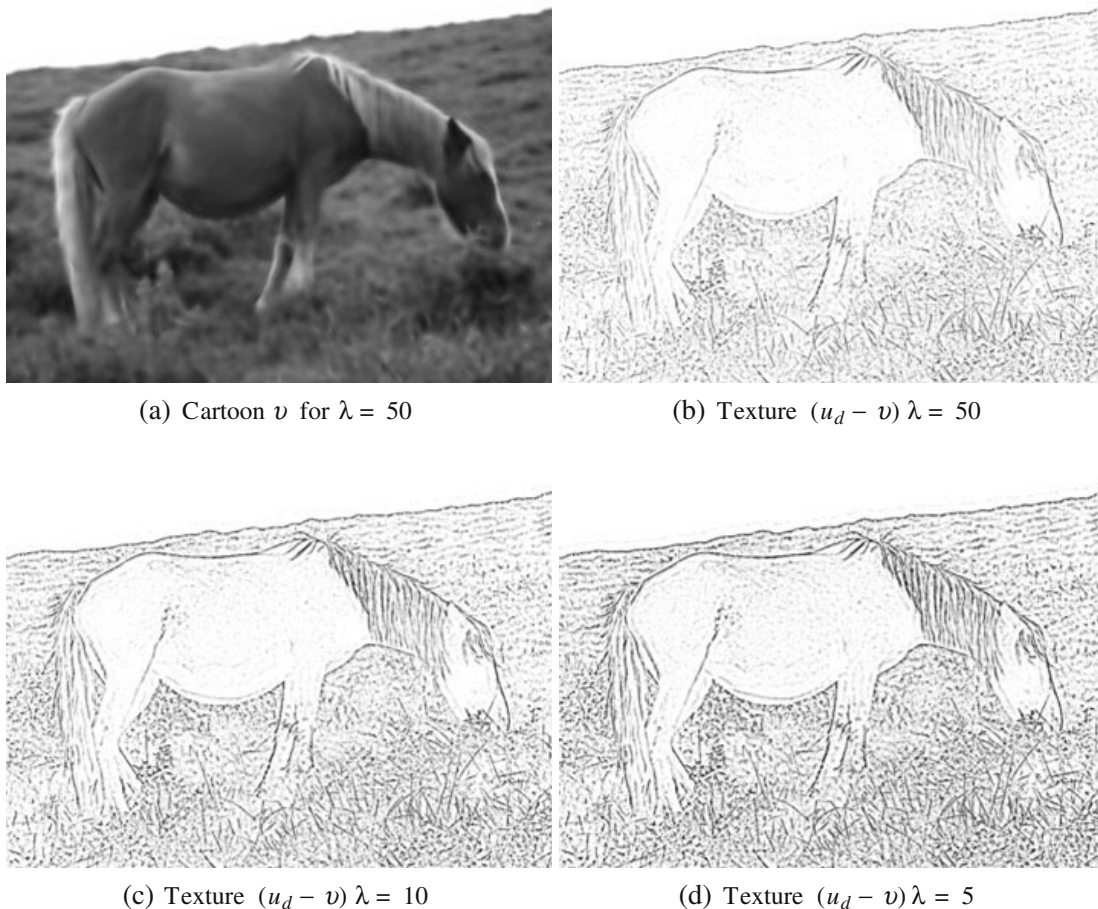


Fig. 14 Texture extraction—rescaled texture image $u_d - v$ at 100 iterations

methods have been developed for texture extraction (see [5, 6] and the references therein). We shall precisely compare the BV^2 method to the existing ones in a forthcoming paper. We present an example in Figs. 13 and 14.

7 Conclusion

The second order approach via the BV^2 space seems promising. Many improvements have to be performed. The algorithm is still slow (though we get acceptable results for quite few iterations $\simeq 30$). We currently investigate the Nesterov- techniques to speed up the method (see Weiss et al. [21] or Fadili and Peyré [13]). Moreover, we have to look for modifications of the variational model using different norms (for example the L^1 norm) for the fitting data term. Furthermore, coupling existing techniques for texture extraction with the second order approach should give quite performing results. This will be done in future works.

References

1. Acar, R., Vogel, C.R.: Analysis of bounded variation penalty methods for ill-posed problems. *Inverse Probl.* **10**(6), 1217–1229 (1994)
2. Ambrosio, L., Fusco, N., Pallara, D.: *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford Mathematical Monographs, Oxford University Press (2000)
3. Attouch, H., Briceño-Arias, L.M., Combettes, P.L.: A parallel splitting method for coupled monotone inclusions. *SIAM J. Control Optim.* **48**, 3246 (2010)
4. Attouch, H., Buttazzo, G., Michaille, G.: *Variational analysis in sobolev and BV spaces: applications to PDEs and optimization*. MPS-SIAM series on optimization. Philadelphia, ISBN 0-89871-600-4 (2006)
5. Aubert, G., Aujol, J.F.: Modeling very oscillating signals, application to image processing. *Appl. Math. Optim.* **51**(2), 163–182 (2005)
6. Aubert, G., Aujol, J.F., Blanc-Feraud, L., Chambolle, A.: Image decomposition into a bounded variation component and an oscillating component. *J. Math. Imaging Vis.* **22**(1), 71–88 (2005)
7. Aubert, G., Kornprobst, P.: *Mathematical problems in image processing, partial differential equations and the calculus of variations*. Applied Mathematical Sciences, vol. 147. Springer Verlag (2006)
8. Aujol, J.F.: Some first-order algorithms for total variation based image restoration. *J. Math. Imaging Vis.* **34**, 307–327 (2009)
9. Chambolle, A.: An algorithm for total variation minimization and applications. *J. Math. Imaging Vis.* **20**, 89–97 (2004)
10. Demengel, F.: Fonctions à hessien borné. *Annales de l'institut Fourier*, Tome **34**(2), 155–190 (1984)
11. Echegut, R., Piffet, L.: A variational model for image texture identification (preprint). <http://hal.archives-ouvertes.fr/hal-00439431/fr/>
12. Ekeland, I., Temam, R.: *Convex Analysis and Variational problems*. SIAM Classic in Applied Mathematics, vol. 28 (1999)
13. Fadili, J., Peyré, G.: Total variation projection with first order schemes (preprint). <http://hal.archives-ouvertes.fr/hal-00380491/fr/>
14. Hinterberger, W., Scherzer, O.: Variational methods on the space of functions of bounded Hessian for convexification and denoising. *Computing* **76**(1–2), 109–133 (2006)
15. Meyer, Y.: *Oscillating patterns in image processing and nonlinear evolution equations*. University Lecture Series, vol. 22. AMS (2002)
16. Osher, S., Fatemi, E., Rudin L.: Nonlinear total variation based noise removal algorithms. *Physica D* **60**, 259–268 (1992)
17. Osher, S., Sole, A., Vese L.: Image decomposition and restoration using total variation minimization and the H^1 norm. *SIAM J. Multiscale Model. Simul.* **1**(3), 349–370 (2003)

18. Osher, S., Vese, L.: Modeling textures with total variation minimization and oscillating patterns in image processing. *J. Sci. Comput.* **19**(1–3), 553–572 (2003)
19. Osher, S.J., Vese, L.A.: Image denoising and decomposition with total variation minimization and oscillatory functions. Special issue on mathematics and image analysis. *J. Math. Imaging Vis.* **20**(1–2), 7–18 (2004)
20. Piffet, L.: Modèles variationnels pour l'extraction de textures 2D. Ph.D. Thesis, Orléans (2010)
21. Weiss, P., Blanc-Féraud, L., Aubert, G.: Efficient schemes for total variation minimization under constraints in image processing. *SIAM J. Sci. Comput.* **31**(n°3), 2047–2080 (2009)
22. Yin, W., Goldfarb, D., Osher, S.: A comparison of three total variation based texture extraction models. *J. Vis. Commun. Image* **18**, 240–252 (2007)

Bibliographie

- [1] Acar, R., Vogel, C.R. : Analysis of bounded variation penalty methods for ill-posed problems. *Inverse Problems* 10, no. 6, 1217–1229 (1994)
- [2] Adams, R.A. : Sobolev spaces. Academic Press, Springer Verlag (1978).
- [3] Aliney, S. : A property of the minimum vectors of a regularizing functional defined by means of the absolute norm. *IEEE Trans. Signal Proc.*, 45(4) :913–917, 1997.
- [4] Ambrosio, L., Fusco, N., Pallara, D. : Functions of bounded variation and free discontinuity problems. Oxford mathematical monographs, Oxford University Press (2000).
- [5] Attouch, H., Briceño-Arias, L.M, Combettes, P.L. : A Parallel Splitting Method for Coupled Monotone Inclusions, *SIAM J. Control Optim.* 48 3246 (2010)
- [6] Attouch, H., Buttazzo, Michaille, G. : Variational analysis in Sobolev and BV spaces : applications to PDEs and optimization. MPS-SIAM series on optimization, 2006
- [7] Aubert, G., Aujol, J.F. : Modeling very oscillating signals, application to image processing. *Applied Mathematics and Optimization*, 51, no 2, 163–182 (2005)
- [8] Aubert, G., Aujol, J.F., Blanc-Feraud, L., Chambolle, A. :Image decomposition into a bounded variation component and an oscillating component. *Journal of Mathematical Imaging and Vision* 22, no 1, 71–88 (2005)
- [9] Aubert, G., Kornprobst, P. : Mathematical Problems in Image Processing, Partial Differential Equations and the Calculus of Variations. Applied Mathematical Sciences 147, Springer Verlag (2006).
- [10] Aujol, J.F. : Some First-Order Algorithms for Total Variation Based Image Restoration. *J Math Imaging Vis* (2009) 34 : 307–327

-
- [11] Aujol, J-F., Chambolle, A. : Dual norms and image decomposition models, *International Journal of Computer Vision*, 63(1) : 85-104 (2005).
- [12] Aujol, J-F., Gilboa, G., Chan, T., Osher, S. : Structure-Texture Image Decomposition—Modeling, Algorithms, and Parameter Selection, *international journal of computer vision*, 67(1) : 111-136 (2005)
- [13] D., Azé , *Eléments d'analyse convexe et variationnelle*, Ellipses, Paris, 1997.
- [14] Barbu V.- Precupanu Th., *Convexity and Optimization in Banach Spaces*, Sijthoff & Noordhoff, Bucarest 1978.
- [15] Bergounioux, M., Piffet, L. : A second-order model for image denoising, *Set-Valued and Variational Analysis*,
- [16] Bergounioux, M., Tran, M., P. : A second order model for 3D-texture extraction, hal-00530816 (2010)
- [17] Blomgren, P., Chan, T., Mulet, P., Wong, C. : Total variation image restoration : numerical methods and extensions. *ICIP*, 3 :384–387 (1997).
- [18] Bredies, K., Kunisch, K., Pock, T. : Total generalized variation. *SIAM Journal on Imaging Sciences*, 3(3) :492-526 (2010).
- [19] Bouman, C., Sauer, K. : A generalized Gaussian image model for edge-preserving MAP estimation. *IEEE Trans. on Image Processing*, 2 :296–310 (1993).
- [20] Brézis H., *Analyse Fonctionnelle*, Masson, Paris, 1987.
- [21] Caselles, V., Chambolle, A., Novaga, M. : The discontinuity set of solutions of the TV denoising problem and some extensions. *Multiscale Model. Simul.* 6(3) :879–894 (2007)
- [22] Chambolle, A. : An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*, 20 89–97 (2004)
- [23] Chambolle, A., Lions, P.-L. : Image recovery via total variation minimization and related problems *Numerische Mathematik*, 76, 167–188 (1997)
- [24] Chan, T., Marquina, A., Mulet, P. : High-order total variation-based image restoration. *SIAM J. Sci. Comput.*, 22(2) :503–516 (2000).
- [25] Chan, T., Esedoglu, S. : Aspects of total variation regularized L1 function approximation. *SIAM Journal of Applied Math.*, 65(5) :1817–1837 (2005).
- [26] Chan, T. F., Esedoglu, S., Park, F. E. : Image decomposition combining staircase reduction and texture extraction. *J. Vis. Comun. Image Represent.*, 18(6) :464–486 (2007).

- [27] J. Darbon, J., Sigelle, M. : Image restoration with discrete constrained total variation part I : Fast and exact optimization. *Journal of Mathematical Imaging and Vision*, 26(3) :277–291 (2006).
- [28] Demengel, F. : Fonctions à hessien borné. *Annales de l'institut Fourier*, tome 34, no 2, pp. 155-190 (1984)
- [29] Dobson, D. C., Santosa, F. : Recovery of blocky images from noisy and blurred data. *SIAM J. Appl. Math.*, 56(4) :1181–1198 (1996).
- [30] Duval, V., Aujol, J-F., Gousseau, Y. : The TV ? L1 model : a geometric point of view. *SIAM Journal on Multiscale Modeling and Simulation*, 8(1) :154-189, (2009).
- [31] Echegut, R., Piffet, L. : A variational model for image texture identification. <http://hal.archives-ouvertes.fr/hal-00439431/fr/>, submitted.
- [32] Ekeland, I., Temam, R. : *Convex Analysis and Variational problems*, SIAM Classic in Applied Mathematics, 28, 1999,
- [33] Fadili, J., Peyré, G. : Total Variation Projection with First Order Schemes, Preprint, <http://hal.archives-ouvertes.fr/hal-00380491/fr/>
- [34] Garnett, J. B., Le, Triet M., Meyer, Y. Vese, L. A. : Image decompositions using bounded variation and generalized homogeneous Besov spaces. *Appl. Comput. Harmon. Anal.*, 23, no. 1, 25–56 (2007)
- [35] Evans, L.C., Gariépy, R. : *Measure theory and fine properties of functions*. CRC Press, 1992
- [36] Hinterberger, W., Scherzer, O. : Variational methods on the space of functions of bounded Hessian for convexification and denoising. *Computing*, 76, no. 1-2, 109–133 (2006)
- [37] Hofmann, B., Kaltenbacher, B., Pöschl, C., Scherzer, O. : A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators. *Inverse Problems*, 23, no. 3, 987–1010 (2007)
- [38] K., Ito, K., Kunisch. : *Lagrange Multiplier Approach to Variational Problems and Applications*. *Advances in Design and Control*, SIAM (2008)
- [39] Le, Triet M., Vese, L. A. : Image decomposition using total variation and div(BMO). *Multiscale Model. Simul.* 4, no. 2, 390–423 (2005)
- [40] Lieu, Linh H., Vese, L. A. : Image restoration and decomposition via bounded total variation and negative Hilbert-Sobolev spaces. *Appl. Math. Optim.* 58 (2008), no. 2, 167–193 (2008)

- [41] Meyer, Y. : Oscillating patterns in image processing and nonlinear evolution equations, Vol. 22 of University Lecture Series, AMS, 2002.
- [42] Louchet C. : Modèles variationnels et bayésiens pour le débruitage d'images : de la variation totale vers les moyennes non-locales. PhD Thesis, Paris, 2008
- [43] Nesterov, Y. : Gradient methods for minimizing composite objective function. CORE discussion paper (2007).
- [44] Nikolova, M. : Estimées localement fortement homogènes. Comptes rendus de l'académie des sciences, t. 325 :665–670 (1997).
- [45] Nikolova, M. : Local strong homogeneity of a regularized estimator. SIAM J. Appl. Math., 61(2) :633–658 (electronic) (2000).
- [46] Nikolova, M. : Weakly constrained minimization : application to the estimation of images and signals involving constant regions. J. Math. Imaging Vision, 21(2) :155–175 (2004).
- [47] Nikolova, M. : A variational approach to remove outliers and impulse noise. JMIV, 20(1-2) :99–120 (2004).
- [48] Nikolova, M. : Analysis of the recovery of edges in images and signals by minimizing nonconvex regularized least-squares. Multiscale Model. Simul., 4(3) :960–991 (electronic) (2005).
- [49] Nikolova, M. : Model distortions in Bayesian MAP reconstruction. Inverse Probl. Imaging, 1(2) :399–422 (2007).
- [50] Osher, S., Fatemi, E, Rudin L. : Nonlinear total variation based noise removal algorithms. Physica D 60, 259–268 (1992)
- [51] Osher, S., Sole, A., Vese L. :, Image decomposition and restoration using total variation minimization and the H^1 norm. SIAM Journal on Multiscale Modeling and Simulation, 1-3, 349–370 (2003)
- [52] Osher, S., Vese, L. : Modeling textures with total variation minimization and oscillating patterns in image processing. Journal of Scientific Computing 19, no 1-3, 553–572 (2003)
- [53] Osher, S. J, Vese, L. A., : Image denoising and decomposition with total variation minimization and oscillatory functions. Special issue on mathematics and image analysis. J. Math. Imaging Vision, 20, no. 1-2, 7–18 (2004)
- [54] Piffet, L. : Modèles variationnels pour l'extraction de textures 2D, PhD Thesis, Orléans, 2010
- [55] Rockafellar, R., T. : Convex Analysis. Princeton University press (1972).

- [56] Ring, W. : Structural properties of solutions of total variation regularization problems. *ESSAIM, Math Modeling and Numerical Analysis*, 34 :799–840 (2000).
- [57] Weiss, P., Blanc-Féraud, L., Aubert, G. : Efficient schemes for total variation minimization under constraints in image processing, *SIAM Journal on Scientific Computing*, 2009, to appear.
- [58] Yin, W., Goldfarb, D., Osher, S. : A comparison of three total variation based texture extraction models. *J. Vis. Commun. Image*, 18, 240–252 (2007)

Loïc PIFFET

Décomposition d'image par modèles variationnels - Débruitage et extraction de texture.

Résumé : Cette thèse est consacrée dans un premier temps à l'élaboration d'un modèle variationnel de débruitage d'ordre deux, faisant intervenir l'espace BV^2 des fonctions à hessien borné. Nous nous inspirons ici directement du célèbre modèle de Rudin, Osher et Fatemi (ROF), remplaçant la minimisation de la variation totale de la fonction par la minimisation de la variation totale seconde, c'est à dire la variation totale de ses dérivées. Le but est ici d'obtenir un modèle aussi performant que le modèle ROF, permettant de plus de résoudre le problème de l'effet *staircasing* que celui-ci engendre. Le modèle que nous étudions ici semble efficace, entraînant toutefois l'apparition d'un léger effet de flou. C'est afin de réduire cet effet que nous introduisons finalement un modèle mixte, permettant d'obtenir des solutions à la fois non constantes par morceaux et sans effet de flou au niveau des détails. Dans une seconde partie, nous nous intéressons au problème d'extraction de texture. Un modèle reconnu comme étant l'un des plus performants est le modèle $TV-L^1$, qui consiste simplement à remplacer dans le modèle ROF la norme L^2 du terme d'attache aux données par la norme L^1 . Nous proposons ici une méthode originale permettant de résoudre ce problème utilisant des méthodes de Lagrangien augmenté. Pour les mêmes raisons que dans le cas du débruitage, nous introduisons également le modèle TV^2-L^1 , consistant encore une fois à remplacer la variation totale par la variation totale seconde. Un modèle d'extraction de texture mixte est enfin très brièvement introduit. Ce manuscrit est ponctué d'un vaste chapitre dédié aux tests numériques.

Mots-clés : espace des fonction à variation bornée BV , espace des fonctions à hessien borné BV^2 , modèle variationnel, algorithmes de projection, variation totale, variation totale seconde, anisotropie.

Variational models for image decomposition - Denoising and texture extraction.

Abstract : This thesis is devoted in a first part to the elaboration of a second order variational model for image denoising, using the BV^2 space of bounded hessian functions. We here take a leaf out of the well known Rudin, Osher and Fatemi (ROF) model, where we replace the minimization of the total variation of the function with the minimization of the second order total variation of the function, that is to say the total variation of its partial derivatives. The goal is to get a competitive model with no *staircasing* effect that generates the ROF model anymore. The model we study seems to be efficient, but generates a blurry effect. In order to deal with it, we introduce a mixed model that permits to get solutions with no *staircasing* and without blurry effect on details. In a second part, we take an interest to the texture extraction problem. A model known as one of the most efficient is the $TV-L^1$ model. It just consists in replacing the L^2 norm of the fitting data term with the L^1 norm. We propose here an original way to solve this problem by the use of augmented Lagrangian methods. For the same reason than for the denoising case, we also take an interest to the TV^2-L^1 model, replacing again the total variation of the function by the second order total variation. A mixed model for texture extraction is finally briefly introduced. This manuscript ends with a huge chapter of numerical tests.

Keywords : space of functions of bounded variation BV , space of bounded hessian functions, variational model, projection algorithms, total variation, second order total variation, anisotropy.



MAPMO, Université d'Orléans
UFR Sciences
Route de Chartres B.P. 6759 - 45067 Orléans cedex 2 FRANCE

