Variable Bandwidth Image Models for Texture-Preserving Enhancement of Natural Images

Noura Azzabou

THÈSE

présentée pour l'obtention du titre de

DOCTEUR DE L'ÉCOLE DES PONTS

Spécialité : Mathématiques, Informatique

par

Noura Azzabou

Variable Bandwidth Image Models for Texture-Preserving Enhancement of Natural Images

Soutenue le 31 Mars 2008 devant le jury composé de :

Rapporteurs	Rachid Deriche	INRIA Sophia Antipolis
	Nir Sochen	Tel Aviv University
	Alan Willsky	Massachusetts Institute of Technology
Examinateurs	Jean Yves Audibert	Ecole des Ponts
	Jean Michel Morel	ENS Cachan
Directeurs de thèse	Frédéric Guichard	DxOLabs
	Nikos Paragios	Ecole Centrale de Paris
Membres invités	Frédéric Cao	DxOLabs
	Pierre Carlier	AFM-CEA, Pitié Salpetrière

Abstract

This thesis is devoted to image enhancement and texture preservation issues. This task involves an image model that describes the characteristics of the recovered signal. Such a model is based on the definition of the pixels interaction that is often characterized by two aspects (i) the photometric similarity between pixels (ii) the spatial distance between them that can be compared to a given scale. The first part of the thesis, introduces novel non-parametric image models towards more appropriate and adaptive image description using variable bandwidth approximations driven from a soft classification in the image. The second part introduces alternative means to model observations dependencies from geometric point of view. This is done through statistical modeling of co-occurrence between observations and the use of multiple hypotheses testing and particle filters. The last part is devoted to novel adaptive means for spatial bandwidth selection and more efficient tools to capture photometric relationships between observations. The thesis concludes with providing other application fields of the last technique towards proving its flexibility toward various problem requirements.

Résumé

Cette thèse s'intéresse aux problèmes de restauration d'images et de préservation de textures. Cette tache nécessite un modèle image qui permet de caractériser le signal qu'on doit obtenir. Un tel model s'appuie sur la définition de l'interaction entre les pixels et qui est caractérisé par deux aspects : (i) la similarité photométrique entre les pixels (ii) la distance spatiale entre les pixels qui peut être comparée à une grandeur d'échelle. La première partie de la thèse introduit un nouveau modèle non paramétrique d'image. Ce modèle permet d'obtenir une description adaptative de l'image en utilisant des noyaux de taille variable obtenue à partir d'une étape de classification effectuée au préalable. La deuxième partie introduit une autre approche pour décrire la dépendance entre pixels d'un point de vue géométrique. Ceci est effectué à l'aide d'un modèle statistique de la co-occurrence entre les observations de point de vue géométrique. La dernière partie est une nouvelle technique de sélection automatique (pour chaque pixel) de la taille des noyaux utilisé au cours du filtrage. Cette thèse est conclue avec l'application de cette dernière approche dans différents contextes de filtrage ce qui montre sa flexibilité vis-à-vis des contraintes liées aux divers problèmes traités.

Acknowledgement

First, I would like to thank my PhD adviser Professor Nikos Paragios for his expert guidance. His patience and unfailing support helped me overcome the difficulties that I faced during this work. His expertise and research insight are key element in this thesis.

My thanks goes also to DxOLabs for funding this thesis and to Frédéric Guichard who gave me the opportunity to have an industrial experience that brought me a wide perspective in the image processing domain. His insightful comments and constructive criticisms at different stages of my research were important and guided the progression of this work. I had also the chance to work with Frédéric Cao his research expertise further enhanced my knowledge about image processing topics. During my visits to DxO I had the opportunity to collaborate on various and stimulating problems with the different members of "l'équipe TI". I highly appreciated their friendship and the convivial atmosphere

I am grateful to Professor: Rachid Deriche, Nir Sochen and Alan Willsky for accepting to review this document in spite their constraints. I appreciated their valuable comments and constructive feedback about my work. I would like to thank Professor Jean Michel Morel and Jean Yves Audibert for accepting to be member of the thesis committee, their presence was an honor for me.

Warm thanks to the members of the vision group in MAS Laboratory, their support and the delightful atmosphere made my days more enjoyable. I would like to specially mention Radhouene for his valuable support through reading my manuscript or his constant encouragement. I also owe to him some material in the DTI section which is the fruit of our collaboration. Special thank to my office-mate Martin, Maxime and Ahmed with whom I had a nice time and the discussion with them was always a pleasure. Finally, my thought goes to my labmate in CERTIS Lab in Ecole des Ponts who make my first year in thesis enjoyable.

Many friends have helped with their support and care to overcome setbacks. I greatly value

their friendship and I deeply appreciate their belief in me. Among them I will cite Afef, Alia, Both, Imene, Amine, Sana, ...

Last but not least, my deepest gratitude goes to my family to whom I dedicate this work. Their love, patience and support were fundamental to the accomplishment of this work.

Table of Contents

1. Introduction en Français 21

1.1 Présentation du problème et Motivations 21

1.2 Les Contributions 22

1.3 Plan de la thèse 24

2. Introduction 27

2.1 Problem Statement and motivation 27

2.2 State of the Art 29

2.2.1 Averaging Based Filters 30

2.2.2 PDE's and Energy Based Image Restoration 33

2.2.3 Image Transform and Compact Representation 38

2.2.4 Statistical Models and Image Denoising 42

2.3 Main Contributions 46

2.4 Thesis Outline 48

3. MPM Adaptive Denoising 51

3.1 Introduction 51

3.2 Unsupervised Information-Theoretic Adaptive Filter 53

3.3 Unsupervised Classification of Image Pixels 55
3.3.1 Texture Feature Extraction 55
3.3.2 Dimensionality Reduction 57
3.3.3 Application to Pixels Classification 59
3.4 Non Parametric Model and Adaptive Denoising 62
3.4.1 Bayesian Formulation of the Problem 62
3.4.2 Non Parametric Density Estimation 67
3.4.3 Variable Bandwidth Selection 69
3.4.4 Marginal Posterior Maximizing 70
3.5 Experimental Results 72
3.6 Conclusion 79

4. Image Reconstruction Using Particle Filters 81
4.1 Introduction 81
4.2 Statistical Description of Image Structure 83
4.3 Overview of Particle Filtering Technique 86
4.3.1 Preliminaries 87
4.3.2 Sampling Importance Resampling Filter (SIR) 88
4.4 Application to image restoration 90
4.4.1 Transition Model 91
4.4.2 Likelihood Measure 92
4.4.3 Intensity Reconstruction 93
4.4.4 Extension to Multiplicative Noise Model 95

4.5 Experimental Results 96
4.5.1 Additive Noise 97
4.5.2 Multiplicative Noise 98
4.6 Conclusion 104

5. TV Based Variable Bandwidth Image Denoising 109
5.1 Introduction 109
5.2 Related Work - Extensions of the TV Model 110
5.2.1 Texture Preserving Denoising Using Spatially Varying Constraints 111
5.2.2 Image Decomposition Models 112
5.2.3 Non Local Functional Based Regularization 114
5.3 Variable Bandwidth Denoising Using Semi Local Quadratic Functional 115
5.3.1 Analysis of the Denoising Model 115
5.3.2 Bandwidth Computation 119
5.4 On the similarity measure between patches 121
5.4.1 New Statistical Similarity Measure Between Patches 122
5.4.2 PCA Based Dictionary 123
5.5 Experimental Results 128
5.5.1 On the weight selection 129
5.6 Conclusion 135

6. Applications 143
6.1 Color Image Denoising 143
6.1.1 Noise Properties from Raw to RGB Images 144

6.1.2 Noise Model Estimation . . . 146
6.1.3 The Denoising Algorithm . . . 147
6.1.4 Experimental Results . . . 149
6.1.5 Discussion . . . 150
6.2 Application to DTI estimation and regularization . . . 150
6.2.1 Introduction . . . 150
6.2.2 DTI Estimation and Regularization . . . 156
Measuring Similarities from diffusion weighted images . . . 157
Semi-Definite Positive Gradient Descent . . . 158
6.2.3 Experimental Validation . . . 160
Artificially Corrupted Tensors . . . 160
DTI towards Understanding the Human Skeletal Muscle . . . 161
6.2.4 Discussion . . . 164
6.3 Speckle suppression in ultrasound sequences . . . 165
6.3.1 Introduction . . . 165
6.3.2 Problem Statement . . . 166
6.3.3 Weights Computation . . . 168
6.3.4 Experimental Results . . . 170
6.3.5 Discussion . . . 172
6.4 Conclusion . . . 174

7. Conclusion . . . 175

8. Conclusion en Français . . . 181

8.1 Perspectives 183

Bibliography 187

List of Figures

2.1 Neighborhhod system and its associated cliques 44

3.1 Example of confusing case for texture: (b) The skin texture (c) A noise patch. . . . 53

3.2 Example of feature distribution for a an image corrupted by Gaussian synthetics noise of standard deviation 20 (b) Joint distribution (skewness, entropy) (c) Joint distribution (entropy, variance) (d) (b) Joint distribution (skewness, variance) . . . 58

3.3 Example of feature distribution for a an image corrupted by real digital camera noise (b) Joint distribution (skewness, entropy) (c) Joint distribution (entropy, variance) (d) (b) Joint distribution (skewness, variance) 58

3.4 Histogram of the projection of the features vector on the first principle component (red) and their approximation with Gaussian mixture models (blue) for (a) Old man image with real camera noise (b) Barbara image 62

3.5 Results of an image partition: (a) original image, and conditional probability function relative to (b) "smooth component" $p(o_x|smooth)$, (c) "texture" $p(o_x|tex)$ and (d) "edges" $p(o_x|edge)$ 63

3.6 Results of an image partition: (a) original noisy image (σ_n=20), and conditional probability function relative to (b) "smooth component" $p(o_x|smooth)$, (c) "texture" $p(o_x|tex)$ and (d) "edges" $p(o_x|edge)$ 64

3.7 Zoom on a detail in the Barbara image (a) original image (b) noisy image (c) restoration results using our technique with variable bandwidth kernels (d) restoration results using our technique with fixed bandwidth kernels (e) restoration result obtained with NL-means algorithm (f)restoration result obtained with UINTA algorithm 75

3.8 Results of our proposed denoising method on real digital camera Noise, (a) original image (b) variable bandwidth function (low intensity ($h_{\mathbf{y}}$=2), high intensity ($h_{\mathbf{y}}$=4)) (c)MPM_{fix} denoising, (d) MPM_{var} denoising. 76

3.9 Results of our proposed denoising method on real digital camera Noise, (a) original image (b) variable bandwidth function (low intensity ($h_{\mathbf{y}}$=2), high intensity ($h_{\mathbf{y}}$=4)), (c)MPM_{fix} denoising, (d) MPM_{var} denoising. 77

3.10 Results of our proposed denoising method on real digital camera Noise, (a) original image (b)MPM_{fix} denoising, (c) MPM_{var} denoising. 78

4.1 Overview of "Random Walks" based image enhancement. 83

4.2 Two pdf distributions $p_{\mu,\sigma}(\mathbf{d})$ for different values of μ and σ (top ($\mu = 39, \sigma = 11.67$), bottom ($\mu = 96, \sigma = 3.55$), and sample generation according to these pdf (red pixel) for two different positions 85

4.3 Illustration of systematic resampling scheme 90

4.4 Example of three particles walks (red pixel) starting from the same origin position (green pixel), the walks follow image structure. 94

4.5 (a) Noise free image (b)Image corrupted with Gaussian noise $\sigma_n = 20$ (c) Bilateral filter restoration (d) Residual of the bilateral filter. 99

4.6 (a) Restoration result obtained with NL-means algorithm (b) NL-mean residual (c) Restoration result obtained with random walks algorithm (L^2 distance) (d) Residual of the random walk based approach 100

4.7 (a) Restoration result obtained with random walks algorithm (Sobolev distance) (b) Residual of the random walk based approach (c) Restoration result obtained with KB06 [76] (d) KB06 [76] residual 101

4.8 Zoom on the Baboon image (a) Original image (b) Noisy image (c) Bilateral filter restoration (d) NL-Means restoration (e) Random walk restoration (f) Adaptive window size restoration. 102

4.9 Zoom on the fingerprint image (a) Original image (b) Noisy image (c) Bilateral filter restoration (d) NL-Means restoration (e)Adaptive window size restoration (f) Random walk restoration. 103

4.10 (a) Original Lena image (b) image corrupted by speckle noise with variance $\sigma_n = 0.025$ (c) Random walk based restoration result (d) Total variation minimizing based restoration [7] 105

4.11 (a) Original Baboon image (b) image corrupted by speckle noise with variance $\sigma_n = 0.025$ (c) Random walk based restoration result (d) Total variation minimizing based restoration [7] 106

5.1 (a) Original image (b) The bandwidth value associated to it 121

5.2 (a) Example of texture (b) Texture corrupted by Gaussian noise $\sigma_n^2 = 20$ – Similarity measure between the central pixel (Red) and the other one using : (c) Using definition (5.16) and the noisy image (d) Using definition (5.33) and the noisy image (patch size $d = 81$) (e) Using definition (5.33) and the noisy image (patch size $d = 25$) 124

5.3 Eigenvectors obtained through PCA decomposition of Barbara image corrupted by Gaussian noise ($\sigma_n = 10$) (patch size 15×15) (Only the first 22 component are significant if T=3) 126

5.4 (a) Projection of the observations set on the first eigenvector of Barbara image (b) Projection of the observation set on the 17^{th} eigenvector (c) Projection of the observations set on the Last eigenvector (49^{th}) 127

5.5 (a) Example of texture (b) Texture corrupted by Gaussian noise $\sigma_n^2 = 40$ – Similarity measure between the central pixel (Red) and the other one using : (c) the L^2 distance between patches (d) feature set based on PCA (e)the L^2 distance between patches using the noise free image 127

5.6 Zoom on the Barbara image (a) Original image, (b) Noisy image (c) Our method using fixed bandwidth (d) Our method using variable bandwidth (e) Non local functional minimizing (5.9) (f) Total variation minimizing with variable local constraint 130

5.7 Zoom on the Barbara image (a) Total variation (b) Anisotropic Diffusion 131

5.8 Zoom on the Lena image (a) Original image, (b) Noisy image (c) Our method using fixed bandwidth (d) Our method using variable bandwidth (e) Non local functional minimizing (5.9) (f) Total variation minimizing with variable local constraint . . . 132

5.9 Zoom on the Lena image (a) Total variation (b) Anisotropic Diffusion 133

5.10 (a) Residual of the Barbara image using the non local convex functional 5.9 (b) Residual of the Barbara image using our approach with fixed bandwidth (c) Residual of the Lena image using the non local convex functional 5.9 (d) Residual of the Lena image using our approach with fixed bandwidth 134

5.11 Zoom on Barbara image and result of our method using different weights definition and fixed bandwidth (a) using distance between pixels features using PCA (b) using the statistical distribution of the L^2 distance between patches and expression (5.33) (c) using L^2 distance between patches and expression 135

5.12 Zoom on Lena image and Result of our method using different weight definition and fixed bandwidth (a) using distance between pixels features using PCA (b) using the statistical distribution of the L^2 distance between patches and expression (5.33) (c) using L^2 distance between patches 136

5.13 Results using NL-means algorithm (a) Original image (b) Noisy image (c) using the statistical distribution of the L^2 distance between patches and expression (5.33) (d) The corresponding residual image 137

5.14 Results using NL-means algorithm (a) using L^2 distance between patches (b) Its corresponding residual image (c) using distance between pixels features using PCA (d) Its corresponding residual image 138

5.15 (a) original image (b) Noisy image. Results using NL-means algorithm (c) using L^2 distance between patches (d) using distance between pixels features using PCA (e) using the statistical distribution of the L^2 distance between patches and expression (5.33) 139

5.16 Evolution of the PSNR value with respect to the retained number of principle component 140

5.17 Evolution of the PSNR with respect to the window radius : blue curve corresponds to the classical weight definition (5.16) , the red curve corresponds to the new definition (5.33) 140

6.1 (a) Macbeth Color Checkers (b) Curve corresponding to the evolution of the standard deviation of the noise with respect to the intensity for each channel Red ,Green and Blue 147

6.2 Example of a raw Image with the Bayer Pattern 147

6.3 (a) Original noisy image (b) Image restored using our method with fixed noise variance (c) Image restored using the NLmean algorithm (d) Image restored using our method with variable noise variance 151

6.4 Difference between the noisy image and the restored one using (a) our method with fixed noise variance (b) Image restored using the NLmean algorithm (c) Image restored using our method with variable noise variance 152

6.5 (a) Original noisy image (b) Image restored using our method with fixed noise variance (c) Image restored using the NLmean algorithm (d) Image restored using our method with variable noise variance 153

6.6 Difference between the noisy image and the restored one using (a) our method with fixed noise variance (b) Image restored using the NLmean algorithm (c) Image restored using our method with variable noise variance 154

6.7 Tensors on a volume slice (Homogeneous tensor field): (a) Noisy tensors (b) Ground-truth (c) Result obtained with [43] (d) Result obtained with our method . . 162

6.8 Tensors on a volume slice (helix): (a) Noisy tensors (b) Ground-truth (c) Result obtained with [43] (d) Result obtained with our method 163

6.9 A slice of the T1-weighted volume, different muscle groups segmented manually . 164

6.10 Estimated tensors without regularization, tensors obtained with our method 164

6.11 (a-1) Image (Sythetic1) corrupted by the speckle $\sigma_n = 0.5$ (a-2) Image (Sythetic2) corrupted by the speckle $\sigma_n = 1$ (b) Result using the anisotropic diffusion [154] (c) Result using the wavelet based technique [116] (d) Result using our algorithm . 172

6.12 Results of restoration on real ultrasound frames. (a) Observed image (b) Result using algorithm [116] (c) Anisotropic diffusion [154] (d) Results of our algorithm . 173

6.13 Results of filtering real ultrasound frames (a) Observed image (b) Result of our algorithm without temporal component ($T_w = 0$) (frame by frame filtering) (c) Result of our algorithm using temporal filtering ($T_w = 2$) (d) Residual obtained with our algorithm without temporal component ($T_w = 0$) (frame by frame filtering)(e) Residual obtained with our algorithm using temporal filtering ($T_w = 2$) . . . 173

7.1 Results of Barbara image restoration (a) using Random walks (b) using MPM estimation with variable kernels bandwidth (c) using functional minimization based approach with variable bandwidth spatial kernel 178

7.2 Results of Baboon image restoration (a) using Random walks (b) using MPM estimation with variable kernels bandwidth (c) using functional minimization based approach with variable bandwidth spatial kernel 179

8.1 Résultat de la restoration de l'image Barbara (a) En utilisant les marches aléatoires (b) En utilisant l'estimation MPM et les noyaux de taille variable (c) Minimisation de fonctionnelle en utilisant des noyaux de taille variable 184

8.2 Résultat de la restoration de l'image Baboon (a) En utilisant les marches aléatoires (b) En utilisant l'estimation MPM et les noyaux de taille variable (c) Minimisation de fonctionnelle en utilisant des noyaux de taille variable 185

Chapter 1

Introduction en Français

La restauration est l'un des composantes fondamentales du traitement des images. Cela consiste à déterminer une image sans bruit à partir d'une observation corrompue. Dans ce contexte, on doit considérer trois aspects importants qui sont (i) une modélisation adéquate de l'image et du bruit (ii) une modélisation de la géométrie de l'image et des interactions entre les pixels (iii) la sélection de la fonctionnelle qui encode les modèles image et bruit ainsi que les interactions entre pixels afin de calculer l'image débruitée. Dans la suite nous présenterons le contexte de ce travail ainsi que les motivations et les principales contributions.

1.1 Présentation du problème et Motivations

On assiste de nos jours à une prolifération d'images numériques dans la vie de tous les jours à travers de nombreuses applications telles que la photographie, l'imagerie médicale, la vidéo surveillance, la navigation et le contrôle industriel, etc. Malgré le progrès important réalisé durant les dernières années par les constructeurs de caméras au niveau de la qualité des images, ces dernières ont besoin d'une étape de prétraitement avant son exploitation. Le besoin d'un compromis entre la qualité de l'image et le prix des caméras, ainsi que le renforcement de la concurrence entre les constructeurs pour produire des caméras à moindre coût rendent les solutions logicielles attractives. Ainsi, les applications de restauration des images telles que le débruitage le déflouage et le inpainting ont attiré beaucoup d'attention dans la communauté de la vision par ordinateur et le traitement des images. Parmi les différents types de dégradations, nous allons particulièrement nous intéresser à la suppression de bruit.
Pour un capteur standard et pour chaque pixel de l'image observée, l'intensité corresponds au

nombre de photons qui atteint un élément de capteur ou photosite durant une période T. Le bruit correspond alors à la fluctuation du nombre de photons par rapport à la moyenne qui correspond à la vraie intensité. La chaleur dégagée par le dispositif électronique est une source de bruit supplémentaire en générant des photons qui affectent chaque photosite et produisent le "bruit d'obscurité". D'autre part, chaque photosite génère du bruit qui peut contaminer les pixels voisins. Plus généralement, on peut approximer la relation entre l'image observée I qui est mesurée par le capteur et l'image sans bruit U pour chaque pixel $\mathbf{x}$ par

$$I(\mathbf{x}) = U(\mathbf{x}) + n(\mathbf{x}) \tag{1.1}$$

où n est un bruit qui dépend du pixel et qu'on assume blanc et de moyenne nulle. Le bruit blanc est caractérisé par l'indépendance entre ses différentes réalisations.
Ainsi, le débruitage est l'estimation de l'image U à partir d'une observation bruitée I et de certaines hypothèses sur le modèle de bruit. Pour résoudre ce problème, diverses méthodes ont été proposées au cours des cinq dernières décennies. Malgré les progrès importants réalisés de nombreux défis doivent encore être soulevés. L'efficacité d'un algorithme de débruitage est liée à sa capacité de préserver les informations incluses dans l'image tels que les bords, la texture et les petits détails. Alors que la contrainte de préservation du contour a été respectée par la plupart des algorithmes non linéaires de débruitage, un effort doit encore être consacré au débruitage de la texture. Cela est dû au fait que la plupart des algorithmes de débruitage supposent que l'image est constante par morceaux. Une telle hypothèse n'est pas adéquate avec les caractéristiques des images naturelles qui peuvent contenir certaines structures aussi oscillantes que le bruit. De nombreuses recherches pour la modélisation de la texture ont été menées, mais leur efficacité est dépendante de la taille de la texture ainsi que de sa structure. Ce dernier point rend pertinents les efforts consentis afin de trouver des modèles d'image plus réalistes et des algorithmes de débruitage qui préservent mieux la texture.

1.2 Les Contributions

Le travail présenté dans cette thèse est un pas vers une meilleure compréhension des images et de la texture dans le but de déruitage. L'objectif principal est d'adapter au contenu de l'image les modèles utilisés dans la restauration. Dans cette dissertation, nous présenterons des modèles mathématiques et les solutions numériques qui utilisent l'information appris à partir de l'image observée afin de construire une approche appropriée de débruitage. Notre principale motivation est de concevoir une technique qui utilise des modèles d'images différents selon le contexte local.

Cette thèse présente de nouvelles approches théoriques dans le but de débruiter des images tout en préservant la texture. Dans un premier temps, nous nous intéresserons aux aspects photométriques du problème dans le but de définir des modèles appropriés capables de décrire l'observation et l'échelle des interactions photométriques entre les pixels. Ensuite, nous allons étudier l'importance de la géométrie dans la définition des dépendances entre les pixels et par conséquents dans les interactions spatiales ente eux. Le dernier volet de la thèse consiste en deux contributions, un terme de régularisation plus général qui permet de prendre en compte la complexité de l'image ainsi qu'une méthode automatique de sélection de la taille de la bande passante spatiale. Dans cette dernière partie nous nous intéresserons également à la définition des poids qui régissent les interactions entre les pixels en fonction du contenu de l'image.

Notre première contribution consiste à introduire la notion de classification dans le processus de débruitage en utilisant des descripteurs locaux. Plusieurs techniques cherchent à s'adapter au contenu de l'image en prenant en compte les caractéristiques de chaque pixel pendant le débruitage. Mais, certaines textures ont les caractéristiques similaires au bruit ce qui empêche l'algorithme de restauration de les détecter. Pour cette raison, nous pensons qu'une étape de pre classification qui fourni un outil plus robuste pour identifier les zones texturées dans l'image est nécessaire. Cette classification consiste à partitionner l'image en des régions localement lisses, des régions texturées et des contours. Ceci est effectué à l'aide d'une classification dans un espace engendré par les caractéristiques locaux qui permettent de décrire les zones homogènes, la texture et les contours. La projection des observations de cet espace dans un autre sous espace est modélisée par une mixture de Gaussienne où chaque composante est associée à chaque classe de pixels. Par la suite, le résultat de l'étape de classification sera intégré dans l'algorithme de débruitage. La technique de filtrage s'appuie sur une technique non paramétrique avec des noyaux afin d'estimer le modèle image. Dans ce contexte, nous allons proposer une méthode automatique de sélection de la taille de ces noyaux qui dépend du résultat de l'étape de classification. En effet, dans le processus de filtrage nous traitons les pixels différemment selon leur degré d'appartenance à l'une des trois composantes de l'image. Ce dernier point est la contribution majeure par rapport aux techniques déjà existante et qui utilisent pour modéliser l'image des techniques non paramétrique d'estimation de densité de probabilité [118, 10]. Contrairement à ces méthodes, nous utilisons des noyaux de tailles variables qui dépendent des propriétés locales du pixel.

La deuxième approche étudiée dans cette thèse est basée sur les marches aléatoires. Notre méthode explore plusieurs ensembles de voisins (ou hypothèses) qui peuvent être utilisés pour le débruitage d'un pixel, à travers une approche de filtrage des particules. L'objectif est de proposer une méthode de sélection des pixels de l'image les plus pertinents qui vont être mis en jeu pour l'estimation de l'intensité d'un pixel donné. En s'appuyant sur une technique de filtre à particules,

la sélection des pixels se fait d'une manière progressive. Nous considérons d'abord un voisinage de taille petite puis on ajoute de plus en plus de pixels en explorant le domaine de l'image tout en étant dirigé par les structures. Contrairement aux filtres de voisinage classiques, le domaine considéré lors du filtrage est adapté à chaque pixel. Le processus de filtrage met en jeu un nombre de particules qui explorent le domaine de l'image en utilisant une distribution statistique qui décrit la géométrie de l'image ainsi l'état d'une particule se réfère à l'état du processus de reconstruction de l'image.

Les deux méthodes précédentes permettent de restaurer l'image pixel par pixel. Bien qu'il s'agisse d'un moyen simple de restauration, un processus d'homogénéisation global où toute l'image est itérativement mise à jour doit être envisagé. De plus, ces deux techniques incluent deux étapes : la caractérisation de la texture et le filtrage. Pour faire face à cette limitation, nous allons considérer un modèle d'image global qui encode implicitement la structure de l'image. Le débruitage sera donc effectué à l'aide de la minimisation d'une fonctionnelle quadratique et convexe qui implique aussi des noyaux à taille variable. Ces noyaux sont utilisés dans le calcul de similarités spatiales et photométriques entre les pixels. Afin de préserver la texture et améliorer la qualité de l'estimation, nous allons considérer que l'échelle des interactions spatiales est variable en fonction du pixel. Ceci permettra de l'adapter au contenu de l'image et à l'échelle de sa texture. La définition d'une mesure de similarité appropriée qui soit plus robuste au bruit que la distance L^2 entre les patches a été également étudiée dans cette dissertation. Cette distance est calculée entre les vecteurs caractéristiques qui sont obtenus par projection des patches de l'image dans un autre sous-espace permettant une meilleure description de la structure des patchs de l'image. De plus, nous avons proposé une nouvelle définition de poids qui est plus cohérente avec la distribution statistique de la distance L^2 entre les patchs de l'image.

Nous avons présenté dans cette thèse plusieurs extensions de la technique de filtrage en minimisant une fonctionnelle d'énergie convexe à d'autres types de bruit et de donnée afin de montrer sa flexibilité. Cette extension concerne (i) la restauration des images en couleur où on doit prendre en considération les propriétés du bruit relatif aux appareils photo numériques (ii) l'estimation et la régularisation des tenseurs de diffusion où les tenseurs doivent être définis positifs (iii) filtrage des séquences Ultrasonores où on adapte la formulation d'énergie à la nature du bruit multiplicatif.

1.3 Plan de la thèse

Cette thèse est organisée comme suit: Le deuxième chapitre est dédié aux modèles non paramétriques d'estimation de densité de probabilité ainsi qu'aux techniques de partition de l'image. Au début,

nous allons présenter un travail lié à cette technique qui est un filtre adaptatif non supervisé (UINTA) et nous pointerons les différences avec ce filtre. Par la suite, nous allons décrire notre approche de partition de l'image en trois classes "régions homogènes", "régions texturées" et "contours". Cette partition s'appuie sur le calcul de descripteurs locaux et la classification à l'aide de mixture de Gaussienne. Après l'étape de classification nous allons nous concentrer sur le débruitage. Nous allons passer en revue la théorie de la décision Bayesienne et les différents types d'estimateurs. Ensuite, nous allons présenter notre technique de débruitage utilisant une technique de Maximum a Posteriori Marginal (MPM) où on estime cette loi a posteriori à l'aide d'une approche non paramétriques. Par la suite, nous nous intéresserons à la sélection de la taille du noyau et au processus de l'optimisation. L'évaluation des performances sera présentée à la fin du chapitre.

Le troisième chapitre sera dédié aux filtres de voisinage et nous utiliserons une stratégie basée sur les marches aléatoires et les filtres à particule afin de sélectionner l'ensemble de pixel le plus adapté pour débruiter un pixel donné. Au début, nous allons présenter un modèle statistique qui permet de décrire la géométrie de l'image et les relations spatiales entre les pixels similaires. Un tel modèle sera utilisé dans le contexte du filtre à particules afin de guider l'évolution de ces dernières. Par la suite, nous présenterons une description brève des techniques des filtres à particules. Nous allons également appliquer cette technique d'estimation pour restaurer des images aussi bien dans le cas de bruit additif que multiplicatif. Nous allons conclure cette section par la validation expérimentale de cette technique.

Dans le quatrième chapitre on va présenter une technique de régularisation qui consiste à minimiser une fonction d'énergie convexe. Au début, on va commencer par présenter l'état de l'art des techniques variationelle. Ensuite, nous allons nous intéresser à la description du modèle que nous utiliserons ainsi que le processus de diffusion sous-jacent. Après la définition du modèle, notre objectif sera la définition des interactions entre les pixels. Cette interaction est régie à l'aide d'une fonction de poids définie par deux noyaux : un qui pénalise la distance spatiale entre les pixels et l'autre la distance photométrique. Dans ce contexte nous avons essayé d'améliorer la définition des poids en optimisant la sélection de la taille du noyau spatial (celui qui gère les interaction spatiales entre les pixels) et en l'adaptant à chaque pixel. Nous nous sommes également intéressés aux similarités photométriques entre les pixels et nous avons proposé deux nouvelles définitions de mesures. La première s'appuie sur une meilleure caractérisation des pixels et la deuxième exploite les propriétés statistiques de la distance L^2 entre patches. Finalement, nous conclurons ce chapitre avec les résultats expérimentaux.

Dans le cinquième chapitre, nous allons étudier plusieurs types de problèmes de régularisation

autres que le problème classique du bruit additif Gaussien et les images de niveau de gris. Au début nous allons considérer le problème de filtrage des images en couleur corrompues par le bruit des appareils photo numériques. Pour cette application, nous allons discuter le modèle de bruit et proposer une technique non paramétrique d'estimation de la fonction de bruit. Cette fonction décrit l'évolution de la variance de bruit en fonction de l'intensité. Par la suite, nous introduirons la procédure de débruitage des images en couleur à l'aide de la minimisation sous contraintes de la fonction d'énergie présentée dans le chapitre précédent. La deuxième application concerne des données de plus grandes dimensions sur des manifolds tels que l'exemple de l'imagerie des tenseurs de diffusion. L'objectif est d'estimer et de régulariser simultanément les tenseurs des diffusions à partir d'images IRM. Les performances de cette méthode ainsi que son impact sur la classification des différents muscles seront évalués. La dernière section est dédiée à la suppression du Speckle dans des séquences ultrasons. Pour atteindre cet objectif nous avons adapté le terme d'attache aux données ainsi que la définition des poids au modèle du bruit. Les validations expérimentales de cette technique seront fournies a la fin de ce chapitre.

La conclusion et la discussion seront présentées dans la dernière partie de ce document. Nous parlerons des limitations des différentes approches proposées ainsi que les perspectives futures de ce travail.

Pour conclure, cette thèse s'articule autour la modélisation des images et du bruit dans le contexte de la restauration des images dans le domaine de la photographie numérique ou l'imagerie médical. Cette thèse a produit (jusqu'a maintenant), un chapitre de livre [109] quatres articles dans des conférences internationales de référence [12, 15, 13, 11] deux articles dans des workshops [14, 110], et des articles de revue en cours (International Journal of Computer Vision, IEEE Transactions on Pattern Analysis and Machine Intelligence, Journal of Mathematical Imaging) ainsi qu'une implication dans un brevet français [30].

Chapter 2

Introduction

Image restoration is one of the most fundamental components of image processing. It consists of recovering a noise-free signal from corrupted observations. Such a problem arises in a number of fields and has been heavily studied in the past decades. In such a context, one has to address three critical aspects, that are (i) appropriate modeling of the image towards understanding the noise level, (ii) appropriate modeling of the geometric dependencies between observations in the image domain and (iii) appropriate selection of a functional that encodes the previously mentioned components towards recovering the noise-free signal. In this chapter, we review the most representative techniques of the field, discuss their strength as well as their limitations in particular with respect to the above mentioned components.

2.1 Problem Statement and motivation

We witness currently the proliferation of digital images in every day life as well as in a number of industrial domains like photography, medical imaging, video surveillance, navigation, industrial inspection etc. Despite important progress made over the past decades from camera manufacturers on the quality of acquisition, these images need often a preprocessing step to be exploitable in a number of fields due to the need of a compromise cost versus quality. This is amplified due to the competition towards providing low cost cameras. Based on the assumption that conventional hardware sensors have reached a certain maturity, the investment needed to improve the quality of the images is disproportional to that improvement. This context favoured the emergence of software/algorithmic approaches toward image quality improvements. Therefore, image restoration application such as denoising, deblurring and inpainting gained a lot of attention from the com-

puter vision and image processing community. Among all types of image degradations, we will focus in this thesis particularly on the noise suppression.

Let us introduce the problem using some standard conventions. Using conventional sensors, the observed image intensity for a given pixel corresponds to the number of photons that reach a cell of the light sensors matrix (called photosite) during a period T. The noise then refers to a fluctuation in the number of photons with respect to a mean value that is the actual intensity of the pixel. The heat generated by the electronic device is an additional noise source; it frees electrons that affects each photosite and gives rise to the "dark noise". On the other hand each photosite itself generates electrical noise that can contaminate its neighbor. In the most general case, one can approximate the relation between the observed image I measured by the sensor and the noise-free image U for each pixel $\mathbf{x}$ by

$$I(\mathbf{x}) = U(\mathbf{x}) + n(\mathbf{x}) \tag{2.1}$$

Where n is a pixel dependent additive noise that is often assumed to be white and of zero mean. The white noise is characterized by the independence between the random variables that correspond to noise realizations for two different pixels.

Hence, denoising refers to estimating the image U given the corrupted observation I and certain assumptions on the noise model. To address this problem various methods were proposed for the past five decades and despite important progress made many challenges are still to be dealt with. The efficiency of a denoising algorithm is related to its ability to preserve image content such as edges, texture and other fine details. While the contour preserving constraint was respected for the most non linear denoising algorithm, some effort has still to be devoted to texture denoising. This is due to the fact that most denoising algorithms are based on mathematical models describing the image as smooth or piecewise smooth. Such an assumption does not comply with natural image characteristics that may contain some structures that are as oscillatory as noise. Many research were also carried toward texture modeling but their efficiency depends on the scale of the texture and on its structure. The latter point makes relevant the effort dedicated to consider more realistic models and to design texture preserving denoising approaches. In this introduction we will briefly review the state of the art in the field and describe their underlying concepts. This review encompass neighborhood filters, Partial derivative, equation based regularization, sparse image representation, and Markov random fields models.

2.2 State of the Art

Image processing and computer vision literature includes different kinds of algorithms to address image enhancement. To this end, a wide variety of mathematical tools was considered such as: signal approximation in sub-spaces, variational calculus, partial derivative equation, probability and statistics, information theory, etc. The definition of frontiers between the different available methods is not straightfoward. In many cases, one can find eminent links between different concepts like PDE's based approaches and total variation minimization or neighborhood filters and wavelet thresholding with the Bayesian decision theory etc. Nevertheless we consider four categories for image restoration techniques : (i) methods related to neighborhood filtering that consists in performing a weighted averaging (ii) PDE's based approaches that are iterative techniques that yield a smoother version of the image with time. (iii) the sparse image representation and domain transforms techniques which decompose the image into a sub-space where the image can be approximated by a few number of coefficients (iv) statistical methods and mainly the Markov Random Fields based approaches.
In order to introduce these methods, some basic definitions of the the underlying image representation are to be considered. We can find in the litterature three distinct forms to represent an image.

- Continuous model where the image is a function defined on a subset $\Omega \subset \mathbb{R}^2$ called the image domain and associates to each element of Ω a value in the set $\{0 \dots 255\}$ for gray level images quantified and coded on 8 bits. In this case the pixel $\mathbf{x} \in \mathbb{R}^2$ is characterized by the couple of its real coordinates.

- Discrete model where the image is defined on a discrete grid $\Omega \subset \mathbb{Z}^2$ called the image domain and with values from the set $\{0 \dots 255\}$ for gray level images quantified and coded on 8 bits. In this case the pixel $\mathbf{x} \in \mathbb{Z}^2$ is characterized by the couple of its integer coordinates.

- Stochastic model where the image is assumed to be a sample of a random variable $U = \{U(\mathbf{x}_k)\}_{1 \leq k \leq |\Omega|}$. $U(\mathbf{x}_k)$ is a random variable that takes values in $\{0 \dots 255\}$ and describes the intensity observed at a pixel $\mathbf{x}_k$.

Using these definitions let us now proceed with a brief review of the state of the art.

2.2.1 Averaging Based Filters

Averaging based filtering approaches are among the most primitive and the most widely used techniques in the field. The central idea is to reduce noise in the image by performing a weighted average of the other pixels in the image. A general formulation of such a filter is

$$U(\mathbf{x}) = \int_{\mathcal{D}} h(\mathbf{x},\mathbf{y}) I(\mathbf{y}) d\mathbf{y} \quad \text{and} \quad \int_{\mathcal{D}} h(\mathbf{x},\mathbf{y}) d\mathbf{y} = 1 \tag{2.2}$$

Where $\mathcal{D}$ is the domain where the averaging is computed. It can be a local neighborhood of $\mathbf{x}$ or the entire image domain.
The selection of the function h is fundamental for the filtering process. The well known Gaussian filter is derived if h is a Gaussian function depending on the spatial distance between pixels. It is well known that the Gaussian filter and other linear ones result in blurred edges and fine details suppression. For this reason, non linear and data driven weighting functions were proposed to ensure better restoration. The underlying concept of these filters is to go further than the spatial distance between pixels to include other image features such as intensity. Despite an enormous volume of research literature in this field, we will focus on the most representative ones:

- **The neighborhood filter**
  These filters performs a weighted average over a local neighborhood of $\mathbf{x}$ (noted $\Pi_{\mathbf{x}}$) where the contribution of each pixel is dependent on the similarity between pixels. Earlier approach refers to the sigma filter [92] where a threshold on the intensity difference between neighboring pixel is considered to discard the irrelevant ones. In [153] a similar idea is expressed through the definition of weights coefficients depending on the gray level difference between pixels. The estimated intensity is written as

$$U(\mathbf{x}) = \frac{1}{Z(\mathbf{x})} \int_{\Pi_{\mathbf{x}}} I(\mathbf{y}) \mathrm{e}^{-\frac{|I(\mathbf{x})-I(\mathbf{y})|^2}{\sigma_{ph}^2}} d\mathbf{y} \quad Z(\mathbf{x}) = \int_{\Pi_{\mathbf{x}}} \mathrm{e}^{-\frac{|I(\mathbf{x})-I(\mathbf{y})|^2}{\sigma_{ph}^2}} d\mathbf{y} \tag{2.3}$$

  σ_{ph} is a parameter that controls the smoothing amount of the filter. Such a definition would preserve edges and other details since the neighboring pixels with an important intensity difference do not contribute in the reconstruction process (when compared with the σ_{ph}). One can see a direct equivalence between this filter and the gaussian filter when σ_{ph} value are important.
  Recent equivalent formulations of this filter are the SUSAN Filter [127] and the Bilateral Filter [134, 112] that combine the spatial distance as well as photometric distance to compute

the weights. These filters are defined as follows

$$U(\mathbf{x}) = \frac{1}{Z(\mathbf{x})} \int_{\Omega} I(\mathbf{y}) \mathrm{e}^{-\frac{\|\mathbf{x}-\mathbf{y}\|^2}{\sigma_s^2}} \mathrm{e}^{-\frac{|I(\mathbf{x})-I(\mathbf{y})|^2}{\sigma_{ph}^2}} d\mathbf{y} \quad Z(\mathbf{x}) = \int_{\Omega} \mathrm{e}^{-\frac{\|\mathbf{x}-\mathbf{y}\|^2}{\sigma_s^2}} \mathrm{e}^{-\frac{|I(\mathbf{x})-I(\mathbf{y})|^2}{\sigma_{ph}^2}} d\mathbf{y} \quad (2.4)$$

σ_s is a spatial parameter that defines the radius of the neighborhood considered for denoising. As in the former case, when σ_{ph} is high this filter is equivalent to the Gaussian kernel smoothing.

This class of approaches combines computational efficiency (local operations) with satisfactory results. The main challenge refers to the appropriate definition of their parameters while at the same time, the computation of weights might be problematic when the images are heavily corrupted. This is due to the fact that simple pair-wise image intensities cannot encode the relation between the local observations. This was addressed by the NL-means algorithm.

- **The Non Local means filter (NL-means)** [26, 25]

  The NL-means is an algorithm that takes advantage of the redundancy and similarity inside an image to perform denoising. Exploiting this natural image property was first used in case of texture synthesis in [51]. Based on this observation, the NL-mean weight function relies on a similarity measure that goes beyond pixel-wise resemblance and is determined through comparison of local image patches. Since similar patches can be found everywhere in the image (long edges, large textured regions with repetitive patterns), the spatial distance between pixels is neglected when considering the definition of the weight. The restored intensity obtained using the NL-means algorithm is given by the following expression

$$U(\mathbf{x}) = \frac{1}{Z(\mathbf{x})} \int_{\Omega} I(\mathbf{y}) \mathrm{e}^{-\frac{d(\mathbf{x},\mathbf{y})}{\sigma_{ph}^2}} d\mathbf{y} \quad Z(\mathbf{x}) = \int_{\Omega} \mathrm{e}^{-\frac{d(\mathbf{x},\mathbf{y})}{\sigma_{ph}^2}} d\mathbf{y} \quad (2.5)$$

$$d(\mathbf{x},\mathbf{y}) = \int_{\mathbb{R}^2} G_{\alpha}(t) \left| I(\mathbf{x}+t) - I(\mathbf{y}+t) \right|^2 dt \quad (2.6)$$

  Where G_{α} is a Gaussian kernel of bandwidth α that defines the size of the window (or image patch) taken for pixel comparison while σ_{ph} plays the same role as in the previously cited filters. The discrete formulation of the NL-means filter requires the definition of a neighborhood $N_{\mathbf{x}}$ and $I_{\mathbf{N}_x}$ the observed intensities within this neighborhood.

$$U(\mathbf{x}) = \frac{1}{Z(\mathbf{x})} \sum_{\mathbf{y} \in \Omega} I(y) e^{-\frac{\left\| I_{\mathbf{N}_x} - I_{\mathbf{N}_y} \right\|_2^2}{\sigma_{ph}^2}} \quad Z(\mathbf{x}) = \sum_{\mathbf{y} \in \Omega} e^{-\frac{\left\| I_{\mathbf{N}_x} - I_{\mathbf{N}_y} \right\|_2^2}{\sigma_{ph}^2}} \quad (2.7)$$

  With this definition the NL-means acts as a basic averaging with equal weights in flat re-

gions, while performing anisotropic filtering for edges or texture.
The computational cost is a serious limitation for the NL-means algorithm, and this was addressed in [98]. Despite the excellent performance of this filter, the selection of the filtering window (pixels contribution in the process), as well as the similarity measure between patches (both in terms of the choice of the metric and the contributing content) are two open issues. This was partially addressed in [76] where the idea of adaptive/variable filter bandwidth was considered.

- **Variable neighborhood size filters**
  These are algorithms where the size of the domain $\mathcal{D}$ over which the mean is computed depends on the pixel position. The optimal window size is determined according to a trade off between variance and bias of the intensity estimator using local averaging [52, 117, 93, 76]. The most recent contribution in this area was presented in [76]. The authors suggest an alternative approach to the NL-means algorithm where the optimal filter bandwidth is computed for each pixel. To give an overview of this method, let us consider a local or semi-local version of the NL-means algorithm. The estimated intensity for a given pixel $\mathbf{x}$ is

$$\hat{U}(\mathbf{x}) = \sum_{\mathbf{y}\in\mathcal{D}_\mathbf{x}} I(y) w_{\mathbf{xy}} \quad \text{with} \quad w_{\mathbf{xy}} = \frac{e^{-\frac{d^2(I_{\mathbf{N}_x}, I_{\mathbf{N}_y})}{\sigma^2_{ph}}}}{\sum_{\mathbf{y}\in\mathcal{D}_\mathbf{x}} e^{-\frac{d^2(I_{\mathbf{N}_x}, I_{\mathbf{N}_y})}{\sigma^2_{ph}}}} \tag{2.8}$$

  $d^2(I_{\mathbf{N}_x}, I_{\mathbf{N}_y})$ is the L^2 distance between patches.
  Now if one considers the error expectation between the estimated intensity and the actual noise free one, the following relation is obtained

$$\mathbf{E}\left[(\hat{U}(\mathbf{x}) - U(\mathbf{x}))^2\right] = b^2_\mathbf{x} + v^2_\mathbf{x} \tag{2.9}$$

  with $b_\mathbf{x}$ [1] being the bias of the estimator that characterizes the accuracy of the approximation of the image at the pixel $\mathbf{x}$ using the local neighborhood $\mathcal{D}_\mathbf{x}$. The variance of the estimator $v_\mathbf{x}$ corresponds to the amount of fluctuation around the mean value of the estimator.
  Ideally the perfect estimator is the one that minimizes both bias and variance which are competing variables. Assuming that the bias is increasing with the size of $\mathcal{D}_\mathbf{x}$ and the variance is decreasing, there exists an optimal size of $\mathcal{D}_\mathbf{x}$ for which a balance between the variance and bias is reached. A direct estimation for this optimal size is not possible since the bias is not available. As an alternative, the author introduced in [76] a data driven window size selector. The underlying key idea is the following : if one considers $\hat{U}^p(\mathbf{x})$ and $\hat{U}^q(\mathbf{x})$ two estimates

[1] $b_\mathbf{x} = U(\mathbf{x}) - \mathbf{E}(\hat{U}(\mathbf{x})) \approx U(\mathbf{x}) - \sum_{\mathbf{y}\in\mathcal{D}_\mathbf{x}} U(y) w_{\mathbf{xy}}$

of $U(\mathbf{x})$ using two windows $\mathcal{D}_{\mathbf{x}}^p$ and $\mathcal{D}_{\mathbf{x}}^q$ of size p and q respectively ($p < q$) smaller than the optimal size value, then it was proved in [75] that

$$\left|\hat{U}^p(\mathbf{x}) - \hat{U}^q(\mathbf{x})\right| \leq \kappa v_{\mathbf{x}}^p \tag{2.10}$$

This observation provides an automatic technique to determine the optimal window size that is defined as:

$$\text{Argmax}_q \left\{ q = |\mathcal{D}_{\mathbf{x}}^q| \quad \text{such that} \quad \left|\hat{U}^p(\mathbf{x}) - \hat{U}^q(\mathbf{x})\right| \leq (2\gamma + \kappa) v_{\mathbf{x}}^p \quad \forall \;\; 1 \leq p \leq q \right\} \tag{2.11}$$

The optimal window is the largest one such that the estimators $\hat{U}^p(\mathbf{x})$ and $\hat{U}^q(\mathbf{x})$ are not too different, for all $1 \leq p \leq q$. Hence, if an estimated intensity using a given window size is far from the intensity value provided by a smaller one, this means that the bias is already too large. In this case a smaller size has to be selected.
A crucial parameter for this method is the threshold κ because a large value of κ yields a large window size when small values favor small windows. For the selection of this parameter the authors use a prior on the image and design their selection process such that high κ values are set for smooth area and smaller one when discontinuities are observed. This algorithm improves substantially the performance of the basic NL-means.

To conclude, we can say that these methods are attractive and simple to implement. Nevertheless one can seek further improvements by adapting in an automatic fashion the shape of the domain $\mathcal{D}$ as well as its size.

2.2.2 PDE's and Energy Based Image Restoration

Over the past decades the image analysis field has seen the emergence of several PDE based models. Image regularization is a field that has largely benefited from these techniques. In such a context, images are considered as evolving functions of time. Within this framework, the final solution of the enhancement process will correspond to the steady-state of the PDE. Such partial differential equations can be either determined through certain expected geometric constraints or can be the outcome of the minimization of a specifically designed cost function. In this section we will review the most classical ones.

- **Geometric Flows and Image Enhancement or axiomatic PDE's**
  The earlier contributions in this field were linear PDE's where the blurring is spatially invariant [151, 85]. For instance, the heat equation is an isotropic filtering and is defined as

$$\frac{\partial U}{\partial t} = c\Delta U \tag{2.12}$$

  Where c is a constant and Δ is the Laplacian operator. In [85] it was proved that performing an isotropic diffusion with $c = 1$ amounts to applying a convolution with a Gaussian kernel $\mathcal{G}_\sigma = \frac{1}{2\pi\sigma}\exp(-\frac{\|\mathbf{x}\|^2}{2\sigma^2})$ with a standard deviation $\sigma = \sqrt{2t}$. The isotropic diffusion blurs the image little by little which is a major drawback in denoising applications.
  This issue was partially addressed by anisotropic diffusion, that is an alternative PDE's based formulation aiming at preserving edges being present in the image. In [115] a paradigm that respects image discontinuities by considering a spatial varying c function instead of the constant one is presented. The underlying PDE is defined using the divergence operator

$$\frac{\partial U}{\partial t} = \text{div}(c(\mathbf{x}, t)\nabla U) \tag{2.13}$$

  To preserve discontinuities the diffusion is conditioned by an appropriate choice regarding the c function. Such a function should favour diffusion in smooth regions and stop it near image boundaries. The use of a function that is decreasing with respect to the norm of the image gradient is a natural selection. In this context, two functions with similar properties and performance were proposed: $c_1(|\nabla U|) = \exp(-\frac{|\nabla U|^2}{K^2})$ and $c_2(|\nabla U|) = \frac{1}{1+\frac{|\nabla U|^2}{K^2}}$ with K a constant that can be assimilated to a gradient threshold. In [22], the author proposed several choices of the c function inspired by the robust estimation framework and discussed the difference between them. An interesting method to understand the behavior of a given PDE is to consider its effect on the gradient direction $\eta = \frac{\nabla U}{|\nabla U|}$ as well as the tangential direction noted ξ. In [87], the authors provided some constraints on the diffusion along these two direction toward better edge preserving. Based on the same observation, other non linear PDE's that aim to restrict the diffusion process only along the tangential direction to the gradient and tuned by the gradient magnitude were suggested [5, 78, 108]. For instance, in [5] a model that considers rather a smoothed version of the image gradient was considered.

$$\frac{\partial U}{\partial t} = g\left(|\mathcal{G} * \nabla U|\right)|\nabla U|\,\text{div}\frac{\nabla U}{|\nabla U|} = g\left(|\mathcal{G} * \nabla U|\right)U_{\xi\xi} \tag{2.14}$$

  Where g is a positive decreasing function and $U_{\xi\xi}$ is the second derivative of the image in the direction ξ. Such a model respects better the image features because it involves larger

neighborhood while performing the diffusion along the tangential direction.
Other group of PDE operators are those that act directly on the level line of the image. The curvature motion is an interesting example of diffusion where the contrast invariance requirement is ensured [4, 155, 101]. The associated PDE is

$$\frac{\partial U}{\partial t} = F(curv(U), t)\,|\nabla U| \tag{2.15}$$

Where $curv(U) = \text{div}\frac{\nabla U}{|\nabla U|}$ refers to the curvature of the level line of U and F is an increasing function with respect to the first argument. The diffusion is performed along the normal direction to the level line and modulated by the curvature which leads to a curve shortening. Diffusion tensors (i.e. symmetric and positive definite $2\times$ 2 matrices) based formalism, provides more generic framework. These formulations rely on the definition of a tensor field that imposes the direction of the smoothing. A general form of such a technique is presented in [145, 146, 149] where the evolution defined by a tensor $\mathbf{D}$ is

$$\frac{\partial U}{\partial t} = \text{div}(\mathbf{D}\nabla U) \tag{2.16}$$

The authors set the diffusion tensor to the structure matrix defined as $\mathbf{D} = (\nabla U \nabla U^T) * \mathcal{G}_\sigma$. The diffusion is done according to the eigenvectors of the matrix $\mathbf{D}$. For homogeneous regions this tensor is isotropic which yields a smoothing in all directions. Along image contours the diffusion is directed by the eigenvector that corresponds to the contour direction. Notice that considering a non local gradient direction estimation maintains coherence in the gradient direction for neighboring pixels which is an important issue in case of noisy images. Another unifying formulation of common diffusion equations that is based on the definition of diffusion tensors was proposed in [135, 137, 48]. It is based on the trace operator and the Hessian matrix $\mathbf{H}$ and is expressed as

$$\frac{\partial U}{\partial t} = \text{trace}(\mathbf{DH}) \tag{2.17}$$

this formalism allows the design of a specific smoothing that respects better the natural regularization properties than the divergence based model. The strength of tensor diffusion based formalism lies in the fact that it separates the design of the diffusion tensors from the smoothing process itself. Based on this property, one can retrieve the geometry of the structure in the image and then perform smoothing based on the computed tensor field.

- **Energy-based Approaches**
  Some PDE based regularization techniques have their roots in functional minimization frame-

work. A smooth version of a noisy image is obtained by minimizing a cost function that penalizes the amount of variation in the image [86, 148, 122]. Earlier work refers to the total variational minimization that was first introduced by Rudin, Osher and Fatemi [122, 121]. They provided an edge preserving restoration approach by minimizing the L^1 norm of the magnitude of the image gradient. This problem is defined in the space of bounded variations functions in Ω ($BV(\Omega) = \left\{ f \in L^1(\Omega) | \int_\Omega |\nabla f| < \infty \right\}$). The interesting aspect of the proposed regularization is the fact that discontinuous and non smooth solutions are possible. The classical formulation of the ROF model is

$$U_{opt} = \text{Argmin} \int_\Omega |\nabla U| \, d\mathbf{x} \tag{2.18}$$

Subject to

$$\int_\Omega (U - I)^2 d\mathbf{x} = \sigma_n^2 \quad \text{and} \quad \int_\Omega (U - I) d\mathbf{x} = 0 \tag{2.19}$$

The estimated image has to be smooth while its difference with observed image (called residual) must have the same properties as noise : zero mean and same variance σ_n^2. One can relax these two constraints and consider instead this unconstrained problem where the minimizer of the following energy function is an estimate of the noise free image.

$$E(U) = \int_\Omega |\nabla U| \, d\mathbf{x} + \lambda \int_\Omega (U - I)^2 d\mathbf{x} \tag{2.20}$$

The second component of the energy is the fidelity term and it constrains the solution to be similar to the observed image I while λ is a parameter that controls the trade off between smoothness and fidelity to the observations. For a given λ value the constraint (2.19) is not necessarily verified unless one devotes special effort to compute λ as done in [74, 122]. The functional (2.20) is strictly convex and in [33] the authors provided a proof of the existence and the uniqueness of the solution. The Euler Lagrange equation associated to this energy is the following:

$$\text{div} \left(\frac{\nabla U}{|\nabla U|} \right) + \lambda (U - I) = 0 \tag{2.21}$$

To solve this equation the authors proposed in [122] the use of artificial time marching which is equivalent to the steepest gradient descent of the energy function. The image is considered as a function of space and time and one has to find the steady state solution for the equation

$$\frac{\partial U}{\partial t} = \text{div} \left(\frac{\nabla U}{|\nabla U|} \right) + \lambda (I - U) \tag{2.22}$$

with $\text{div} \left(\frac{\nabla U}{|\nabla U|} \right) = curv(U)$ where $curv(U)$ is the curvature of the level line of U [122].

Based on this flow, the amount of the smoothing applied to an observation is related to the curvature of the level line passing through it. The edges are preserved because they have small curvature. Nevertheless details and other oscillatory components different from noise may also be suppressed.
A more general formulation for the total variation minimization can be expressed if one considers only the regularization flow as follows :

$$\hat{U} = \text{Argmin}_U \left(\int_\Omega \Phi(|\nabla U|) d\mathbf{x} \right) \tag{2.23}$$

where Φ is a positive and increasing function defined on $\mathbb{R}$. We recall that the connection of this formulation with PDE's is established by the Euler-Lagrange equation and the introduction of an artificial time parameter so that at the convergence the energy gradient is equal to zero. The resultant PDE to problem (2.23) is of the form

$$\frac{\partial U}{\partial t} = \text{div} \left[\frac{\Phi' \left(|\nabla U| \right)}{|\nabla U|} \nabla U \right] \tag{2.24}$$

Hence, for a particular choice of the Φ function, we find again some models introduced above. The isotropic diffusion is obtained when $\Phi(s) = s^2$ [133], The Perona Malik corresponds to $\Phi(s) = 1 - \exp(-\frac{s^2}{K^2})$ and the Rudin Osher Fatemi model discussed above is defined by $\Phi(s) = s$. Contrarely to axiomatic PDE's where one can enforce the diffusion only on the tangential direction of the gradient, the functional minimization is less flexible. More explicitly, when considering the decomposition of the diffusion process according to the gradient direction and the tangential one, we can notice that the coefficients are dependent. To this end, a careful choice of the Φ function to favour diffusion on the tangential direction with respect to the normal one has to be made.
Total variation minimizing based approaches are among the most popular denoising techniques that gained a lot of attention and were extended over the years to address many image processing problems such as deblurring, inpainting, optical flow estimation, stereo reconstruction, etc. For a more comprehensive theoretical analysis of total variational minimizing using PDE's we refer the reader to [86].

- **The Beltrami Flow**
  The Beltrami flow is another tool towards image analysis and regularization which relies on representing the image as a 2D surface in the 3D space for gray scale image and higher dimensional spaces for multi valued images [80, 82, 81, 129]. Thus considering the image as an embedding map $\mathbf{M} : \mathbf{x} \rightarrow (\mathbf{x}, U(\mathbf{x}))$, the authors propose to minimize with respect to U a measure of this map which is defined by the generic Polyakov measure. This measure

corresponds to the image surface and in case of gray scale images it could simply be written as

$$\hat{U} = \text{Argmin}_U \left(E(U) = \int_\Omega d^2\sigma \sqrt{1 + |\nabla U|^2} \right) \tag{2.25}$$

Performing a gradient descent for the area of the graph representing U leads to the following diffusion equation where each point of the image is moving according to the normal of the surface which enables image edges preservation

$$\frac{\partial U}{\partial t} = \text{div} \left(\frac{\nabla U}{\sqrt{1 + |\nabla U|^2}} \right) \tag{2.26}$$

In [129] the author provided a general framework that uses measures on maps between Riemannian manifolds and showed the connections between this approach and other anisotropic diffusion PDE's.

At the end of this overview we want to point out that this is not an exhaustive presentation of PDE's based image analysis and other interesting work can be found in [31, 135]. Multi-valued images and constrained PDE are also an active research area and one can refer to [135, 48] for a review of these approaches.

Finally to conclude we can say that PDE's are an interesting tools for image analysis able of extract image content at different scales. Nevertheless, as far as denoising application is concerned their performance is limited. This is due to the presence of fine texture and structure in the image that has similar scale to the noise and cannot, in the most general case, be characterized by a simple feature as gradient direction or other limited local interactions. One can overcome this limitation through the introduction of more global methods, like for example image decomposition/representation in sub-spaces.

2.2.3 Image Transform and Compact Representation

Image decomposition on a specific basis or dictionary offers powerfull tool used in many image processing applications (compression, restoration, inpainting etc). Their aim to provide a compact image model where an image can be reconstructed using few number of elements. In this section we will present two main approaches

- **Image transform and wavelet basis**
  Several image transforms was proposed in the litterature [57, 102, 37, 41] toward image fea-

ture extraction and modelling. Denoising is among the applications that could take advantage of this representation. The first attempts in this direction were presented in [152] where the spectral image content is analyzed in a local window (Discrete Cosine Transform or local Discrete Fourier Transform) then modified and the inverse transform gives an estimate of the noise free intensity.The major breakthrough in this direction came with the introduction of the wavelet decomposition. Wavelet transform provides a signal representation that encodes space and frequency. Moreover, it has a high capacity of approximating a given signal by a small set of coefficients. It refers to an image decomposition on a set of functions composed of shifted and dilated versions of the wavelet noted ψ. A wavelet is a function in $L^2(\mathbb{R}^2)$ localized in space and that satisfies $\int_{\mathbb{R}^2} \psi(\mathbf{x})d\mathbf{x} = 0$ and $\|\psi\| = 1$. Hence, if we note $\psi_{\mathbf{v},s}$ the wavelet obtained after a shift $\mathbf{v}$ and dilation s, the definition of the wavelet coefficients of I is

$$I_{\mathbf{w}}(\mathbf{v}, s) = \langle I, \psi_{\mathbf{v},s} \rangle = \int_{\mathbb{R}^2} I(\mathbf{x})\psi_{\mathbf{v},s}(\mathbf{x})d\mathbf{x} \text{ where } \psi_{\mathbf{v},s}(\mathbf{x}) = \frac{1}{\sqrt{s}}\psi\left(\frac{\mathbf{x}-\mathbf{v}}{s}\right) \tag{2.27}$$

A particular and interesting case of wavelets are the orthogonal wavelets that form an orthogonal basis of $L^2(\mathbb{R}^2)$ defined as $\left\{\psi_{\mathbf{n},j} = \frac{1}{\sqrt{2^j}}\psi\left(\frac{\mathbf{x}-2^j\mathbf{n}}{2^j}\right)\right\}_{j\in\mathbb{Z},\mathbf{n}\in\mathbb{Z}^2}$. Since the image is defined on a finite domain, it can be expressed through an orthogonal basis using a finite set of coefficients [103]. The image is projected on two orthogonal subspaces, the first one corresponds to the details of the image at different scales, the second to the lower resolution of the image. This decomposition is expressed as follows

$$I = \sum_{j=L+1}^{J} \sum_{\mathbf{n}\in[0,2^{-j}]^2} \langle I, \psi_{\mathbf{n},j} \rangle \psi_{\mathbf{n},j} + \sum_{\mathbf{n}\in[0,2^{-J}]^2} \langle I, \phi_{\mathbf{n},j} \rangle \phi_{\mathbf{n},j} \tag{2.28}$$

This decomposition involves two main terms the first refers to the detail being present in the image at different scales while the second corresponds to a coarse version of the image represented by the basis vector $\{\phi_{\mathbf{n},j}\}_{\mathbf{n}}$.

Since the wavelet transform concentrates a lot of energy on few coefficients that correspond to the useful information in the image while the noise is represented by small wavelet coefficients, a natural way to perform denoising in the wavelet domain is to perform a thresholding [103]. An estimate of U can be expressed as

$$\hat{U} = \sum_{j=L+1}^{J} \sum_{\mathbf{n}\in[0,2^{-j}]^2} \rho_T\left(\langle I, \psi_{\mathbf{n},j} \rangle\right) \psi_{\mathbf{n},j} + \sum_{\mathbf{n}\in[0,2^{-J}]^2} \rho_T\left(\langle I, \phi_{\mathbf{n},j} \rangle\right) \phi_{\mathbf{n},j} \tag{2.29}$$

Where ρ_T is the thresholding function and it could be a hard thresholding or soft thresh-

olding. While the hard thresholding technique causes small oscillation near the edges of the reconstructed image (Ringing effect), the soft one reduces considerably this artefact. The threshold value T is a critical parameter for the denoising process and it must be carefully selected to reach a balance between details preservation and noise suppression. In [45] through the minimization of a quadratic Bayesian risk function, the author showed that the optimal threshold that minimizes the upper bound of the risk is $T = \sigma_n\sqrt{2\log N}$ where N is the number of wavelet coefficients.

Notice that this threshold value as well as the reconstruction process based on thresholding holds for any other decomposition on orthogonal basis different from wavelets and for an extensive review and details the reader can refer to [103]. The wavelet basis is particularly attractive because of its ability to encode an important amount of energy through few coefficients. Nevertheless the wavelet decomposition reaches its limits when dealing with textured images where small coefficients correspond also to texture. To overcome this limitation an adaptive basis selection based on wavelet packet decomposition was presented in [88]. In addition to that a lot of attention was devoted to sparse image representation using alternative choice of basis like bandelet, contourlet etc [90, 130] during the past year. However, these methods fail to account for textured image content because they are more adapted to a geometric description of the image than to texture description. Furthermore, despite the fact that such an approach comes with a rather powerful representation, choosing a specific wavelet function able to capture the wide variability of natural images is a rather challenging task. The use of image-driven sparse representations can address the above mentionned limitation.

- **Image-driven dictionary and sparse representations**

  These techniques aim at determining the optimal image representation (sub-space) through a direct analysis of the image content [123, 94, 24]. In order to facilitate the introduction of the method, let us consider an overcomplete dictionary matrix $\mathbf{D}$ of size $L \times K$ containing K columns that refer to the dictionary elements or "atoms". A sparse representation of a vector $\mathbf{v}$ (of size L) is a linear combination of a few number of the atoms $\{\mathbf{d}_i\}_{1\leq i\leq K}$. Using such a basis, one would like to be able to optimally reconstruct the image using the smallest possible set of atoms. This amounts to solving the following problem

$$\min_{\mathbf{a}} \|\mathbf{a}\|_0 \quad \text{subject to} \quad \|\mathbf{v} - \mathbf{D}\mathbf{a}\|_2 \leq \epsilon \tag{2.30}$$

  Where $\|\mathbf{a}\|_0$ refers to the number of non zero entries of a vector.

  The most recent work in this area refers to the K-SVD algorithms for designing dictionaries and sparse representations [3]. The K-SVD algorithm aims at finding simultaneously the

dictionary $\mathbf{D}$ and the coefficient vector $\mathbf{a}$. This is achieved using a set of training samples containing several vectors $\{\mathbf{v}_p\}_{1\leq p\leq M}$ and by minimizing the following objective function

$$min_{\mathbf{D},\mathbf{A}} \|\mathbf{V}-\mathbf{D}\mathbf{A}\|_F \quad \text{subject to} \quad \forall\, 1\leq p\leq M,\ \|\mathbf{a}_p\|_0 \leq T \tag{2.31}$$

Where $\mathbf{a}_p$ is a K element vector associated to the coefficients of $\mathbf{v}_p$, $\mathbf{V}$ is a $L\times M$ matrix with vector column $\mathbf{v}_p$ and $\mathbf{A}$ is a $K\times M$ matrix with vector column $\mathbf{a}_p$. The Frobenius norm is denoted $\|.\|_F$. Computing the solution of the problem (2.31) is achieved in two steps: sparse coding and dictionary update. The sparse coding refers to finding the optimal coefficient $\mathbf{a}_p$ for each signal given the dictionary and is performed using matching pursuit algorithms [19]. The dictionary optimization is done one atom at a time assuming the other atoms constant. Each atom is obtained through a Singular Value Decomposition (SVD) of a matrix that corresponds to the approximation error when this atom is removed [3].
This concept was considered to address denoising in [53, 100, 99] where the undelying idea was to design a dictionary specific for each image based on observed noisy patch. More explicitly they assume that image patches can be approximated by a sparse representation. The denoising problem in [53] is formulated as follows

$$\min_{\mathbf{a}_\mathbf{x},U,\mathbf{D}} \left\{ \lambda \|U-I\|_2 + \sum_{\mathbf{x}\in\Omega} \mu_\mathbf{x} \|\mathbf{a}_\mathbf{x}\|_0 + \sum_{\mathbf{x}\in\Omega} \|\mathbf{D}\mathbf{a}_\mathbf{x} - R_\mathbf{x}U\|_2 \right\} \tag{2.32}$$

where $R_\mathbf{x}U$ is the image transform that extracts an image patch at location $\mathbf{x}$.
The first term is a fidelity to data term that constraint the restored image to be close to the observed one. The second term is the sparseness constraint that imposes to each coefficients vector $\mathbf{a}_\mathbf{x}$ to be sparse. The third term refers to the decomposition of the image patches on the designed dictionary $\mathbf{D}$. The solution of the problem is obtained in iterative fashion by fixing two variables and optimizing with respect to the third one. Numerous extensions of this concept were reported recently in the literature like for example color image denoising as well as restoration based on multi-scale image representation [100, 99].

The use of sparse sub-space representations carries certain strengths as well as certain limitations. If we recall the three aspects being associated with the denoising problem, then for example it is hard to imagine how one can encode prior knowledge on the noise model in the process. In a number of applications (like for example in medical), such a knowledge is available and could be used to improve the performance of the process. Statistical methods are a convenient tool to include such a prior.

2.2.4 Statistical Models and Image Denoising

The central idea behind such a concept is to reformulate image reconstruction as a statistical estimation problem over a random variable. In such a context, based on certain assumptions on the properties of this variable, the task of enhancement consists in recovering the noise-free image that minimizes a cost function given the observed sample. Different models can be considered regarding this density with the most popular ones being exposed in the remainder of this section.

- **The Wiener filter**
  The Wiener filter is the optimal linear filter, noted $\hat{U}$, that minimizes the quadratic Bayesian risk function defined as

$$R(V) = \mathbf{E}\left(\|U - V\|_2^2\right) \tag{2.33}$$

  The orthogonality principle states [103] that a Wiener estimate verifies

$$\mathbf{E}\left[(\hat{U} - U)I^T\right] = 0 \tag{2.34}$$

  Now if we assume that the Wiener estimate can be expressed as $\hat{U} = \mathbf{W}I$ and that U and n are independent, based on equation (2.34), the linear filter matrix can be expressed as

$$\mathbf{W} = R_U(R_U + R_n)^{-1} \tag{2.35}$$

  with $R_U = \mathbf{E}\left([U - \mathbf{E}(U)]\,[U - \mathbf{E}(U)]^T\right)$ the covariance matrix of U while R_n is the one relative to noise.
  Now in order to compute the Wiener filter on the discrete Fourier basis, the covariance matrix must verify

$$\mathbf{E}[U(i)U(j)] = R_U(i-j) \quad \text{and} \quad \mathbf{E}[n(i)n(j)] = R_n(i-j) \tag{2.36}$$

  Where R_U and R_n are periodically extended. This constraint means that the correlation between the variables $U(i)$ and $U(j)$ only depends on distance between them. Now we will consider the discrete Fourier transform of the R_U and R_n which refers to the power spectrum noted S_U and $S_n = \sigma_n^2$ (in the case of a white Gaussian noise). Hence, the coefficients of the optimal filter in the transform domain are expressed as

$$W_F(m) = \frac{S_U(m)}{S_U(m) + \sigma_n^2} \tag{2.37}$$

  In the frequencies domain the Wiener filter attenuates each element of the Fourier transform

of the Image I by a coefficient that depends on the signal to noise ratio. The attenuation is stronger when the noise amount is important with respect to the signal.

- **Maximum a Posteriori estimation and Markov Random Fields**
  The Maximum a Posteriori estimation is among the most popular statistical tools used in image restoration problems. The Maximum a Posteriori estimator refers to computing an image $\hat{U}$ that maximizes the conditional probability $P(U|I)$. Using the Bayes rule the posterior probability can be expressed as

$$P(U|I) = P(I|U)P(U) \tag{2.38}$$

  $P(I|U)$ refers to observing the noisy image given the true image, this posterior corresponds to the noise model. Thus under the assumption of white noise where each pixel is affected independently from the others, the conditional probability of the noisy image is

$$P(U|I) = \prod_{\mathbf{x}\in\Omega} P(I(\mathbf{x})|U(\mathbf{x})) = \prod_{\mathbf{x}\in\Omega} P(n(\mathbf{x})) \tag{2.39}$$

  Once noise model is defined the MAP estimation relies on the knowledge of the prior probability $P(U)$ which refers also to the image model. A direct estimation of this probability is intractable because of the enormous dimensionality of the image space which is equal to the number of the pixel in this image. Toward reducing the complexity of the problem, many researchers considered lower dimension probability density functions to model the statistical dependence between pixels which is restricted to be local. For instance, Markov random fields (MRFs) are stochastic models that characterize the local spatial interactions in data. They were used to compute prior image models [58]. Within this framework the image is assumed to be a Markov field and verifies two constraints:

$$P(U) > 0 \tag{2.40}$$

$$P\left(U(\mathbf{x})|\left\{U(\mathbf{y})\right\}_{\mathbf{y}\in\Omega\setminus\mathbf{x}}\right) = P\left(U(\mathbf{x})|\left\{U(\mathbf{y})\right\}_{\mathbf{y}\in\mathbf{N}_x}\right) \tag{2.41}$$

  The second constraint amounts to say that the conditional probability of an intensity at a given pixel $\mathbf{x}$ given the whole image does only depend on the set of observation in a local Neighborhood of $\mathbf{x}$ (noted $\mathbf{N}_x$). The strength of the Markovian models relies on the Hammersely-Clifford theorem that establishes the equivalence between Markov random field and Gibbs random fields (GRF) [71]. The definition of a GRF requires the notion of a clique. A clique C, associated with a neighborhood system $\mathcal{N}$, is a subset of pixels where each pixel is the neighbor of every other pixel according to the neighborhood system

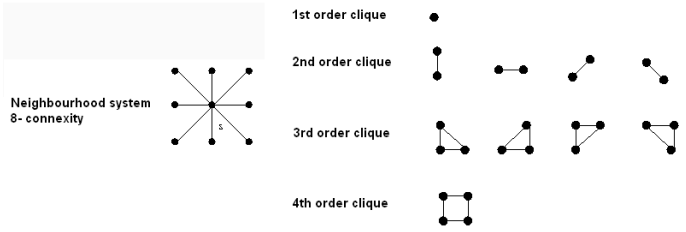


Fig. 2.1: Neighborhhod system and its associated cliques

$\mathcal{N}$. An illustration of the clique definition and neighborhood system is presented in figure [Fig.(2.1)]. Now a Gibbs random field is characterized by its joint probability defined as

$$P(U) = \frac{1}{Z}\exp(-\frac{E(U)}{\tau}) \quad \text{with} \quad E(U) = \sum_{c\in C} E_c(U) \tag{2.42}$$

Where $Z = \sum_{I\in\Gamma}\exp(-\frac{E(I)}{\tau})$ is a normalization constant computed on the space of all possible configurations of the random field associated to the image. E is the energy function that is the sum of local energy functions or the clique-potential while τ is the temperature parameter that comes from the analogy with statistical physics. Many efforts were devoted to define appropriate energy functions towards accurate natural image modeling. The first image model comes from statistical physics field and is associated to binary images and known as the Ising Model. The extension of the Ising model to an image with more then two gray level is The Pott model which is generally used in segmentation. As far as image regularization is concerened very basic potential definition is the quadratic Gaussian model where the clique potential is defined by intensity difference between neighboring pixels

$$E(U) = \beta \sum_{c_2=(\mathbf{x},\mathbf{y})} (U(\mathbf{x}) - U(\mathbf{y}))^2 \tag{2.43}$$

Based on this energy formulation the most probable images are those having a low intensity difference between neighboring pixels which is a reasonable assumption for natural images. Notice here that one can establish a connection between this model and the minimization of the L^2 norm of the image gradient introduced in the previous section. These models are fairly simple and involve a small number of parameters. However they are not able to model texture like image content. This limitation is due to the small clique size that can hardly capture the image structure at relatively large scale. Furthermore, the prior probability has a pre-specified form that do not comply with the reality unless some training is performed.
In this context, MRF models that are based on more global image priors and computed based on parameter learning from a training sample set, are alternative approaches that offer a more

realistic image representation. One of these approaches is the FRAME (Filter, Random field, And Minimax Entropy) model [157] designed for texture description and analysis. This approach aims at characterizing the ensemble of images U with the same texture appearance by a probability distribution $P(U)$ by making inference using a set of observed texture samples. This is performed in two steps: (1) A set of filters is applied to the texture samples to capture their features, then histograms of the filtered images are computed. These histograms are viewed as estimates of the marginal distributions of $P(U)$ (2) The maximum entropy principle is employed to derive an estimate of $P(U)$ noted $\hat{P}(U)$ and which is restricted to have the same marginal distribution. The FRAME model is a step toward texture modeling but as it was stated by the author this approach is efficient for building models based on texture similar in appearance. Its limited ability for generalization makes it more appropriate for texture synthesis than denoising. More data driven approaches where all the parameters of the potential function is learned during the training set is the Fields of Experts model [120]. Contrarily to the FRAME model the linear filters involved in this framework are computed during the training process. This approach bears similarities with the sparse image coding and relies on a training step where the filters are the element of the dictionary. The ability of this approach to model various data types is conditioned by the size of the neighborhood system used in the definition of the image potential and considering large neighborhood will increase the complexity level of the approach.

Multiscale random fields are powerful techniques to provide image models. This framework relies on a multiresolution image representation where image sites at given scale are conditionally independent given the immediate coarser level. This yields a scale-recursive models that are more efficient computationally on one hand and provide multiscale models that are more interesting then conventional MRF models when dealing with texture [97]. Multiscale image representations were also used jointly with Gaussian scale mixture [119, 141] toward image restoration and texture preservation. Before concluding we want to refer to [114, 95] where the reader can find extended reviews of Markovian approaches toward image restoration and to [150] for multiscale Markov models and their applications.

Image models are critical components of the denoising process, in particular when referring to texture. For instance, texture can either refer to micro or macro-patterns and therefore can sometimes be considered as noise. For this reason, one should be able to account for varying degrees of complexity in the observed content and be able to adjust the denoising process according to it. The most prominent aspects to be considered mainly are : the definitions of the interaction between pixels and the scale or the extent of these interactions. In this thesis we would like to address the

above concerns

2.3 Main Contributions

The work carried in this thesis is a step toward better natural image and texture understanding and hence denoising. Our main focus was to adapt the image models inherent to the restoration process to the image content. In this dissertation we present mathematical models and their computational solutions that learn from the image itself to build an appropriate denoising approach. Our main motivation was to design a technique that uses different image models according to the local context. This thesis introduces novel theoretical approaches to deal with image enhancement and texture preservation. First, we focus on the photometric aspect of the problem towards defining appropriate image models to express the observations or the photometric scale. Then, we study the importance of geometry towards capturing co-dependencies between observations and therefore defining the most appropriate spatial scale. The last theoretical component of the thesis consists in two contributions, a more general regularization term which can encode complex interactions between observations and a unified framework that involves the optimal automatic selection of spatial bandwidth as well as more appropriate definitions of photometric component.

Our first contribution consists in introducing the notion of classification in the process of denoising using filter operators. Several techniques are based on models that aim to cope with image structure implicitly by considering some features during the denoising process. Unfortunately, some texture exhibit similar behavior to the noise so that the restoration algorithm fails to detect them. For this reason we believe that a pre classification step provides a more robust tool toward textured region identification. This classification consists in an image partition to locally smooth area, edges and texture. This is done through a soft classification on a feature space that can encode edgeness, texturness and smoothness. The projection of this feature space to an appropriate sub-space yields a Gaussian mixture model of observations where each component corresponds to one of the above hypotheses. Once the classification step is achieved, its result is encoded in the denoising algorithm. The filtering technique is based a non parametric density estimation of image priors. In this context, we propose an automatic kernel bandwidth selection that relies on the image partition previously done. Hence, in the filtering process we process pixels differently according to their membership to one of three image components. The latter point is the major contribution with respect to prior technique using non parametric density estimation [118, 10]. Contrarily to these methods, we use variable bandwidth kernels that depends on local pixel's properties.

The second approach studied in this thesis is based on random walks. Our method explores multiple neighbor sets (or hypotheses) that can be used for pixel denoising, through a particle filtering approach. Our objective was to provide a selection method of the relevant pixels participating in intensity estimation of a given one. Based on particle filtering techniques, this is done in a progressive and sequential manner by considering a very small and local neighborhood and going further and further while being guided by the image geometry. Contrarily to classical neighborhood filters, the domain on which we perform filtering is adapted to each pixel. The filtering process involves a number of particles that explore the image domain using a statistical distribution that describes the image geometry and the state of a particle refers to the state of the image reconstruction process.

The two previous methods restore the image pixel by pixel. Although it is a simple way for restoration, a global homogenization process where the whole image is iteratively updated to obtain the restored image should be considered. Moreover these techniques are composed of two steps: texture characterizing and then filtering. To overcome this limitation, we will consider a global image model where image structure is encoded implicitly. Hence in this chapter we will perform denoising through the minimization of a quadratic functional that involves variable bandwidth kernels. These kernels reflect spatial and photometric similarities between pixels. Toward texture preservation and estimation accuracy improvement, we considered a spatial bandwidth dependent on the pixel position to adapt the interaction between pixels to the local scale. The definition of an appropriate similarity measure more robust to noise than the L^2 distance between patches was also addressed in this dissertation. This measure is computed between pixel features that are determined after the projection of image patches on another sub-space that encodes better the image content. In addition to that, we proposed a weight definition that is more coherent with the statistical distribution of the L^2 distance between patches.

We presented in this thesis several extensions of the convex functional minimization framework to other noise types or data types to show its flexibility. These extensions concern : (i) color image restoration where we take into consideration the actual noise properties in digital cameras (ii) diffusion tensor field estimation and regularization where the constraint of being definite positive have to be verified by these tensors. (iii) ultrasound sequences filtering where the energy formulation was adapted to the address the multiplicative noise model.

2.4 Thesis Outline

The thesis is organized as follows: The second chapter is dedicated to the non parametric models based estimation and to the image partition technique. In the beginning we will give an overview of a closely related work that is the unsupervised information-theoretic adaptive filter and we point out the difference with this algorithm. Next, we will describe our strategy toward image partition to three main classes ”smooth regions”, ”textured regions”’ and ”edges”. This partition involves feature extraction and clustering through Gaussian mixture models. Once the classification step is performed we will focus on the denoising scheme. We will briefly review the Bayesian decision theory and the various estimators' types. Next we will present our denoising technique based on Marginal Posterior Maximizing where this posterior is estimated using non parametric density estimation techniques. Then we discuss the kernel bandwidth selection and we describe the optimization process. The performance evaluation is presented at the end of the chapter.

The third chapter is devoted to the neighborhood filtering technique and the use of random walks and particle filtering strategy to select the most appropriate domain and the set of pixels to denoise a given one. In the beginning, we will present a statistical model that aims to describe the geometry of the image and the spatial relationships between similar pixels. Such a model will be used in the context of multiple hypotheses testing to guide the particle evolution. then, we will provide a brief overview of particle filter techniques , and amend this statistical estimation tool to image denoising in the case of additive noise as well as multiplicative one. We will conclude this section with the experimental validation of this technique.

In the fourth chapter we will present the technique that refers to functional minimizing based regularization. In the beginning we will start by a review of the state of the art for total variation minimization techniques. Then we focus on the model description and we provide the inherent diffusion process. Once the model is established we will be concerned by the interaction between pixels within this model. This interaction is governed by weights defined by two kernels one that penalizes spatial distance between pixels, the other the photometric difference. In this context, we tried to improve the weight definition by optimizing the selection of the spatial kernel bandwidth and adapting it to each pixel position. We focused also on the definition of new photometric similarity measures between pixels. In this context, two new weight definitions were suggested at the end of the section: one based on better feature extraction the second based on the statistical properties of the L_2 distance between patches. Finally we will conclude this chapter with the experimental results.

In the fifth chapter we will study the various types of regularization problems beyond the clas-

sical additive Gaussian noise on gray scale images. In the beginning we will focus on the problem of color image corrupted by real digital camera noise. To this end we discuss the noise model and we provide a non parametric technique to estimate the noise function that is the evolution of the noise variance with respect to the intensity. After that, we will present a procedure of color image denoising based on minimizing under several constraints the functional presented in chapter 4. The second application concerns data in higher dimension/defined on constrained manifolds like for example the case of diffusion tensors in DTI imaging. The goal is to jointly estimate and regularize the diffusion tensors using several observed diffusion weighted MRI images. The performance of this method as well as its impact on tensor classification was evaluated. The final section, was dedicated to speckle suppression in ultrasound sequences. To achieve this objective an adequate fidelity to data term as well as an appropriate weight definition were introduced. The experimental validation of this technique is provided at the end of the chapter.

Conclusions and discussions are part of the last section of the document presenting the main shortcomings and limitations of the approaches as well as the potential future perspectives of this work.

To conclude, this thesis evolves around image/noise modeling in the context of reconstruction and image inference from corrupted data as well as their applications in digital photography and medical image analysis. It has produced (up to now), one book chapter [109] three major conference publications [12, 15, 13, 11], two major workshop papers [14, 110] and has pending journal submissions (International Journal of Computer Vision, IEEE Transactions on Pattern Analysis and Machine Intelligence, Journal of Mathematical Imaging) as well as an involvement in a Frecnch patent [30].

Chapter 3

MPM Adaptive Denoising

The selection of a statistical model towards explaining the observations and recovering the noise-free signal is critical in the image restoration process. Conventional approaches often considered local parametric models exploiting the local smoothness assumption that fails in case of texture. In this chapter, we introduce a novel approach where the approximation makes no assumption on the nature of distributions and is determined from the data itself. To compute our observation model, we consider a variable bandwidth non-parametric approach to better reflect data variability when considering observations related with texture. The bandwidth selection is done according to a soft classification through a conventional mixture model. Such a classification is performed on image pixels based on their features vector that characterizes textureness. The outcome of this modeling is used within an MPM (Marginal Posterior Maximum) estimation framework, with aim to restore each pixel's intensity given the noisy image. This results in a novel approach for image enhancement towards better texture preservation. Promising results demonstrate the potentials of the proposed framework.

3.1 Introduction

Texture restoration is a key feature for image denoising and the challenge is the trade off between noise suppression and fine details and texture preservation. State of the art techniques rely on various image priors and parametric models with various complexity [122, 156, 120, 58]. The majority of such methods assume a static prior that is not adapted to the different components of an image. Natural images often consist of smooth, textured regions and contours. These elements have different properties and a unique prior that models in the same fashion all these element is

inaccurate.
Many algorithms provided satisfactory restoration quality [26, 76, 10, 53, 100, 88, 45], but their performance depends mainly on the texture present on the image. In other words some texture patterns can be efficiently and easily restored because they correspond to patterns where structure, orientation and scale could be detected. Other texture types are more difficult to recover because their properties are close to the noise ones. An example of such texture is given in figure [Fig. 3.1]. This figure shows that there are not major differences between the skin texture and noise at least in terms of variance.
In order to handle properly the texture in the image one should identify regions that correspond to noise and those to the texture and structure. Such a separation could be useful to many image processing applications such as compression, demosaicing, deblurring etc. Indeed, it provides a tool for parameters selection and offers the possibility to use different image priors for each region. Under the assumption that the textured and texture-free regions have been identified, we can design an adaptive algorithm for denoising which takes into account the particularities of the textured regions in the image toward better preserving them. In order to address the above demand one should be able to answer the following fundamental questions:

- is it possible to create a soft image partition into smooth regions edges and textures ?
- if such partition is available, What is the appropriate way to encode such prior into a denoising algorithm towards better image reconstruction?

First, we introduce a soft-classification technique based on the analysis of local features under the assumption that the image consists of three classes: (i) smooth regions, (ii) edges (iii) textured regions. Such classification is the outcome of a Gaussian mixture model where the observations are the projection of the local features on a convenient sub-space. Therefore, it assigns to each pixel a probability measure that reflects its membership degree to each class.
As far as denoising is concerned, we adopt an MPM (Marginal Posterior Mode) estimation based technique. This technique aims to maximize the probability of observing a particular intensity value $U(\mathbf{x})$ given the noisy image I. The posterior is computed using a non parametric image model where the bandwidth of the selected Gaussian kernel depends on the classification result and in particular is a function of membership probabilities. The central idea is to consider Kernels with small bandwidth/variance in texture and edges while augment their bandwidth when referring to smooth regions.

(a) (b) (c)

Fig. 3.1: Example of confusing case for texture: (b) The skin texture (c) A noise patch.

The chapter is organized as follows: In section 2 we briefly review related work that is based on modeling patch distribution using non parametric density estimation. Section 3 is devoted to pixels classification. In section 4, our new non-parametric image model toward adaptive denoising is detailed.

3.2 Unsupervised Information-Theoretic Adaptive Filter

Statistical image model is a fundamental component of various types of image processing applications. To this end many models was provided in the literature such as MRF based techniques [120, 156, 58] and generalized Laplacian distributions that describe the joint intensity statistical distribution [91, 68] or the joint distribution of wavelet coefficients [126]. Some of these models describe interaction between pixels at very local scale and would therefore fail to catch texture properties. Furthermore, the cited models are parametric they require a training step for parameter selection. Non parametric models are an alternative solution toward image content modeling. This technique was recently considered in [10]. The unsupervised information-theoretic adaptive filter (UINTA) for image relies on high dimension statistical distribution of the image patches. Under the assumption that the noise increases the randomness of the signal, the UINTA strategy of regularization consists in minimizing the entropy of the image pattern distribution. Let us consider a random variable Z that corresponds to the image patch realization. Then in order to perform

denoising, one has to estimate the probability density function relative to Z in order to compute the entropy. To this end, the authors define a set of observations $\mathcal{S} = \{U_{\mathbf{N}_x} \ \mathbf{x} \in \mathcal{A}\}$ where $\mathcal{A}$ is a region of Ω (the image domain), $U_{\mathbf{N}_x}$ is a noise-free patch relative to $\mathbf{x}$. These observations can be used to determine an approximation of the patch distribution using a non parametric density estimator (that will be developed in section 3.4.2). The kernels used are Gaussian and isotropic where the covariance matrix is proportional to the identity matrix with diagonal elements equal to σ. Hence an estimate of $p(Z)$ is

$$p(Z) = \frac{1}{|\mathcal{A}|} \sum_{\mathbf{x} \in \mathcal{A}} G_\sigma(Z - U_{\mathbf{N}_x}) \quad G_\sigma(X) = \frac{1}{\sqrt{2\pi}^d \sigma^d} \exp\left(-\frac{\|X\|^2}{2\sigma^2}\right) \tag{3.1}$$

where d is the cardinality of the patch (number of observations). Assuming such distribution/image model, one can address denoising through the minimization of the disorder of the system that is equivalent to the entropy. Knowing that the entropy is the expectation of the negative log-probability, it can be approximated by the mean of log probability of the samples

$$H(Z) = -\frac{1}{|\Omega|} \sum_{\mathbf{y} \in \Omega} \log \left[\frac{1}{|\mathcal{A}_y|} \sum_{\mathbf{x} \in \mathcal{A}_y} G_\sigma(U_{\mathbf{N}_y} - U_{\mathbf{N}_x}) \right] \tag{3.2}$$

where $\mathcal{A}_y$ is a region in the image containing the image patches where the probability density of the patch $U_{\mathbf{N}_y}$ is computed. Towards the lowest potential of the entropy, the image is updated according to a gradient descent method.

The expression of H indicates that the UINTA algorithm enforces the similarity between group of patches with the same content in the observed image. Such a method is global, where all the patches are updated at the same time and experimental results show its efficiency. Nevertheless, the selection of the kernel bandwidth is critical, because it determines the ability of the model to preserve data variability. For instance high σ values will reduce the number of modes in the distribution to be approximated and this leads to loss of details for patches observed less frequently in the image.

Motivated by the ability of the non parametric density estimation to model image patches distribution, we adopted this tool to compute the pdf underlying to our denoising technique. But for our approach, we used spatially varying kernel bandwidth to account for the variability of the data. Therefore appropriate importance can be also given to patches that arise less frequently (like edges) or being present in the image in a non-uniform distribution (like textures). The soft classification of the image content provides us with an efficient tool for the bandwidth selection. We have also to emphasize that our technique involves a local cost function. For this reason, the optimization process is simpler because each pixel is updated independently from the others.

3.3 Unsupervised Classification of Image Pixels

Image/texture segmentation/classification is an on-going research area from the very early years of computer vision fields. The aim is to attribute to the observations labels which create some consistency on a predefined feature space that can be intensities, color or texture. Such a process involves two aspects: (i) the selection of the feature space, and (ii) the selection of the classification method which is seeking for statistical separation of the features. We should recall that the aim of this process in the context of our research is to not to provide a texture classification technique, but an indicator on the membership of pixel to edges, textured regions and smooth regions.

3.3.1 Texture Feature Extraction

Texture analysis and characterization are active research topics in computer vision where a big effort was devoted to provide models toward texture segmentation or synthesis. The most prominent wok in this direction refers to features extraction using various types of methods that will be briefly reviewed here

- Transform analysis: The human visual system analyzes the textured images by decomposing them into a frequency and an orientation component [69]. These properties inspired many methods that are based on analyzing the orientation and frequency component of texture elements. Early approaches have focused on energy distribution using the Fourier transform [16] but the need for localized information resulted in the emergence of local analysis technique. Gabor filters [57, 23, 50] are frequency and orientation selective filters widely used in texture analysis. Wavelet transform and filter banks [37, 41, 66, 102] provide also a multiresolution representation that offers local measure of energy toward specific texture representation.

- Model Based methods: the aim of such approaches is to provide an image model useful not only to describe texture but also to synthesize it. In this category one can refer to statistical models used in the design of a probability distribution functions to characterize texture such as Markov random fields (MRF). MRF do depend on a potential function that defines the interaction between a pixel and its neighborhood. The design of such a potential must satisfy a good balance between simplicity and generalization ability. Simple models include small parameters number and limited clique size [58]. Sophisticated ones include more parameters and large clique size to adequately describe texture of rich content [156]. Multiresolution statistical description are proposed in [150]. Other methods that describe interaction between

pixels are the Grey Level Co-occurrence Matrix [54], second order statistics [35] and linear transformation [139].

The characterization of texture using such complex methods is important when seeking for a high-level understanding of the image content. Texture classification often relies on these representations. Since in our case, the objective is to perform a classification into textured regions, smooth regions and edges, one can consider simple representations as well as descriptors determined from the observed image without going through the complex process of texture characterization. In other words, if we consider a single pixel and the local observation being associated with it (a local image patch), we want to know whether it corresponds to a textured pattern, to a uniform one corrupted by noise or to a part of an edge. Statistical moments are well established indicators for texture characterization and therefore, one can consider a classification that relies on fairly simple local image statistics. We considered statistical descriptors rather than filter response to characterize the texture because some textures have random structure and geometric descriptor can fail to describe it. Hence we will consider a pixel $\mathbf{x}$, a neighborhood $\mathbf{N}_x$ relative $\mathbf{x}$ of size $w_p \times w_p$ and the patch $I_{\mathbf{N}_x}$ (resp $U_{\mathbf{N}_x}$) which is the restriction of the observed image I (resp U) to $\mathbf{N}_x$. Considering the intensity inside a patch as a random variable we can compute

- **Entropy** which is a measure of uncertainty of a distribution. If we call GL the set of observed intensities inside the patch $I_{\mathbf{N}_x}$ the entropy is defined as $e_I(\mathbf{x}) = -\sum_{i \in GL} p_i \log_2(p_i)$. p_i is the probability of observing the intensity i in the patch $I_{\mathbf{N}_x}$. In the case of additive Gaussian noise the entropy of a noisy homogeneous patch is equivalent to the entropy of the Gaussian noise. We assume that a patch belonging to a smooth region in the noise free image has almost constant intensity and the only source of variation is the noise. Therefore, we expect that all pixels belonging to smooth regions have similar local entropy.

- **Variance** which is a measure of dispersion of random variable or a distribution with respect to the mean. For the patch $I_{\mathbf{N}_x}$ the variance is defined as $v_I(\mathbf{x}) = \frac{1}{w_p^2} \sum_{\mathbf{y} \in \mathbf{N}_x} [I(\mathbf{y}) - m_I(\mathbf{x})]^2$ with $m_I(\mathbf{x}) = \frac{1}{w_p^2} \sum_{\mathbf{y} \in \mathbf{N}_x} I(\mathbf{y})$. The noise being independent from signal, the variance inside a noisy patch $I_{\mathbf{N}_x}$ is the sum of the noise variance and the variance of the noise free one $U_{\mathbf{N}_x}$. One can conclude that the variance inside a noisy textured patch is more important than the variance inside a noisy smooth patch.

- **Skewness** is the ratio between the third moment and the second moment of a distribution. It is an indicator of the symmetry of a given distribution an estimator of the skewness inside a that patch $\mathbf{N}_x$ is defined as $s_I(\mathbf{x}) = \frac{\sum_{\mathbf{y} \in \mathbf{N}_x} [I(\mathbf{y}) - m(\mathbf{x})]^3}{v(\mathbf{x})^{\frac{3}{2}}}$. This criterion could be discriminatory

because the noise distribution is symmetric while the distribution of pixels intensities inside patches belonging to image texture does not obey to any symmetry constraints.

Based on these local descriptors we can discriminate noisy patches textured noisy one and those corresponds to image structure and edges. The choice of these features is motivated by the fact that the only available information about the noise is its statistical model. The image patches that belong to smooth noisy regions have variance, entropy and skewness values close to those of the noise. On the other hand the patches relative to image texture or structure have richer content and will exhibit more variability in their statistical behavior. In figure [Fig.(3.2),Fig.(3.3)] is represented the joint distribution of the couple (entropy, variance) (variance, skewness) and (skewness, entropy) for two different images one corrupted by synthetic Gaussian noise and the other by real digital camera noise.
Once the feature space has been determined and a feature vector was computed for every image pixel, the classification consists in assigning to each pixel a probability value measuring its membership to a textured pattern a uniform one or other image structure. For this purpose, we have to find a statistical model that is able to describe the distribution of the observations (feature vector of all image pixels) and provides an automatic classification tool of them. Let us consider that the three hypotheses (smooth, edges, texture) are independent and normally distributed. Then, the observed density on joint space should form three clusters describing the different properties of the different classes. In such a context, one can deduce from the global distribution the behavior of the individual components using a mixture model. Under this assumption, we can consider that each component can be characterized using a single Gaussian model. Given that condition and the fact that statistical inference in the original 3D space (entropy, variance, skewness) involves important correlation between the distributions, we first proceed with a dimensionality reduction of the problem. The idea is to consider a projection into a subspace where the separation between classes could be easily done.

3.3.2 Dimensionality Reduction

Principal component analysis (PCA) is a tool of describing data structure and its dispersion in specific directions. It provides an orthonormal basis where the projection of the data set on the basis vectors has a decreasing variance. When considering, the projection of an observation on a sub-space formed by the directions relative to the most prominent variance, the PCA refers to a dimensionality reduction technique. For instance, one can retain a small subset of projections according to the importance of the corresponding direction towards describing the data. Hence, the PCA reduces the dimensionality of the data set or observed samples while keeping the maximum

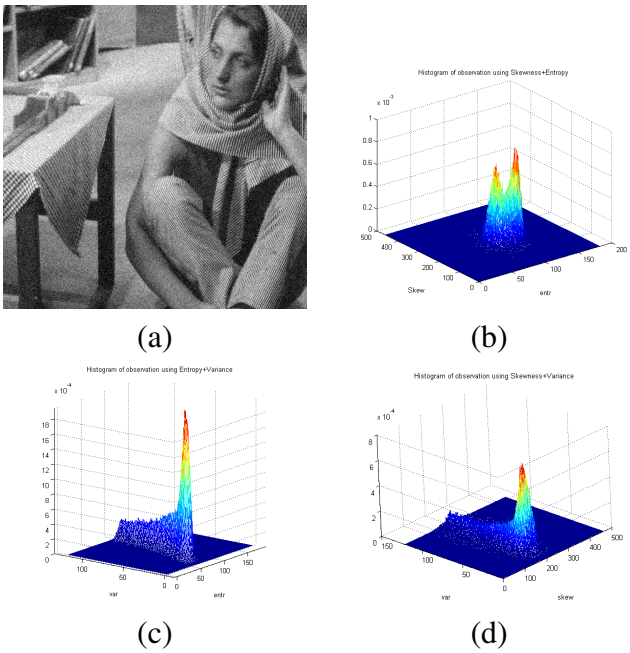


Fig. 3.2: Example of feature distribution for a an image corrupted by Gaussian synthetics noise of standard deviation 20 (b) Joint distribution (skewness, entropy) (c) Joint distribution (entropy, variance) (d) (b) Joint distribution (skewness, variance)

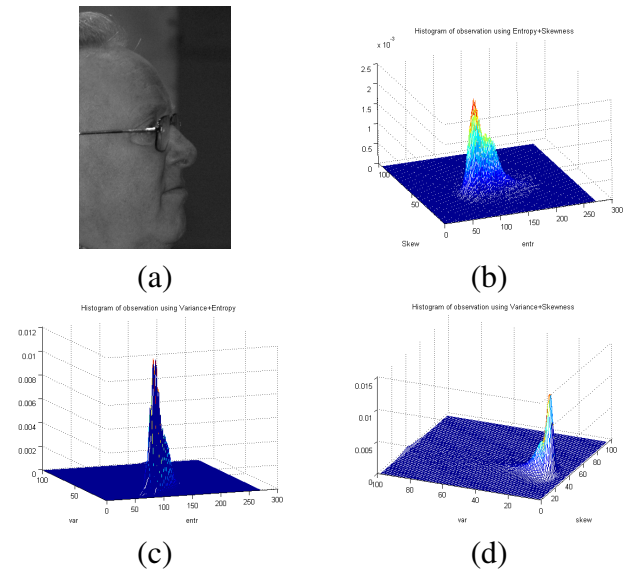


Fig. 3.3: Example of feature distribution for a an image corrupted by real digital camera noise (b) Joint distribution (skewness, entropy) (c) Joint distribution (entropy, variance) (d) (b) Joint distribution (skewness, variance)

of variance information. This is a desirable property when performing classification since it leads to a tractable problem in comparison with higher dimension original one under the condition that the retain directions form well separated clusters.
Now to describe PCA we will consider the specific case of pixel classification the set of observation corresponds to $\mathcal{F} = \{f_{\mathbf{x}} \in \mathbb{R}^3\}_{\mathbf{x}\in\Omega}$ and $f_{\mathbf{x}}$ is the feature vector described in the previous paragraph. In order to build the sub-space that optimally describes the data variation, we first have to arrange these observations on a matrix $\mathbf{M}$ of dimension $|\Omega| \times 3$ where each line of the matrix corresponds to a feature vector. Hence as proved in [44], the most prominent directions of projection corresponds to the eigenvectors of the covariance matrix of the column vector of $\mathbf{M}$. Thus if we note the column vectors $\{\mathbf{M}^i\}_{1\leq i\leq 3}$, the covariance matrix $\mathbf{C}$ defined as

$$\begin{aligned} \mathbf{C}(i,j) &= \text{cov}(\mathbf{M}^i, \mathbf{M}^j) \\ \text{cov}(\mathbf{M}^i, \mathbf{M}^j) &= \text{E}\left[\left(\mathbf{M}^i - \bar{\mathbf{M}^j}\right)\left(\mathbf{M}^j - \bar{\mathbf{M}^j}\right)\right] \quad \bar{\mathbf{M}^j} = \text{E}\left[\mathbf{M}^j\right] \end{aligned}$$

The eigenvectors of the matrix $\mathbf{C}$ form an orthonormal basis and the associated eigenvalues are the variance of the observation when projected on these eigenvectors. Note also that the projection of the data set on each direction of the new basis is uncorrelated. Hence for our classification process, we considered the first principle component since it describes better the data point distribution and preserves their variability. Such a choice allows us to reduce the problem dimensionality which can speed up the classification process and makes it simpler. Using the projection on the first principle component (noted $\mathbf{e_1}$) we associate to each pixel $\mathbf{x}$ a feature $o_{\mathbf{x}} = \langle f_{\mathbf{x}}, \mathbf{e_1} \rangle$.
Now, once the set of observations determined, we will present the classification framework.

3.3.3 Application to Pixels Classification

After the feature extraction step, our aim is to perform a soft classification of the different image pixels. Knowing that a natural image is composed of three main components that are: edges texture and smooth regions, we will consider a Gaussian mixture models. Gaussian mixture models are examples of multi-modal densities that make the assumption that the probability density function of observed samples is a sum a Gaussian distributions [49]. They have been used in clustering where one makes the assumption that the distribution inside each class is a Gaussian distribution. In this context we will associate a Gaussian for each class: texture, smooth region, and edges. To present the classification process, let us introduce $\mathbf{O} = \{o_{\mathbf{x}}\}_{\mathbf{x}\in\Omega}$ the set of N unlabeled observations corresponding to the feature vectors of the image pixels and $\mathcal{C} = \{smooth, edge, tex\}$ the set of the clusters being present in the image. Thus, using the empirical distribution of these observations,

we can consider the following model

$$\begin{aligned} p(\mathbf{o_x}|\Theta) &= \sum_{k\in\mathcal{C}} \pi_k \mathcal{G}(\mathbf{o_x}, \mu_k, \sigma_k^2) \\ \mathcal{G}(\mathbf{o_x}, \mu_k, \sigma_k^2) &= \frac{1}{\sqrt{2\pi}\sigma_k} \exp^{-\frac{\|\mathbf{o_x}-\mu_k\|^2}{2\sigma_k^2}} \quad \text{and} \quad \sum_{k\in\mathcal{C}} \pi_k = 1 \end{aligned} \tag{3.3}$$

with $\Theta = \{\pi_k, \mu_k, \sigma_k, \}_{k=1}^{|\Omega|}$ is the set of parameters defining the mixture model. Each cluster k is defined by three parameters: π_k the a priori probability of the cluster, μ_k and σ_k are respectively the mean and standard deviation of the Gaussian distribution ($\mathcal{G}$) they represent the center of the cluster k as well as a measure of its spread. An illustration of the approximation of the empirical distribution with Gaussian mixtures is presented in [Fig.(3.4)]. The estimation of the model's parameters is done through a Maximum Likelihood approach. Then assuming that the set of observations $\mathbf{O}$ was generated using the model (3.3) then the likelihood of drawing these samples is

$$L(\mathbf{O}|\Theta) = \log\left[\prod_{\mathbf{x}\in\Omega}\sum_{k\in\mathcal{C}} \pi_k \mathcal{G}(o_\mathbf{x}, \mu_k, \sigma_k^2)\right] \tag{3.4}$$

and an optimal estimation of the parameter vector assuming no prior knowledge or constraints will be the one that maximizes the likelihood of the observations, given the model parameters,

$$\hat{\Theta} = \text{Argmax}_\Theta L(\mathbf{O}|\Theta) \tag{3.5}$$

Many algorithms can be used to compute the maximum likelihood estimate of Θ, it was shown that the Expectation Maximization based approach is well adapted to solve this problem[106, 42] . The derivative of the likelihood function with respect to the model parameters is defined as:

$$\frac{\partial L(\mathbf{O}|\Theta)}{\partial \mu_k} = \sum_{\mathbf{x}\in\Omega} \frac{p(k|o_\mathbf{x})}{\sigma_k^2}(\mu_k - o_\mathbf{x}) \tag{3.6}$$

$$\frac{\partial L(\mathbf{O}|\Theta)}{\partial \sigma_k} = \sum_{\mathbf{x}\in\Omega} p(k|o_\mathbf{x})\left(-\frac{D}{\sigma_k} + \frac{\|\mu_k - o_\mathbf{x}\|}{\sigma_k^3}\right) \tag{3.7}$$

$$p(k|o_\mathbf{x}) = \frac{\pi_k \mathcal{G}(o_\mathbf{x}, \mu_k, \sigma_k^2)}{\sum_{i\in\mathcal{C}} \pi_i \mathcal{G}(o_\mathbf{x}, \mu_i, \sigma_i^2)} \tag{3.8}$$

$p(k|o_\mathbf{x})$ is the conditional probability of selecting the cluster k given the observation $o_\mathbf{x}$.
As far as the a priori cluster probability π_k is concerned, it must verify two constraints that are:

$\pi_k \geq 0$ and $\sum_{k=1}^{M} \pi_k = 1$. In order to impose these constraints one can set:

$$\pi_k = \frac{e^{\lambda_k}}{\sum_{k \in \mathcal{C}} e^{\lambda_k}} \quad \text{In this case} \quad \frac{\partial L(\mathbf{O}|\Theta)}{\partial \lambda_k} = \sum_{\mathbf{x} \in \Omega} (p(k|o_{\mathbf{x}}) - \pi_k) \tag{3.9}$$

Setting the derivative of the likelihood function to zero, results in the following equations for the Gaussian mixture model

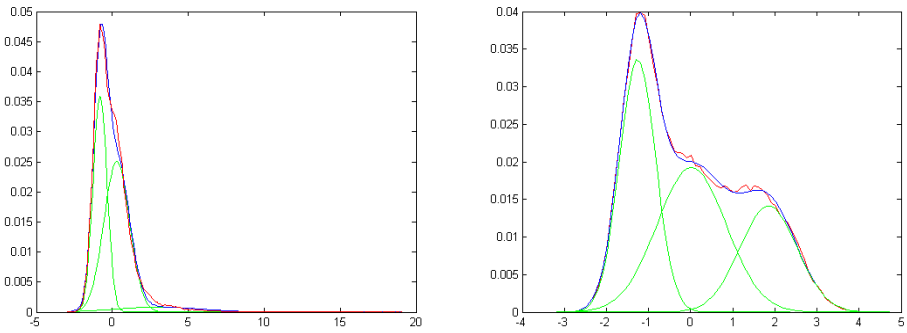
$$\mu_k = \frac{\sum_{\mathbf{x} \in \Omega} p(k|o_{\mathbf{x}}) o_{\mathbf{x}}}{\sum_{\mathbf{x} \in \Omega} p(k|o_{\mathbf{x}})} \tag{3.10}$$

$$\sigma_k^2 = \frac{\sum_{\mathbf{x} \in \Omega} p(k|o_{\mathbf{x}}) \left\| o_{\mathbf{x}} - \mu_k \right\|}{\sum_{\mathbf{x} \in \Omega} p(k|o_{\mathbf{x}})} \tag{3.11}$$

$$\pi_k = \frac{1}{|\Omega|} \sum_{\mathbf{x} \in \Omega} p(k|o_{\mathbf{x}}) \tag{3.12}$$

These expressions demonstrate that μ_k and σ_k are respectively the conditional mean and standard deviation of the samples with respect to the cluster k. These equations are dependent to each other and it is not possible to find an exact, solution when seeking model's parameters estimation. Therefore, the most common approach to overcome this constraint is through an iterative optimization process. Thus starting from an initial guess π_k^0,μ_k^0,σ_k^0, one have to perform iteratively the update using the equations (3.10)(3.11)(3.12). It is important to note that assuming the number of components known, the initial conditions of the system play an important role to the quality of the final approximation when using the EM approach.

Starting from the set of observations and using the expectation maximization algorithm we were able to determine the parameters of the model. Therefore one question remains: which Gaussian for which class. It is important to point out that we perform a fully unsupervised classification. In other words, at the end of our classification process, we have three Gaussians modeling three data clusters but we ignore which Gaussian corresponds to texture and which ones represent the smooth regions or edges. Therefore, a label has to be assigned to each Gaussian. To address this issue inference based labeling was adopted. Under the assumption that an image consists mostly of smooth regions we can use the prior density of each class to assign labels to them according to the following relation $\pi_{smooth} > \pi_{tex} > \pi_{edge}$. Parallel to that, knowing that pixels that belong to uniform noisy patches have similar descriptor we expect that the Gaussian that represents the cluster of smooth regions will have a small variance compared to the others, an assumption that has been validated by experimental results. For the example shown in [Fig.(3.5)], the Gaussian component with the smallest variance and its dominance over the other hypotheses corresponds to the column (b). It is clear that this component refers to the class describing the uniform assumption with noise. The classification result presented in image [Fig.(3.6)] show that texture and structure clusters can be mixed together. Since we process the edges and the texture in the same fashion misclassifica-



(a) (b)

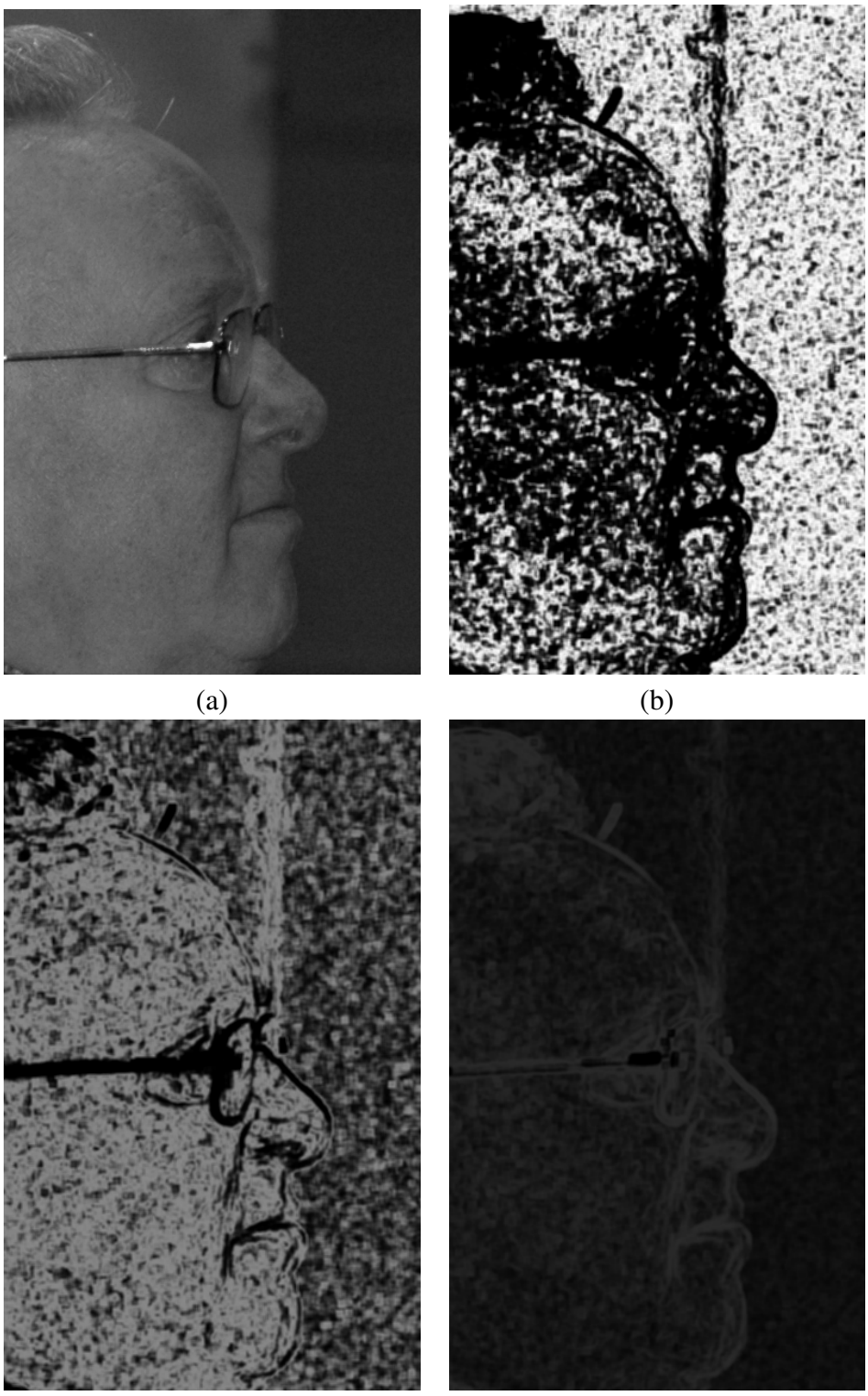
Fig. 3.4: Histogram of the projection of the features vector on the first principle component (red) and their approximation with Gaussian mixture models (blue) for (a) Old man image with real camera noise (b) Barbara image

tion wont have a great impact on the algorithm. Once the classification of image pixels performed, we will be concerned with the use of such information to build an adaptive denoising technique where particular attention is paid on how we treat texture and edges. The detailed description of the proposed algorithm will be focus of the next section.

3.4 Non Parametric Model and Adaptive Denoising

3.4.1 Bayesian Formulation of the Problem

As stated before the Bayesian formulation is a classical tool in image restoration problems. The use of a specific estimator depends on the cost function that one aims to minimize as well as the application. To introduce our denoising method, we will assume that I, U and n are three random variables defined on a discrete partition $\Omega \in \mathbb{Z}^2$. We recall that n is an additive noise independent from the signal, I is the observed noisy image and U is the noise-free image. To estimate the noise-free image, one needs to define a decision function that associates to the noisy observation the image that can be considered as the noise-free one. Such a function is noted Φ_I and defined on the space Γ that corresponds to all possible image realizations. Note that for gray scale images where the gray level ranges between 0 and 255 the set Γ contains $256^{|\Omega|}$ elements. To evaluate the pertinence of the decision Φ_I, one has also to define a loss function that evaluates a risk when a decision is taken. Hence to introduce the Bayesian risk and the decision function we will consider a loss function L defined on $\Gamma \times \Gamma$ and having values in $\mathbb{R}^+$. To compute an estimate of the noise-free image U given the noisy observations, one has to minimize a Bayesian risk that is the posterior



(a) (b)

(c) (d)

Fig. 3.5: Results of an image partition: (a) original image, and conditional probability function relative to (b) "smooth component" $p(o_x|smooth)$, (c) "texture" $p(o_x|tex)$ and (d) "edges" $p(o_x|edge)$

(a) (b)

(c) (d)

Fig. 3.6: Results of an image partition: (a) original noisy image (σ_n=20), and conditional probability function relative to (b) "smooth component" $p(o_x|smooth)$, (c) "texture" $p(o_x|tex)$ and (d) "edges" $p(o_x|edge)$

expected value of a loss function and defined as

$$\mathcal{R}(\Phi_I) = \mathbf{E}\left[L(J, \Phi_I)|I\right] = \sum_{J\in\Gamma} L(J, \Phi_I)p(J|I) \tag{3.13}$$

In the Bayesian decision theory, three types of loss functions and therefore three estimators are considered.

- **Maximum a Posteriori Estimator (MAP)**
  The Maximum a Posteriori estimator is associated to a global loss function L defined as

  $$L(I_1, I_2) = 1 \quad \text{if} \quad I_1 = I_2 \quad \text{and} \quad L(I_1, I_2) = 0 \quad \text{otherwise}$$

  In this case

  $$\mathcal{R}(\Phi_I) = \mathbf{E}\left[L(J, \Phi_I)|I\right] = 1 - p(J = \Phi_I|I)$$

  Consequently the optimal estimator Φ_I corresponds the image that maximizes the posterior probability $p(J|I)$. Such an estimate relies on the appropriate definition of the posterior probability that involves knowledge about the noise model and the natural image model. As we pointed out in the first chapter the image model can refer to MRF's based parametric models [58, 156, 120, 97, 119] or non parametric ones [10, 77].

- **Maximum Posterior Mode estimator (MPM)**
  The Maximum Posterior Mode estimator is associated to the following local loss function

  $$L(I_1, I_2) = \sum_{\mathbf{x}\in\Omega} l(I_1(\mathbf{x}), I_2(\mathbf{x})) = \sum_{\mathbf{x}\in\Omega} \mathbf{1}_{\{I_1(\mathbf{x})\neq I_1(\mathbf{x})\}}$$

  The Bayes risk in this case corresponds to

  $$\begin{aligned}
  \mathcal{R}(\Phi_I) = \mathbf{E}\left[L(J, \Phi_I)|I\right] &= \sum_{J\in\Gamma}\left(\sum_{\mathbf{x}\in\Omega} l(J(\mathbf{x}), \Phi_I(\mathbf{x}))\right) p(J|I) \\
  &= \sum_{\mathbf{x}\in\Omega}\sum_{J\in\Gamma} l(J(\mathbf{x}), \Phi_I(\mathbf{x}))p(J|I) \\
  &= \sum_{\mathbf{x}\in\Omega} E\left[l(J(\mathbf{x}), \Phi_I(\mathbf{x}))|I\right]
  \end{aligned} \tag{3.14}$$

  The Bayes risk is the sum of the conditional expectation of the local cost function. The criteria to be minimized is sum of independent positive terms thus the optimal estimator is the one that minimizes each term $E\left[l(J(\mathbf{x}), \Phi_I(\mathbf{x}))|\, I\right]$. This result is valid for any loss function defined as the sum of local loss functions. In the particular case of the loss function

introduced before

$$E\left[l(J(\mathbf{x}),\Phi_I(\mathbf{x}))|I\right] = 1 - p(J(\mathbf{x}) = \Phi_I(\mathbf{x})|I)$$

In such a case the optimal estimator is the one that is defined as $\Phi_I(\mathbf{x}) = \text{Argmax}_{J(\mathbf{x})} p(J(\mathbf{x})|I)$. Thus for each pixel in the image, one has to compute the most probable intensity given the noisy observation I. This can be achieved through the definition of an appropriate posterior density. The definition of such a posterior will be addressed in the next section.

- **Conditional Mean Estimator**
  The associated loss function is also local and defined as

$$L(I_1, I_2) = \sum_{\mathbf{x}\in\Omega} \|I_1(\mathbf{x}) - I_2(\mathbf{x})\|^2$$

using the result obtained in equation (3.14)

$$\mathcal{R}(\Phi_I) = \mathbf{E}\left[L(J,\Phi_I)|I\right] = \sum_{\mathbf{x}\in\Omega} E\left[(J(\mathbf{x}) - \Phi_I(\mathbf{x}))^2|I\right] \tag{3.15}$$

With this local formulation of the estimator, one has to minimize each term of the sum to obtain the optimal estimator. By introducing $m_{\mathbf{x}} = E\left[p(J(\mathbf{x})|I)\right]$ and considering the relation

$$\begin{aligned} E\left[(J(\mathbf{x}) - \Phi_I(\mathbf{x}))^2|I\right] &= \sum_{J(\mathbf{x})\in E} (J(\mathbf{x}) - m_{\mathbf{x}})^2 p(J(\mathbf{x})|I) \qquad (3.16)\\ &+ \sum_{J(\mathbf{x})\in E} (\Phi_I(\mathbf{x}) - m_{\mathbf{x}})^2 p(J(\mathbf{x})|I) \\ &= Cte + (\Phi_I(\mathbf{x}) - m_{\mathbf{x}})^2 \qquad (3.17)\end{aligned}$$

The minimum of the risk where the estimate value relative to $\mathbf{x}$ is defined as

$$\Phi_I(\mathbf{x}) = E\left[p(J(\mathbf{x})|I)\right] \tag{3.18}$$

In the context of image restoration the MAP estimator is a very popular tool to estimate the noise free image that relies on global models. The MPM and conditional mean estimators are local and estimate the image pixels in an independent fashion. Regarding their loss functions, the MPM estimator penalizes the number of differences between the images regardless their values contrarily to the conditional mean estimator. In [25], a proof that the earlier presented NL-means algorithm is consistent with the conditional mean estimator is presented. In the present work we consider the

MPM estimator because its loss function is more restrictive for each pixel. Further more it allows us to define local image models, according to the soft classification results. Based on the MPM estimation technique, the restoration process consists in computing the intensity for each pixel in the image by maximizing the conditional marginal posterior probability. Thus, the estimate $\hat{U}(\mathbf{x})$ of the original observation at a given position $\mathbf{x}$ satisfies,

$$\hat{U}(\mathbf{x}) = \text{Argmax}_{U(\mathbf{x})} \left[p\left(U(\mathbf{x})|I\right) \right]$$

To estimate the marginal posterior, we will make the hypothesis that the observation $U(\mathbf{x})$ is conditioned on image observations defined at a restricted neighborhood of the image. Without loss of generality such assumption is valid for natural images where image content is independent at a large scale. Thus, if we introduce $\mathbf{N}_x$, a $w_p \times w_p$ square neighborhood of $\mathbf{x}$ and $I_{\mathbf{N}_x}$ the set of intensity observed within the neighborhood, the marginal posterior is defined as:

$$p\left(U(\mathbf{x})|I\right) \approx p\left(U(\mathbf{x})|I_{\mathbf{N}_x}\right) = \frac{p\left(U(\mathbf{x}), I_{\mathbf{N}_x}\right)}{p\left(I_{\mathbf{N}_x}\right)} \tag{3.19}$$

$p\left(I_{\mathbf{N}_x}\right)$ being constant, one has to determine a model relative to the joint probability $p\left(U(\mathbf{x}), I_{\mathbf{N}_x}\right)$. In the absence of a prior on the noise free image we can not obtain a direct estimate of the joint probability. The only available information being the observed image I, one have to exploit the fact that natural images are redundant which means that an image could contain several copy of the same patch [26]. This suggests that one can use the different noisy patches in the image to build a statistical model [77, 10]. For our case, to estimate the joint probability $p\left(U(\mathbf{x}), I_{\mathbf{N}_x}\right)$ we will consider the following set of samples

$$\mathcal{S} = \left\{ [I(\mathbf{y}), I_{\mathbf{N}_y}] \quad \text{where} \quad \mathbf{y} \in R_{\mathbf{x}} \right\} \tag{3.20}$$

with $R_{\mathbf{x}}$ is the domain where is selected the set of samples patches used to estimate the joint probability. This domain can be the whole image domain or just a rectangular window centered on $\mathbf{x}$ and of size $w_R \times w_R$. Once the sample set is defined, the joint probability can be approximated using non parametric density function estimation technique.

3.4.2 Non Parametric Density Estimation

The non parametric density function estimation is a technique that aims at finding the pdf associated to a random variable $\mathbf{V}$ (defined on $\mathbb{R}^d$) given an arbitrary set of observations $\{\mathbf{V}_i\}_{i=1}^M$. Since parametric density functions such as Gaussian or Laplacian distributions are restrictive and make

explicit assumptions on the expected distribution of the observed samples, more general techniques have been considered [113, 70]. The most intuitive tool is the histogram based estimation. However, this approach is highly sensitive to the selection of the bins, while being discontinuous with derivatives ill-defined, etc. An alternative to the histogram and a generalization of it are the kernel based estimators which can model arbitrary distribution and are defined as the following

$$f(\mathbf{V}) = \frac{1}{M}\sum_{i=1}^{M}\mathbf{K_H}\left(\mathbf{V}-\mathbf{V_i}\right) \tag{3.21}$$

Such an approach associates to each observation a continuous estimator (kernel) centered on it. Then, it approximates the probability of a sample by measuring the distance from each continuous estimator. The advantages of such a method are explicit use of observations, and certain continuity properties in terms of low and higher order derivatives. The kernel $\mathbf{K_H}$ is a function defined on $\mathbb{R}^d$ that must satisfy certain properties so that f can be considered as a probability density function and provides a good estimation of the actual distribution. These properties are:

$$\int_{\mathbb{R}^d}\mathbf{K_H}(\mathbf{V})d\mathbf{V}=1 \qquad \int_{\mathbb{R}^d}\mathbf{V}\mathbf{K_H}(\mathbf{V})d\mathbf{V}=0 \qquad \int_{\mathbb{R}^d}\mathbf{V}.\mathbf{V}^T\mathbf{K_H}(\mathbf{V})d\mathbf{V}=\mathbf{I_d} \tag{3.22}$$

$\mathbf{I_d}$ is the $d\times d$ identity matrix. $\mathbf{H}$ is the bandwidth of the kernel it controls the smoothness degree of the estimation and the efficiency of the method is highly dependent on the selection of this parameter as well as the kernel expression. Various types of kernels can be used such as the uniform kernel that refers also to Parzen window, the Epanechnikov kernels and Gaussian kernels. For our application we selected an isotropic Gaussian kernel with bandwidth defined as $\mathbf{H}=\mathbf{hI_d}$ (where $\mathbf{h}$ is a positive scalar). Our choice of the Gaussian kernel is motivated by its nice properties of continuity and derivability that are also inherited by the estimated density function.

$$\mathbf{K_H}(\mathbf{u}) = \frac{1}{\mathbf{h}^d\sqrt{(2\pi)^d}}\exp\left(-\frac{\|\mathbf{u}\|^2}{2\mathbf{h}^2}\right)$$

The quality of the kernel density estimator is measured by the mean squared error between the actual density function and its estimate. As far as the point wise error is concerned, it is defined by

$$\begin{aligned} MSE(\mathbf{V}) = E\left[\hat{f}(\mathbf{V})-f(\mathbf{V})\right]^2 &= var\left(\hat{f}(\mathbf{V})\right)+\left[E\left(\hat{f}(\mathbf{V})\right)-f(\mathbf{V})\right]^2 \\ &= var\left(\hat{f}(\mathbf{V})\right)+\left[Bias\left(\hat{f}(\mathbf{V})\right)\right]^2 \end{aligned} \tag{3.23}$$

The bias of the estimator gives an indicator of the closeness of the mean of the estimate to the real distribution and the variance is a measure of how precise is the estimator. Ideal estimator is the one

who reaches a balance between variance and bias.
In [142], it was shown that the bias and variance could be approximated by

$$Bias\left(\hat{f}(\mathbf{V})\right) \approx A\mathbf{h}^2\Delta f(\mathbf{V}) \quad \text{and} \quad var\left(\hat{f}(\mathbf{V})\right) \approx B\mathbf{h}^{-d}f(\mathbf{V}) \tag{3.24}$$

This relation show that there exists a trade off between variance and bias: small bandwidth value will result in less biased estimator however it will affect its accuracy. In practice, it is not possible to directly compute the optimal value of the bandwidth because the lower bound of the square error depends on the unknown density distribution. To address this limitation, many data driven approaches was proposed such as the plug in rule [125]. These methods rely on a fixed bandwidth kernel and are based on a smoothness assumption of the underlying density. Their main limitation is that an over-smoothing is often observed in particular for the peaks of the density. Variable bandwidth is an alternative when the data exhibit scale variations.

3.4.3 Variable Bandwidth Selection

We recall that our purpose is to estimate a probability density function of patches distribution on the observed noisy image. This is done through non parametric density estimation technique involving isotropic Gaussian kernels. In this context, the estimation process depends only on one parameter that is $\mathbf{h}$. The effect of broadening the kernel (high value of $\mathbf{h}$) will result in a loss of resolution by merging multiple modes, that may represent some data properties. On the other hand, small values of $\mathbf{h}$ may decrease contribution of the neighboring samples and lead to a rather inaccurate probability density estimate. To reflect the data diversity and the variability of concentration of observations, we have to associate to each data point $\mathbf{V}_i$ a differently scaled kernels $\mathbf{h}_i$. Under these considerations the formulation of the non parametric density estimator is:

$$f(\mathbf{V}) = \frac{1}{M}\sum_{i=1}^{M}\frac{1}{\mathbf{h}_i^d\sqrt{(2\pi)^d}}\exp\left(-\frac{\|\mathbf{V}-\mathbf{V_i}\|^2}{2\mathbf{h}_i^2}\right) \tag{3.25}$$

As far as our application is concerned, the observation set $\mathcal{S}$ is composed of image patches as defined in (3.20). We can assume that image patches belonging to smooth regions does not exhibit high amount of variability. Hence towards decreasing the importance of variation in the distribution when referring to homogeneous patches, we have to increase the bandwidth values. Thus, all the samples will inherit equal importance with such a selection during the pdf approximation process. Regarding, the image patches belonging to various types of texture and other structure being present in the image, the choice of a high bandwidth value will results in a probability density

function that don't take into account data diversity. Using such an estimate will affect the quality of restoration leading to over smoothed images. In this case, small bandwidth values should be considered in order to obtain multi-modal probability density functions. Furthermore, the selection of small bandwidth is equivalent to be restrictive on the choice of samples to be used for the density estimation and hence for the pixel intensity restoration. So, for a given pixel, only pixels having high amount of photometric similarity with it will be considered in the denoising process. To sum up and in order to have a probability density function of the image patches distribution (3.19), we need to define an appropriate bandwidth function to take into account data diversity. To this end, we exploit the result of the pixels classification step presented in the previous section to build a pixel-specific bandwidth. Since the bandwidth values must be small for textured patches and high for the homogeneous one, the bandwidth function should be decreasing with respect to the conditional probability relative to textured region or edges and increasing for homogeneous patches. Therefore for a given observation $\mathbf{y}$ in $\mathcal{S}$ we associate a bandwidth value $\mathbf{h_y}$ that can be defined using the following function

$$\mathbf{h_y} = \sigma_0 \left(\frac{p_{smooth}(o_\mathbf{y})}{p_{tex}(o_\mathbf{y}) + p_{edge}(o_\mathbf{y}) + 1} + c \right) \tag{3.26}$$

$$p_k(o_\mathbf{y}) = \mathcal{G}(\mathbf{o_y}, \mu_k, \sigma_k^2) \quad \text{for} \quad k \in \{smooth, tex, edge\} \tag{3.27}$$

We recall that $o_\mathbf{y}$ is the feature vector in the position $\mathbf{y}$ and $\mathcal{G}$ a Gaussian distribution, while p_{smooth},p_{tex} and p_{edge} are respectively the conditional probability for $o_\mathbf{y}$ to be in a noisy smooth region, textured region or edges. σ_0 and c are parameters to be fixed according to noise level. With such a choice, we adopt for pixels that belong to smooth regions with high value of p_{smooth}, high kernel bandwidth. For image component identified as texture or edges, p_{smooth} tends to be close to zero and implies smaller bandwidth values. An example of such bandwidth estimation is shown in [Fig.(3.8-b),Fig.(3.9-b)] where clearly texture regions are associated with low bandwidth which produces a density that preserves image structure.

3.4.4 Marginal Posterior Maximizing

Let us recall that our denoising consists in computing an estimate of the pixel intensity that maximizes the conditional posterior defined in equation (3.19). It amounts also to maximizing a function $E(U(\mathbf{x}))$ that corresponds to the joint probability $p\left(U(\mathbf{x}), I_{\mathbf{N}_x}\right)$. Based on the set $\mathcal{S}$ of patches observed in the noisy image and by applying the sample based density estimation technique using

variable bandwidth Gaussian kernels an estimate of $p\left(U(\mathbf{x}), I_{\mathbf{N}_x}\right)$ is

$$p\left(U(\mathbf{x}), I_{\mathbf{N}_x}\right) = E(U(\mathbf{x})) \tag{3.28}$$

$$= \frac{1}{|R_{\mathbf{x}}|} \sum_{\mathbf{y} \in R_{\mathbf{x}}} \frac{1}{(\mathbf{h_y}\sqrt{2\pi})^d} \exp - \left(\frac{\left\| [U(\mathbf{x}), I_{\mathbf{N}_x}] - \left[I(\mathbf{y}), I_{\mathbf{N}_y}\right] \right\|_2}{2\mathbf{h_y}^2} \right)$$

$$= \frac{1}{|R_{\mathbf{x}}|} \sum_{\mathbf{y} \in R_{\mathbf{x}}} \frac{1}{(\mathbf{h_y}\sqrt{2\pi})^d} \exp - \left(\frac{\left\| I_{\mathbf{N}_x} - I_{\mathbf{N}_y} \right\|_2}{2\mathbf{h_y}^2} \right) \exp - \left(\frac{\left[U(\mathbf{x}) - I(\mathbf{y})\right]^2}{2\mathbf{h_y}^2} \right) \tag{3.29}$$

d is the dimension of vector in the set of sample $\mathcal{S}$ defined in (3.20) and $|R_{\mathbf{x}}|$ refers to the cardinality of $R_{\mathbf{x}}$. The bandwidth of the isotropic Gaussian kernel depends on the image content and is defined by equation (3.27). Let us introduce now the variable $w_{\mathbf{xy}}$ defined by

$$w_{\mathbf{xy}} = \frac{1}{(\mathbf{h_y}\sqrt{2\pi})^d} \exp - \left(\frac{\left\| I_{\mathbf{N}_x} - I_{\mathbf{N}_y} \right\|_2}{2\mathbf{h_y}^2} \right) \tag{3.30}$$

Equation (3.29) becomes

$$E(U(\mathbf{x})) = \frac{1}{|R_{\mathbf{x}}|} \sum_{\mathbf{y} \in R_{\mathbf{x}}} w_{\mathbf{xy}} \exp - \left(\frac{\left[U(\mathbf{x}) - I(\mathbf{y})\right]^2}{2\mathbf{h_y}^2} \right) \tag{3.31}$$

We have to point out that the coefficients $w_{\mathbf{xy}}$ express the photometric similarity between the neighborhood relative to $\mathbf{x}$ and $\mathbf{y}$. This shows that the contribution of any sample patch centered on $\mathbf{y}$ in the set $\mathcal{S}$ is proportional to its similarity to the observed patch centered on $\mathbf{x}$. In order to minimize the expression (3.31), we have to compute its derivative with respect to $U(\mathbf{x})$

$$\frac{dE}{dU(\mathbf{x})} = \frac{1}{|R_{\mathbf{x}}|} \sum_{\mathbf{y} \in R_{\mathbf{x}}} \frac{w_{\mathbf{xy}}}{\mathbf{h_y}^2} \left[I(\mathbf{y}) - U(\mathbf{y})\right] \exp - \left(\frac{\left[U(\mathbf{x}) - I(\mathbf{y})\right]^2}{2\mathbf{h_y}^2} \right) \tag{3.32}$$

Accessing a direct value of the optimal value of $U(\mathbf{x})$ that verifies $\frac{\partial E}{\partial U(\mathbf{x})} = 0$ is not feasible. Hence, the gradient ascent algorithm was chosen to detect the optimal intensity which corresponds to the mode of the distribution. Starting from an initial observation $U^0(\mathbf{x}) = I(\mathbf{x})$ the update of intensity is done according to

$$U^{t+1}(\mathbf{x}) = U^t(\mathbf{x}) + dt \sum_{y \in R_x} \frac{w_{\mathbf{xy}}}{\mathbf{h_y}^2} \left[I(\mathbf{y}) - U^t(\mathbf{x})\right] \exp - \left(\frac{\left[U^t(\mathbf{x}) - I(\mathbf{y})\right]^2}{2\mathbf{h_y}^2} \right) \tag{3.33}$$

We have to precise that the gradient ascent algorithm provides local maximum. The pixel intensity is attracted by the closest mode of the distribution $p\left(U(\mathbf{x}), I_{\mathbf{N}_x}\right)$. We have to point out that the

convergence to local maximum is not a major drawback of the algorithm since it consider the closest intensity to the observed one leading to better preserving to image details and texture.
Note that maximizing the function E is equivalent to the mean shift based mode detection [36, 38, 39]. To introduce this technique let us consider the derivative expression (3.32) that is equivalent to

$$\begin{aligned} \frac{dE}{dU(\mathbf{x})} &= \frac{1}{|R_{\mathbf{x}}|}\left[\sum_{\mathbf{y}\in R_{\mathbf{x}}} \frac{w_{\mathbf{xy}}}{{\mathbf{h_y}}^2} \exp - \left(\frac{[U(\mathbf{x})-I(\mathbf{y})]^2}{2{\mathbf{h_y}}^2}\right)\right] \\ &\times \left[\frac{\sum_{\mathbf{y}\in R_{\mathbf{x}}} \frac{w_{\mathbf{xy}}}{{\mathbf{h_y}}^2} I(\mathbf{y}) \exp - \left(\frac{[U(\mathbf{x})-I(\mathbf{y})]^2}{2{\mathbf{h_y}}^2}\right)}{\sum_{\mathbf{y}\in R_{\mathbf{x}}} \frac{w_{\mathbf{xy}}}{{\mathbf{h_y}}^2} \exp - \left(\frac{[U(\mathbf{x})-I(\mathbf{y})]^2}{2{\mathbf{h_y}}^2}\right)} - U(\mathbf{x})\right] \end{aligned} \tag{3.34}$$

The derivative is equal to zero when the second term in the product is equal to zero. Thus the optimal estimate of $U(\mathbf{x})$ is the one that verifies

$$\frac{\sum_{\mathbf{y}\in R_{\mathbf{x}}} \frac{w_{\mathbf{xy}}}{{\mathbf{h_y}}^2} I(\mathbf{y}) \exp - \left(\frac{[U(\mathbf{x})-I(\mathbf{y})]^2}{2{\mathbf{h_y}}^2}\right)}{\sum_{\mathbf{y}\in R_{\mathbf{x}}} \frac{w_{\mathbf{xy}}}{{\mathbf{h_y}}^2} \exp - \left(\frac{[U(\mathbf{x})-I(\mathbf{y})]^2}{2{\mathbf{h_y}}^2}\right)} - U(\mathbf{x}) = 0 \tag{3.35}$$

When considering the sequence $\{U^t(\mathbf{x})\}_{t=0,2..}$ defined as

$$U^{t+1}(\mathbf{x}) = \frac{\sum_{\mathbf{y}\in R_{\mathbf{x}}} \frac{w_{\mathbf{xy}}}{{\mathbf{h_y}}^2} I(\mathbf{y}) \exp - \left(\frac{\left[U^t(\mathbf{x})-I(\mathbf{y})\right]^2}{2{\mathbf{h_y}}^2}\right)}{\sum_{\mathbf{y}\in R_{\mathbf{x}}} \frac{w_{\mathbf{xy}}}{{\mathbf{h_y}}^2} \exp - \left(\frac{[U^t(\mathbf{x})-I(\mathbf{y})]^2}{2{\mathbf{h_y}}^2}\right)} \tag{3.36}$$

A convergence proof of the following sequence was provided in [38] if the used kernel is convex and has monotically decreasing profile which is the case of a Gaussian kernel.
Our denoising algorithm bears some similarities with the well known mean shift filtering algorithm. The main difference between our approach and the mean shift filtering consists in the mean shift computation where we take into account the similarity measure between local patches. Therefore, we are able to improve the performance in the case of textured regions because the process goes beyond simple pixel-wise comparisons between pixels to be extended to local neighborhood.

3.5 Experimental Results

The efficiency of denoising technique is related to its ability to remove noise while preserving the details and being as close as possible to the original noise-free image. Hence, to assess the quality

of a technique we consider two criteria

- Subjective criteria that refer to the visual quality of the restored image like absence of artifacts and detail preservation. One can also compute the residual image that is the difference between the noisy image and the restored one. Ideally this image corresponds to the noise added to the original image. Thus the residual should not exhibit image structure.

- Objective criteria that refer to the degree of similarity between the restored image and the noise free one. To evaluate this, several measures were considered such as the Mean Square Error, the Mean Absolute Error, The Signal to Noise Ratio, and the Peak Signal to Noise ratio. The computation of one of these criteria is possible when the noise free image is available. Hence, the objective evaluation is only possible for synthetic noise. In this thesis we considered the PSNR criterion that is defined as

$$PSNR = 10log_{10}\frac{255^2}{MSE} \quad \text{with} \quad MSE = \frac{1}{|\Omega|}\sum_{\mathbf{x}\in\Omega}(U(\mathbf{x}) - \hat{U}(\mathbf{x}))^2 \tag{3.37}$$

  We recall that U is the noise free image and $\hat{U}$ is its estimation by the denoising method.

To evaluate the performance of our method, comparisons with two state-of-the art techniques are presented. The first is the NL-means algorithm [25] because it also aims at reducing the quadratic Bayes risk (3.18) using a different cost function when compared to the MPM estimator that we used. The second is the unsupervised information-theoretic adaptive filter (UINTA) [10] that relies also on non parametric kernel based probability density estimation. We considered the fixed bandwidth kernel-estimator (MPMfix) as well towards evaluating the importance of soft classification and non-parametric density approximation driven from the image content.
To this end we considered a set of images frequently used in the context of restoration and we added to them a synthetic Gaussian noise of known standard variation ($\sigma_n = 20$). The parameter setting was selected to insure a good balance between noise suppression and details preserving. For instance the NL-means filter [1] was used with $\sigma = 15$, the patch size used for comparison is 7×7 and the weighted average was computed over a local neighborhood. Our tests proved that a non local smoothing is not necessary the most efficient technique regarding texture preservation. The size of the window that provides better results is dependent on the choice of σ_{ph} and in this particular case we computed the weighted average on a local window of size 11×11. For the UINTA algorithm, we considered the set of parameters provided by the author in [10] as well as his MATLAB implementation. Parameters that are involved in our denoising approach are the size

[1] using our implementation

	Barbara	Boat	Fingerprint	House	Lena	Baboon
MPM-Fix	29.18	28.84	26.38	31.16	31.19	25.26
MPM-var	28.9	29.11	26.68	31.02	31.25	25.39
UINTA [10]	29.53	28.66	26.59	31.85	30.82	24.09
NL-mean	28.78	28.92	26.45	30.86	31.13	25.18

Tab. 3.1: PSNR values for denoised image (The PSNR of the image corrupted by Gaussian noise of std=20 is equal to 22.15)

of the two neighborhoods $R_{\mathbf{x}}$ and $N_{\mathbf{x}}$ as well as the bandwidth of the kernels that are selected according to expression (3.27). $R_{\mathbf{x}}$ is the size of the neighborhood used for the posterior estimation. Using an important size of $R_{\mathbf{x}}$ allows a better estimation, but it is computationally complex. In this experiments we selected a 19×19 neighborhood size. $N_{\mathbf{x}}$ is the size of the noisy patch around the pixel that we want to recover, in our experiments we selected a 7×7 patch size. Regarding the choice of the photometric bandwidth, it is dependent on noise level. In case of additive Gaussian noise with standard deviation $\sigma_n = 20$, we considered this couple of parameter ($\sigma_0 = 4$ and $c = 4$). Regarding computation time we implemented our algorithm using C++ and for these set of parameter it takes 4 minutes to process a 512×512 image on a Pentium IV -2GHz machine (1mn for classification and 3 mn for the denoising) . We reported in table (Tab.3.1) the PSNR values of each method. We can notice that our approach outperforms the two other algorithms in terms of PSNR. As far visual quality is concerned, we present in figure [Fig.(3.7] a zoom of a textured part in Barbara image as well as the associated residual. These images show that in terms of detail preservation the three approaches are equivalent. Nevertheless, the UINTA result provides more pleasant visual results and better texture reconstruction. This is natural outcome because it is a global homogenization algorithm where all the image pixels are updated at the same time which enforces consistency between them contrarily to our algorithm. The later aspect makes also this algorithm more complex and computationally inefficient. The comparison with the NL-means result show that in case of additive Gaussian noise, there an assessment regarding the selection of the estimator that can be used in denoising is not straightforward. For instance, the MPM estimator and the TPM one can yield to comparable results.

Now if we consider the impact of using a variable kernel size according to the pixel position, we can notice that it preserves more details even though the textured part contains some times more noise. The tests carried on images corrupted by real camera noise and illustrated in figures [Fig.(3.8),Fig.(3.9),Fig.(3.10)] show the advantage of using a variable bandwidth since yields more natural aspect of the texture than the fixed one.

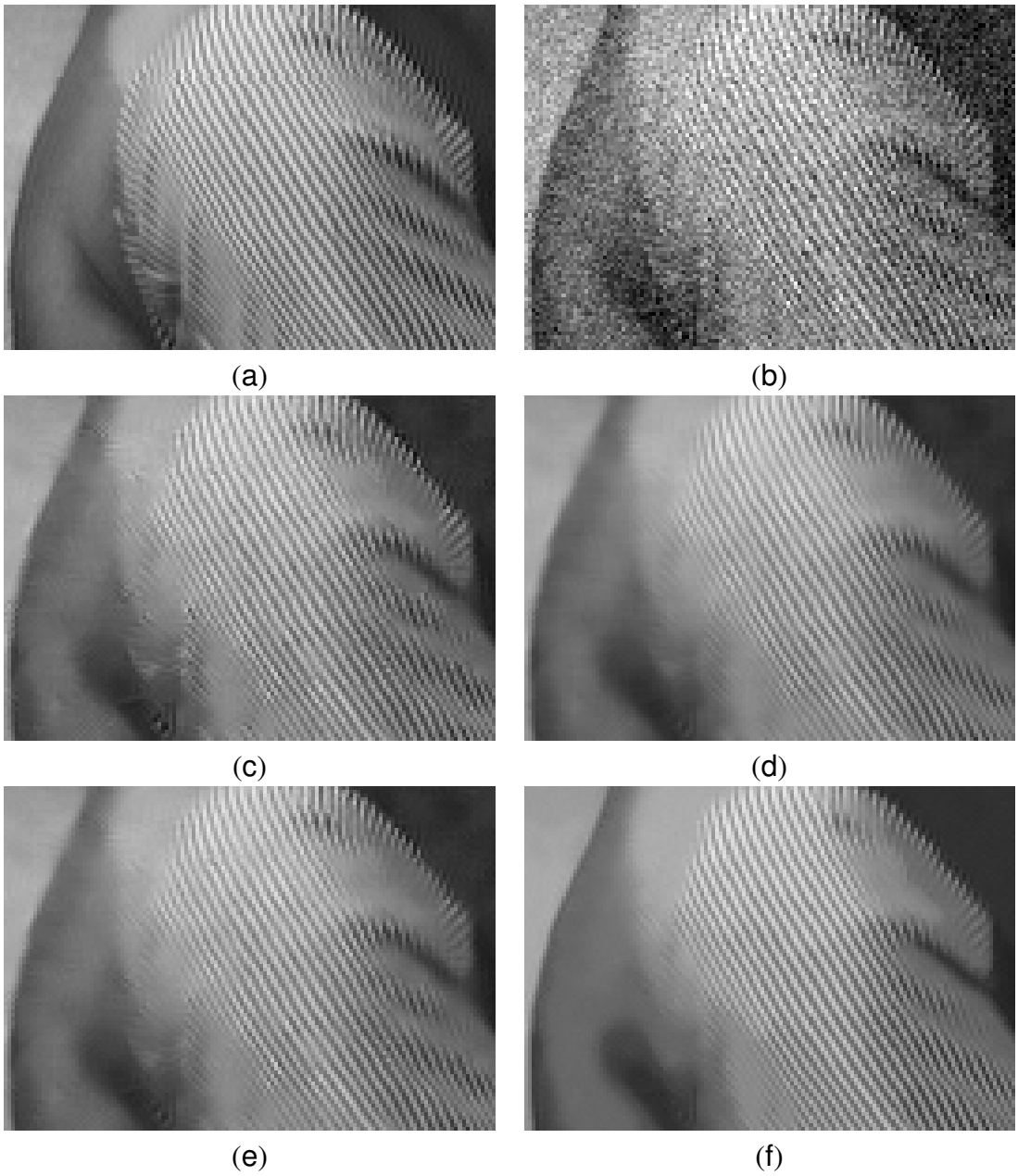
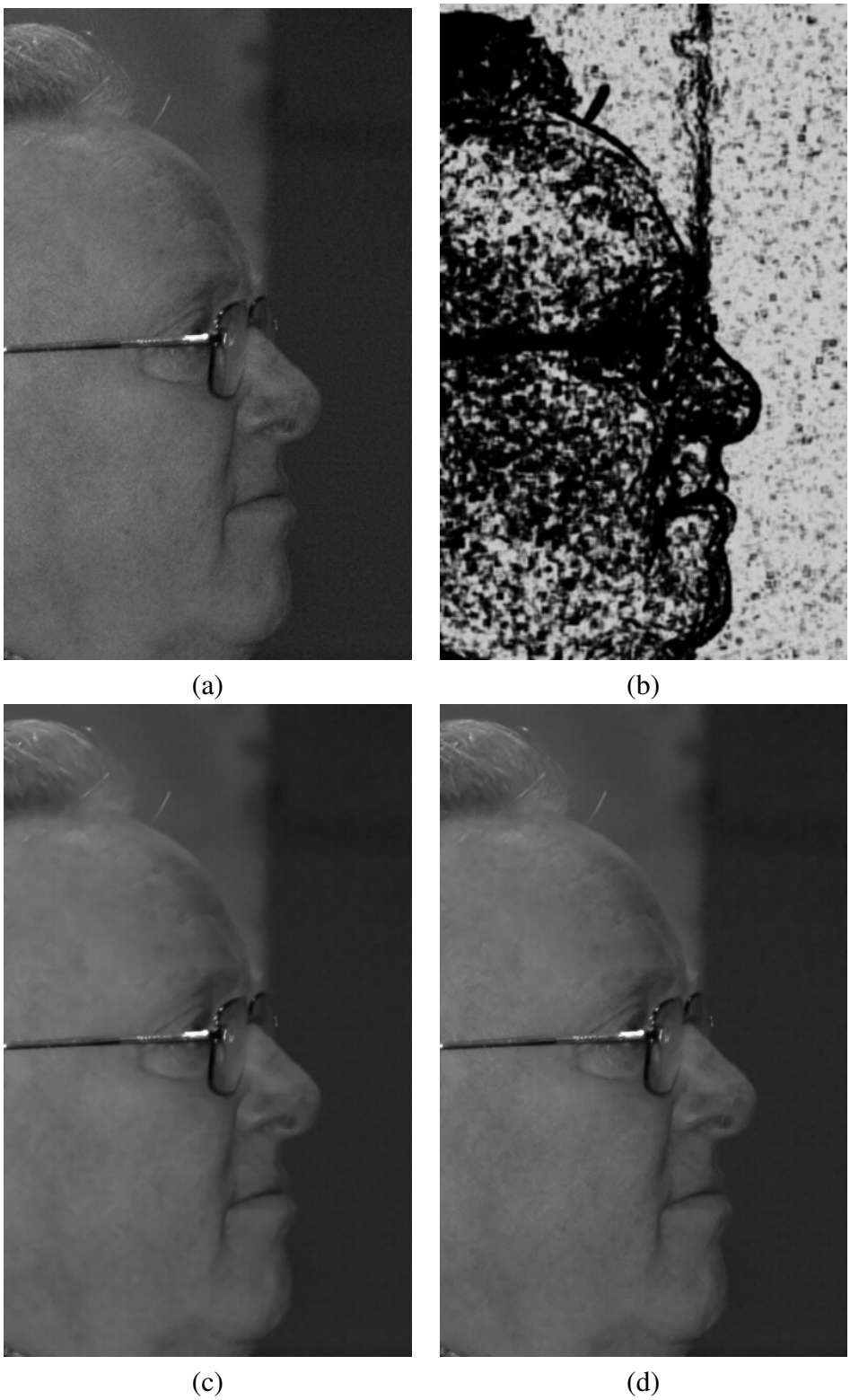


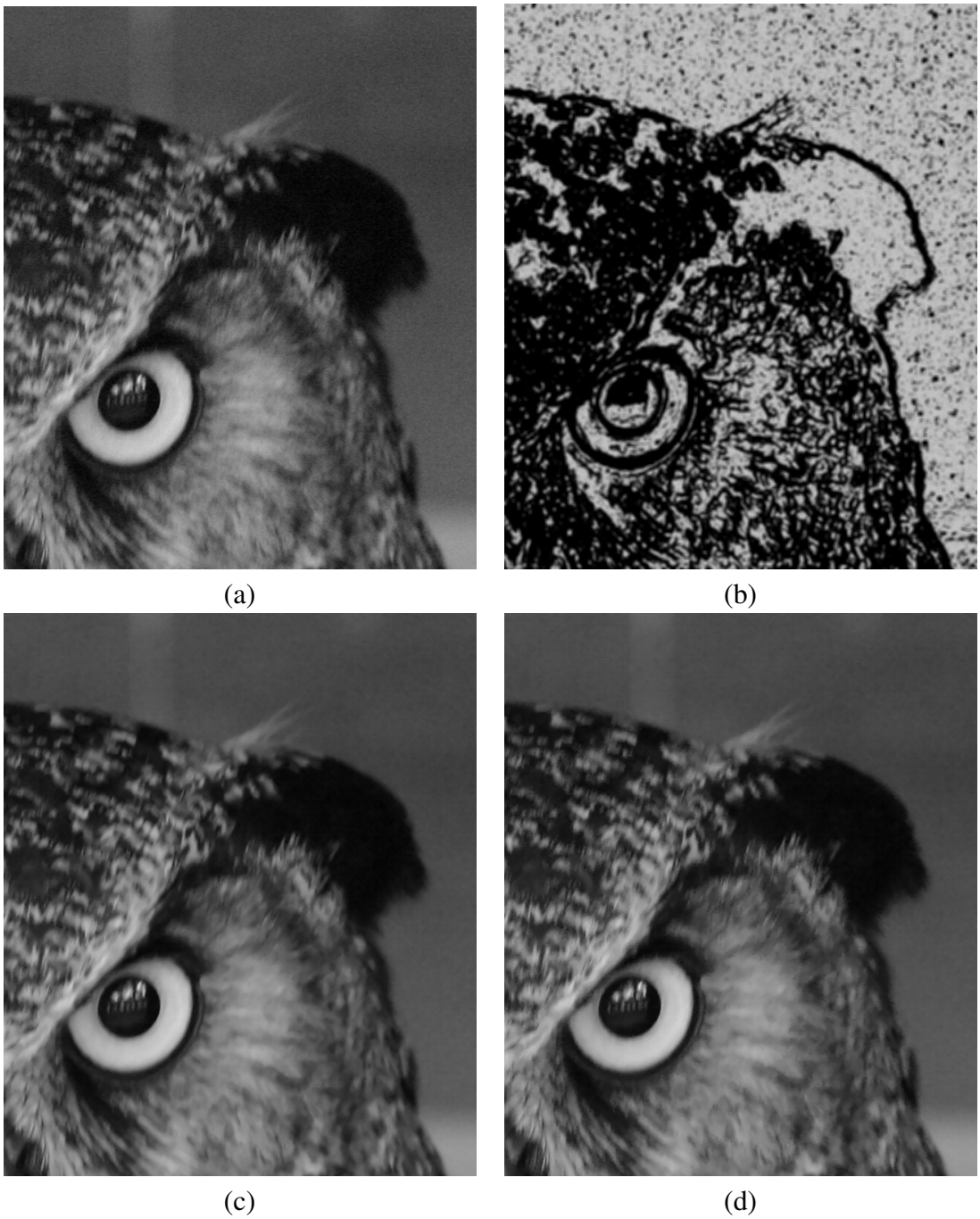
Fig. 3.7: Zoom on a detail in the Barbara image (a) original image (b) noisy image (c) restoration results using our technique with variable bandwidth kernels (d) restoration results using our technique with fixed bandwidth kernels (e) restoration result obtained with NL-means algorithm (f)restoration result obtained with UINTA algorithm



(a) (b)

(c) (d)

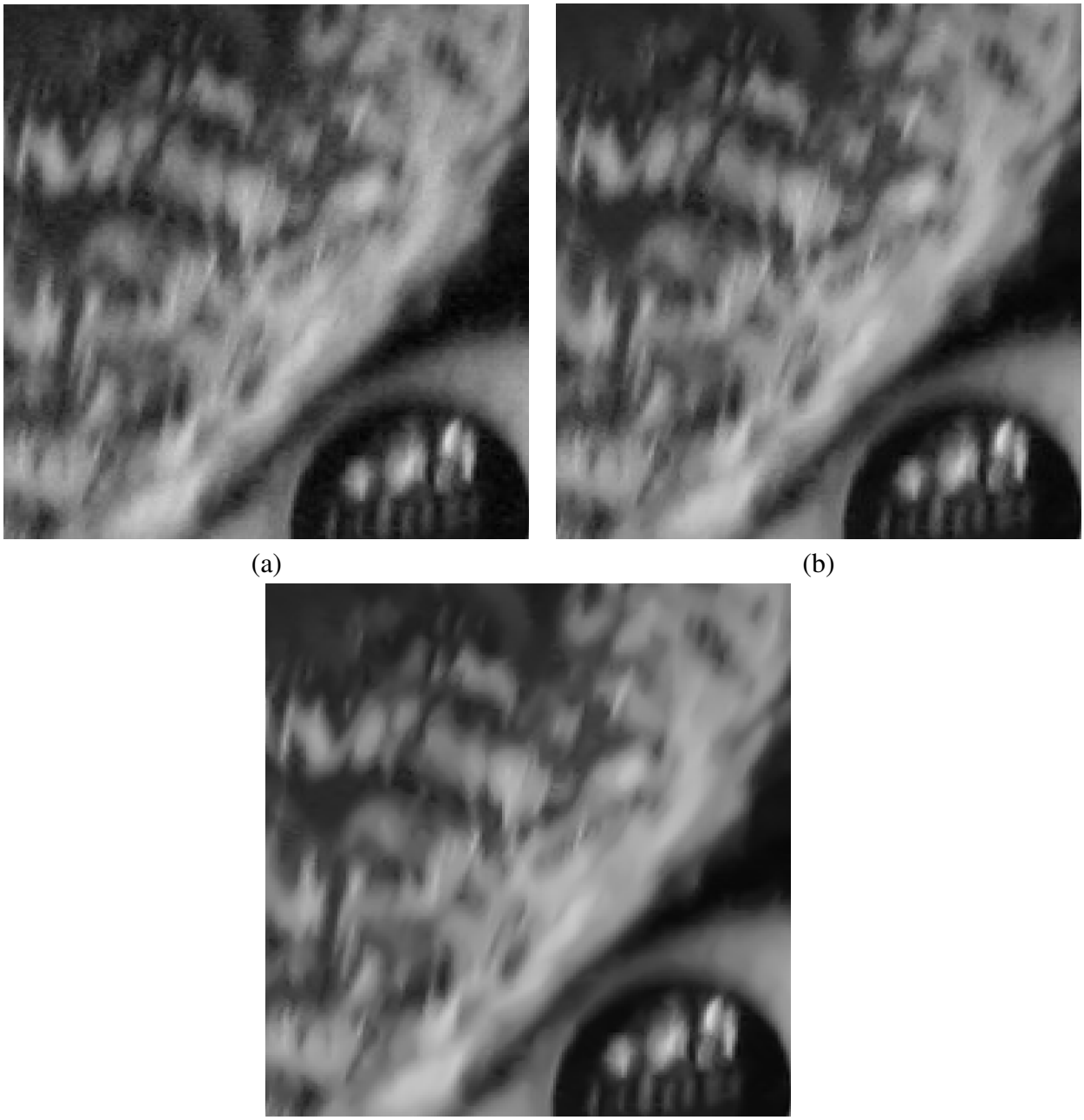
Fig. 3.8: Results of our proposed denoising method on real digital camera Noise, (a) original image (b) variable bandwidth function (low intensity ($h_{\mathbf{y}}$=2), high intensity ($h_{\mathbf{y}}$=4)) (c)MPM_{fix} denoising, (d) MPM_{var} denoising.



(a) (b)

(c) (d)

Fig. 3.9: Results of our proposed denoising method on real digital camera Noise, (a) original image (b) variable bandwidth function (low intensity ($h_{\mathbf{y}}$=2), high intensity ($h_{\mathbf{y}}$=4)), (c)MPM_{fix} denoising, (d) MPM_{var} denoising.



(a) (b)

(c)

Fig. 3.10: Results of our proposed denoising method on real digital camera Noise, (a) original image (b)MPM_{fix} denoising, (c) MPM_{var} denoising.

3.6 Conclusion

In this chapter we have mostly focused on deriving appropriate local image models towards expressing the observed density. This model, aims to express the co-dependencies of the observations at fixed geometric scale. The algorithm is composed of two steps :

- image partition to identify the smooth regions and the others containing texture or details
- denoising using an MPM (maximum posterior mode) estimator

In this chapter our main contribution was providing an automatic tool for soft classification using Gaussian Mixture model where observations are combination of simple statistical descriptor of local patches. In addition to that, a new formulation of the denoising problem using MPM estimator was considered. The posterior marginal pdf was approximated through non parametric density estimation techniques based on variable bandwidth kernels. Based, on the observation that the kernel bandwidth is playing a key role in the accuracy of the model we designed a bandwidth function that takes advantage from the classification step to better reflect the data variability. The experimental results and comparison with state of the art techniques demonstrated the efficiency of this method in the denoising. Further improvement can be gained using other priors on the image content. Eventually the use of more complex non-linear functions to determine the bandwidth of the kernels according to the soft classification probabilities is also a promising direction.

Improving the image approximation model is a critical aspect in image restoration. Another important component of the process is the geometric model used to determine the image description. The spatial bandwidth of the kernel and the photometric one are two critical components in the process. Therefore, adopting proper geometric models to guide interactions between observations will be the focus of our next chapter.

Chapter 4

Image Reconstruction Using Particle Filters

Exploiting geometric image structure is a promising direction towards better reconstruction. In this chapter, we introduce a reconstruction framework that explicitly accounts for image geometry when defining the spatial interaction between pixels in the filtering process. To this end, image structure is captured using local co-occurrence statistics and is incorporated to the enhancement algorithm in a sequential fashion using the particle filtering technique. In this context, the reconstruction process is modeled using a dynamical system where its evolution is guided by the prior density describing the image structure. Towards optimal exploration of the image geometry, an evaluation process of the state of the system is performed at each iteration. Promising results using additive and multiplicative noise models demonstrate the potentials of such an explicit modeling of the geometry.

4.1 Introduction

Let us recall that the core components of an image enhancement approach consist in: (i) the selection of the bandwidth or the scale of interaction between pixels, (ii) the selection of the weights contributing to the reconstruction. These issues were the focus of some averaging based filters that consist in performing a weighted sum of noisy pixels over a domain (filtering window). Local methods include the sigma filter [92], the bilateral filter [134]. Non local approaches refer to the Non Local mean algorithm [26]. Such methods use a fixed size as well as a fixed shape neighborhood for restoration and do not take into account the image structure in the selection of the neighborhood. For some of them, this is done implicitly through the weight definition, but adapting the window where one can search for the candidate pixel is as important as a good weight

definition.
Examples of variable size filtering windows are introduced [76, 52, 117] where the authors propose an iterative algorithm to select the most appropriate bandwidth for each pixel. Their experimental results show that adapting the bandwidth for each pixel improves the quality of restoration compared to a fixed size window.
The use of non-parametric densities with variable bandwidth being a function of the textureness of the pixel, earlier presented in this thesis, addresses mostly the selection of bandwidth relative to gray level similarity. A step further in this direction consists of defining more appropriate means for selecting the spatial scale of interaction between pixels. In the context of this chapter, we would like to study the geometric structure of the image and if possible define novel mechanisms for pixels selection that explicitly account for the observed image structure. Such a process involves two steps, (i) a learning stage where the image structure is modeled, and (ii) a reconstruction step that consists of a novel mathematical approach to encode measurements driven by the geometric image model. In this chapter, we will focus on making the best possible selection of pixels contributing to the reconstruction process through exploring the observed image geometry. Such geometry is expected to be anisotropic. More explicitly, we take into account the image structure to retrieve similar pixels to a reference one (that we want to reconstruct) without any exhaustive scan of the image domain. To this end, the reconstruction is done in a sequential fashion using notions of dynamical systems evolution. The idea is to consider multiple random walks/trajectories as different candidates for the filtering window and evaluate the pertinence of each trajectory. Such a process involves two aspects, (i) the selection of the trajectory, (ii) and the evaluation of the trajectory appropriateness. Each walk is composed of a number of possible neighboring sites/pixels in the image which are determined according to the observed image structure. Towards optimizing the selection of candidate pixels within a walk as well as the overall performance of the method, image structure at local scale is considered through a learning stage. It consists in computing a probability density function that describes the spatial relation between similar image patches in a local scale. Random perturbations according to these densities guide the ”trajectories” of a discrete number of walkers, while a weighted integration of the weighted ”average” values of these walks leads to the image reconstruction. Image similarities, geometric smoothness as well photometric consistency for the samples are the criteria used to determine the quality of a given walk. The overview of our method is presented in [Fig.(4.1)].

The remainder of the chapter is organized as follows: in section 2 we present the learning of image structure at a local scale. Random walks and particle filters are presented in section 3. The section 4 is devoted to the application of the particle filtering to image denoising for the case of additive Gaussian noise. In section 5, we present experimental results and comparisons with other

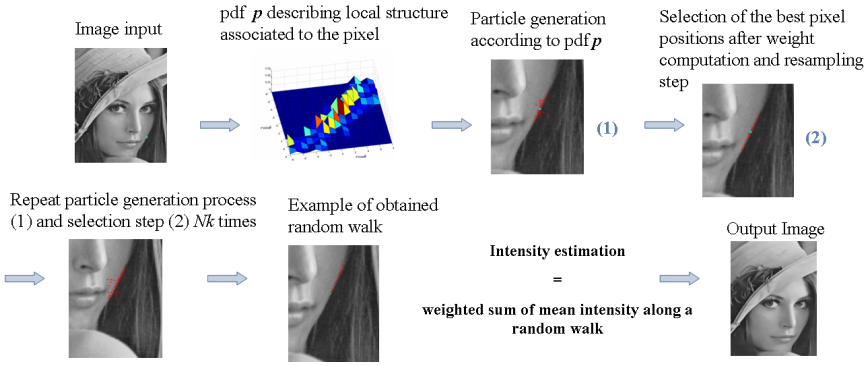


Fig. 4.1: Overview of "Random Walks" based image enhancement.

filtering methods.

4.2 Statistical Description of Image Structure

A constant effort was devoted to provide an efficient tool towards characterization of natural images content. In the context of computer vision and in particular image reconstruction, the performance of the designed solutions is highly dependent on the accuracy of such models. Image models can be either global or local. Global models focus on the statistics over the entire domain, while local ones aim to capture the local co-dependencies within the observed image structure. The most popular work in this field refers to the generalized Laplacian distribution of the wavelet coefficients distribution [102] as well as the joint statistics of pairs or triplets of pixels in the wavelet domain [91, 126, 68]. These statistical models are generic for all natural images but they are global and they do not give specific information on how the information is spatially distributed in the image. The Gray Level Co-occurrence Matrix (GLCM) [65] is a similar tool for statistical description and characterization of texture. It is defined for a displacement vector $\mathbf{d} = (\mathbf{dx}, \mathbf{dy})$ by the $G \times G$ matrix $M_{\mathbf{d}}$ (where G is the number of gray levels in the image). The coefficient $M_{\mathbf{d}}(I_1, I_2)$ is the number of occurrence of the pair of gray level I_1 and I_2 after a displacement $\mathbf{d}$. In other words, it is proportional to the joint probability of the intensity of two pixels that are separated by the displacement vector $\mathbf{d}$

$$M_{\mathbf{d}}(I_1, I_2) = |\{\mathbf{x} \in \Omega \text{ such that } I(\mathbf{x}) = I_1, I(\mathbf{x}+\mathbf{d}) = I_2\}| \tag{4.1}$$

I is the observed image, Ω the spatial domain and $|.|$ refers to the number of elements in a set. This matrix is of significant importance since implicitly it provides information on the geometric structure of the image.

As far as our denoising algorithm is concerned, we aim to find a description of the spatial rela-

tionship between similar pixels belonging to the same structure in the image. Incorporating such information will help the algorithm to cope with different image components leading to an adapted filtering window. More explicitly, given the position of an observation, our objective is to know what is the most likely pixel position in the vicinity of the observation with similar image content. This could help to define the transitions on the image lattice toward finding similar pixels within a window of a fixed radius s. Such a move is by definition anisotropic and is deduced from the observations. For this purpose, we introduce a probability function with respect to the displacement vector $\mathbf{d} \in \mathbb{R}^2$. For a given intensity I_0 it is defined as

$$p_{I_0}(\mathbf{d}) \propto |\{\mathbf{x} \in \Omega \text{ such that}, I(\mathbf{x}) = I(\mathbf{x}+\mathbf{d}) = I_0\}| \tag{4.2}$$

which corresponds to the probability of observing the intensity I_0 at the position $(\mathbf{x}+\mathbf{d})$ knowing that $I(\mathbf{x}) = I_0$. Such co-occurrences will likely be sensitive to noise and therefore, the constraint of exact matching could be relaxed, leading to:

$$p_{I_0}(\mathbf{d}) \propto |\{\mathbf{x} \in \Omega \text{ such that}, \quad I(\mathbf{x}) = I_0 \text{ and } |I(\mathbf{x}) - I(\mathbf{x}+\mathbf{d})| < \epsilon\}| \tag{4.3}$$

with ϵ a constant determined according to the observed noise level.
With this formulation, and for each gray level in the image we associate to each spatial transition $\mathbf{d}$ the probability of leading to the similar pixels. It is important to note that this probability measure relies only on gray level similarity which is not sufficient to characterize image structure especially in the presence of noise. Based on this, we will rather consider the local mean instead of the intensity at a pixel level. Furthermore, we can find in an image different structures with the same gray level. Therefore, we need an additional constraint such as the local variance which is a simple primitive capable to describe texture at local scale. Considering both local mean μ and local variance σ, would result in a more specific pdf describing spatial relation only between patches having close mean intensity and variance. The probability measure that describes the spatial relationship between similar patches knowing their local mean and variance is then

$$\begin{aligned} p_{\mu,\sigma}(\mathbf{d}) \propto |\{\mathbf{x} \in \Omega \text{ such that } m(\mathbf{x}) = \mu \quad |m(\mathbf{x}) - m(\mathbf{x}+\mathbf{d})| < \epsilon_1 \\ \text{and } std(\mathbf{x}) = \sigma \quad |std(\mathbf{x}) - std(\mathbf{x}+\mathbf{d})| < \epsilon_2\}| \end{aligned} \tag{4.4}$$

where $m(.)$ (resp $std(.)$) corresponds to the function that computes the local mean (resp local standard deviation) in a given position. ϵ_1, ϵ_2 are constants to be fixed according to the noise level. We must point out that we compute the probabilities of displacement values included in a window of radius s which corresponds to the local scale. The selection of such a parameter is important and must be adapted to the image content. Choosing a high value for s increases computation

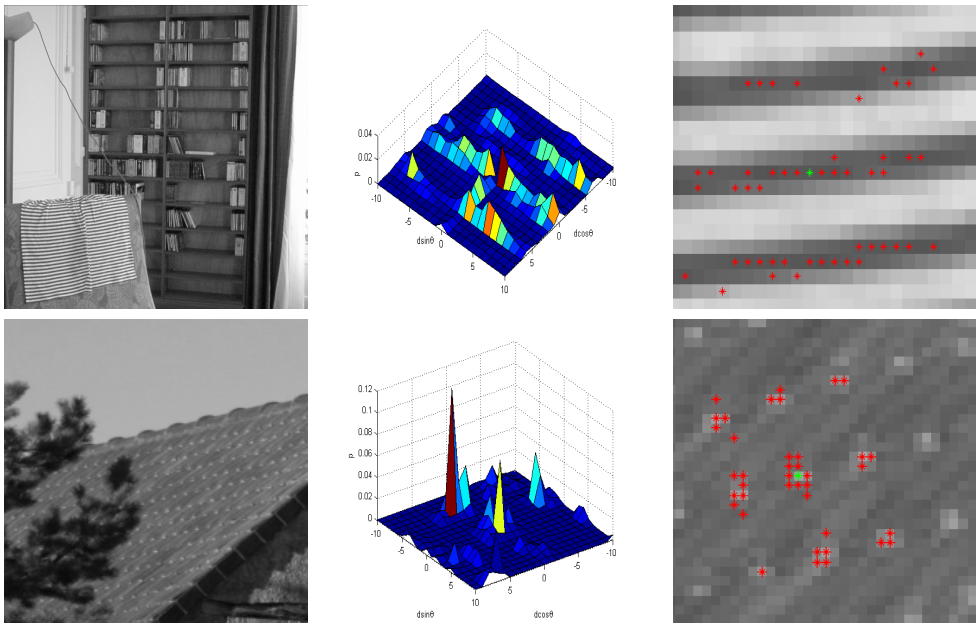


Fig. 4.2: Two pdf distributions $p_{\mu,\sigma}(\mathbf{d})$ for different values of μ and σ (top ($\mu = 39, \sigma = 11.67$), bottom ($\mu = 96, \sigma = 3.55$), and sample generation according to these pdf (red pixel) for two different positions

complexity while a small value for the scale parameter would result in a less accurate description of the structure.

The outcome of this process consists in a probability density function $p_{\mu,\sigma}(\mathbf{d})$ that aims to find a spatial representation of different structures through the computation of the relative position of similar patches. It is obvious that such a density is far from being parametric due to the randomness of the observed image geometry. In practice the estimation of the pdf is based on a non-parametric kernel density approximation [142] like Parzen windows.

Let $\Omega_{\mu\sigma} = \{\mathbf{x} \in \Omega, m(\mathbf{x}) = \mu, \sigma(\mathbf{x}) = \sigma\}$ and $\mathbf{V}(\mathbf{x}) = [m(\mathbf{x}), std(\mathbf{x})]$. Then, the probability density function estimator is given by the following expression

$$\hat{p}_{\mu,\sigma}(\mathbf{d}) = \frac{1}{|\Omega_{\mu\sigma}|} \sum_{\mathbf{x} \in \Omega_{\mu\sigma}} \mathrm{K}_{\mathbf{H}} \left(\mathbf{V}(\mathbf{x}) - \mathbf{V}(\mathbf{x} + \mathbf{d}) \right) \tag{4.5}$$

with $\mathbf{H}$ being a symmetric definite positive kernel. Gaussian kernels are the most common selection of such an approach and they were considered in our case to approximate $p_{\mu,\sigma}(\mathbf{d})$. The bandwidth $\mathbf{H}$ is a diagonal matrix with coefficients equal to $\epsilon_1 = 3$ and $\epsilon_2 = 1.5$. In terms of the mean and variance quantization values, we considered a quantification step $q = 2$. Examples of these densities are shown in [Fig. (4.2)]

Once the structural model has been constructed from the image, we are able starting from a given image position $\mathbf{x}$ characterized by the couple (μ,σ) to sample using the learned density a number

of displacement vectors towards the pixels that belong to the same structure as $\mathbf{x}$. This property is illustrated in [Fig.(4.2)] where several positions (in green) similar to the origin (in red) are drawn without scanning its entire neighborhood. The displacements with important probability are those that guarantee local statistics preservation in terms of appearance. We recall that the learning of the probability density function that describes the spatial interactions between similar pixels was restricted to a local neighborhood of size s. The generalization to larger scale will be done in a sequential fashion and this will be explained in the following section.
In the context of denoising with explicit geometric modeling of the spatial relationships between observations, our aim is to determine the most appropriate set of neighbors to estimate the noise-free intensity of a given pixel without an exhaustive scan of a large image domain. To this end, one needs to define a strategy to generate neighborhood candidate windows that takes into account the image content. In the present work, this is done through particle filters technique and this will be the focus of the following section.

4.3 Overview of Particle Filtering Technique

As stated before, our purpose is to determine the most appropriate window or neighborhood shape and size to estimate the image intensity in a given position. One attempt to do that was presented in [128] where the authors perform filtering by selecting the neighboring pixels in a random fashion but without taking image structure into account. The image lattice is represented using a graph where the weights of vertices encode the gray level difference between pixels. The transition of a walker to another is based only on these weights without considering the previous positions or local image statistics. In our algorithm, we explore the image domain using a sequential procedure taking into account the image structure as well as the previous observation and positions. To this end, we will consider the trajectory of a particle that starts from the pixel that we want to recover and moves to other neighboring ones. Each trajectory corresponds to a statistical hypothesis and will result in a reconstructed value. One should address two aspects in such a context: (i) the trajectories themselves or transitions between stages, and (ii) the "appropriateness" of each trajectory given the origin and the associated observations. Such appropriateness should depend on the consistency of the observations, the ability to encode local image structure and the relationship between the origin value and the contributing pixels. Toward optimizing this motion, we have to estimate the probability density function of visiting a position $\mathbf{x}$ at time t knowing all the past observations that refer to the intensities relative to the previously visited positions. This pdf is noted $p(\mathbf{x}_t|\mathbf{z}_{1:t})$.
In order to estimate such a pdf, we use particle filtering techniques that are an efficient sequential

Monte-Carlo methods introduced in [64]. We will give a brief overview of the use of particle filtering as a way of sequential estimation of posterior pdf and focus on the Sequential Importance Resampling technique. For an extended review the reader can refer to [6, 47, 46].

4.3.1 Preliminaries

Let us introduce first the problem of Bayesian filtering where one has to estimate the evolution of a hidden state of a system given a set of observations. Such a system involves a function that relates states with observations and a function that relates the present state with the past ones. To introduce this notion, let us note $(\mathbf{x}_{1:k})$ the state sequence and $(\mathbf{z}_{1:k})$ the sequence of measurements that verify

$$\begin{aligned} \mathbf{x}_k &= f_k(\mathbf{x}_{k-1}, m_{k-1}) \\ \mathbf{z}_k &= h_k(\mathbf{x}_k, n_k) \end{aligned}$$

where $f_k : R^{n_x} \times R^{n_v} \rightarrow R^{n_x}$ is the transition function, $h_k : R^{n_x} \times R^{n_n} \rightarrow R^{n_z}$ is the observation function while $(m_{1:k})$ and $(n_{1:k})$ are independent and identically distributed noise sequences.
Filtering is then equivalent to determine an estimation of the state vector $\mathbf{x}_k$ knowing the set of all available observations $(\mathbf{z}_{1:k})$. In the Bayesian framework, one has to compute the posterior pdf $p(\mathbf{x}_k|\mathbf{z}_{1:k})$ which can be obtained recursively in two stages: prediction and update. The main idea is to suppose that the pdf $p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})$ is available and calculate $p(\mathbf{x}_k|\mathbf{z}_{1:k})$ via the Bayes rule (4.6)

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) = \frac{p(\mathbf{z}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{z}_{1:k-1})}{p(\mathbf{z}_k|\mathbf{z}_{1:k-1})} \tag{4.6}$$

where the prior pdf is computed via the Chapman-Kolmogorov equation

$$p(\mathbf{x}_k|\mathbf{z}_{1:k-1}) = \int p(\mathbf{x}_k|\mathbf{x}_{1:k-1})p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})d\mathbf{x}_{k-1} \tag{4.7}$$

where $p(\mathbf{z}_k|\mathbf{z}_{1:k-1}) = \int p(\mathbf{z}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{z}_{1:k-1})d\mathbf{x}_k$, $p(\mathbf{z}_k|\mathbf{x}_k)$ is defined by the observation model and $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ by the transition model. One has then to evaluate the expression (4.7) in order to determine the probability of having the state $\mathbf{x}_k$ given the observations $(\mathbf{z}_{1:k})$. The recursive computation of the prior and the posterior pdf leads to the exact computation of the posterior density. Nevertheless, in practice, it is impossible to compute explicitly the posterior pdf $p(\mathbf{x}_k|\mathbf{z}_{1:k})$, and therefore an approximation method needs to be introduced.

4.3.2 Sampling Importance Resampling Filter (SIR)

The key idea is to approximate the posterior pdf by N_p random state samples $\{\mathbf{x}_k^m, m = 1..N_p\}$ with N_p associated weights $\{w_k^m, m = 1..N_p\}$ that reflect the importance of the associated sample in the pdf. Thus the posterior pdf can be approximated by a discrete weighted sum:

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) \approx \sum_{m=1}^{N_p} w_k^m \delta(\mathbf{x}_k - \mathbf{x}_k^m) \quad \text{and} \quad \sum_{m=1}^{N_p} w_k^m = 1$$

where δ is the Kronecker Delta function defined as $\delta(x) = 1$ if $x = 0$ and $\delta(x) = 0$ if $x \neq 0$.
The generation of the samples is done through the principle of Importance Sampling [20, 46]. This concept relies on the use of another density function (q) called the proposal and from which it's easier to draw samples than the posterior p. In this case and if we assume that the samples $\mathbf{x}_{0:k}^m$ are generated from the Importance density q, we have:

$$w_k^m = \frac{p(\mathbf{x}_{0:k}^m)}{q(\mathbf{x}_{0:k}^m)} \tag{4.8}$$

Returning to the posterior pdf and using the Bayes rule for each particle one can obtain

$$p(\mathbf{x}_{0:k}^m|\mathbf{z}_{1:k}^m) = \frac{p(\mathbf{z}_k^m|\mathbf{x}_k^m)p(\mathbf{x}_k^m|\mathbf{x}_{k-1}^m)}{p(\mathbf{z}_k^m|\mathbf{z}_{1:k-1}^m)} p(\mathbf{x}_{0:k-1}^m|\mathbf{z}_{0:k-1}^m) \tag{4.9}$$

By substituting (4.8) in (4.9) and in case of choosing a proposal that can be factorized as

$$q(\mathbf{x}_{0:k}^m|\mathbf{z}_{1:k}^m) = q(\mathbf{x}_k^m|\mathbf{x}_{0:k-1}^m, \mathbf{z}_{1:k}^m) q(\mathbf{x}_{0:k-1}^m|\mathbf{z}_{1:k}^m)$$

the weight update equation is

$$w_k^m \propto \frac{p(\mathbf{z}_k^m|\mathbf{x}_k^m)p(\mathbf{x}_k^m|\mathbf{x}_{k-1}^m)p(\mathbf{x}_{0:k-1}^m|\mathbf{z}_{0:k-1}^m)}{q(\mathbf{x}_k^m|\mathbf{x}_{0:k-1}^m, \mathbf{z}_{1:k}^m) q(\mathbf{x}_{0:k-1}^m|\mathbf{z}_{1:k}^m)} \tag{4.10}$$

$$= w_{k-1}^m \frac{p(\mathbf{z}_k^m|\mathbf{x}_k^m)p(\mathbf{x}_k^m|\mathbf{x}_{k-1}^m)}{q(\mathbf{x}_k^m|\mathbf{x}_{0:k-1}^m, \mathbf{z}_{1:k}^m)} \tag{4.11}$$

Now if we consider a proposal that depends only on $\mathbf{x}_{k-1}^m$ and $\mathbf{z}_k^m$ such that

$$q(\mathbf{x}_k^m|\mathbf{x}_{0:k-1}^m, \mathbf{z}_{1:k}^m) = q(\mathbf{x}_k^m|\mathbf{x}_{k-1}^m, \mathbf{z}_{1:k}^m)$$

the modified weight for the particle generated from the proposal q is then equal to

$$w_k^m = w_{k-1}^m \frac{p(\mathbf{z}_k^m|\mathbf{x}_k^m)p(\mathbf{x}_k^m|\mathbf{x}_{k-1}^m)}{q(\mathbf{x}_k^m|\mathbf{x}_{k-1}^m, \mathbf{z}_{1:k}^m)} \tag{4.12}$$

This equation shows that particle weights are updated using two main information: the observation pdf which reflects the likelihood of having an observation $\mathbf{z}_k$ knowing the state $\mathbf{x}_k$ and the transition model which controls the evolution of a particle state.
The choice of the proposal is important for the efficiency of the algorithm and we can find in the literature various Importance Sampling functions. A simple and efficient choice is proposed in the Sampling Importance Resampling (SIR) algorithm where the proposal $q(\mathbf{x}_k|\mathbf{x}_{1:k}^m, \mathbf{z}_k)$ is equal to the prior density $p(\mathbf{x}_k|\mathbf{x}_{k-1})$. In this case, equation (4.12) becomes simply

$$w_k^m \propto w_{k-1}^m p(\mathbf{z}_k|\mathbf{x}_k^m), \tag{4.13}$$

To summarize, the concept of particle filtering consists of three main steps:

- hypothesis generation according to the transition law $p(\mathbf{x}_k^m|\mathbf{x}_{k-1}^m)$
- computation of the likelihood of observations generated by this hypothesis $p(\mathbf{z}_k^m|\mathbf{x}_k^m)$
- weight updating according to $w_k^m \propto w_{k-1}^m p(\mathbf{z}_k|\mathbf{x}_k^m)$

A common problem in this estimation strategy is the samples degeneration, where many particles carry on less information and the variance in the weight increases. In such a situation, many particles have their weight close to zero and do not contribute to the approximation of $p(\mathbf{x}_k|\mathbf{z}_{1:k})$. To overcome this problem, a resampling step is necessary. The basic idea is to eliminate the particles that have limited capacity to approximate the density and focus on the ones which substantially contribute to the reconstruction of the density (important weights). The resampling refers to the generation of a new set of samples $\left\{\mathbf{y}_k^j\right\}$ that approximates the pdf $p(\mathbf{x}_k|\mathbf{z}_{1:k})$ so that $P(\mathbf{y}_k^j = \mathbf{x}_k^i) = w_k^i$ which means that the resulting samples are identically distributed and so the weights are reset to $w_k^j = \frac{1}{N_p}$. Many algorithms were proposed to perform particle selection. In the present work, we considered the systematic resampling scheme [84] which consists in eliminating particle with small weight and making several copies of the other particles according to their weight. In figure [Fig.(4.3)], an illustration of the idea of resampling as well as a detailed description of this process are presented.

Now towards the application of such an algorithm in the context of image denoising we have to define the following elements

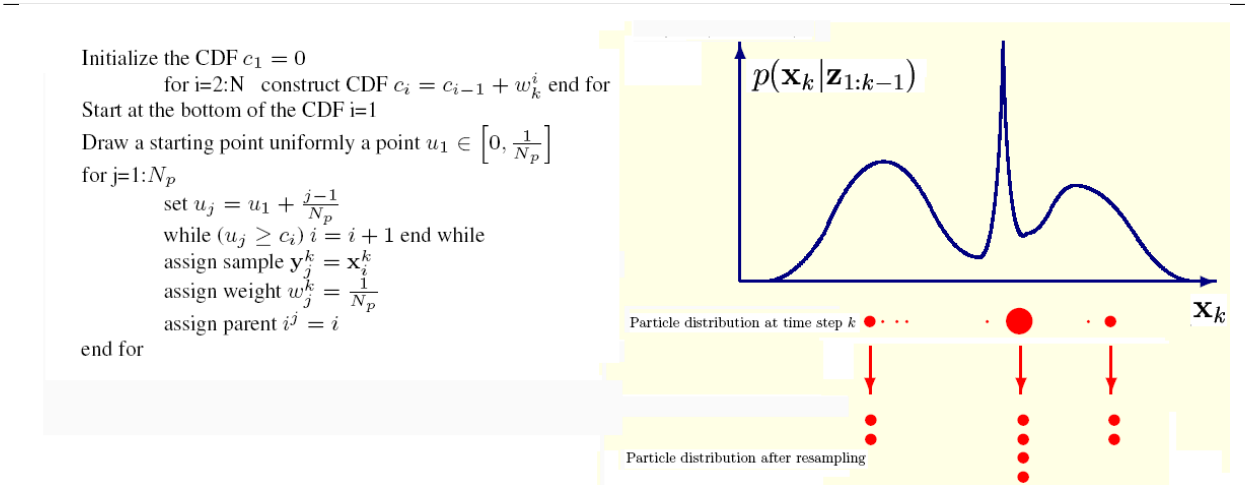


Fig. 4.3: Illustration of systematic resampling scheme

- A state vector $\mathbf{x}_k$ at a given time k that refers to the state of the reconstruction process. It corresponds to the position of the new pixel added to the filtering window, as well as the currently reconstructed value.

- An observation which corresponds to local image statistics (patch, local mean, local standard deviation) relative to a new candidate pixel to the current state vector.

- A measure that evaluates the contribution of each new candidate pixel to the estimation process.

- A systematic way of generating new samples given the existing ones able to account for the expected image geometry.

4.4 Application to image restoration

Thus, given an origin pixel $(\mathbf{x}_0)$ the process aims to recover a number of "random" positions $W(\mathbf{x}_0) = (\mathbf{x}_0, \mathbf{x}_1, ..., \mathbf{x}_T)$ with similar properties to $\mathbf{x}_0$ to reconstruct the corrupted origin value $I(\mathbf{x}_0)$. To this end we have to estimate in a sequential fashion the posterior $p(\mathbf{x}_k|\mathbf{z}_{1:k-1})$ using the previously described algorithm. One can imagine multiple particles that are moving randomly on the image domain. The path (or random walk) of a particle corresponds to a candidate filtering window. The weighted mean of the intensities along a given particle trajectory is used to restore the intensity of $\mathbf{x}_0$. As far as observations are concerned, they correspond to the local patch around the pixel $\mathbf{x}_0$ and around the candidate $\mathbf{x}_k$. To perform filtering, our approach requires the definition

of a perturbation model as well as a likelihood measure that reflects the contribution of a trajectory to the denoising process.

4.4.1 Transition Model

An important issue for the proposed algorithm consists in defining an appropriate strategy for samples perturbation. An intuitive solution is to perform Gaussian perturbations. Such a choice, suggests an isotropic way for particle transitions which can be justified in the case of homogeneous regions but not in the presence of some texture or edges.
An efficient manner to adapt the particle transition to the image content is to rely on the statistical model for image structure introduced in the section (4.2). The distribution $p_{\mu,\sigma}$ determines the transition model between particle at position $\mathbf{x}_k$ and $\mathbf{x}_{k+1}$. We recall that $p_{\mu,\sigma}(\mathbf{d})$ is a probability density function that indicates for a pixel $\mathbf{x}_k$(whose local mean and variance are equal to μ and σ) the probability that the destination position $\mathbf{x}_{k+1}$ according to the transition vector $\mathbf{d}$ has close local statistics to $\mathbf{x}_k$. This pdf is computed in the beginning of the process for each couple (mean, variance). We recall here its definition when considering local statistics of $\mathbf{x}_k$. Using $\Omega_{\mu\sigma} = \{\mathbf{x} \in \Omega, m(\mathbf{x}) = m(\mathbf{x}_k) = \mu, \sigma(\mathbf{x}) = \sigma(\mathbf{x}_k) = \sigma\}$ and $\mathbf{V}(\mathbf{x}) = [m(\mathbf{x}), std(\mathbf{x})]$, transition model is

$$\hat{p}_{\mu,\sigma}(\mathbf{d}) = \frac{1}{|\Omega_{\mu\sigma}|} \sum_{\mathbf{x}\in\Omega_{\mu\sigma}} \mathrm{K}_{\mathbf{H}} \left(\mathbf{V}(\mathbf{x}) - \mathbf{V}(\mathbf{x}+\mathbf{d})\right) \tag{4.14}$$

Thus making transitions according to this pdf guarantees that a maximum number of particles will explore sites that are similar to the origin pixel $\mathbf{x}_0$. Such a strategy makes possible the use of a reduced number of particles and could also prevent their degeneration. The essence of the proposed algorithm is to explore the image domain in a random fashion and to find the maximum number of similar sites to the origin one. A way to do that is to use an important number of particles which is equivalent to examine all the pixels situated on a circle around $\mathbf{x}_k$, nevertheless it is computationally expensive. With a reduced number of particles we have to explore the image domain with a smart fashion such that particles trajectories follow image structure without an exhaustive scan of the pixel neighborhood. In this context, the statistical structural description of the image geometry is a convenient transition model. Once such transitions have been identified, the next step consists of defining a likelihood measure for each particle based on its trajectory and the intensity observation relative to the particle position.

4.4.2 Likelihood Measure

Within the proposed approach, the random walk that is the set of the image sites contributing to the reconstruction of a given pixel is determined using a sequential approach. To evaluate the relevance of a new particle state, a likelihood measure must be introduced. This measure evaluates the added value of a new position. The first criterion that is considered is the similarity between the image content observed in the origin pixel $\mathbf{x_0}$ and the current path position $\mathbf{x}_k$. We adopt the image patches comparison, motivated by the fact that measuring similarities between image patches is an efficient and popular metric used in many computer vision tasks such as texture synthesis, restoration etc [51, 26]. In such a context we define the observation model as

- The L^2 distance between the local patch around the origin pixel $\mathbf{x}_0$ and the one corresponding to a current particle position $\mathbf{x}_k$. Thus if we call Ds_k this similarity measure, we have

$$Ds_k = \frac{\sum_{\mathbf{v}\in[-W,W]^2} \left|I(\mathbf{x}_0+\mathbf{v}) - I(\mathbf{x}_k+\mathbf{v})\right|^2}{(2W+1)^2} \tag{4.15}$$

  where W is the bandwidth which must be carefully selected to get a reliable measure of similarity while being computationally efficient.

- Another alternative to the L^2 distance is the Sobolev distance between patches defined as:

$$\begin{aligned} Ds_k &= \frac{\sum_{\mathbf{v}\in[-W,W]^2} \left|I(\mathbf{x_0}+\mathbf{v}) - I(\mathbf{x}_k+\mathbf{v})\right|^2}{(2W+1)^2} \\ &+ \frac{\sum_{\mathbf{v}\in[-W,W]^2} \left\|\nabla I(\mathbf{x}_0+\mathbf{v}) - \nabla I(\mathbf{x}_k+\mathbf{v})\right\|^2}{(2W+1)^2} \end{aligned} \tag{4.16}$$

  Such a distance accounts also for the image gradient which is an important information on image structures and their orientation. Thus similar patch have to be similar both in sense of their intensity and their gradient.

In addition to the measure of the quality of a new position in the particle walk, we consider a measure of its consistency within the whole trajectory. This measure is the intra-variability of intensities in the different positions visited by the particle during its walk. This constraint gives more importance in the filtering process to the pixel set composed of element having homogeneous properties. This is mainly interesting in the case of edges or other linear structures in the image, where walks that follow exactly the structures would be favored compared to other trajectories that fail to take into consideration this structure during their evolution. This feature can enhance

in an efficient manner image structures. Hence, in order to account for the intra-variability of the trajectories, we consider as an observation the walk variance, with respect to the origin value.

$$Dv_k = \frac{1}{k}\sum_{p=0}^{k}\left(I(\mathbf{x}_p) - I(\mathbf{x}_0)\right)^2 \tag{4.17}$$

Note that such an observation provides a measure of the "uniformity" of the trajectory and could also be determined within a larger neighborhood (not at the pixel level).
Finally, the importance of a new sample given the prior state of the walk and the observation that is the patch around $\mathbf{x_0}$ is then defined as an exponential function of the two metrics.

$$w_k = e^{-\left(\frac{Ds_k}{2\sigma_g^2} + \frac{Dv_k}{2\sigma_v^2}\right)} \tag{4.18}$$

where σ_g and σ_v are constants that determine the bandwidth of the weight computation function. The choice of these values depends on the noise level of the image. High values of σ_g and σ_v would relax the constraints on trajectories evolution and would therefore result in a loss of image details.

4.4.3 Intensity Reconstruction

After having defined the transition model as well as the observation model, we focus on the intensity reconstruction process. To this end, for each pixel $\mathbf{x_0}$ of the image, we generate N_p particles by applying N_p perturbations starting from the initial position $\mathbf{x}_0$ and according to the transition law $p_{\mu,\sigma}$ $[\mu = m(\mathbf{x}_0)\ ,\ \sigma = std(\mathbf{x}_0)]$. Iterating the transition process k time for each particle m yields a walk that is the set of successive positions of the particle $\left\{\mathbf{x}_0^m, \mathbf{x}_1^m \ldots \mathbf{x}_{k-1}^m\right\}$. We define then the walk value $\hat{I}_k^m(\mathbf{x})$ as the average weighted value along the walk. It corresponds to the value used to reconstruct the original pixel according to the walk of the particle m :

$$\hat{U}_k^m(\mathbf{x}) = \frac{1}{k}\sum_{p=0}^{k} I(\mathbf{x}_p^m) \tag{4.19}$$

As we discussed above the pertinence of a walk as well as the associated reconstructed value is computed using the likelihood measure w_k^m defined in expression (4.18). Once the weights of all the particles are determined, we normalize them so that their sum is equal to one. Next, we perform resampling in order to discard the particles with insignificant weights. These particles are not consistent with the reference pixel $\mathbf{x}_0$ and using them will introduce a bias in the reconstruction process. Finally, the reconstructed value is a weighted average of the mean intensity of each walk

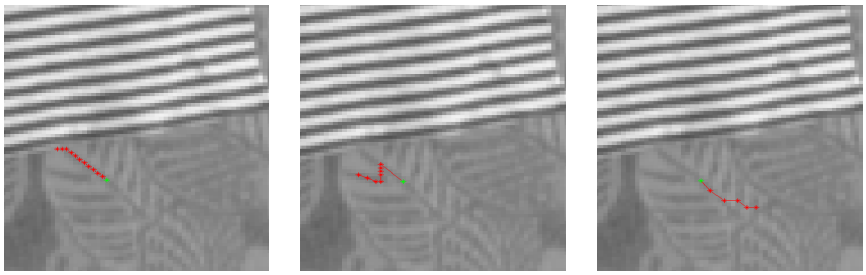


Fig. 4.4: Example of three particles walks (red pixel) starting from the same origin position (green pixel), the walks follow image structure.

for a pixel $\mathbf{x}_0$
compute local mean μ and local standard deviation σ
determine structural model $p_{\mu,\sigma}(\mathbf{d})$
for $m = 1 : N_p$ set $\mathbf{x}_0^m = \mathbf{x}_0$
for $k = 0 : T-1$
generate N_p transition samples $\mathbf{d}_k^m$ from the pdf $p_{\mu,\sigma}(d)$
set $\mathbf{x}_{k+1}^m = \mathbf{x}_k^m + \mathbf{d}_k^m$
compute an intensity for each walk according to equation(4.19)
compute the weight of each particle according to equation (4.18)
normalize the weight of the particle by setting $w_k^m = \frac{w_k^m}{\sum_{m=1}^{N_p} w_k^m}$
compute the estimated intensity $\hat{I}_k(\mathbf{x})$ using expression (4.20)
perform resampling according to the algorithm described in [Fig.(4.4)]
$\left\{\mathbf{x}_{k+1}^m\right\}_{1\leq m\leq N_p} = Resampling\left(\left\{\mathbf{x}_{k+1}^m\right\}_{1\leq m\leq N_p}, \left\{w_{k+1}^m\right\}_{1\leq m\leq N_p}\right)$
end for
The final pixel intensity is defined as $\hat{I}_{T-1}(\mathbf{x})$(4.20)
end for

Tab. 4.1: Description of the proposed denoising algorithm

and defined as

$$\hat{U}_k(\mathbf{x}) = \sum_{m=0}^{N_p} w_k^m \hat{U}_k^m(\mathbf{x}) \tag{4.20}$$

The whole process (transition, weight computation, and resampling) is repeated (T) times. An example of particle walks is shown in [Fig. (4.4)], where the pixel in green refers to the origin pixel we want to denoise and each trajectory is a candidate filtering window. In practice we use (N_p=60) particles, with (T=6) pixels contributing to each walk. To illustrate the random walks filtering an overview of the whole process is presented in [Fig.(4.1)] and the denoising algorithm is given in table (Tab.4.1).

The model described above and mainly the likelihood measure assumes an additive zero mean

noise model, but this framework can be used with different noise models. The extension of the method to cope with different noise models should involve both the statistical modeling of co-dependencies between pixels as well as the likelihood used to determine the importance of a new sample. In order to show how such an approach can be generalized, we will consider the case of multiplicative white noise.

4.4.4 Extension to Multiplicative Noise Model

In the computer vision field many images are rather corrupted by multiplicative noise called speckle. This is the case of ultrasound images or Synthetic Aperture Radar(SAR) ones. The interpretation and exploitation of such images require a pre-processing step that consists in speckle removal. In the literature we can find various types of speckle models and we will focus in this section on the case of SAR images. The classical modeling in this case is $I = U \times n$. The speckle n has a mean equal to one and follows a Gamma distribution defined as [138]

$$G_L(n(\mathbf{x})) = \frac{L^L}{\Gamma(L)} n(\mathbf{x})^{L-1} e^{-Ln(\mathbf{x})}$$

Based on this noise model we have to determine the appropriate transition function as well as the likelihood definition in case of multiplicative noise.

- The transition model is based on the structural model presented in the first section. This model can also be retained in case of multiplicative noise. If we consider two similar patches in the original image with the same mean and variance their local statistics (mean and variance) remain close to each other when corrupted by noise. To justify this choice, we will call (μ,σ^2) the mean and variance of two similar patches in the noise free image and ($\mu_n = 1$,σ_n^2)are those relative to noise. The resulting noisy patch will have local mean and local variance that can be approximated by μ and $\sigma^2\sigma_n^2 + \mu^2\sigma_n^2 + \mu_n^2\sigma^2$ respectively.

- As far as likelihood measure is concerned we cannot use the same similarity definition as the additive one. The difference between pixels intensities ranges in an interval that is proportional to the real intensity observation. That is not the case for the additive noise where the difference is only dependent on the noise model. To prove our statement, let us consider the L^2 distance between two patches centered in $\mathbf{x}_0$ and $\mathbf{x}_k$ in the noisy image I

$$d(\mathbf{x}_0, \mathbf{x}_k) = \sum_{\mathbf{v} \in [-W,W]^2} (I(\mathbf{x}_0 + \mathbf{v}) - I(\mathbf{x}_k + \mathbf{v}))^2$$

In case the observed patches are derived form the same patch in the noise free image noted U, we can write:

$$d(\mathbf{x}_0, \mathbf{x}_k) = \sum_{\mathbf{v}\in[-W,W]^2} U^2(\mathbf{x}_0 + \mathbf{v}) \left[n(\mathbf{x}_0 + \mathbf{v}) - n(\mathbf{x}_k + \mathbf{v})\right]^2$$

This shows that the L^2 distance is not relevant in the case of multiplicative noise because it is not only dependent on the noise model but also on pixels intensities. Thus, a new expression of distance between patches has to be introduced to deal with this issue:

$$d(\mathbf{x}_0, \mathbf{x}_k) = \sum_{\mathbf{v}\in[-W,W]^2} \frac{\left[I(\mathbf{x}_0 + \mathbf{v}) - I(\mathbf{x}_k + \mathbf{v})\right]^2}{U(\mathbf{x}_0 + \mathbf{v})U(\mathbf{x}_k + \mathbf{v})}$$

This similarity measure involves the noise free observations that can be approximated by a local average of the noisy observations over a neighborhood of size $(2W + 1) \times (2W + 1)$. Furthermore the distance has to be symmetric. In this context, we selected the following distance definition between two image patches around $\mathbf{x}_0$ and $\mathbf{x}_k$

$$Ds_k = \frac{\sum_{\mathbf{v}\in[-W,W]^2} \left|I(\mathbf{x}_0 + \mathbf{v}) - I(\mathbf{x}_k + \mathbf{v})\right|^2}{\left(\sum_{\mathbf{v}\in[-W,W]^2} I(\mathbf{x}_0 + \mathbf{v})\right)\left(\sum_{\mathbf{v}\in[-W,W]^2} I(\mathbf{x}_k + \mathbf{v})\right)}. \tag{4.21}$$

Following the same idea, the trajectory's variability in this case is expressed as

$$Dv_k = \frac{1}{k}\sum_{p=0}^{k} \frac{\left[I(\mathbf{x}_p) - I(\mathbf{x}_0)\right]^2}{I(\mathbf{x}_0)I(\mathbf{x}_p)} \tag{4.22}$$

Finally we keep the same choice of weight particle computation as defined in equation (4.18). Once the transition model and the likelihood measure are defined for the particular case of speckle noise, we perform denoising as described in table (Tab.4.1).

4.5 Experimental Results

In order to evaluate the performance of the proposed framework, the cases of additive and multiplicative noise were considered and compared with some of the state of the art methods with assumptions/aims similar to the ones of our approach.

4.5.1 Additive Noise

In order to evaluate the performance of the proposed algorithm, we consider the same validation procedure used in the previous chapter. This evaluation will be based on a quantitative criterion (the PSNR defined in 3.37) and a qualitative one (visual assessment). The latter quantifies the ability of the algorithm to suppress noise while preserving image details. In terms of competitive approaches we have considered filters that encode implicitly the image structure through the definition of the weights (bilateral [134] , NL-means [26]). Furthermore we have considered an approach that also performs bandwidth selection [76]. Our simulation consists in adding a synthetic white Gaussian noise of known standard deviation $\sigma_n = 20$ to the set of images frequently adopted by the image restoration community for validation. Their parameters were tuned manually to achieve optimal performance in terms of PSNR and visual quality. Regarding our approach, each walk consists of 6 steps and for each pixel we have considered 60 particles (T=6, Np = 60). Given this choice we are able to explore 6×60 positions in the image for each pixel. The radius of the explored region is of the order $T \times s$ where s is the maximum distance that a particle can achieve in a transition (in our case $s = 8$). The window size used for likelihood computation is 7×7 ($W = 3$) the parameters of the Gaussian used for likelihood computation (4.18) are set to $\sigma_g = \sigma_v = 15$. The same set of parameters was considered when the Sobolev distance between patches was used. We have to precise that in this case, we estimated the image gradient on a smoothed version of the image by a Gaussian kernel. For the bilateral filter, the parameters are $\sigma_s = 3$ and $\sigma_{ph} = 20$. The NL-means filter was used with $\sigma_{ph} = 15$, the patch size used for comparison is 7×7 and the weighted average was computed over a local neighborhood of size 11×11. The processing time for these experiments is around 2mn for 512 $\times$ 512 images using a C++ program and a Pentium IV-2GHz machine. As far as the variable bandwidth method is concerned, the results were provided by the authors and one can find more details in [76]. In table (Tab.4.2), we provided the PSNR values relative to each method. From this table we can conclude that in general, computing similarity on image patches is better than the pixel-wise comparison. The approach we presented and the NL-mean algorithm have about the same performance, while the method presented in [76] outperforms all the others which is natural since the underlying idea of this approach is to minimize a bound of the quadratic error mean. Moreover, our algorithm selects the patches for comparison randomly while being guided at the same time by the image structure. The number of pixels contributing to the walks has a tremendous impact on the performance of the method. When only a few samples have an important weight in the filtering process, it produces under smoothing in homogeneous areas and fails to recover texture. These problems are not observed when considering the variable bandwidth approach [76] since it involves all the pixels within a neighborhood in the filtering process while we can miss some relevant ones. In terms of photometric distance comparison, one

	Bilateral	NLmean	R.Walks (L^2)	R.Walks (sob)	KB06 [76]
Barbara	26.75	28.78	29.29	29.4	30.37
Boat	27.82	28.92	28.72	28.93	30.12
Fingerprint	24.12	26.45	26.4	26.96	28.16
House	29.18	30.86	31.28	31.53	32.90
Lena	29.28	31.18	30.7	31.07	32.64
Baboon	24.95	25.18	25.51	25.45	26.29

Tab. 4.2: PSNR values for the denoised images (The PSNR of the image corrupted by additive Gaussian noise of std=20 is equal to 22.15)

can observe that the Sobolev norm has a slightly better performance than the L^2 distance because it involves other features than the image intensity. The qualitative results presented in figures [Fig.(4.5),Fig.(4.6),Fig.(4.7)] show that the restoration quality of our algorithm when compared to the one using an adaptive window size as well as the NL-means algorithm is quite satisfactory. In figure [Fig.(4.8)], are presented restoration results emphasizing an area with a very fine texture scale. This figure (as well as [Fig.(4.7)-c]) shows the presence of some artifacts in the denoised image using [76] due to over smoothing. Such a result is due to the large window size selected in this region that reduced excessively the variance of the estimator. This is not the case for our algorithm, as it yields results with a more natural aspect even though the fine texture in the face is not completely recovered. In terms of qualitative results, our approach does equally well if compared to the NL-mean algorithm. Note that in case of important noise level the efficiency of the learning step of the image structures is compromised. In such a situation the transition model is not too robust with respect to noise. Considering such a model for sampling new particle positions could result in small numbers of relevant pixels which may not be sufficient for the restoration of some structures. The figure [Fig.4.9] shows another example where similar interpretation of the results can be deduced.

4.5.2 Multiplicative Noise

In order to test performance with respect to multiplicative noise, the same examples were used and we corrupted them by a synthetic speckle. The speckle signal has a Gamma distribution and two variants have been considered. Both of them have a mean equal to 1, and different standard deviation values: $\sigma_n = 0.1$ and $\sigma_n = 0.25$. We compared our approach to the one presented in [7] where a new variational formulation adapted to multiplicative noise like speckle was introduced. In this model the regularization term is inspired by the TV model [122] while the data term was modified in order to cope with the nature of the corruption. Regarding the parameters of our method, as in the case of additive noise, 60 particles per pixel were considered and their evolution

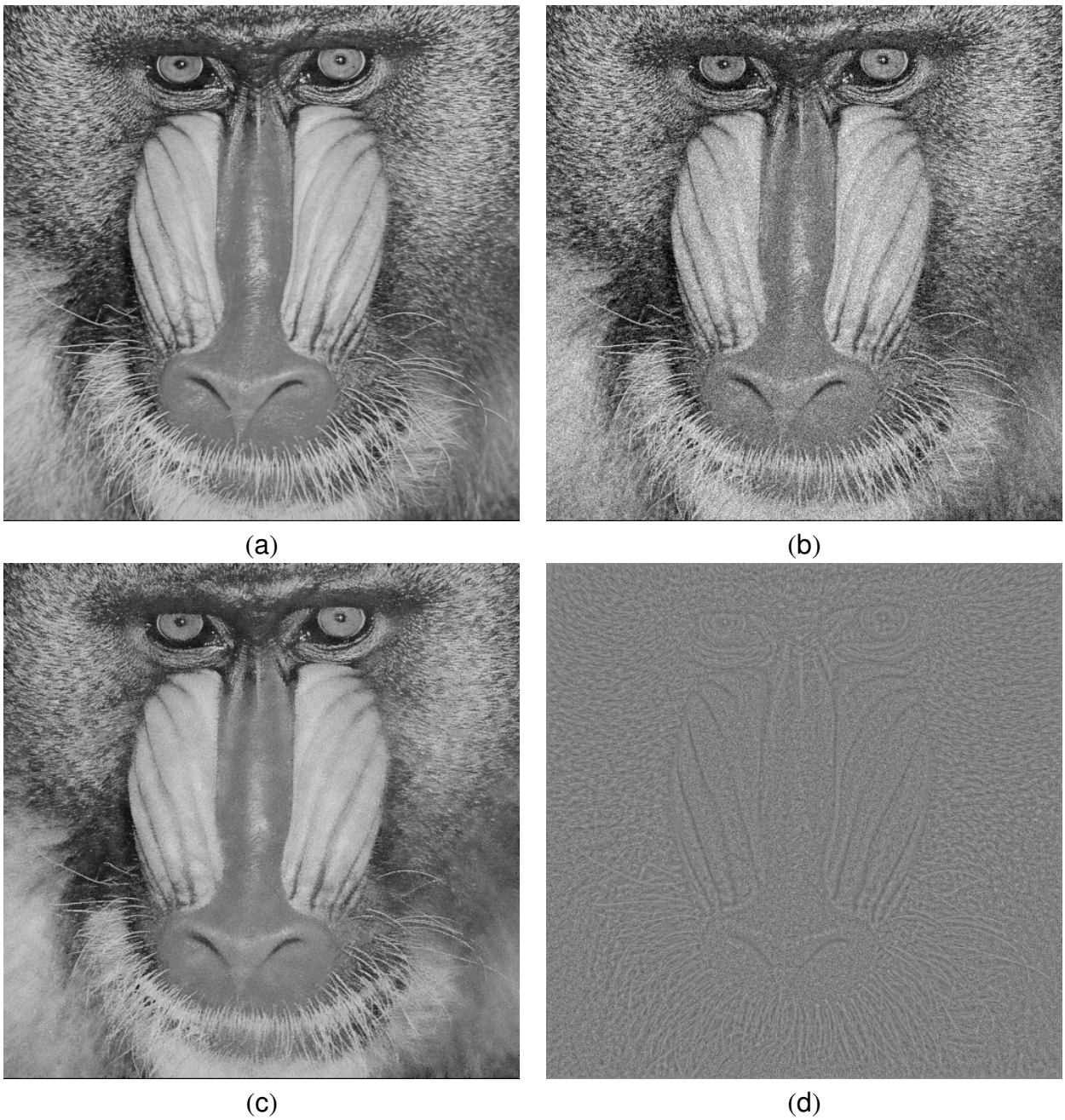


Fig. 4.5: (a) Noise free image (b)Image corrupted with Gaussian noise $\sigma_n = 20$ (c) Bilateral filter restoration (d) Residual of the bilateral filter.

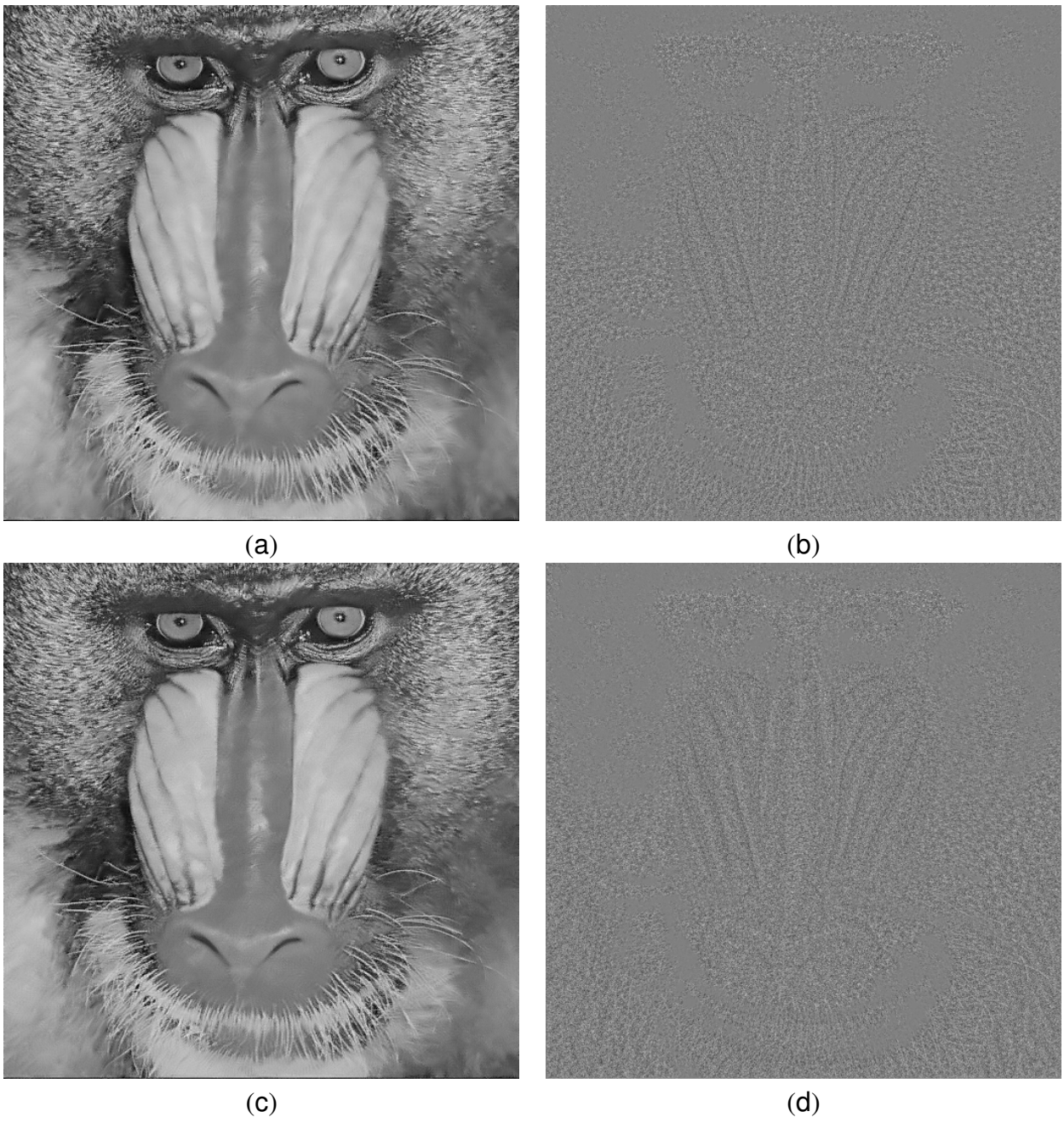


Fig. 4.6: (a) Restoration result obtained with NL-means algorithm (b) NL-mean residual (c) Restoration result obtained with random walks algorithm (L^2 distance) (d) Residual of the random walk based approach

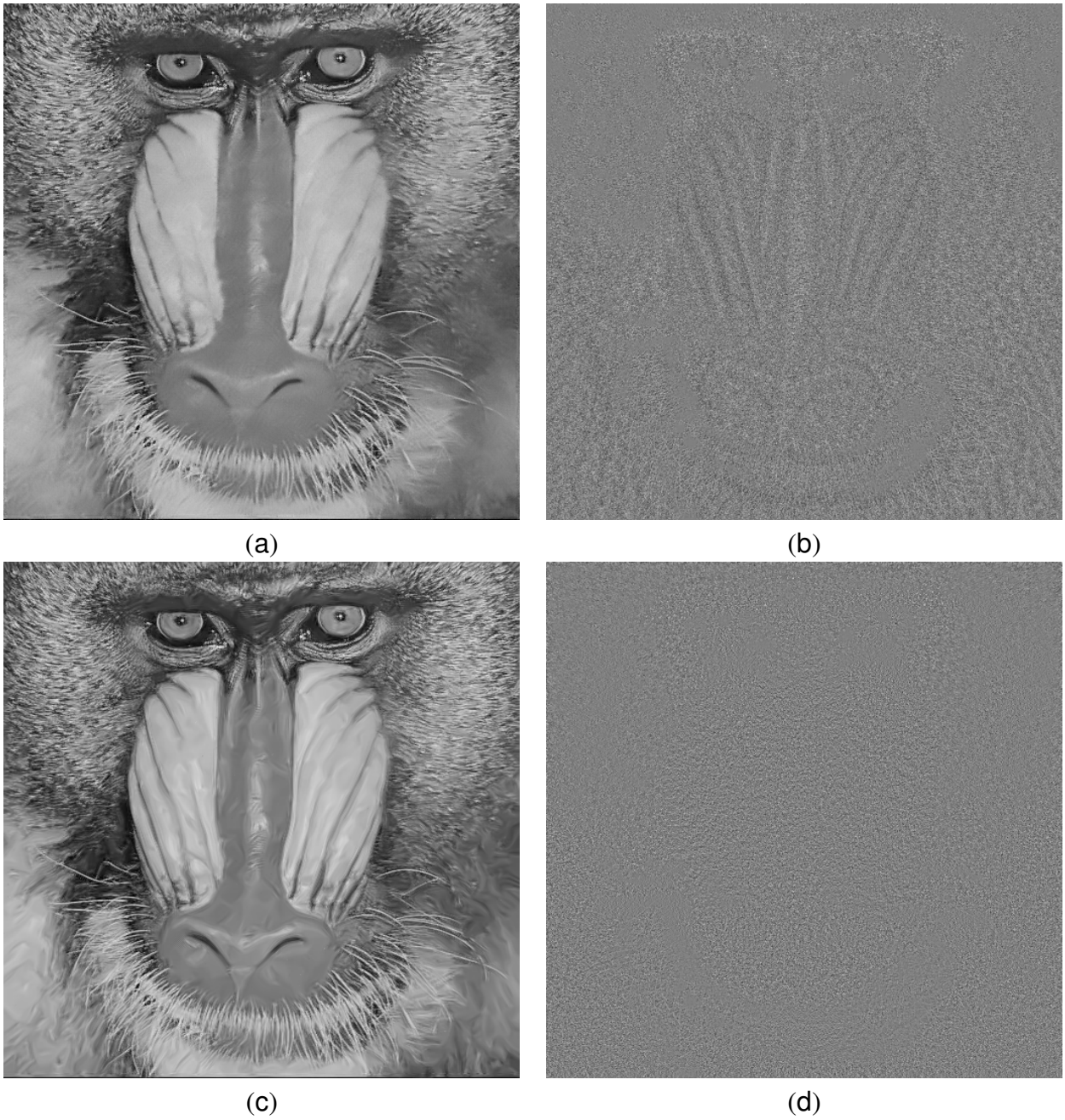


Fig. 4.7: (a) Restoration result obtained with random walks algorithm (Sobolev distance) (b) Residual of the random walk based approach (c) Restoration result obtained with KB06 [76] (d) KB06 [76] residual

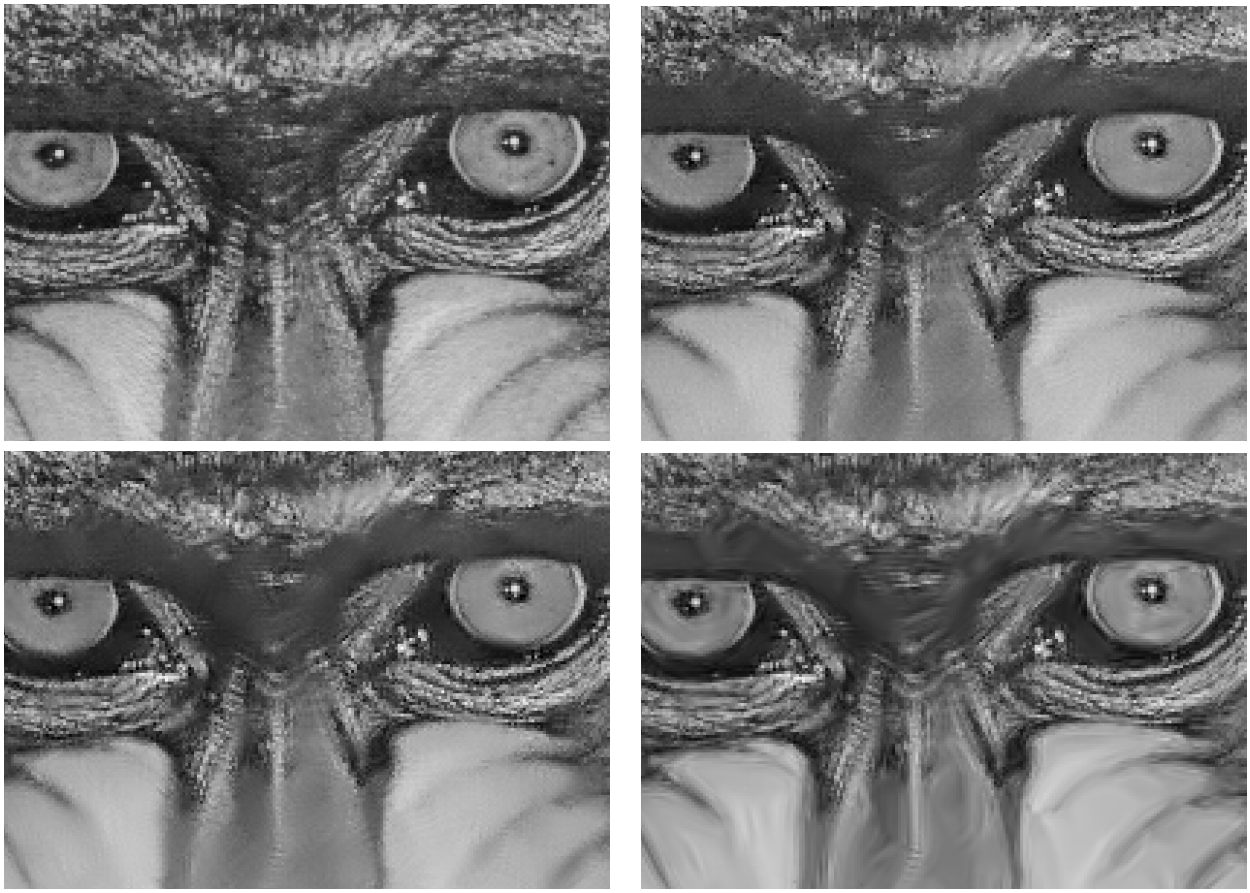


Fig. 4.8: Zoom on the Baboon image (a) Original image (b) Noisy image (c) Bilateral filter restoration (d) NL-Means restoration (e) Random walk restoration (f) Adaptive window size restoration.

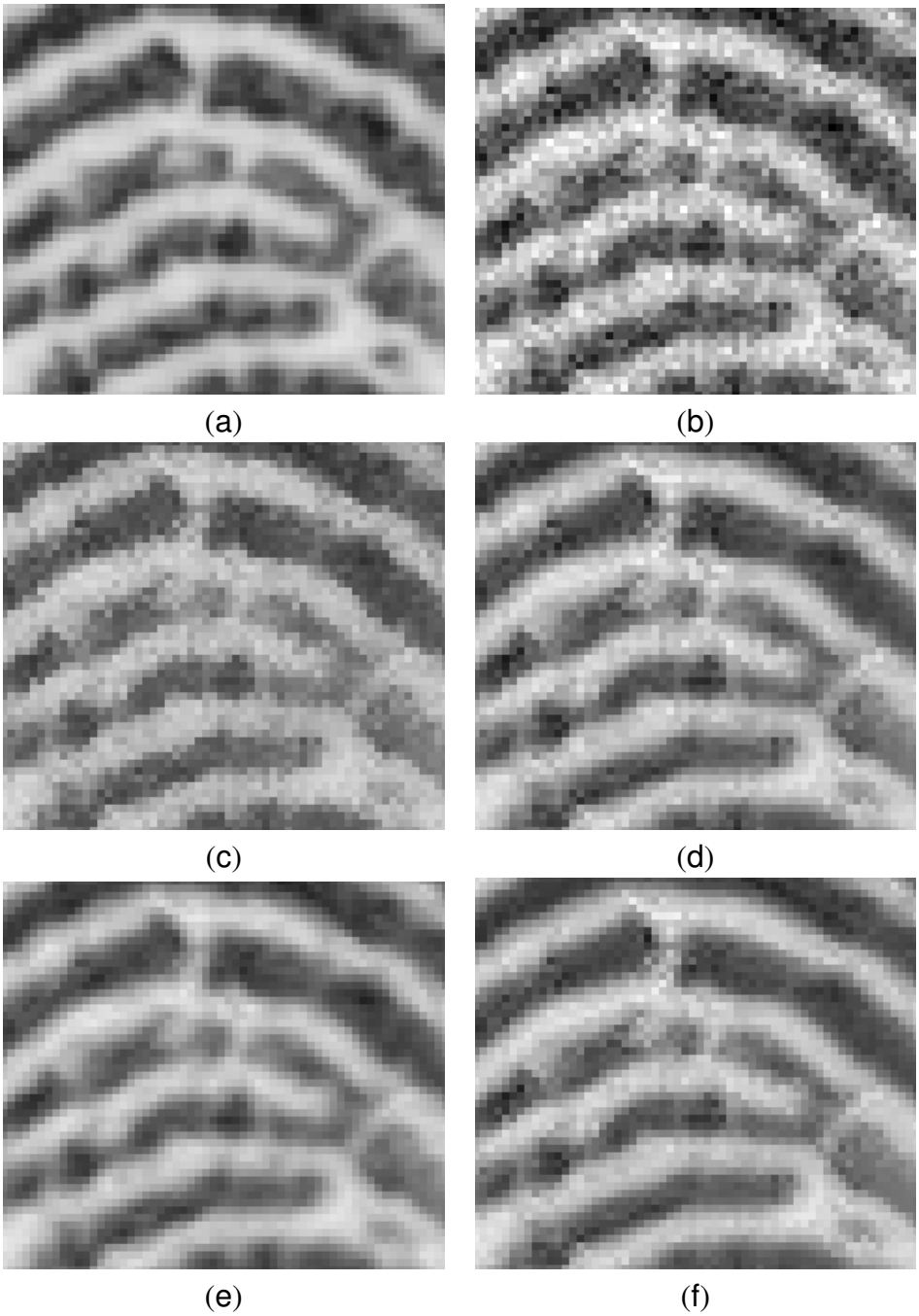


Fig. 4.9: Zoom on the fingerprint image (a) Original image (b) Noisy image (c) Bilateral filter restoration (d) NL-Means restoration (e)Adaptive window size restoration (f) Random walk restoration.

	Barbara		Boat		Fingerprint		Baboon		Lena	
σ_n	0.1	0.25	0.1	0.25	0.1	0.25	0.1	0.25	0.1	0.25
Noisy	25.93	18.51	25.37	17.73	24.76	17.58	25.43	17.86	25.66	18.15
R. Walks	31.09	25.95	30.66	25.97	28.45	22.84	27.51	22.95	33.35	27.54
TV [7]	29.27	24.31	31.06	26.73	28.22	23.72	27.85	22.90	32.75	28.69

Tab. 4.3: PSNR values for the denoising of images corrupted by multiplicative noise

involves 6 steps ($T = 6$, $Np = 60$). The weight computation was done using the following set of parameters: ($\sigma_v = \sigma_g = 0.2$) for the case of noise variance equal to $\sigma_n = 0.25$, while ($\sigma_v = \sigma_g = 0.1$) for the noise variance equal to $\sigma_n = 0.1$ which are values of the same range as the noise standard deviation. In table (Tab.4.3), we report comparative experimental results. Based on these results, we can conclude that the proposed approach can handle efficiently images corrupted by multiplicative noise regardless its model. If we consider comparison with the TV based model in terms of PSNR we can say that globally they are equivalent. Nevertheless, restoration results in [Fig.(6)] show that our method provides better image quality and is able to preserve texture and fine details. The variational approach is based on the piecewise smooth image model for the regularization term which is not very efficient in texture restoration and creates some artifacts like the spurious edges.

4.6 Conclusion

In this chapter, we introduced a new denoising technique that explicitly accounts for geometric relationships of image content towards anisotropic filtering domain. This was achieved using particle filters technique with the perturbation model being determined from the image. The structural model of the image was learned in a first step in order to guide the particle evolution toward image structure such as edges. Based on such a model, the approach acts like an isotropic filter when considering smooth areas, and anisotropic when dealing with edges. Besides, when we consider the case of regular texture with a repetitive pattern, the presented technique is able to recover similar patches in a large neighborhood using a small number of particles while visiting a limited number of positions using a learning step at local scale (section 4.2). However, in the case of random texture, generalizing at larger scale the spatial relationship between similar pixels learned at local scale is not obvious. In such a case the evolution of particles is not optimal and a bigger number of particles is needed. However, both the theoretical framework and the experimental validation suggest that the use of the image structure to determine the perturbation model and the consideration of multiple hypotheses testing can yield interesting results. We have also shown that our method

Fig. 4.10: (a) Original Lena image (b) image corrupted by speckle noise with variance $\sigma_n = 0.025$ (c) Random walk based restoration result (d) Total variation minimizing based restoration [7]

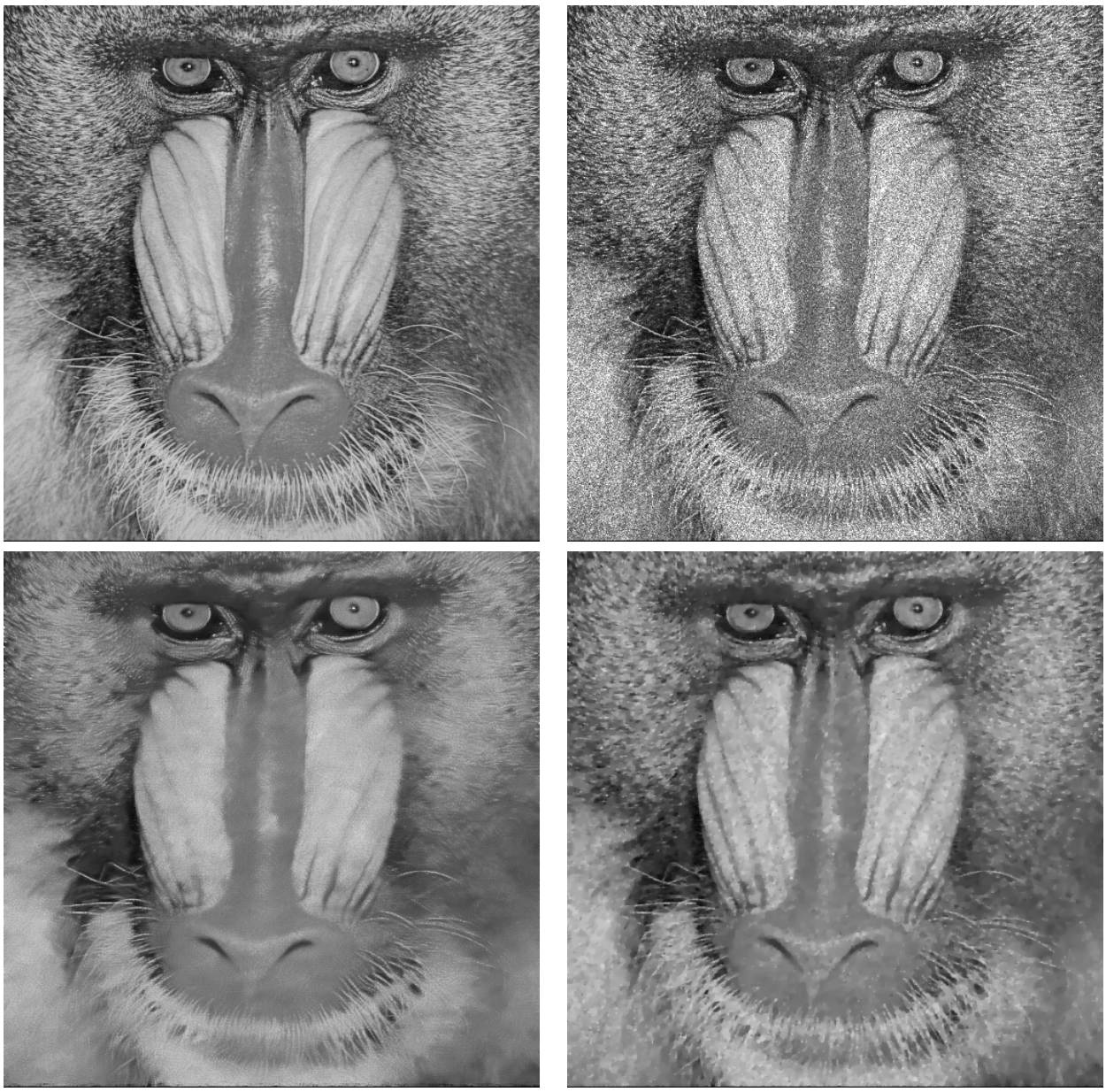


Fig. 4.11: (a) Original Baboon image (b) image corrupted by speckle noise with variance $\sigma_n = 0.025$ (c) Random walk based restoration result (d) Total variation minimizing based restoration [7]

is flexible with respect to the choice of the similarity measure between image patches. Besides, it can handle other types of noise models than the additive Gaussian one.

Improving the learning stage of the image structural model and guiding the particles to the most appropriate directions could be a step toward increasing the efficiency of particle transitions. On the other hand, an automatic technique to capture the scale of texture should be considered. It enables to determine the patch size considered when computing the likelihood of a new particle state on one hand and the size of the maximum radius of transitions on the other hand. The likelihood measure could be also modified to be more specific to each noise distribution and more robust.

The neighborhood size as well as the appropriate weight definition are issues that will be addressed in the next chapter

Chapter 5

TV Based Variable Bandwidth Image Denoising

Image denoising is often addressed through the optimization of specifically designed objective functions. Such functions consists of two terms, a regularization and a data fidelity one. The regularization component corresponds to the image model that defines the interactions between the image pixels through the selection of a bandwidth (spatial and photometric). In this chapter, inspired from the Total Variation approach we consider a regularization approach that can better express image complexity. Furthermore, we investigate new means of defining appropriate photometric distances that account to certain extent for the properties of the noise model. Last, but not least we introduce the notion of automatic spatial bandwidth selection towards better capturing the image structure. The outcome consists of a natural way of building image models based on the noisy observation. Promising quantitative and qualitative results as well as comparisons with the state of the art methods demonstrate the potentials of the proposed framework.

5.1 Introduction

Our contributions up to now consists of (i) a novel approach to determine the interaction between pixels contributing to the reconstruction using non-parametric density models. The bandwidth of these densities is defined using a soft-classification between texture and non-texture with aim to be very selective when referring to texture. The second contribution (ii) refers to a novel approach of explicit modeling for the geometric dependencies between image content towards more selective contribution of observations in the reconstruction process. It is natural to consider that an integrated approach that addresses both tasks at the same time could have extreme potentials. In other words, one should (i) be able to model either explicitly or implicitly the geometric dependencies

between observations through the definition of the most appropriate distance between observations, while (ii) having a mean of automatic selection for the filter bandwidth according to the observed image content. To this end, we have to design an appropriate cost function where a global image model that reflects the structure of the observed image is used. Such an idea can be found in the total variation minimization based approaches which are a dominant formulation in the field [122, 121]. The underlying assumption of such modals is that the image is piece-wise constant, a condition that cannot be satisfied in the context of texture images. Therefore, one has to consider an appropriate regularization functional that is more realistic in this context when compared with the piecewise smoothness assumption used in total variational approaches. The underlying image model assumes linear relations between pixels with interactions that are governed by weights reflecting photometric similarities and geometric distances. The proper definition of these weights allows the implicit modeling of geometry subject to the constraint that insignificant weights are assigned to pixels with irrelevant content. For photometric similarity, the classical L^2 distance between patches is a very popular measure, Nevertheless it is computed on the noisy image and did not take into account the actual statistical distribution of the L^2 distance. As an alternative, we will present two methods for similarity computation between patches. The first one relies on a projection of the observations into a dictionary/subspace (principal component analysis). The second is based on the statistical distribution of the L^2 distance between patches. Regarding the spatial interactions between pixels, we considered for the weight definition a variable spatial kernel bandwidth. Such a choice is motivated by the works that demonstrate the necessity of varying the neighborhood size in the filtering process. In other words, the filter domain depends on the pixel location and is also a variable of the objective function to be minimized. In this context, we propose an automatic technique to select the level of direct interaction (bandwidth) between pixels. Such bandwidth definition aims to optimally encode local relations of the image content while at the same it is constrained to be spatially smooth.

The present chapter will be organized as follows: in the second section we will briefly review some recent denoising approaches based on total variational minimization. The variable bandwidth denoising approach will be the topic of the third section. Next, we will present the new similarity measure definition. The last section will be devoted to the experimental results.

5.2 Related Work - Extensions of the TV Model

As we stated before, total variation minimization based approaches provided elegant mathematical tool toward image restoration. However, these methods have serious limitation to deal with tex-

tured images. For this reason, many variants were proposed in order to overcome the limitations of this model. Spatially varying constraints were proposed in [62] to better preserve texture. The data fidelity and its impact on better texture model were the focus of many researchers [140, 9]. Non local regularization functional were also considered as an alternative to the classical L^1 gradient norm [60, 61, 59].

5.2.1 Texture Preserving Denoising Using Spatially Varying Constraints

The idea of imposing local constraints was introduced in [121]. In this section we will present a recent approach to texture preservation that is based on spatially varying denoising parameters [62]. The basis idea of this technique is imposing local constraints to the difference between the observed image and the restored one instead of the global one (eq 2.19). The drawback of a global constraint is the difficulty to reach a compromise between textured image part and cartoon image part in terms of denoising quality. The denoising problem is formulated in the following manner:

$$U_{opt} = \text{Argmin}\left(\int_{\Omega} \Phi(|\nabla U|)d\mathbf{x}\right) \quad \text{Subject to} \quad P_R(\mathbf{x}) = S(\mathbf{x}) \tag{5.1}$$

Φ is an increasing function defined on $\mathbb{R}$. $P_R(\mathbf{x})$ is a local measure of power (energy) of the residual image $I_R = I - U$ and defined as follows

$$P_R(\mathbf{x}) = \frac{1}{\|\Omega\|}\int_{\Omega} \left[I_R(\mathbf{x}_1) - \mu(I_R)\right]^2 w_{\mathbf{x}}(\mathbf{x}_1)d\mathbf{x}_1 \tag{5.2}$$

where $w_{\mathbf{x}}(\mathbf{x}_1) = w\left(\|\mathbf{x}_1 - \mathbf{x}\|\right)$ is a normalized radial smoothing kernel. $\mu(I_R)$ is the expected value. $S(\mathbf{x})$ is a positive function that depends on the prior and have an influence on the degree of smoothness of the final image.

The problem (5.1) is solved using Lagrange multipliers leading to the following energy to be minimized

$$E = \int_{\Omega} \Phi(|\nabla U|) + \lambda(\mathbf{x})P_R(\mathbf{x})d\mathbf{x} \tag{5.3}$$

The derivative of E with respect to U is

$$\bar{\lambda}(\mathbf{x})(U - I) - \text{div}\left(\Phi'\frac{\nabla U}{|\nabla U|}\right) = 0 \;\text{ with }\; \bar{\lambda}(\mathbf{x}) = \int_{\Omega} \lambda(\mathbf{x}_1)w_{\mathbf{x}}(\mathbf{x}_1)d\mathbf{x}_1 \tag{5.4}$$

The equation is solved in an alternative manner: using a gradient descent algorithm to compute U, then the λ function is updated using the new value of U. After multiplying equation (5.4) by $(U - I)$ and taking into account the constraint of the problem, the following relations of $\lambda(\mathbf{x})$ can

be obtained

$$\int_{\Omega} [\lambda(\mathbf{x})S(\mathbf{x}) - Q(\mathbf{x})]\, d\mathbf{x} = 0 \quad Q(\mathbf{x}) = (U - I)\text{div}\left(\Phi' \frac{\nabla I}{|\nabla U|}\right) \tag{5.5}$$

A sufficient condition for $\lambda(\mathbf{x})$ to satisfy (5.5) is $\lambda(\mathbf{x}) = \frac{Q(\mathbf{x})}{S(\mathbf{x})}$. Given that textured images will be degraded with high denoising level (small $\lambda(\mathbf{x})$ values), the variance or power of the residual in textured regions will be important. An appropriate choice of the prior function should take into account this observation and assign high values of $S(\mathbf{x})$ if $\mathbf{x}$ is a pixel that belongs to the cartoon part (homogeneous regions) of the image and low $S(\mathbf{x})$ values for pixels belonging to the non cartoon part (refers to texture and details). To capture a measure about the textureness of a pixel, the authors considered a residual image I_{R_0} that is the difference between the observed image and the reconstructed one using fixed and small λ value so that enough information on the image content can be found on I_{R_0}. Given the residual image I_{R_0}, the author proposed as prior function $S(\mathbf{x}) = \frac{\sigma_n^4}{P_{R_0}(\mathbf{x})}$. With such a choice of the prior, the parameter $\lambda(\mathbf{x})$ is proportional to $P_{R_0}(\mathbf{x})$ leading to more or less smoothing according to the image content. Thus for two pixels $\mathbf{x}$ in the cartoon part of the image and $\mathbf{y}$ in a textured region they satisfy $S(\mathbf{x}) > S(\mathbf{y})$ which is equivalent to less smoothing of the texture.

This algorithm, while being able to outperform the ROF model, presents a major limitation. Considering the piecewise smooth image model as a prior could result in a loss of texture even though local constraints are applied. The considered criterion for texture and cartoon part detection that is local variance is not a robust one when handling fine structure that cannot be detected in the residual image.

A more elegant approach to image restoration based on total variation minimizing consists of separating the image into three components: cartoon part (piecewise constant), the texture and the noise. The theoretical framework will be briefly described in the following section.

5.2.2 Image Decomposition Models

Texture modeling and image decomposition into cartoon and oscillatory part gained lot of attention during the last years. The first work in this direction was introduced in [107] where a new space was proposed to model oscillating patterns toward decomposing an image into smooth regions and oscillatory component that correspond to texture (or noise). This space is noted G and is the dual space of the bounded variation functions space. The formal definition of the G space is

$$G = \left\{V \;\text{ such that }\; \exists \;\; g_1, g_2 \in L^{\infty}(\mathbb{R}^2) \;\mid\; V = \partial_x g_1 + \partial_y g_2\right\} \tag{5.6}$$

The norm associated to the space G is noted $\|.\|_*$ and is defined as

$$\|V\|_* = \inf_{\mathbf{g}=(g_1,g_2)} \left\{ \left\| \sqrt{g_1^2 + g_2^2} \right\|_{L^\infty} \mid V = \partial_x g_1 + \partial_y g_2 \right\} \tag{5.7}$$

It has been shown that oscillatory patterns have smaller $\|.\|_*$ then the classical L^2 used to measure the variations of the oscillatory signal in the classical ROF model [107, 9]. Hence using such a norm in a minimization functional process would help to capture the textured content of an image. The first contributions using the introduced norm [140] focused mainly on separating the oscillatory component from the piecewise constant one. These models are not efficient when used for denoising because the noise and the texture will belong to the same component. Such a limitation is addressed in [9] where the authors propose to decompose the image into three components $I = U + V + W$ that correspond to: the cartoon component, the texture and the noise ones. To this end three spaces were considered the Bounded Variation space for image structure , the G space for texture and a dual space to the Besov one which is more adapted to noise. The model formulation is stated as following

$$\inf_{(U,V,W)} E = \int_\Omega |\nabla U| + J^*(\frac{V}{\mu}) + B^*(\frac{W}{\delta}) + \frac{1}{2\lambda} \|I - U - V - W\|^2 \tag{5.8}$$

where $B^*(\frac{W}{\delta}) = \chi\{\|W\|_E \leq \delta\}$ and $J^*(\frac{V}{\mu}) = \chi\{\|V\|_* \leq \mu\}$. χ being the indicator function (0 if the condition is satisfied ∞ if not). The parameters μ and δ correspond to the maximum value of the norms of the extracted components. The norm $\|.\|_E$ is the norm associated to the dual space of Besov. The objective function consists of four terms (i) a regularization term on the cartoon component, (ii) a term that restrains the norm $\|\|_*$ of the texture component to be smaller then a given threshold (iii) a constraint on the norm associated to the noise component (iv) a fidelity term that aims to minimize the distance between the observation and the decomposition (cartoon+texture+noise). The minimization problem can be solved by an alternate minimization with respect to each variable

- with V and W fixed $U = \text{Argmin}_{U \in L^1(\Omega)} \left(J(u) + \frac{1}{2\lambda} \|I - U - V - W\|^2 \right)$
- with U and W fixed $V = \text{Argmin}_{V \in \mu B_G} \|I - U - V - W\|^2$ where $\mu B_G = \{\|V\|_* \leq \mu\}$
- with U and V fixed $W = \text{Argmin}_{V \in \delta B_E} \|I - U - V - W\|^2$ where $\delta B_E = \{\|W\|_E \leq \delta\}$

In the absence of an explicit expression for the E norm as well as the norm *, the author proved that this problem can be solved using orthogonal projection technique. For all implementation details and proofs the reader can refer to [9, 8]. For denoising application, the extraction of the component

W is of particular interest. The main idea behind the process is that the noise component is obtained by a soft thresholding of the wavelet coefficients of the residual image $I - U - V$ that contains a mixture of noise and details. The efficiency of the technique is then highly dependent on wavelet basis and the threshold choice. Many experimental results show the potential of this decomposition method, nevertheless its performance in terms of texture and noise separation depends on the noise model and the texture nature.

5.2.3 Non Local Functional Based Regularization

In the context of image restoration the regularization term plays a key role since it corresponds to the image model. Non local extensions of the classical TV formulation gained a lot of interest and were introduced few years ago in [83, 59, 61]. The most recent work in this direction is [59, 61] where a convex functional was studied. The corresponding regularization function is a weighted sum of the difference between pixels.

$$E_{reg}(U) = \int_{\Omega\times\Omega} \Phi\left[|U(\mathbf{x}) - U(\mathbf{y})|\right] w_{\mathbf{xy}} d\mathbf{x} d\mathbf{y} \tag{5.9}$$

Where Φ is a convex positive function defined on $\mathbb{R}$ that verifies $\Phi(0) = 0$. $w_{\mathbf{xy}}$ are positives and symmetric weights $w_{\mathbf{xy}} = w_{\mathbf{yx}}$ depending on the image features and corresponding to similarity between pixels $\mathbf{x}$ and $\mathbf{y}$. The weights definition is based on an L^2 distance between features (but various definitions of features can be considered and some examples were provided in [60]).
In the context of denoising application, one have to add the constraint $\|U - I\|^2 = |\Omega|\,\sigma_n^2$, so that the equivalent unconstrained problem is

$$E(U) = \int_{\Omega\times\Omega} \Phi\left[|U(\mathbf{x}) - U(\mathbf{y})|\right] w_{\mathbf{xy}} d\mathbf{x} d\mathbf{y} + \lambda \|U - I\|^2 \tag{5.10}$$

The corresponding Euler Lagrange descent flow and using an artificial time variable is

$$\frac{\partial U}{\partial t} = \int_{\Omega} \Phi'\left[|U(\mathbf{x}) - U(\mathbf{y})|\right] \frac{U(\mathbf{x}) - U(\mathbf{y})}{|U(\mathbf{x}) - U(\mathbf{y})|} d\mathbf{y} - \lambda(U - I) \tag{5.11}$$

like the basic TV formulation the parameter λ can be viewed as fixed parameter and thus the optimal solution does not obey necessarily the constraint. It can be viewed as the Lagrangian multiplier in this case the optimal λ have to be estimated as in [122, 74, 59].
We will give in this section an overview of the particular case where Φ is a quadratic function ($\Phi(s) = s^2$) studied in [60]. For this model, the regularization flow amounts to applying a linear

operator L_U defined as

$$L_U = \frac{\partial U}{\partial t} = \int_\Omega [U(\mathbf{x}) - U(\mathbf{y})]\, w_{\mathbf{xy}} d\mathbf{y} \tag{5.12}$$

This equation can be seen as a diffusion equation on weighted graph, where the nodes are the image pixel and the vertices are weighted by $w_{\mathbf{xy}}$ [60, 132]. The operator L_U satisfies also basic properties that show its effectiveness toward noise suppression. If we note U^t the image obtained at time t starting from the observed image $I = U^0$, the following properties hold

- mean preserving $\int_\Omega U^t(\mathbf{x})d\mathbf{x} = \int_\Omega I(\mathbf{x})d\mathbf{x}$
- the extremum principle holds $min_{\mathbf{x}}(I(\mathbf{x})) \leq U^t(\mathbf{x}) \leq max_{\mathbf{x}}(I(\mathbf{x}))\ \ \forall x \in \Omega\ \ \forall t \geq 0$
- Under the constraint that zero is an eigenvalue of multiplicity one of the linear operator L_U, $U(\mathbf{x}, t \to \infty) = const = \int_\Omega I(\mathbf{x})d\mathbf{x}$
- The L^2 norm of $U^t(\mathbf{x})$ is decreasing with time.

Numerical scheme and experimental results was provided in [60] using the same weights definition as the NL-means algorithm. The authors show that their quadratic non local regularization functional is an efficient tool for image restoration when compared to the ROF model and the NL-means algorithm. This is not surprising because the TV based flow is local and relies on the assumption of spatial piecewise smoothness which is not necessarily verified for textured image. In the following we will focus on the analysis of another quadratic functional toward regularization that relies on a different assumption about the image.

5.3 Variable Bandwidth Denoising Using Semi Local Quadratic Functional

5.3.1 Analysis of the Denoising Model

The previously presented methods consider the same image prior which is piecewise smoothness. The classical TV formulation prior suggests that the image is piecewise constant at local scale. Opposite to that, the non local models tend to minimize the intensity difference between similar pixels which assumes certain constancy for similar pixels regardless their position. In the presented approach we will consider a different prior that suggests that the image verifies the mean property (a property verified by harmonic functions). This model assumes that a pixel can be approximated by a weighted average of other image pixels. Therefore the regularization term based on this prior

is defined as:

$$E_{reg}(U) = \int_{\Omega} \left(\left[\frac{1}{Z(\mathbf{x})} \int_{\Omega} w_{\mathbf{xy}} U(\mathbf{y}) d\mathbf{y} \right] - U(\mathbf{x}) \right)^2 d\mathbf{x} \tag{5.13}$$

where, $w_{\mathbf{xy}}$ are the weighted mean coefficients and $Z(\mathbf{x}) = \int_{\Omega} w(\mathbf{x}, \mathbf{y}) d\mathbf{y}$ is a normalization coefficient.

The definition of the weights $w_{\mathbf{xy}}$ is crucial for the model because it refers to the interactions between pixels. A possible choice could be the linear model where the weights are a Gaussian function of the spatial distance between pixels. Such a definition does not take into account the image content and could result in a loss of image details because photometric consistency between pixels is not considered. For this reason we consider a non linear weighted average, that encodes both similarity between pixels and the spatial distance between them. A possible choice for the similarity measure is the L^2 distance between image patches. Thus the weight definition is

$$w_{\mathbf{xy}} = \exp\left(-\frac{\left\| U_{\mathbf{N}_x} - U_{\mathbf{N}_y} \right\|^2}{2\sigma_{ph}^2} \right) \exp\left(-\frac{\|\mathbf{x} - \mathbf{y}\|^2}{2\sigma_s^2} \right) \tag{5.14}$$

The parameter σ_{ph} and σ_s are variables that determine the bandwidth of the range (photometric) kernel and the spatial one. The spatial bandwidth determines the radius of interaction between pixels and the selection of this parameter will be the focus of the section (5.3.2) Such an image model is more accurate then the piecewise constant one, because first it relies on a prior where interaction between pixel goes beyond a local neighborhood traditionally considered in the basic TV formulation. Furthermore, the image model that we considered is more general and could be a further step toward modeling image complexity.

This regularization term can be jointly considered with a data-fidelity term that aims to preserve the distance between the observed and the reconstructed image. To this end we consider the classical L^2 distance between the noisy observation and the restored one. The denoising is based on the minimization of the following cost function

$$E(U) = \int_{\Omega} \left(\left[\frac{1}{Z(\mathbf{x})} \int_{\Omega} w_{\mathbf{xy}} U(\mathbf{y}) d\mathbf{y} \right] - U(\mathbf{x}) \right)^2 d\mathbf{x} + \lambda \int_{\Omega} \left(I(\mathbf{x}) - U(\mathbf{x}) \right)^2 d\mathbf{x} \tag{5.15}$$

λ is a coefficient that controls the trade off between the fidelity to data component and the regularization one.

The dependence of $w_{\mathbf{xy}}$ on U makes the minimization of the function (5.15) intractable because it is a non convex function. To address this difficulty we consider an alternative weights definition.. Assuming that the structure of the noise-free image is preserved in the noisy version we can

compute the weights on the noisy image so that

$$w_{\mathbf{xy}} = \exp\left(-\frac{\left\|I_{\mathbf{N}_x} - I_{\mathbf{N}_y}\right\|^2}{2\sigma_{ph}^2}\right)\exp\left(-\frac{\|\mathbf{x}-\mathbf{y}\|^2}{2\sigma_s^2}\right) \tag{5.16}$$

With such a definition, we obtain a quadratic and thus convex cost function. This regularization expression was also considered in [28] in the context of image deblurring but without a spatial distance constraint. This method combines the reverse heat equation with an additional term to preserve more details. In this dissertation we present further analysis of the model and we will propose a different numerical scheme then the one used in [28]. Besides we provide an automatic technique to select the spatial bandwidth.

This cost function admits a unique minimum that could be obtained using a gradient descent algorithm starting from the noisy observation $U^0 = I$.

$$U^{t+1} = U^t - dt\nabla E \tag{5.17}$$

where

$$\begin{aligned}\nabla E_{\mathbf{x}} &= 2\int_\Omega\left(\int_\Omega \frac{w_{\mathbf{yz}}}{Z(\mathbf{z})}U(\mathbf{y})d\mathbf{y} - U(\mathbf{z})\right)\frac{w_{\mathbf{xz}}}{Z(\mathbf{z})}d\mathbf{z} \\ &+ 2\left(U(\mathbf{x}) - \int_\Omega \frac{w_{\mathbf{xy}}}{Z(\mathbf{x})}U(\mathbf{y})d\mathbf{y}\right) + 2\lambda(U(\mathbf{x}) - I(\mathbf{x}))\end{aligned} \tag{5.18}$$

The derivative expression shows the fundamental difference with the model presented in [60]. To fix the idea, we will call residual the term $R_{\mathbf{x}} = \frac{1}{Z(\mathbf{x})}\left[\int_\Omega w_{\mathbf{xy}}U(\mathbf{y})d\mathbf{y}\right] - U(\mathbf{x})$. Contrarily to the model in [60], the regularization flow goes beyond the residual information at given position to include neighboring residual values and this enforces the coherence of neighboring pixels. Now if we consider the case where $\sigma_{ph} = \infty$, thus the residual is consistent with the Laplacian operator, and the regularization flow can be written as:

$$\frac{\partial U}{\partial t} = \Delta U - \mathcal{G} * \Delta U \tag{5.19}$$

$\mathcal{G}$ is a Gaussian kernel with standard deviation σ_s. This expression shows that the regularization flow is a combination of two processes: (i) regularization based on the first term that corresponds to the heat equation) (ii) sharpening based on the second term which is the filtered Laplacian value. Now we will present a matrix formulation of the problem where the images U and I are presented by an array of size $|\Omega| \times 1$. We will introduce the weight matrix Σ of size $|\Omega| \times |\Omega|$ that is defined

as

$$\Sigma(i,j) = \frac{w_{\mathbf{x}_i\mathbf{x}_j}}{\sum_j w_{\mathbf{x}_i\mathbf{x}_j}} \quad \Sigma(i,i) = -1$$

The restoration problem is equivalent to

$$E(U) = \|\Sigma U\|^2 + \lambda \|U - I\|^2 = U^T\Sigma^T\Sigma U + \lambda(U^T - I^T)(U - I) \tag{5.20}$$

$$\nabla E = \Sigma^T\Sigma U + \lambda(U - I) = \left(\Sigma^T\Sigma + \lambda\mathbf{I_d}\right) U + \lambda I \tag{5.21}$$

$\Sigma^T\Sigma$ is a symmetric positive matrix thus the matrix $\Sigma^T\Sigma + \lambda\mathbf{I_d}$ can be inverted. An explicit expression of the optimal solution of the problem (5.15) given a fixed λ is

$$U^{\lambda}_{opt} = \left(\Sigma^T\Sigma + \lambda\mathbf{I_d}\right)^{-1} \lambda I \tag{5.22}$$

Furthermore, by introducing an artificial time variable we can obtain the following regularization flow

$$\frac{\partial U}{\partial t} = -\left(\Sigma^T\Sigma\right) U \tag{5.23}$$

Starting from the observed image $U^0 = I$, the associated numerical scheme is

$$U^{t+1} = U^t - dt\Sigma^T\Sigma U^t = \left(\mathbf{I_d} - dt\Sigma^T\Sigma\right) U^t = \left(\mathbf{I_d} - dt\Sigma^T\Sigma\right)^t I \tag{5.24}$$

which satisfies certain desirable properties:

- The mean of the image U is constant where the variance is decreasing with time:
  An estimate of the mean of the image m_U is

  $$m_U = \frac{1}{|\Omega|}\langle C, U\rangle \quad \text{with} \quad C = (1, 1, \ldots 1)^T$$

  Thus

  $$\frac{\partial m_U}{\partial t} = \frac{1}{|\Omega|}\left\langle C, \frac{\partial U}{\partial t}\right\rangle = \frac{1}{|\Omega|}\left\langle C, -\Sigma^T\Sigma U\right\rangle = -\frac{1}{|\Omega|}C^T\Sigma^T\Sigma U = 0$$

  $C^T\Sigma^T = 0$ because the sum of the columns of Σ are equal to zero.

  $$\frac{\partial \|U\|^2}{\partial t} = \frac{\partial}{\partial t}\langle U, U\rangle = \left\langle U, \frac{\partial U}{\partial t}\right\rangle = \left\langle U, -\Sigma^T\Sigma U\right\rangle \leq 0 \tag{5.25}$$

  The variance of the image is defined as $V_U = \|U\|^2 - m_U^2$, and therefore with the above minimization schema, we can conclude that the variance of the image is decreasing with time.

- For the convergence of such a scheme the time step have to be chosen such that the eigenvalues of the matrix $\left(\mathbf{I_d} - dt\Sigma^T\Sigma\right)$ have an absolute value smaller than one. If we note $(\mu_i)_{1\leq i\leq|\Omega|}$ the eigenvalues, then such condition is satisfied when $dt \leq \frac{2}{max(\mu_i)}$.

- The schema converges to the constant image if Σ is not decomposable into blocks.
  The constant vector is an eigenvector associated to the zero eigenvalue of the matrix Σ. The matrix $\Sigma^t\Sigma$ have the same eigenvectors set associated to zero eigenvalue as Σ. To prove this, let V such that $\Sigma^t\Sigma V = 0$ then $\langle\Sigma^t\Sigma V, V\rangle = 0 = \langle\Sigma V, \Sigma V\rangle = \|\Sigma V\|$ thus V is also an eigenvalue of Σ. Hence, the constant eigenvector associated to eigenvalue zero of the matrix $\Sigma^t\Sigma$ is unique if the matrix Σ is not a block decomposable which mean that there exists a path between two pixels $\mathbf{x}$ and $\mathbf{y}$ (Perron-Frobenius theorem).

- Relation between the regularization flow and the exact solution given in equation (5.22).
  To this end, we will consider the projection of the image I on the orthogonal basis $(V_i)_{1\leq i\leq|\Omega|}$ composed of the eigenvectors of the matrix $\Sigma^T\Sigma$.
  The optimal solution (5.22) can be written as

$$U^{\lambda}_{opt} = \sum_{i=1}^{|\Omega|} \frac{\lambda}{\lambda+\mu_i}\alpha_i V_i \;\text{ with }\; I = \sum_{i=1}^{|\Omega|}\alpha_i V_i \tag{5.26}$$

  While considering only the regularization flow we have

$$U^t = \sum_{i=1}^{|\Omega|}(1-dt\mu_i)^t\alpha_i V_i \approx \sum_{i=1}^{|\Omega|}(1-tdt\mu_i)\alpha_i V_i \quad \text{(if } dt \text{ small)} \tag{5.27}$$

  Therefore, when $t_0 \approx \frac{1}{dt(\lambda+\mu_i)}$ an approximation of the exact solution U^{λ}_{opt} can be provided by U^{t_0}

5.3.2 Bandwidth Computation

Our model involves weights that reflect the similarity between "neighborhood" pixels and the interaction between them. The selection of the spatial bandwidth is crucial in the regularization process. In [76], it has been proved using a statistical estimation technique that the choice of the convolution kernel bandwidth is important for the accuracy of the intensity estimation in case of NL-means algorithm [26]. The selection of the neighborhood size during the filtering process has been also studies in a number of papers [52, 117]. In general, increasing the kernel size reduces the variance of the estimator while increasing its bias.

In the context of our estimation framework, an expression of the bias and variance of the estimator is not straightforward. Besides, the notion of image-based variable bandwidth is very interesting and promising idea towards improving the performance of the method. An intuitive motivation for that is the following: for a fixed value of σ_{ph}, large spatial kernels will result in an integration domain that involves an extended neighborhood a nice property when considering smooth areas. On the other hand, the same bandwidth could lead to texture over smoothing. Therefore, kernel bandwidth needs to be adapted to each pixel in the image in order to obtain an optimal balance between noise suppression and details preserving. The most natural approach to account for the above mentioned requirements is to introduce the bandwidth selection in the optimization process. In other words, we would like to replace the constant σ_s value by a pixel dependent one $\sigma_s(\mathbf{x})$, and minimizing the cost function (5.15) with respect to the spatial bandwidth. Besides, under the assumption that image content at the very local scale is coherent, one can also suppose that the spatial bandwidth function is smooth. For this reason we consider an additional smoothness term (the L_2 norm of the gradient magnitude of $\sigma_s(\mathbf{x})$) during the estimation of the bandwidth $\sigma_s(\mathbf{x})$. The objective function to be minimized with respect to $\sigma_s(\mathbf{x})$ is

$$\begin{aligned} E(U,\sigma_s) &= \int_\Omega \left(\left[\frac{1}{Z(\mathbf{x})} \int_\Omega w_{\mathbf{xy}}(\sigma_s(\mathbf{x}))U(\mathbf{y})d\mathbf{y} \right] - U(\mathbf{x}) \right)^2 d\mathbf{x} \\ &+ \lambda \int_\Omega \left(I(\mathbf{x}) - U(\mathbf{x})\right)^2 d\mathbf{x} + \mu \int_\Omega \left\|\nabla\sigma_s(\mathbf{x})\right\|^2 d\mathbf{x} \end{aligned}$$

with μ is a parameter that controls the trade-off between the smoothness of the bandwidth and the constraint that have to be verified. The minimum of this cost function is computed using a gradient descent algorithm. The gradient of the energy E with respect to σ_s is

$$\nabla E|_{\sigma_s(\mathbf{x})} = \left(\frac{1}{Z(\mathbf{x})} \int_\Omega w(\mathbf{x},\mathbf{y})I(\mathbf{y})d\mathbf{y} - I(\mathbf{x}) \right) \mathcal{G}(\mathbf{x}) + \mu\Delta\sigma_s \tag{5.28}$$

with $\mathcal{G}(\mathbf{x})$ being,

$$\begin{aligned} \mathcal{G}(\mathbf{x}) &= \frac{\partial}{\partial\sigma_s} \left[\frac{1}{Z(\mathbf{x})} \int_\Omega w(\mathbf{x},\mathbf{y},\sigma_s)I(\mathbf{y})d\mathbf{y} - I(\mathbf{x}) \right] \\ &= \frac{\left[\int_\Omega \left\|\mathbf{x}-\mathbf{y}\right\|^2 w_{\mathbf{xy}}(\sigma_s(\mathbf{x}))U(\mathbf{y})d\mathbf{y}\right]\left[\int_\Omega w_{\mathbf{xy}}(\sigma_s(\mathbf{x}))d\mathbf{y}\right]}{\sigma_s^3\left[\int_\Omega w_{\mathbf{xy}}(\sigma_s(\mathbf{x}))d\mathbf{y}\right]^2} \\ &- \frac{\left[\int_\Omega w(\mathbf{x},\mathbf{y})U(\mathbf{y})d\mathbf{y}\right]\left[\int_\Omega \left\|\mathbf{x}-\mathbf{y}\right\|^2 w_{\mathbf{xy}}(\sigma_s(\mathbf{x}))d\mathbf{y}\right]}{\sigma_s^3\left[\int_\Omega w_{\mathbf{xy}}(\sigma_s(\mathbf{x}))d\mathbf{y}\right]^2} \end{aligned} \tag{5.29}$$

(a) (b)

Fig. 5.1: (a) Original image (b) The bandwidth value associated to it

Starting from a fixed bandwidth value $\sigma_s = 3\mathbf{I_d}$, the bandwidth function is updated according to

$$\sigma_s^{t+1}(\mathbf{x}) = \sigma_s^t(\mathbf{x}) - dt\nabla E|_{\sigma_s(\mathbf{x})} \tag{5.30}$$

The function E is non convex with respect to σ, the algorithm will then converge to a local minimum. When σ is close to infinity the energy function is still well defined and has a finite value. Contrarily to that when $\sigma_s(\mathbf{x})$ is close to zero, $Z(\mathbf{x}) \to 0$. Thus it is appropriate to impose a minimum value to $\sigma_s(\mathbf{x})$. This is done by stopping $\sigma_s(\mathbf{x})$ evolution when it becomes smaller than a critique value $\sigma_{inf} = 1$.
Image denoising and bandwidth selection are addressed in an iterative fashion through an alternation of the minimization processes with respect to U then with respect to σ_s. However, in practice and due to computational complexity we perform the bandwidth optimization once on a smoothed version of the image. An example of bandwidth selection is shown in [Fig. (5.1)] where certain qualitative interpretation regarding the behavior of the process can be extracted. Scale selection in edges and texture regions refer to lower values than smooth regions which is a natural outcome. Given such a general framework, one has to still address the task of distance selection towards capturing implicitly the image structure.

5.4 On the similarity measure between patches

The definition of the photometric similarity between pixels is also important. In the following section, we provide some ideas in order to improve this measure using other image feature or a new statistical measure. The distance between image patches is a very common choice towards defining photometric pixel interactions. However, the weight definition based on this distance (5.16) is

rather heuristic. Hence, when specific information about the noise model is available, one can design an alternative weight definition based on the statistical distribution of the distance between patches. Furthermore, this distance is computed on the observed image its appropriateness could be questioned in particular when the noise level is important. This limitation could be improved by considering an alternative set of features to characterize a given pixel. The basic idea is to consider a subspace and the projection of the observations on this subspace forming a feature vector. This vector could better describe the data structure if the subspace is powerful enough while being robust to noise (due to the projection).

5.4.1 New Statistical Similarity Measure Between Patches

In this section, we will focus on the statistical distribution of the L^2 distance between patches. To this end, let us consider two patches $I_{\mathbf{N}_x}$ and $I_{\mathbf{N}_y}$ (of size d) that correspond to two different observations of the noise free patch $U_{\mathbf{N}_x}$ then

$$\begin{aligned} d^2(I_{\mathbf{N}_x}, I_{\mathbf{N}_y}) = \sum_{i=1}^{d} (I_{\mathbf{N}_x}(i) - I_{\mathbf{N}_y}(i))^2 &= \sum_{i=1}^{d} (U_{\mathbf{N}_x}(i) + n(i) - U_{\mathbf{N}_x}(i) - n(j))^2 \\ &= \sum_{i=1}^{d} (n_{\mathbf{x}}(i) - n_{\mathbf{y}}(i))^2 \end{aligned} \quad (5.31)$$

$n_{\mathbf{x}}$ (resp $n_{\mathbf{y}}$) refers to the noise patch associated to $\mathbf{x}$ (resp $\mathbf{y}$). In case of Gaussian noise $[n_{\mathbf{x}}(i) - n_{\mathbf{y}}(i)]$ follows a Gaussian distribution of Zero mean and variance $2\sigma_n^2$. Now if we consider the following distance

$$d_{\mathbf{xy}} = \frac{d^2(I_{\mathbf{N}_x}, I_{\mathbf{N}_y})}{2\sigma_n^2} = \sum_{i=1}^{d} \frac{(n_{\mathbf{x}}(i) - n_{\mathbf{y}}(i))^2}{2\sigma_n^2} \quad (5.32)$$

Given that the sum of the square of d random variables having a normal distribution is distributed according to a χ_2 law with mean equal to d and variance equal to $2d$, we can conclude that, $d_{\mathbf{xy}}$ follows χ_2 distribution. Based on this assumption, a new similarity measure that depends only on the noise variance information can be introduced. Unlike the similarity value relative to photometric consistency introduced in (5.16), where one have to tune the parameter σ_{ph} this measure is automatic.

The χ_2 distribution can be approximated by a Gaussian one with the same mean and variance when d is high. We will assume that for patches larger than 5×5, $d_{\mathbf{xy}}$ has a Gaussian distribution of

mean d and variance $2d$. The similarity measure can be expressed as

$$w_{\mathbf{xy}} = \frac{1}{2\sqrt{d\pi}} \exp - \left(\frac{(d_{\mathbf{xy}} - d)^2}{4d} \right) \exp \left(- \frac{\|\mathbf{x} - \mathbf{y}\|^2}{2\sigma_s^2} \right) \tag{5.33}$$

To understand the impact of the choice of the patch size d on the quality of restoration, we will consider the mean square distance between patches $\bar{d}_{\mathbf{xy}} = \frac{1}{d} d_{\mathbf{xy}}$. $\approx$ In case $d_{\mathbf{xy}}$ has a Gaussian distribution, $\bar{d}_{\mathbf{xy}}$ has also a Gaussian distribution of constant mean (equal to one) and variance equal to $\frac{2}{d}$. This result provides us with an interpretation of the link between the size of patches considered for comparison and the amount of regularization. When considering patches of large size, the weight decrease is fast because the Gaussian variance is inversely proportional to the patch size. This is synonymous of better selection of similar patches and also small amount of smoothing (for a given $\mathbf{x}$ several $\mathbf{y}$ have weights close to zero). Contrarily to that, when considering small d values the weights have a smaller decrease rate and several pixels $\mathbf{y}$ in the neighborhood of $\mathbf{x}$ have a considerable weight which leads to high level of noise reduction. Based on these observations one understands why the Non Local mean filter acts better than the bilateral filter in texture denoising. Beyond the new weights definition different from the one traditionally used in similarity computation, one can also use different patch size for comparison according to the image context. A possible way to do that is to consider the pre-classification step introduced in chapter 2 and assign a high patch size to pixels belonging to the texture and smaller one for those belonging to the smooth area.

In figure [Fig.(5.2)], an example of weight values reflecting the similarity between a textured patch and the remaining patches in the image is given. We considered for weights computation only the photometric distance between patches. This experiment shows that the new definition provides weights with more disparity than the classical definition which is equivalent to better patches selection. The effect of the patch size is illustrated in [Fig.(5.2-d),Fig.(5.2-e)]. Considering large patches for comparison ensures better texture preservation.

5.4.2 PCA Based Dictionary

Image decomposition in subspaces, like wavelets, Fourier, etc. reduces the dimensionality of the problem and often associate noise to the least important components. The central idea of our approach is to decompose image content into a dictionary, or a subspace and then use this subspace (in particular the projection of the image patch into the subspace) to define the distance between pixels. One can claim that such an image representation will remove certain amount of noise, which is critical when determining similarities between observations of neighborhood pixels.

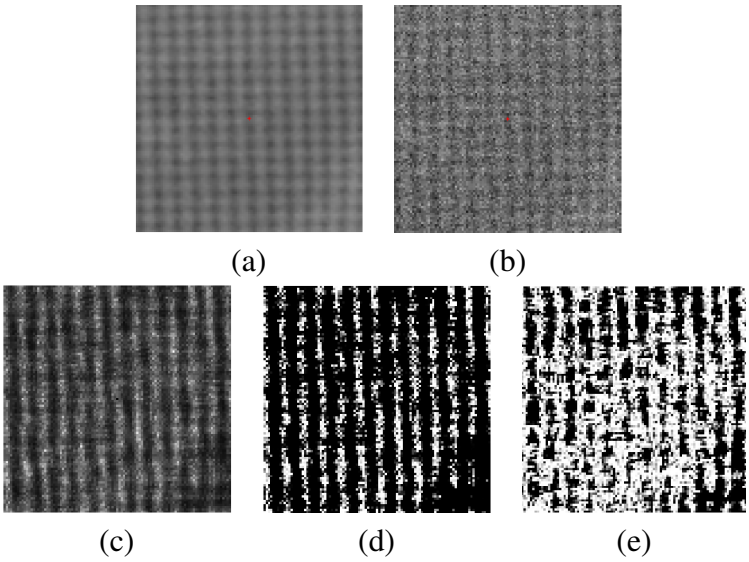


Fig. 5.2: (a) Example of texture (b) Texture corrupted by Gaussian noise $\sigma_n^2 = 20$ – Similarity measure between the central pixel (Red) and the other one using : (c) Using definition (5.16) and the noisy image (d) Using definition (5.33) and the noisy image (patch size $d = 81$) (e) Using definition (5.33) and the noisy image (patch size $d = 25$)

Several image filters can be used to perform local features extraction of the image mainly filter banks [57, 102, 66]. In this work, we preferred a data driven approach that consists in analyzing the image patches and to provide a reliable description of them.

A simple way to perform image patches description is principle component analysis (PCA) which is a powerful tool that provides a compact representation of the most prominent element in the image. We will then consider the set of all image patches present in the image and using PCA, we will identify the principle directions of the data variation which correspond to the relevant information about the image. It is important to note that we do not make an explicit assumption on the linearity of the image model and the PCA is just considered to produce a projection subspace.

Let us now introduce $I_{\mathbf{N}_x}$ which is a one dimensional vector composed of the pixels intensities observed in a patch around $\mathbf{x}$ and noted as $I_{\mathbf{N}_x} = \{I_{\mathbf{N}_x}(i)\}_{1 \leq i \leq d}$. The patch width is w_p and the dimension of the associated vector is $d = w_p \times w_p$. The principle component analysis will be applied on the observations set relative to all patches within the image $\mathcal{S} = \{I_{\mathbf{N}_x} \ \ \mathbf{x} \in \Omega\}$. First we have to define the matrix of observations M of size $|\Omega| \times d$ where each row correspond to the observation vector $I_{\mathbf{N}_x}$. The principle component analysis is performed through the eigenvectors and eigenvalues computation of the $d \times d$ covariance matrix C_r that is defined as follows

$$\begin{aligned} C_r(i,j) &= cov(V_i, V_j) \quad V_j \text{ is the } j^{th} \text{ colon vector of the matrix } M \\ cov(V_i, V_j) &= \sum_{k=1}^{|\Omega|} (V_i(k) - \bar{V_i})(V_j(k) - \bar{V_j}) \ \text{ and } \ \bar{V_i} = \frac{1}{|\Omega|} \sum_{k=1}^{|\Omega|} V_i(k) \end{aligned} \tag{5.34}$$

The principle components correspond to the eigenvectors noted $\{e_k\}_{1\leq k\leq d}$ of the matrix C_r the variance of the data according to each direction is given by the eigenvalues $\{\lambda_k\}_{1\leq k\leq d}$. Hence, we obtain an orthonormal basis on which we can decompose the image to generate a new set of features. The projection of the observations set on the most prominent direction allows the extraction of smooth image content as well as texture and edges. The remaining components that correspond to small amount of variability contain limited information about the image structure due to the dominance of the noise component.

The eigenvectors can be interpreted as a dictionary where the elements correspond to image characteristics (cf figure [Fig.(5.3)]). The projection of the set of observations in $\mathcal{S}$ on each vector on the new basis is equivalent to a correlation between an image patch and an element of the dictionary learned using PCA. The projection on some principle components is illustrated in figure [Fig.(5.4-a),Fig.(5.4-b),Fig.(5.4-c)]. We can notice that the first principle component acts as a low pass filter that captures all smooth contents of the image. The projection on other components extracts the texture and small details in the image. It's important also to point out that vectors associated to small eigenvalues that are close to the noise variance emphasize the random component introduced by noise.

A natural choice for computing a new set of features for each image pixel is to perform the projection of the image patch on the new orthonormal basis provided by PCA. Nevertheless, considering the entire basis is not relevant because the L^2 distance between the new features vectors remains unchanged. On top of that, the main argument for subspace consideration is to discard the directions of the projection where the noise variance is higher than the signal variance. Knowing that the projection of an image patch that is corrupted by a Gaussian noise of variance σ_n^2, on an orthonormal basis will result in feature vector corrupted by a Gaussian noise with the same variance, the signal to noise ratio for a given direction of projection k is defined by $\frac{\lambda_k}{\sigma_n^2}$. To select the directions of projection that are reliable for feature vector computation, a threshold for the SNR could be considered. Taking into consideration that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_d$, we will restrain the patches projection to the subspace E_q generated by the first q vector $(e_l)_{1\leq k\leq q}$ such that $\frac{\lambda_k}{\sigma_n^2} > T$ for all $1 \leq k \leq q$. Under these considerations, we will measure the similarity between pixels in the new subspace (E_q) of dimension $q \leq d$.

Finally, if we note $\mathbf{F_x}$ the new feature vector obtained by the projection of $I_{\mathbf{N}_x}$ on the subspace E_q , where

$$\mathbf{F_x}(k) = \sum_{i=1}^{d} I_{\mathbf{N}_x}(i) e_k(i) \qquad 1 \leq k \leq q \tag{5.35}$$

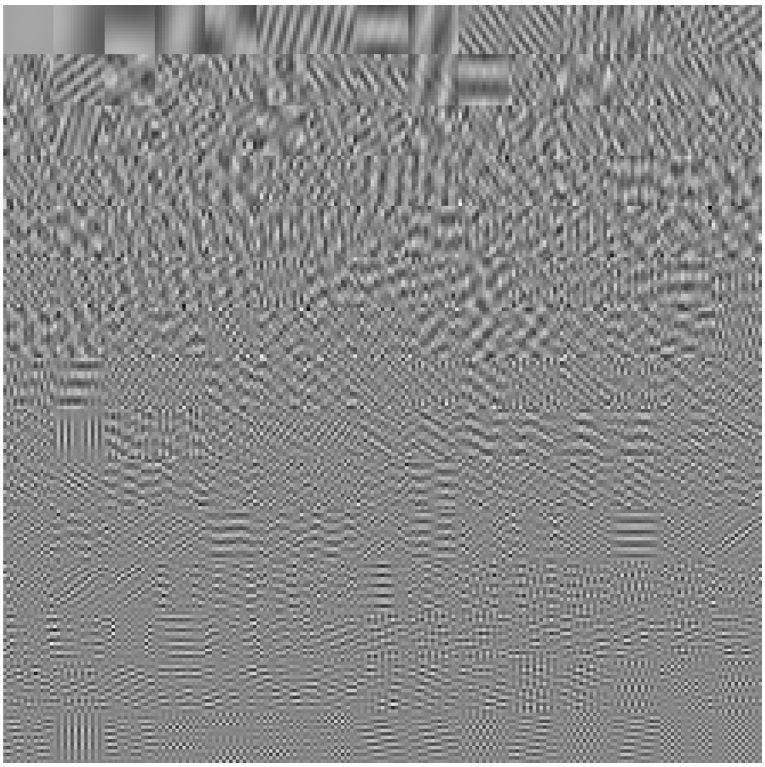


Fig. 5.3: Eigenvectors obtained through PCA decomposition of Barbara image corrupted by Gaussian noise ($\sigma_n = 10$) (patch size 15×15) (Only the first 22 component are significant if T=3)

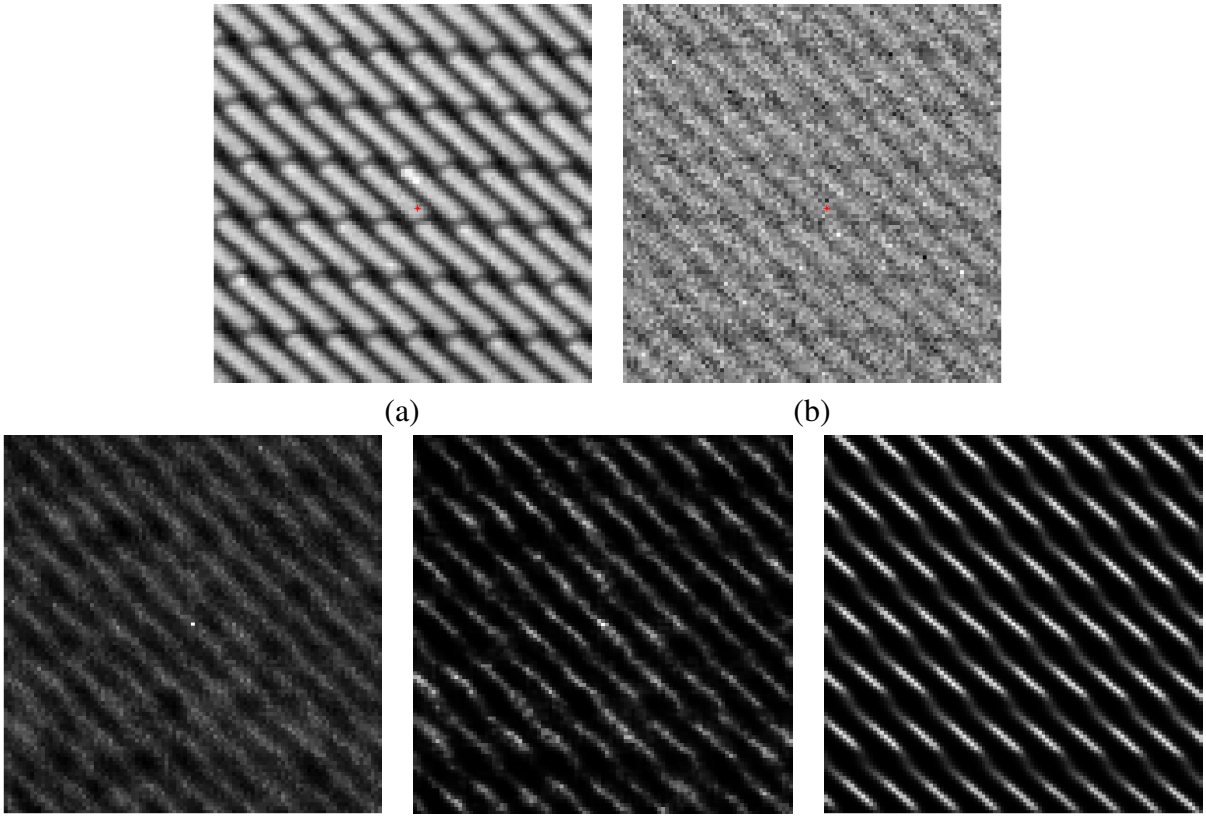
The definition of the similarity measure between two pixels $\mathbf{x}$ and $\mathbf{y}$ is defined as

$$w_{\mathbf{xy}} = \exp\left(-\frac{\|\mathbf{F_x} - \mathbf{F_y}\|^2}{2\sigma_{ph}^2}\right) \exp\left(-\frac{\|\mathbf{x} - \mathbf{y}\|^2}{2\sigma_s^2}\right) \tag{5.36}$$

Some idea on the appropriateness of this weight definition is given in [Fig.(5.5)] where the similarity measure between the central pixel (in Red) and the other pixels in the case of noisy texture is represented. Two weight definitions were considered, one based on the L^2 distance between patches the other on the new set of features computed using PCA. On the other hand, we computed the weight based on the L^2 distance between noise free image patches. It is clear that the new similarity measure is better selective toward pixels and closer to the one obtained using a noise free image.

(a) (b) (c)

Fig. 5.4: (a) Projection of the observations set on the first eigenvector of Barbara image (b) Projection of the observation set on the 17^{th} eigenvector (c) Projection of the observations set on the Last eigenvector (49^{th})



(a) (b)

(c) (d) (e)

Fig. 5.5: (a) Example of texture (b) Texture corrupted by Gaussian noise $\sigma_n^2 = 40$ – Similarity measure between the central pixel (Red) and the other one using : (c) the L^2 distance between patches (d) feature set based on PCA (e)the L^2 distance between patches using the noise free image

5.5 Experimental Results

We will focus on performance evaluation of the method that we presented in this chapter. In this context, we will consider the case of synthetic additive white Gaussian noise with known standard deviation ($\sigma_n = 20$). For comparison, we have considered the following methods

- Total variation minimizing [122] with data term $\lambda = 0.1$
- Anisotropic diffusion [115] with an edge stopping function $(1 + |\nabla I|^2 / K^2)^{-1}$ $(K = 5)$
- Total variational with local constraints on the residual variance introduced in [62][1].
- Non local extension of the TV model [60] [2] using $\lambda = 0.1$. The weights considered in this functional are computed according to expression (5.16) using patches of size 7×7, $\sigma_s = 4$ and $\sigma_{ph} = 15$ (We did not use a fully non local version of this algorithm using).

In terms of variants of our method, we have considered one with fixed bandwidth and one with variable bandwidth selection through optimization. In terms of parameters, the following set was used to compute the weights according to (5.16): patch size 7×7, $\sigma_{ph} = 15$ and $\sigma_s = 4$. The same set of parameters was used for the variable bandwidth (σ_s was updated for each pixel). The coefficient that controls the balance between the regularization and the fidelity to data was set to $\lambda = 0.1$. Regarding time computation using this set of parameters is 5 mn to minimize the energy with respect to U and 1mn to compute the optimal spatial bandwidth. The convergence is obtained after few iterations. We used a C++ code and Pentium IV-2GHz machine. The time computation is a limitation of our method but optimizing the code would reduce considerably the computation time.

In table (Tab.5.1), the PSNR values show that the regularization model that we presented outperforms all the other techniques. The total variation minimization, the anisotropic diffusion and the total variation with local constraints are local methods based on the image gradient. They rely on a piecewise constant image model and it is known that such an assumption is a not realistic for natural images. The figure [Fig.(5.7), Fig.(5.9)] provides certain visual evidence for this claim since the TV based approaches creates some spurious contours inside homogeneous regions while the anisotropic diffusion yields to many flat area and destroys completely the texture in the image. Note also that our approach yields better results than the non local extension of the total variation [60] defined by the equation (5.9). In addition to PSNR values, figures [Fig. (5.6),Fig. (5.8)] as

[1] We used the Matlab implementation of the author
[2] Using our implementation

	Barbara	Boat	Fingerprint	House	Lena	Baboon	Peppers
ourMethod-Fix	30.19	29.76	27.99	32.11	31.95	26.33	30.51
ourMethod-Var	30.46	29.94	27.65	32.34	32.12	26.02	30.67
Method in [60]	29.01	28.75	26.95	31.12	30.82	25.72	29.41
TV-ROF [122]	26.18	27.72	26.08	28.43	28.45	25.18	28.51
TV-Adaptive [62]	27.20	28.79	26.34	30.92	30.64	25.53	29.49
Anis Diff [115]	26.45	28.06	24.81	29.41	29.27	23.68	-

Tab. 5.1: PSNR values for denoised image (The PSNR of the image corrupted by gaussian noise of std=20 is equal to 22.15)

well as residual images [Fig. (5.10)] show that our regularization process (fixed or variable bandwidth) performs better in terms of fine details preservation. We recall that our regularization flow had an additional term that can be seen as a sharpening term that involves a weighted average of the other "residual" values. This term enforces the coherence between neighboring pixels.
Now if we consider the performance of the variable bandwidth version with respect to the fixed one, we can notice certain improvement in image details preservation (see for example the figure [Fig. (5.8)] that contains details in Lena's hat). Nevertheless, for some images the fixed bandwidth performs better than the variable one. These images are composed mainly of textured content and the majority of pixels have similar local properties for this reason considering the same spatial bandwidth for all of them could be more interesting. On top of that, the term considered for the bandwidth optimization is not convex with respect to σ_s and therefore, the estimation cannot be optimal.

5.5.1 On the weight selection

The selection of the photometric distance is a crucial component of our approach and in this chapter we have introduced novel means of defining this distance. In order to evaluate their impact on the process we will compare three distance definitions: (i) the L^2 distance between patches 5.16 (ii) distance between new feature vector computed using PCA and threshold value $T = 1.5$ (iii) The normalized L^2 distance between patches by the noise variance and using the observation that this distance follows a χ_2 distribution (5.33). The PSNR values in table (Tab.5.3) show that the three weight definition are equivalent. Therefore, the proposed approach is not too sensitive to the weight definition as far as they reflect faithfully the image structure. Figures [(Fig.5.11),(Fig.5.12)] confirm visually this observation, where we can see, for both examples, equivalent amount of details in the three images.

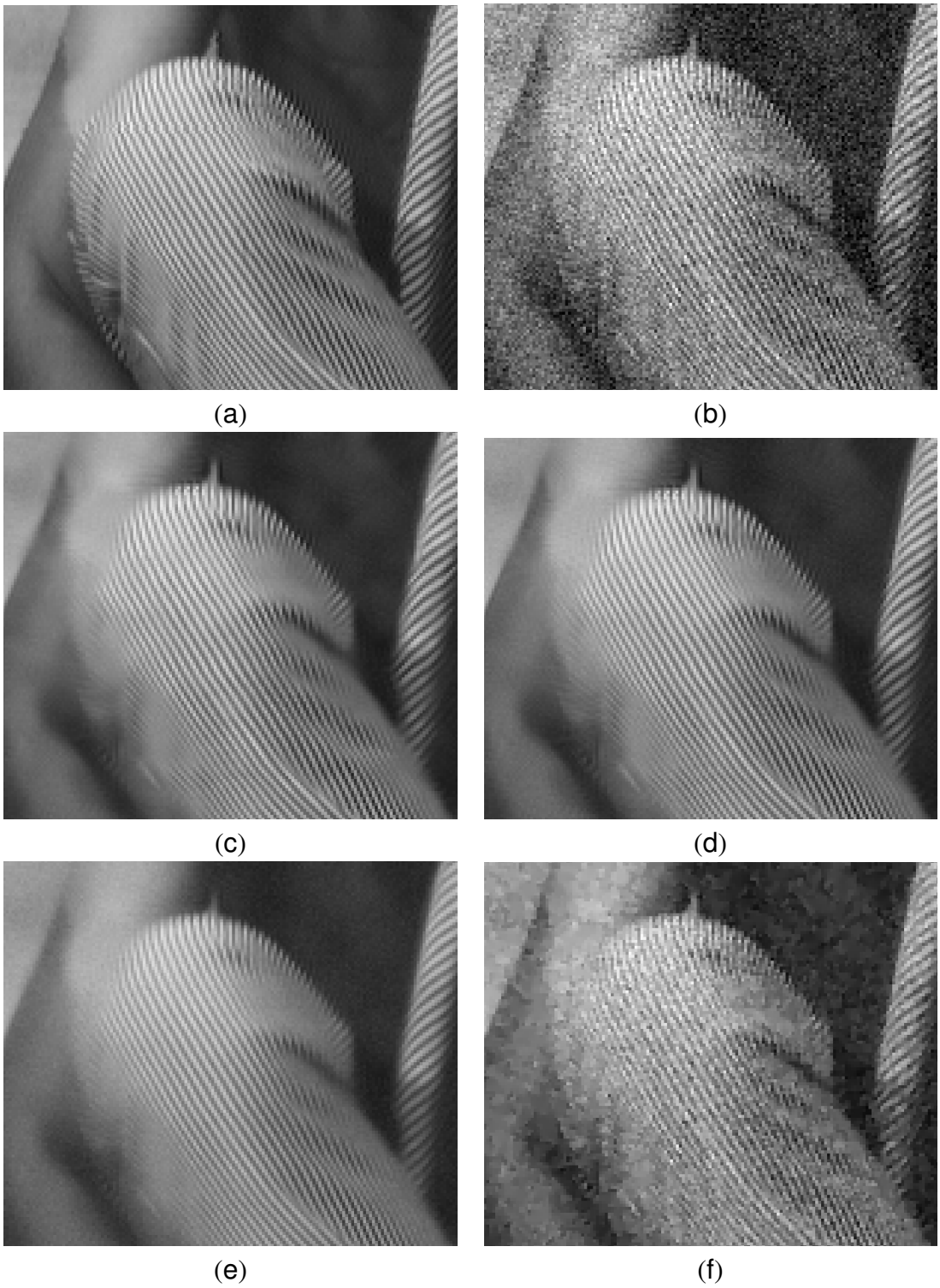
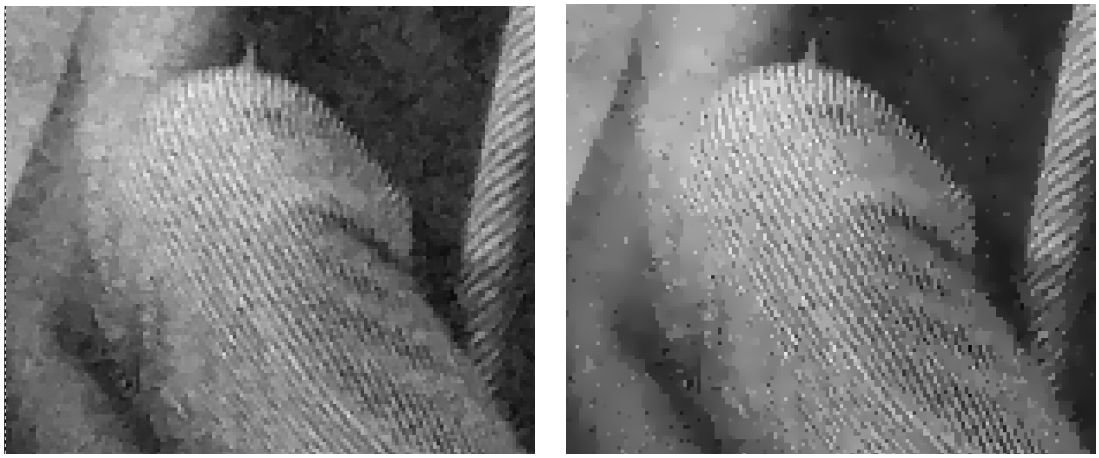


Fig. 5.6: Zoom on the Barbara image (a) Original image, (b) Noisy image (c) Our method using fixed bandwidth (d) Our method using variable bandwidth (e) Non local functional minimizing (5.9) (f) Total variation minimizing with variable local constraint



(a) (b)

Fig. 5.7: Zoom on the Barbara image (a) Total variation (b) Anisotropic Diffusion

	Barbara	Boat	Fingerprint	House	Lena	Baboon	Peppers
ourMethod (L^2)	30.19	29.76	27.99	32.11	31.95	26.33	30.51
ourMethod (PCA)	30.05	29.70	27.8	32.34	31.6	26.08	30.54
ourMethod (χ_2)	30.07	29.61	27.14	32.34	31.8	26.19	30.61

Tab. 5.2: PSNR values for denoised image (The PSNR of the image corrupted by gaussian noise of std=20 is equal to 22.15)

We will perform the same test with respect to the NL-means algorithm. To perform comparison, we selected the following set of parameters: $\sigma_{ph} = \sigma_n = 20$, $\sigma_s = \infty$, and a patch sizeof 5×5 to perform the PCA. The local filtering window was set to 11×11. The PSNR values in table (Tab.5.4) show that using the new set of features for pixels comparison improves the performance of the NL-means algorithm. This is illustrated by the residual images in figures [Fig.(5.13),Fig.(5.14)] that contains less structure (mainly stripes in the pants and scarf). On the other hand a zoom on the image detail in [Fig.(5.15)] show that texture is better reconstructed. The NL-means algorithm is simple but it is very sensitive to the parameters selection. Hence, an accurate measure of similarity between pixels insures better quality restoration. The impact of the parameter q which is the number of retained PCA components is illustrated in figure [Fig.(5.16)]. This graph shows that the performance depends on the number of components considered in the feature vector computation. One can claim that few principal directions are not sufficient to incorporate local image information in pixel characteristics vector. On the other hand, retaining an important number of directions will impact the amount of noise being considered in the reconstruction. Consequently, the accuracy of the similarity computation will decrease. The automatic selection of the optimal q value was not addressed in this thesis but could be an interesting issue.
Now if we compare the original NL-means algorithm and the modified version using the weight

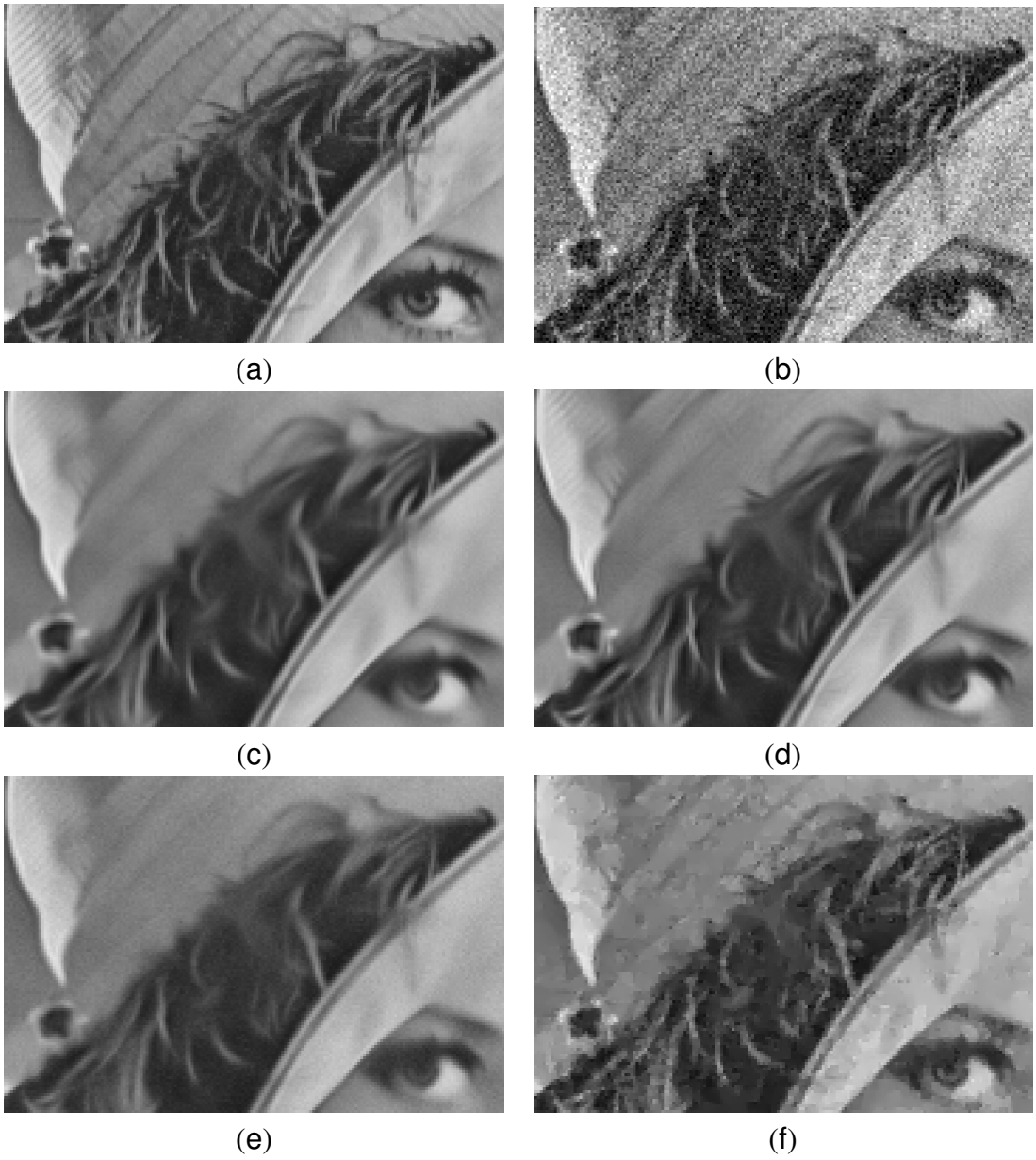
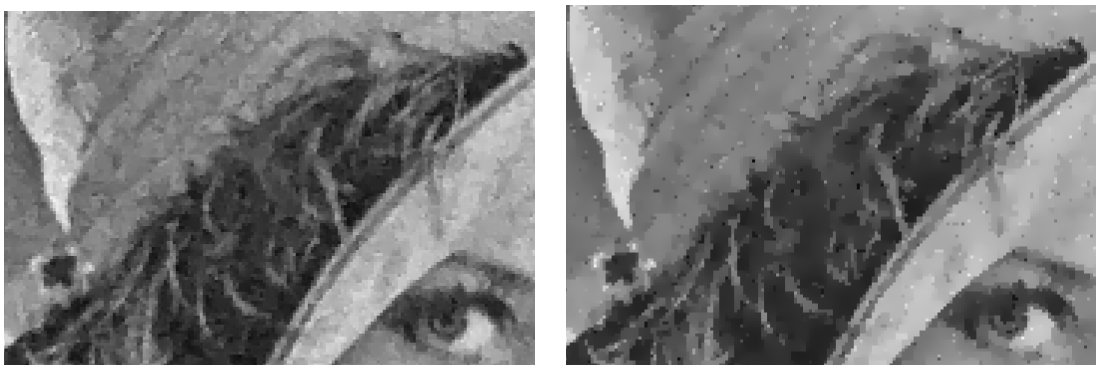


Fig. 5.8: Zoom on the Lena image (a) Original image, (b) Noisy image (c) Our method using fixed bandwidth (d) Our method using variable bandwidth (e) Non local functional minimizing (5.9) (f) Total variation minimizing with variable local constraint



(a) (b)

Fig. 5.9: Zoom on the Lena image (a) Total variation (b) Anisotropic Diffusion

	Barbara	Boat	Fingerprint	House	Lena	Baboon	Peppers
ourMethod (L^2)	30.19	29.76	27.99	32.11	31.95	26.33	30.51
ourMethod (χ_2)	30.07	29.61	27.14	32.34	31.8	26.19	30.61
NL-Means (L^2)	28.73	28.49	26.27	31.37	30.95	24.5	29.94
NL-Means (χ_2)	29.33	28.9	26.04	31.64	31.29	23.99	29.68

Tab. 5.3: PSNR values for denoised image (The PSNR of the image corrupted by gaussian noise of std=20 is equal to 22.15)

definition (5.33), we can notice that we also obtain better results. Furthermore, the visual results show that details are better preserved when considering the actual distribution of the L^2 distance between patches. This weight definition is particularly attractive when the noise variance is known so one can simply select $\sigma_{ph} = \sigma_n$ without any further tuning of the parameter $\sigma_p h$. This is not the case with the classical weight definition used in the NL-means algorithm, which seems to be quite sensitive to the selection of parameters. This experimental conclusion is supported from some theoretical insights presented in [76]. In particular, it was proven that the computation of a weighted mean over a very large neighborhoods does not insure a high quality restoration unless σ_{ph} is carefully selected. Contrarily to that considering the actual distribution of the distance between patches in case of additive Gaussian noise allows an automatic selection of σ_{ph}. Besides, its performance is not too sensitive to the selection of the window size used for computing the weighted average. These observations are illustrated by the graph in figure [Fig.(5.17)] that represents the PSNR evolution of the NL-means algorithm with the size (equal to $2T_f + 1 \times 2T_f + 1$) of the domain used to perform the weighted average. We recall that we selected $\sigma_{ph} = \sigma_n = 20$, $\sigma_s = \infty$ and 5×5 patch size for weight computation. With the classical weights definition the performance drops rapidly when the window size becomes bigger then the optimal one. This is not the case when considering the actual distribution of the normalized L^2 distance between patches where the performance increases with the window size. Furthermore, after reaching an optimal window the decrease of the

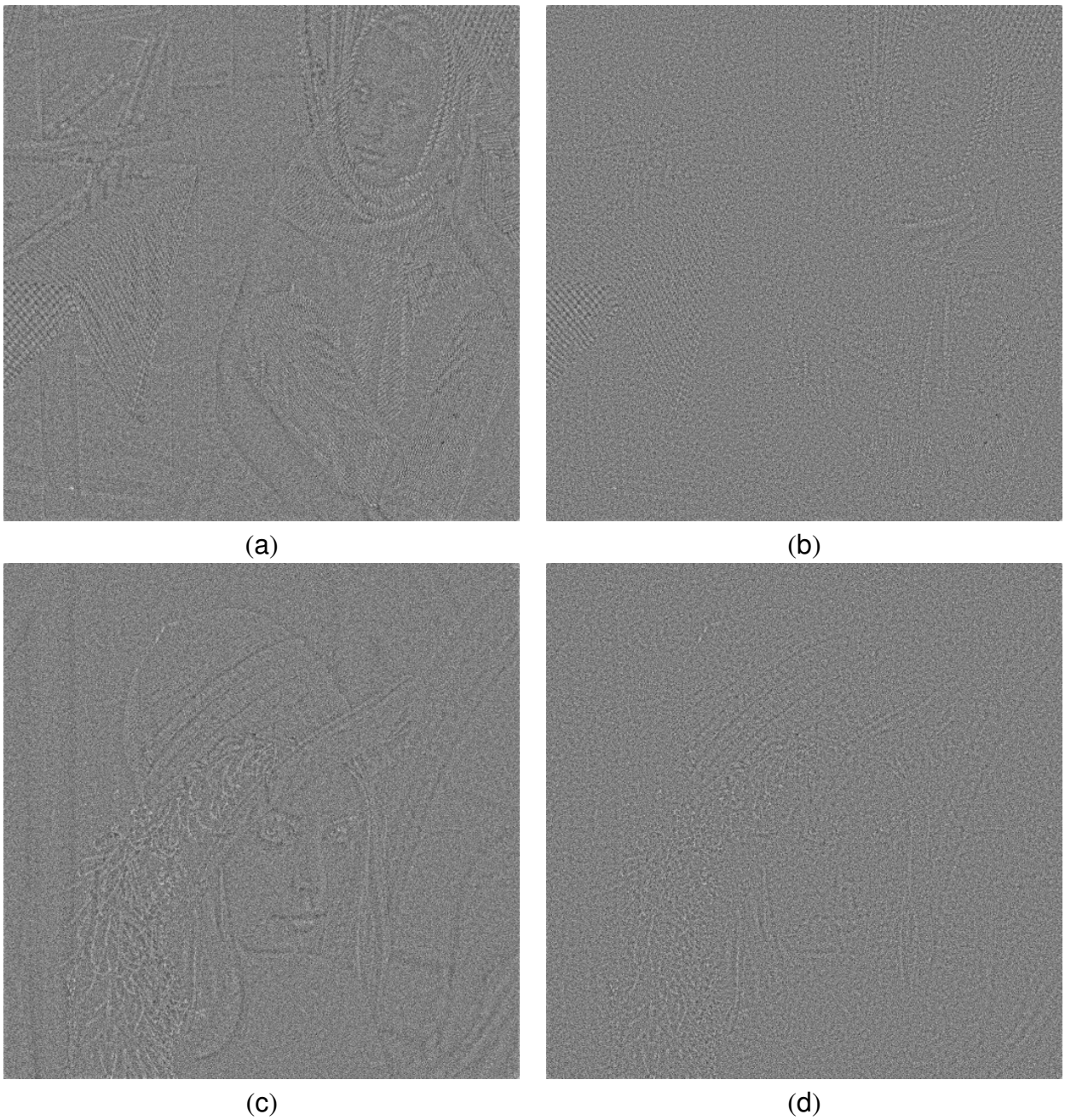


Fig. 5.10: (a) Residual of the Barbara image using the non local convex functional 5.9 (b) Residual of the Barbara image using our approach with fixed bandwidth (c) Residual of the Lena image using the non local convex functional 5.9 (d) Residual of the Lena image using our approach with fixed bandwidth

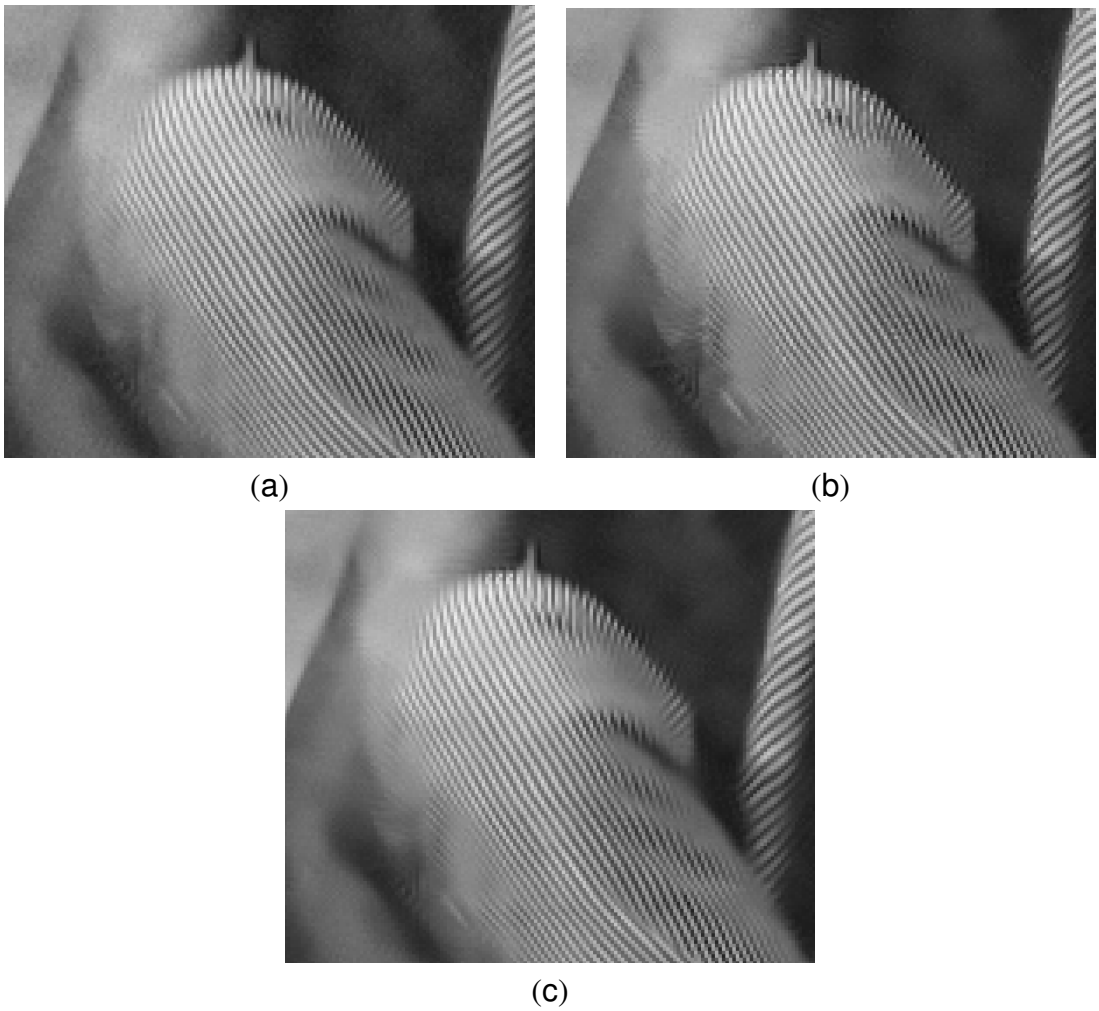


Fig. 5.11: Zoom on Barbara image and result of our method using different weights definition and fixed bandwidth (a) using distance between pixels features using PCA (b) using the statistical distribution of the L^2 distance between patches and expression (5.33) (c) using L^2 distance between patches and expression

PSNR is slow.

5.6 Conclusion

In this chapter, we were interested in total variational methods and their extensions. Our contribution consists in considering new local quadratic regularization energy in the context of denoising. The underlying image model is more realistic than the piecewise smooth one and the solution is easily tractable. The experimental results showed that the regularization based on this energy out-

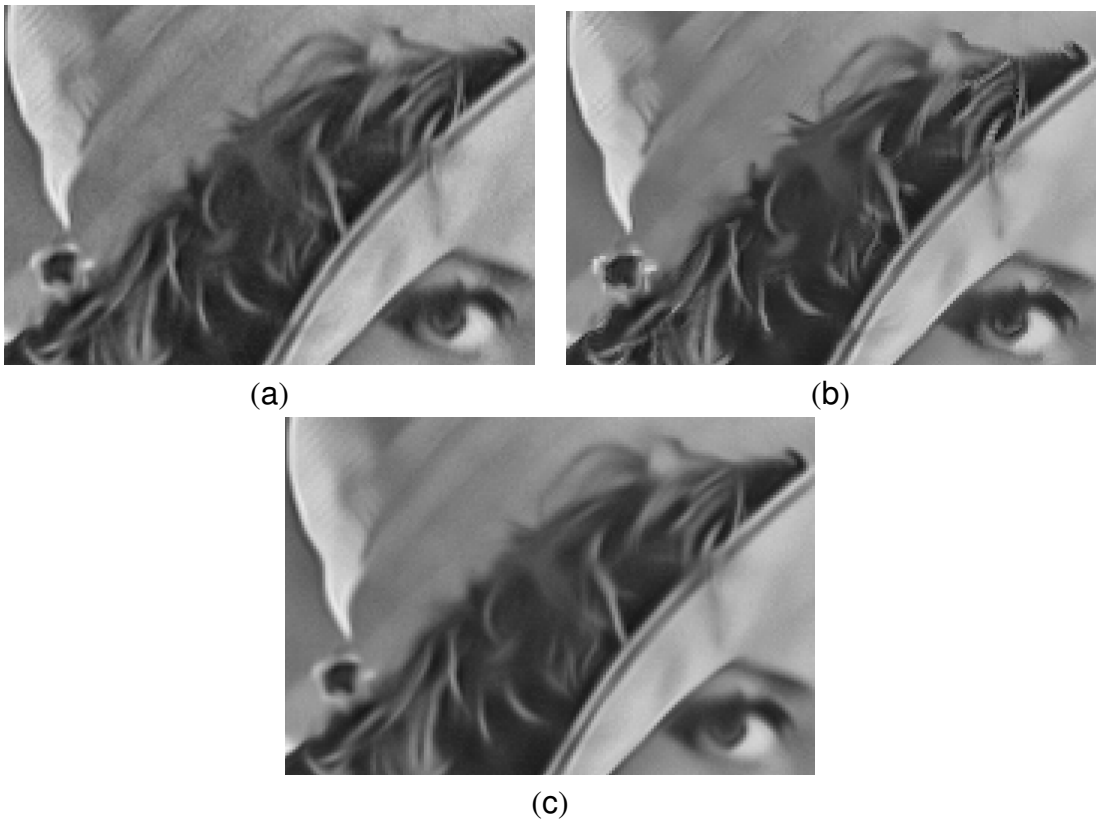


Fig. 5.12: Zoom on Lena image and Result of our method using different weight definition and fixed bandwidth (a) using distance between pixels features using PCA (b) using the statistical distribution of the L^2 distance between patches and expression (5.33) (c) using L^2 distance between patches

performs all the other techniques based on total variation minimization and anisotropic diffusion

The interaction between pixels and new weight definition were also the focus of this chapter. First, an effort was devoted to compute an adapted spatial bandwidth for each pixel. Contrarily to non local approaches, our approach entails the spatial distance which is as important as photometric one. In natural image similar patches are more likely to be found in local or semi local neighborhood. The function being minimized with respect the spatial bandwidth is not convex further improvement can be achieved by using appropriate optimization technique toward finding global minimum. In addition to that, including a prior on the spatial bandwidth could improve the

	Barbara	Boat	Fingerprint	House	Lena	baboon	Peppers
NL-Means (L^2)	28.73	28.49	26.27	31.37	30.95	24.5	29.94
NL-mean (PCA)	29.83	29.29	27.09	31.87	31.71	25.56	30.41
NL-Means (χ_2)	29.33	28.9	26.04	31.64	31.29	23.99	29.68

Tab. 5.4: PSNR values for denoised image (The PSNR of the image corrupted by Gaussian noise of std=20 is equal to 22.15)

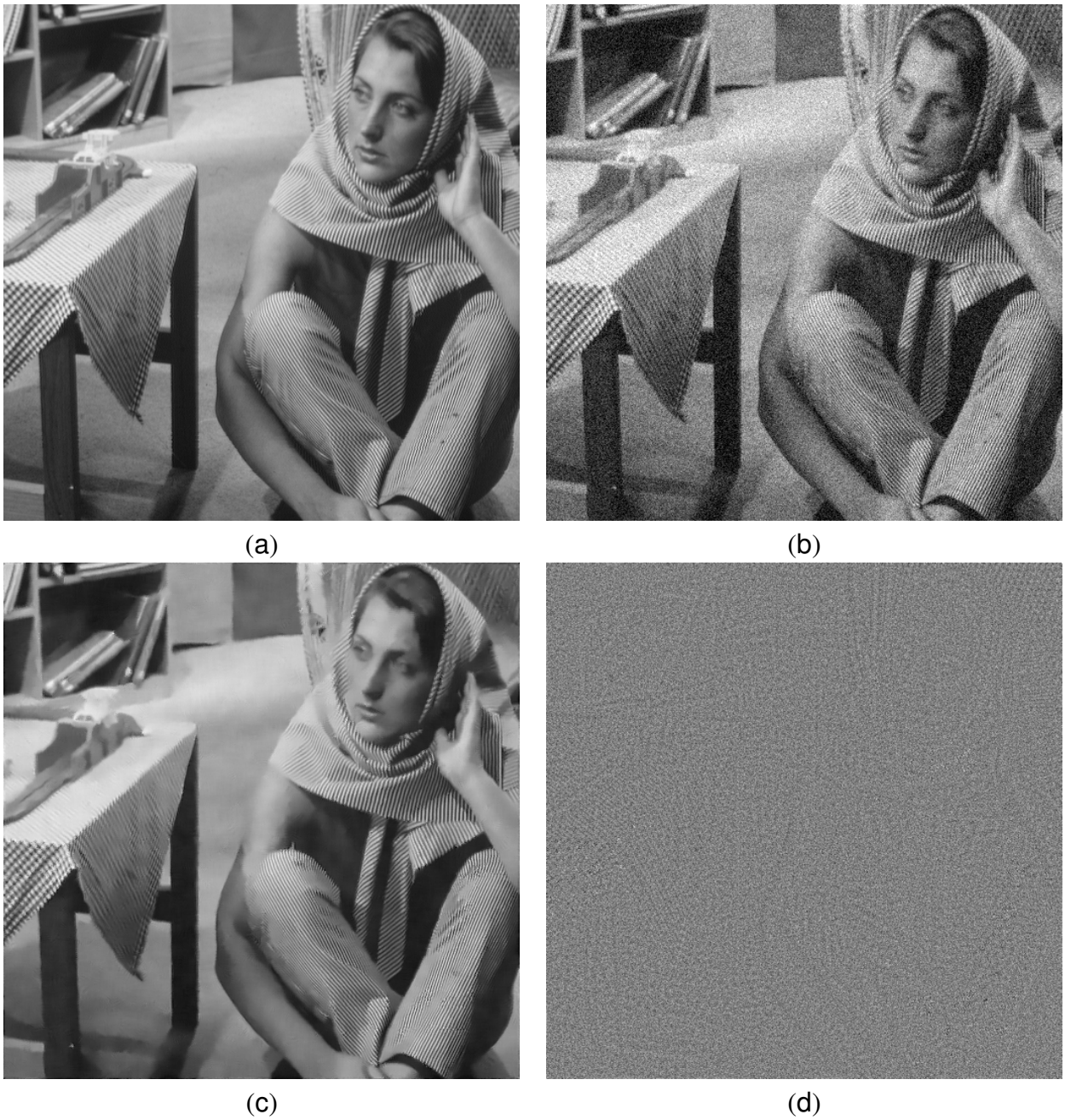


Fig. 5.13: Results using NL-means algorithm (a) Original image (b) Noisy image (c) using the statistical distribution of the L^2 distance between patches and expression (5.33) (d) The corresponding residual image

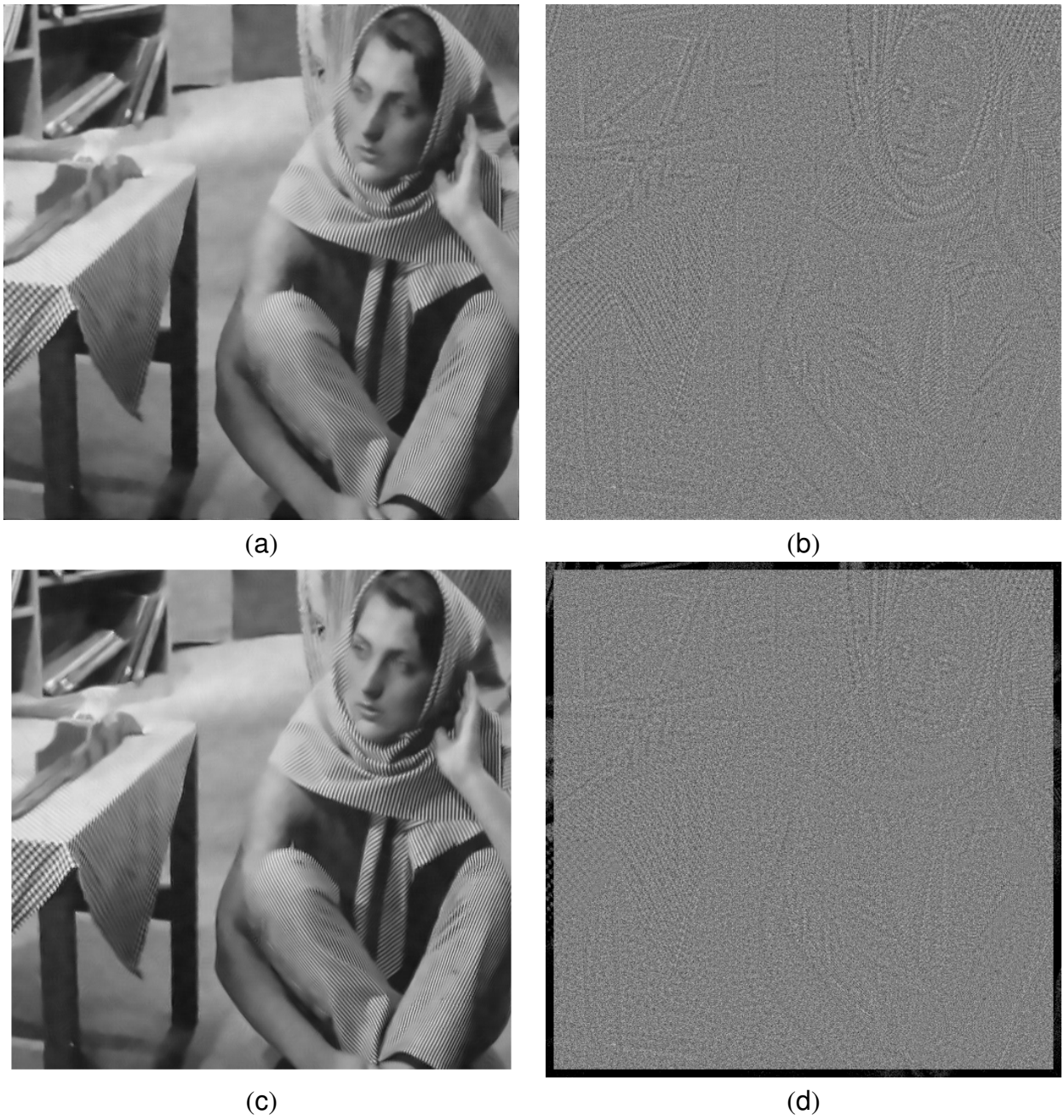


Fig. 5.14: Results using NL-means algorithm (a) using L^2 distance between patches (b) Its corresponding residual image (c) using distance between pixels features using PCA (d) Its corresponding residual image

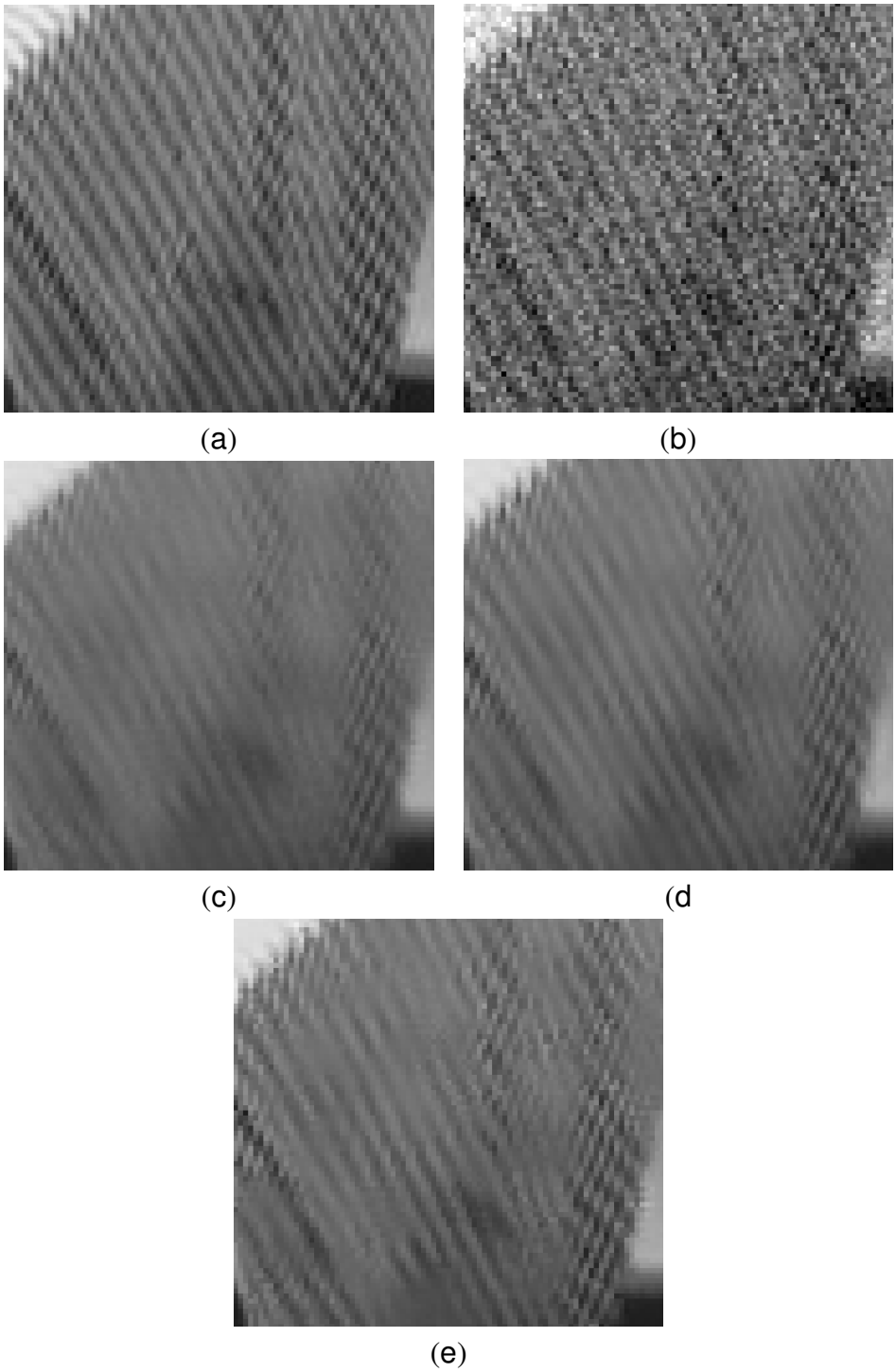


Fig. 5.15: (a) original image (b) Noisy image. Results using NL-means algorithm (c) using L^2 distance between patches (d) using distance between pixels features using PCA (e) using the statistical distribution of the L^2 distance between patches and expression (5.33)

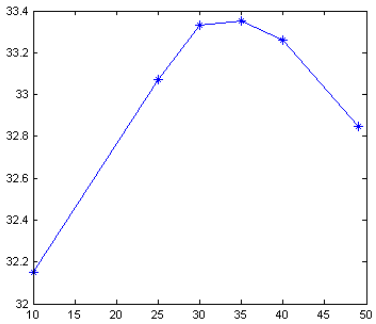


Fig. 5.16: Evolution of the PSNR value with respect to the retained number of principle component

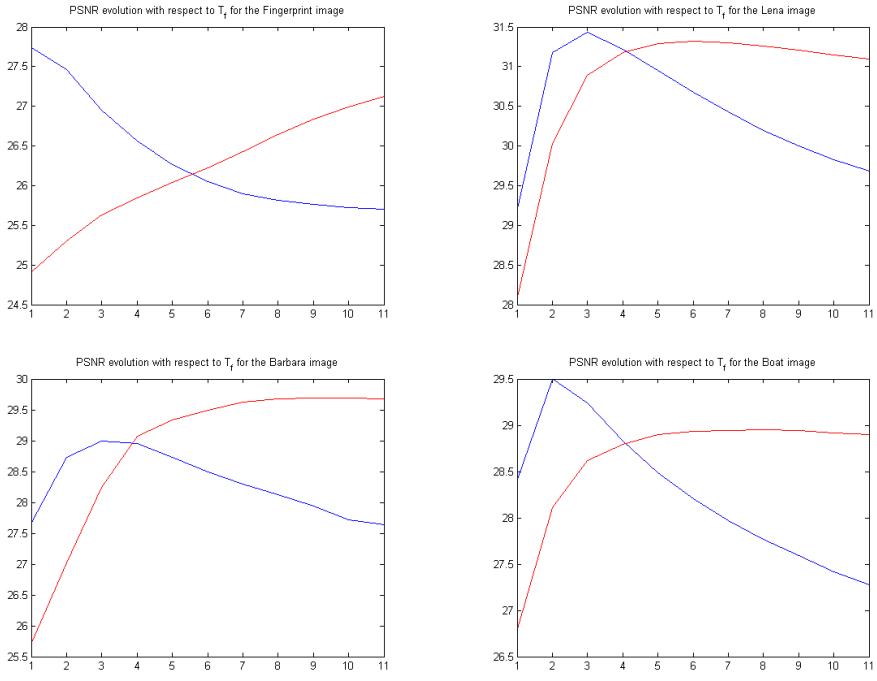


Fig. 5.17: Evolution of the PSNR with respect to the window radius : blue curve corresponds to the classical weight definition (5.16) , the red curve corresponds to the new definition (5.33)

performance of the approach.

As far as photometric similarity is concerned, we proposed new weights definition one based on PCA analysis of image patches the other based on the statistical behavior of the L^2 distance between noisy patches. These weights definitions are designed to better reflect similarity between pixels which is very crucial in the denoising process. For the statistical new measure the choice of the patch size is important and an automatic selection of this parameter should be addressed.

In the following chapter we will consider different context and other applications for the presented total variation approach mainly for real digital camera noise and medical imaging application.

Chapter 6

Applications

The proliferation of image sensors has led to an explosion of the nature of images to be considered either in terms of lower or higher level tasks. Denoising is among the most fundamental with great importance in a number of fields where either the sensors are not so well developed or a compromise is to be find between cost and acquisition quality. In this chapter, we show that the convex functional minimizing framework that we presented can tackle other different restoration problem and can deal with both scalar and vectorial observations. First, we addressed RGB images denoising using more realistic noise models towards defining more appropriate data-fidelity terms. Next, we extend our framework to non-linear manifolds that is the case of semi-definite positive matrices explaining the diffusion of water in human skeletal muscle as determined from DTI images. Last but not least, we are concerned with speckle reduction in ultrasound images with aim to facilitate the process of automatic organ delineation. The variety of fields, noise models and image properties demonstrate the potentials of the convex functional minimizing approach presented in this thesis.

6.1 Color Image Denoising

Classical denoising approaches assume that images are corrupted by an additive white Gaussian noise with a fixed variance. In this section, we will show that such an assumption is unrealistic for digital camera noise. Furthermore, we will introduce an extension of the previous total variation based approach for color images that is based on realistic noise assumptions. To this end, we will focus on analyzing the noise produced by the digital camera. In this context, we provide an empirical and non parametric estimation technique for the evolution of the noise variance with

respect to the intensity. Finally, denoising is performed by minimizing a regularization function under a set of constraints derived from the noise model.

6.1.1 Noise Properties from Raw to RGB Images

The raw image is the one obtained from the impact of light photons on the camera sensor. It is a one channel image composed of color filter arrays (Bayer pattern). A given pixel in the image corresponds to a color filter cell, while its intensity refers to the intensity of the color associated to it. An example of a raw image is shown in figure [Fig.(6.2)]. The observed image is corrupted by noise due to three perturbation sources [143]: the photon noise, the dark noise and the spatial noise. The photon noise refers to the fluctuation of the number of photons that reache the sensor and its variance is a linear function of this number and thus linear with respect to the pixel intensity. The dark noise is generated by the leakage current and independent on the pixel intensity. The spatial noise has a quadratic dependency on the pixel intensity and its existence is due to the fact that pixels are not perfectly similar and behave in a different fashion. Under all these considerations, the noise model can be approximated by a Gaussian noise with a variance that is a quadratic function of the intensity. Thus if we note I^r (respectively U^r) the observed raw image (respectively the noise free-image), the noise variance for a given pixel depends on its intensity $U^r(\mathbf{x})$

$$\sigma_n^2(\mathbf{x}) = \alpha(U^r(\mathbf{x}))^2 + \beta U^r(\mathbf{x}) + \gamma \tag{6.1}$$

Starting from the raw observations, the final RGB image is obtained after a processing chain composed of different low level vision algorithms. These algorithms will impact the noise distribution as well as its variance behavior. Let us consider the noise-model observed at the RAW space and try to understand how it propagates through the processing chain that entails three major steps

- **White Balance Correction**
  It consists in multiplying each pixel in the raw image by a constant coefficient according to its color (Red, Green or Blue) to homogenize intensity in gray regions. We will define I^w the image obtained after white balance correction, thus for a given pixel $\mathbf{x}$ that corresponds to a color $c \in \{R, G, B\}$ the intensity and the noise variance are modified as

  $$I_c^w(\mathbf{x}) = \alpha_c I^r(\mathbf{x}) \quad \text{and} \quad \sigma_{n,w}^2(\mathbf{x}) = \alpha_c^2 \sigma_n^2(\mathbf{x})$$

  After this step, the noise variance depends not only on the intensity value of the pixel but also on its color. More explicitly, two pixels having the same intensity and same variance before

white balance application will have different variance if they represent different colors in the Bayer pattern.

- **Demosaicing**
  The demosaicing step consists in recovering the missing color information for each pixel. It computes a three-channel image (RGB) starting from a one channel image composed of the Bayer pattern. If one takes the example of a green pixel, the corresponding red and blue intensities must be interpolated. Demosaicing is also a challenging low-level vision task and some interpolation artifacts have to be addressed. To this end a lot of effort was devoted to design efficient algorithms that could preserve high frequencies and avoid color artifacts [79, 18, 29]. A general expression of the interpolation of the missing color value is

$$I_c^d(\mathbf{x}) = \int_{\Omega_c} w_{\mathbf{xy}} I^w(\mathbf{y}) d\mathbf{y} \tag{6.2}$$

  where I_c^d refers to the channel c of the image I^d obtained after demosaicing, Ω_c is the set of pixels where the intensity associated to the color c is available and measured by the sensor, and $w_{\mathbf{xy}}$ are weights that depend on the demoisaicing algorithm that is not necessarily linear. The equation (6.2) shows that after the demosaicing step, the noise variance depends also on pixel position and the interpolation coefficients. The demosaicing introduces a spatial correlation in the noise model. An exact noise model is therefore intractable unless one uses a linear algorithm of demosaicing which is not the case in practice.

- **Color transform and contrast enhancement**
  The color transform is applied to the image in order to obtain a color rendering close to the photographed scene. This transform is linear and consists in multiplying each vector $V_{\mathbf{x}} = (I_R^d(\mathbf{x}), I_G^d(\mathbf{x}), I_B^d(\mathbf{x}))^T$ by a 3×3 color matrix $\mathbf{M}$. Thus the covariance matrix associated to the noise affecting a given pixel $\mathbf{x}$ is

$$\Sigma_n(\mathbf{x}) = \mathbf{M} \cdot \Sigma_n^d(\mathbf{x}) \cdot \mathbf{M}^T \qquad \Sigma_n^d(\mathbf{x}) = \begin{pmatrix} \sigma_{n,R}^2(\mathbf{x}) & 0 & 0 \\ 0 & \sigma_{n,G}^2(\mathbf{x}) & 0 \\ 0 & 0 & \sigma_{n,B}^2(\mathbf{x}) \end{pmatrix} \tag{6.3}$$

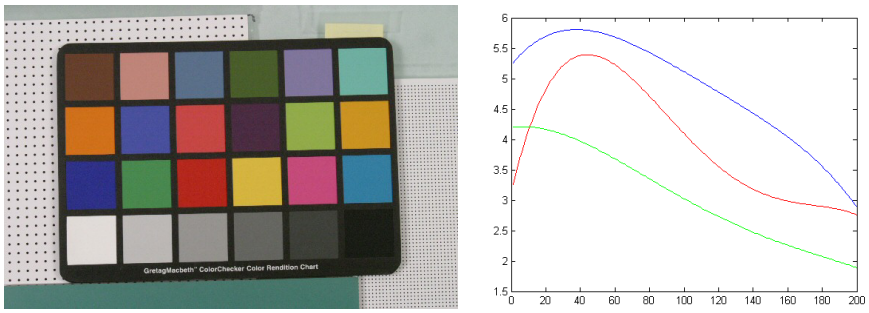
  Where $\sigma_{n,R}^2(\mathbf{x}), \sigma_{n,G}^2(\mathbf{x}), \sigma_{n,B}^2(\mathbf{x})$ are the noise variance relative to each color channel in the pixel $\mathbf{x}$. Σ_n^d is a diagonal matrix because we assume that in the raw image neighboring pixels (and thus color) are independent, which is a reasonable assumption in practice. Nevertheless, the resultant covariance matrix is not diagonal which means that the color transform introduces inter channel correlation between noise components.
  The final processing step is the Gamma correction to map the histogram of the image and

enhance the contrast mainly in dark regions. This is a non linear transform that adds a complexity level to the noise model.

The description of elementary steps in the image chain shows that the noise model is far from being additive white Gaussian. This hypothesis holds in the raw space, which makes it the ideal space to perform denoising. Nevertheless, the raw image as well as the transformation process are not always available and one has to find the noise model relying on the RGB image.

6.1.2 Noise Model Estimation

To perform a denoising task one needs a prior knowledge about the noise model. Among various research work toward noise modeling, we can cite as examples [96, 56] where the noise level function was computed based on one image. The noise level function associates to each intensity value the corresponding noise variance. The problem of noise estimation based on RGB images was addressed in [96] where a learning step using several camera transfer functions is performed to describe the space of noise level functions. This is done through a PCA analysis to determine an orthogonal basis of the space of noise level functions. To estimate the noise model, the image is divided into segments and for each segment the mean and the variance are computed to obtain a sample set of couples (intensity, variance). Next, based on these samples a maximum likelihood technique is used to estimate the coefficients of the noise level function in the learned basis. The experimental validation shows the potential of such an approach, nevertheless the estimation is reliable if the color distribution spans the full range of the spectrum and the image is composed of piecewise constant regions without textured content.
In our application, we propose a simpler approach that involves a weak calibration stage. To this end, we assume that some information on the image acquisition parameters is available (basically the camera gain value). The process consists in shooting the Macbeth Color Checker [Fig.(6.1)] that is composed of rectangular homogeneous regions with various colors and intensities. The obtained calibration image is composed of 24 patches corresponding to 24 different intensities per channel. For each calibration pattern the mean and the variance are computed. Contrarily to a natural image, finding the segment is a very simple task. Furthermore, the segment has a constant intensity and the only source of variation is the noise. An accurate estimation of the variance is thus possible mainly in case of patches of big size (about 100×100 pixel in practice). Using this image, one has 24 samples of the noise level function for each channel. To derive the missing noise variance relative to all possible intensities in the range $[0..255]$, an interpolation has to be performed. An example of the noise level function relative to the Canon 10D digital camera is provided for each channel in [Fig. (6.1)]. Note that the used technique for noise function



(a) (b)

Fig. 6.1: (a) Macbeth Color Checkers (b) Curve corresponding to the evolution of the standard deviation of the noise with respect to the intensity for each channel Red ,Green and Blue

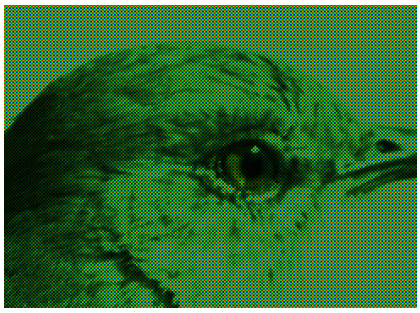


Fig. 6.2: Example of a raw Image with the Bayer Pattern

estimation requires only little knowledge about shooting conditions. Also, it does not require information about the camera transfer function which makes it more flexible.

6.1.3 The Denoising Algorithm

In this section, we assume that the image U and I are RGB images composed of 3 channels that are respectively (U_R, U_G, U_B) and (I_R, I_G, I_B). The regularization process is formulated in the following fashion

$$E_{reg}(U) = \int_\Omega \left\| \left[\frac{1}{Z(\mathbf{x})} \int_\Omega w_{\mathbf{xy}} U(\mathbf{y}) d\mathbf{y} \right] - U(\mathbf{x}) \right\|^2 d\mathbf{x} \tag{6.4}$$

$$
\begin{aligned}
E_{reg}(U) &= \int_\Omega \left(\left[\frac{1}{Z(\mathbf{x})} \int_\Omega w_{\mathbf{xy}} U_R(\mathbf{y}) d\mathbf{y} \right] - U_R(\mathbf{x}) \right)^2 d\mathbf{x} \\
&+ \int_\Omega \left(\left[\frac{1}{Z(\mathbf{x})} \int_\Omega w_{\mathbf{xy}} U_G(\mathbf{y}) d\mathbf{y} \right] - U_G(\mathbf{x}) \right)^2 d\mathbf{x} \\
&+ \int_\Omega \left(\left[\frac{1}{Z(\mathbf{x})} \int_\Omega w_{\mathbf{xy}} U_B(\mathbf{y}) d\mathbf{y} \right] - U_B(\mathbf{x}) \right)^2 d\mathbf{x} \qquad (6.5)
\end{aligned}
$$

The fidelity to data is ensured by imposing a constraint during the regularization process. This constraint relies on the non-parametric noise model that we estimated for each channel. Thus for any color channel $c \in \{R, G, B\}$, the constraints are

$$
\int_{\Gamma_c(GL)} (I_c(\mathbf{x}) - U_c(\mathbf{x}))^2 d\mathbf{x} = \sigma_n^2(GL) \qquad (6.6)
$$

Where GL is a gray level that ranges between 0 and 255, $\Gamma_c(GL)$ is the level line of the image U_c associated to the intensity GL, and $w_{\mathbf{xy}}$ are weights that are computed on the luminance image $Y = 0.3I_R + 0.6I_G + 0.1I_B$ and defined in (5.16). The choice of the luminance image to compute the weights is motivated by the fact that the most important image content and structure is present in this component.

The equivalent Euler Lagrange formulation using the Lagrange multipliers λ_i

$$
E_U = E_{Reg}(U) + \sum_{c \in \{R,G,B\}} \sum_{i=0}^{255} \lambda_i^c \int_{\Gamma_c(i)} (I_c(\mathbf{x}) - U_c(\mathbf{x}))^2 d\mathbf{x} \qquad (6.7)
$$

This energy formulation can be split in three independent terms (one for each color channel). Therefore we have three identical constrained optimization problems that are solved in an iterative fashion. Now if we consider a particular channel U_c ($c \in \{R, G, B\}$), and a set of $\{\lambda_i^c\}_{0 \le i \le 255}$, the optimal image is the one that corresponds to the steady state of the following PDE :

$$
\frac{\partial U_c}{\partial t}(\mathbf{x}) = \nabla E_{Reg}^c(\mathbf{x}) + \sum_{i=0}^{255} \lambda_i^c \delta(U_c(\mathbf{x}) - i) \left[U_c(\mathbf{x}) - I_c(\mathbf{x}) \right] \qquad (6.8)
$$

Where δ is Kronecker Delta function that is defined as $\delta(x) = 1$ if $x = 0$ $\delta(x) = 0$ otherwise. The gradient of the regularization term in the direction $U_c(\mathbf{x})$ is defined as

$$
\nabla E_{Reg}^c(\mathbf{x}) = 2 \int_\Omega \left(\int_\Omega \frac{w_{\mathbf{yz}}}{Z(\mathbf{z})} U_c(\mathbf{y}) d\mathbf{y} - U_c(\mathbf{z}) \right) \frac{w_{\mathbf{xz}}}{Z(\mathbf{z})} d\mathbf{z} + 2 \left(U_c(\mathbf{x}) - \int_\Omega \frac{w_{\mathbf{xy}}}{Z(\mathbf{x})} U_c(\mathbf{y}) d\mathbf{y} \right)
$$

Now in order to impose the constraint (6.6) on the final solution we have to update each λ_i^c value using the same technique used in [122, 62]. To this end we consider the steady state of equation

(6.8) where the following is verified

$$\begin{aligned}
\nabla E^c_{Reg}(\mathbf{x}) &+ \sum_{i=0}^{255} \lambda^c_i \delta(U_c(\mathbf{x}) - i)\,[U_c(\mathbf{x}) - I(\mathbf{x})] = 0 \\
[U_c(\mathbf{x}) - I_c(\mathbf{x})]\,\nabla E^c_{Reg}(\mathbf{x}) &+ \sum_{i=0}^{255} \lambda^c_i \delta(U_c(\mathbf{x}) - i)\,[U_c(\mathbf{x}) - I_c(\mathbf{x})]^2 = 0 \\
\int_{\Gamma_c(k)} [U_c(\mathbf{x}) - I_c(\mathbf{x})]\,\nabla E^c_{Reg}(\mathbf{x}) d\mathbf{x} &+ \lambda^c_k \int_{\Gamma_c(k)} [U_c(\mathbf{x}) - I_c(\mathbf{x})]^2\, d\mathbf{x} = 0
\end{aligned}$$

λ^c_k is then updated according to

$$\lambda^c_k = \frac{\int_{\mathbf{x}\in\Gamma_c(k)} [I_c(\mathbf{x}) - k]\,\nabla E^c_{Reg}(\mathbf{x}) d\mathbf{x}}{2\sigma^2_n(k)} \tag{6.9}$$

The data-fidelity term depends on the intensity and inversely proportional to the noise standard deviation. This is coherent with the idea that content with high noise level require more regularization than the one with small variance.

6.1.4 Experimental Results

In order to evaluate the performance of our method when applied to real digital camera noise, we have considered examples of noisy images produced by different digital cameras. The denoising process consists in alternating two steps: (i) minimizing the energy (6.7) using a gradient descent (ii) updating the λ_i values according to (6.9). The following set of parameters was selected for denoising : the time step for the gradient descent algorithm was set to $dt = 0.01$ and for each $0 \leq i \leq 255$, $\lambda_i = 0.05$. The weight definition is the distance between patches defined in (5.16) with a patch size $w_p = 5$, a fixed spatial bandwidth $\sigma_s = 3$, and $\sigma_{ph} = 3$.
We compared the result of the proposed method with the NLmean algorithm. For the RGB images, each color vector is restored by a weighted averaging of the neighboring ones using weights computed on the luminance image. For the weight computation, we considered the same definition and parameters used for our algorithm. Due to computational complexity, the neighborhood size was restrained to a 15×15 window. We must note that the parameters setting of each method was fixed in order to suppress the maximum amount of noise while preserving details and texture. For comparison, we considered also the result of our proposed approach for the unconstrained formulation (without updating the values of λ_i).
In figure [Fig.(6.3), Fig.(6.4)], we can see that the fixed noise parameter approach results on an oversmoothing the white texture as shown by the details present in the residual image [(6.4)]. Such

a limitation is dealt with when considering the noise variation with respect to the intensity. The NLmean restoration result is not homogeneous: some texture regions are well preserved but other are not. Other examples were presented in figure [Fig.(6.5), Fig.(6.6)] with a different camera (Nikon D70). We noticed that with a variable noise variance, the textured white strips in the image are preserved. This behavior is related to expression (6.9) where λ_i values are inversely proportional to the noise standard deviation. This confirms that a better denoising must take into account the evolution of noise with image intensity.

6.1.5 Discussion

In this section, it was shown that the denoising formulation presented in chapter (5) can be extended to deal with more realistic noise characteristics. Considering the appropriate noise variance ensures better details and texture preservation. We did not consider a variable bandwidth kernel for weight computation due to computational complexity. Investigating alternative color spaces and color representations like the YCrCb to account for correlation between channels instead of treating them in an independent manner is a promising direction. Furthermore, such an approach would naturally deal with color artifacts. As far as noise properties are concerned, considering other statistics like the autocorrelation function could bring an important added value to the denoising process.

6.2 Application to DTI estimation and regularization

6.2.1 Introduction

Diffusion tensor imaging (DTI) is an emerging non-invasive modality allowing the quantitative investigation of water protons diffusion within biologic tissues. Since diffusion is sensitive to the presence of organized structures, DTI is used mostly in brain studies and has become a tool to infer white matter connectivity [89]. Such a modality offers measurements of the amount of diffusion of water molecules in several different directions. One then can infer the estimation of a tensor which is a 3×3 symmetric positive definite matrix representing the uncertainty on the position of water protons with a Gaussian model of displacement. The (DTI) imaging protocol produces 3D weighted images for different gradient directions and based on these image one can estimate the underlying diffusion tensors. The DTI experimental protocol yields noisy observations due to the diffusion-sensitizing magnetic gradient. Furthermore, the clinical protocols refer to relatively low

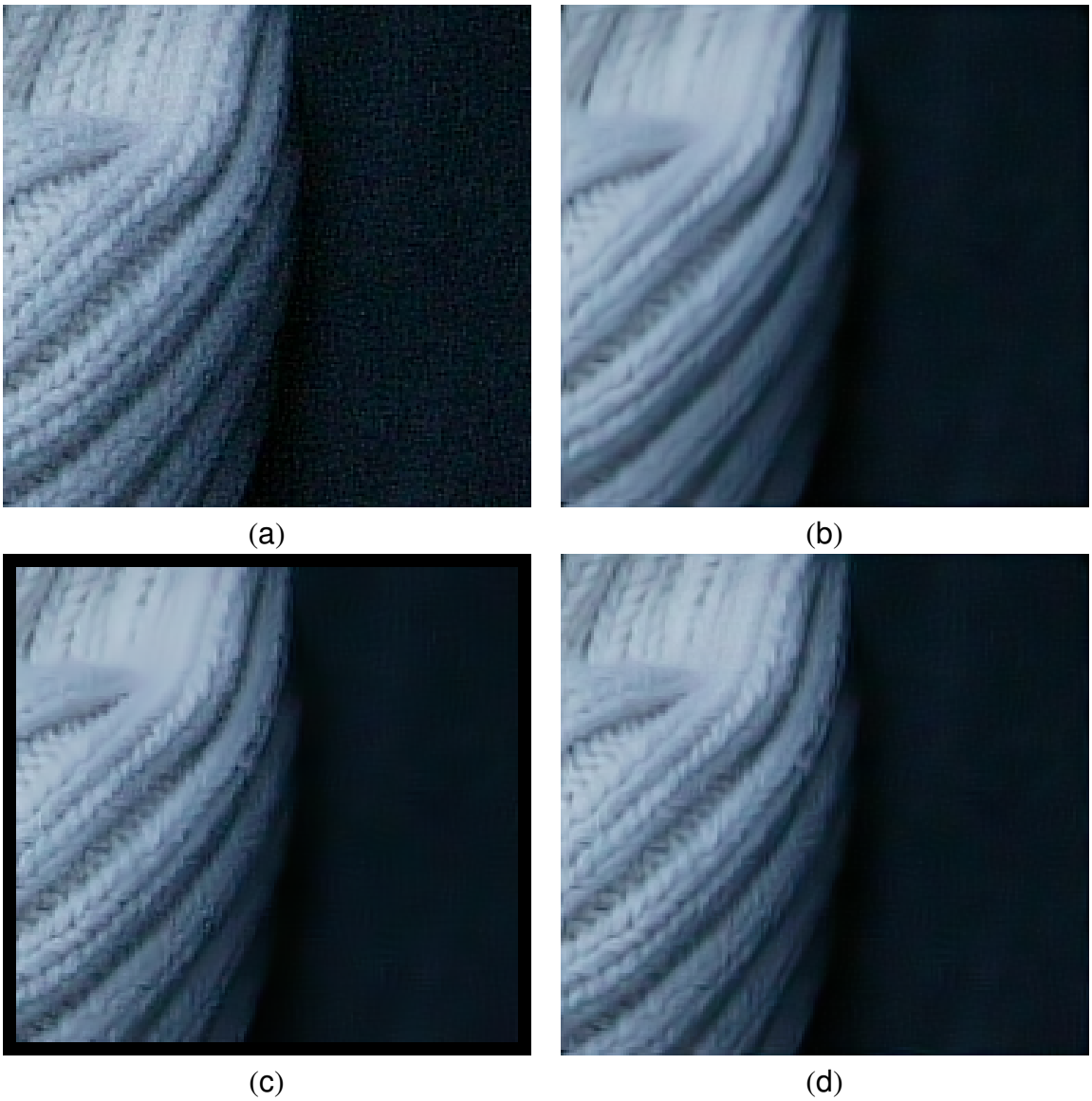


Fig. 6.3: (a) Original noisy image (b) Image restored using our method with fixed noise variance (c) Image restored using the NLmean algorithm (d) Image restored using our method with variable noise variance

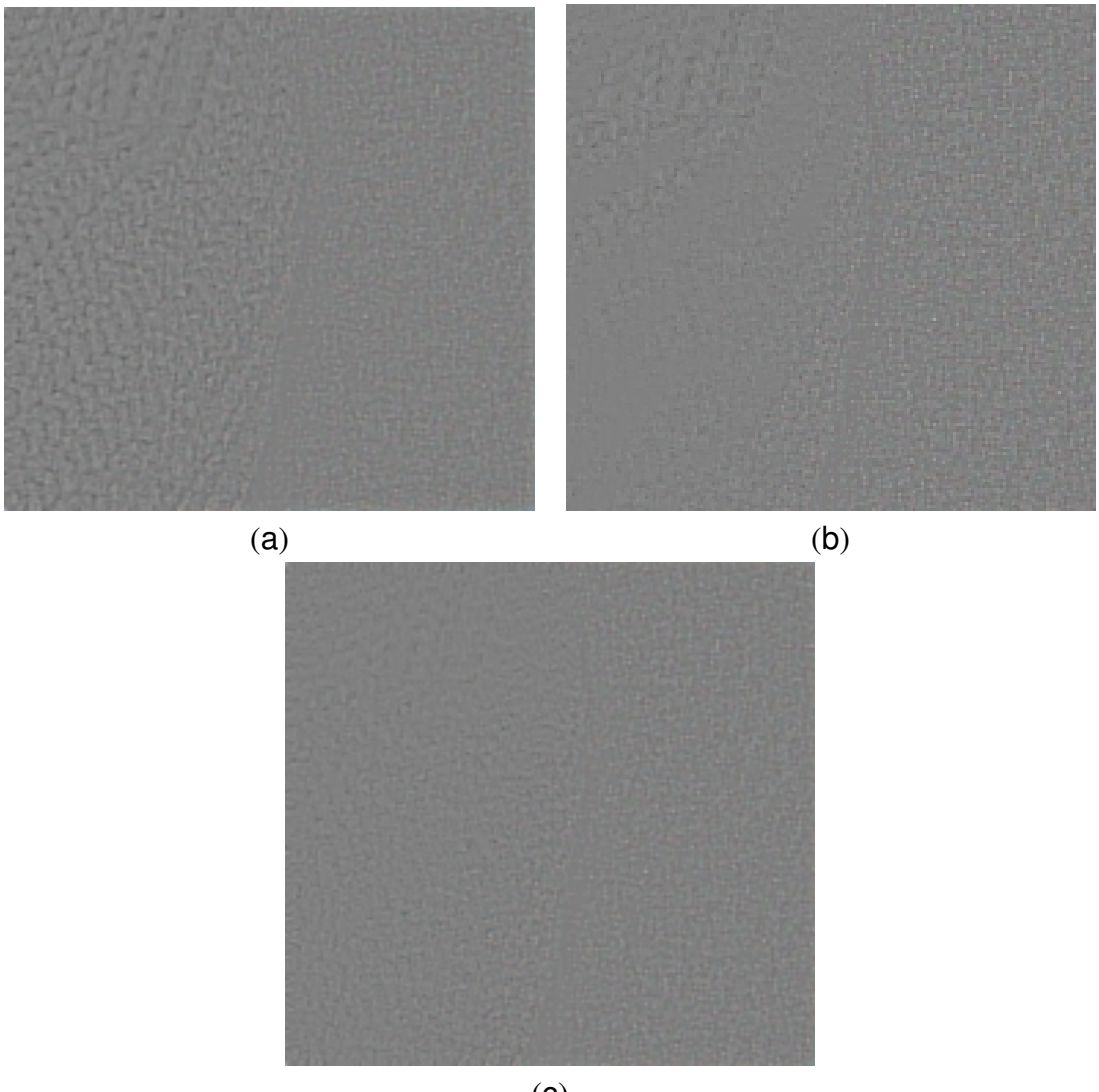


Fig. 6.4: Difference between the noisy image and the restored one using (a) our method with fixed noise variance (b) Image restored using the NLmean algorithm (c) Image restored using our method with variable noise variance

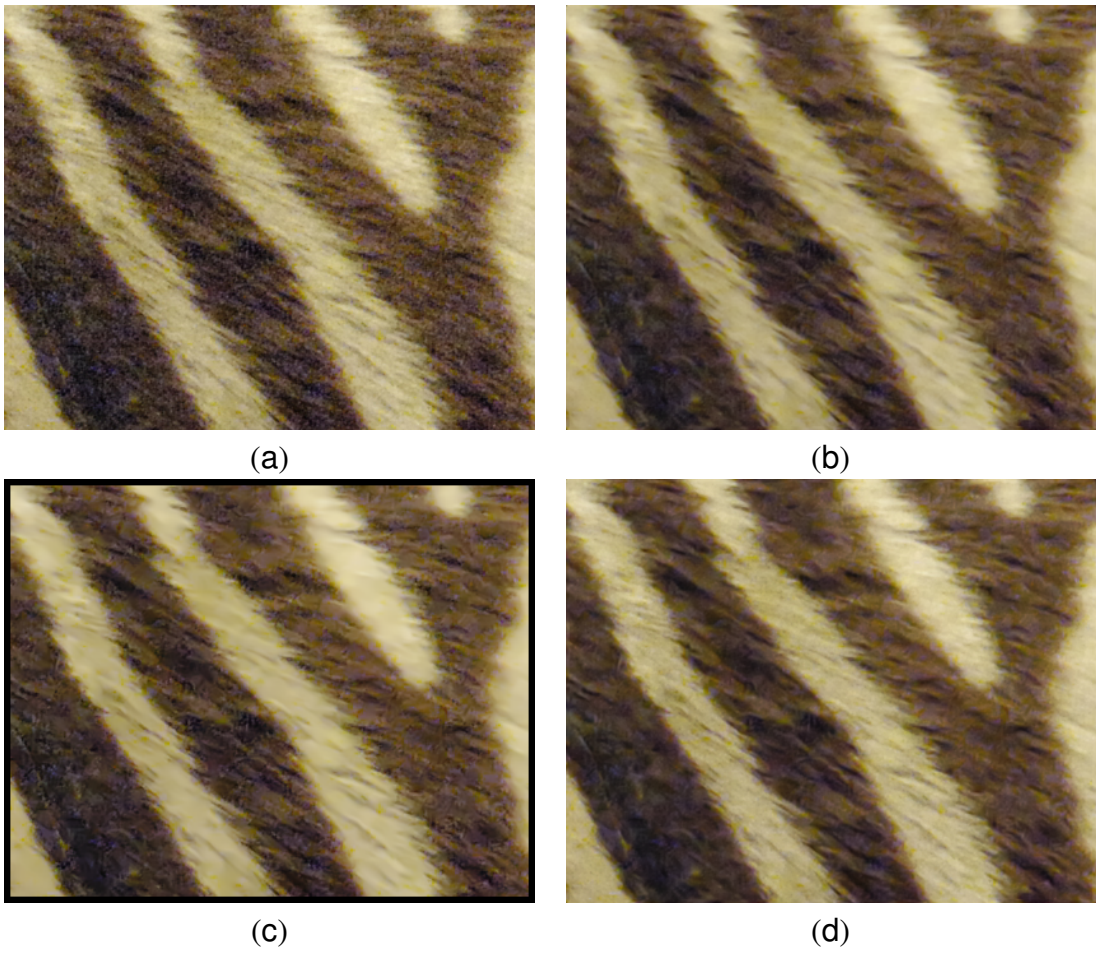


Fig. 6.5: (a) Original noisy image (b) Image restored using our method with fixed noise variance (c) Image restored using the NLmean algorithm (d) Image restored using our method with variable noise variance

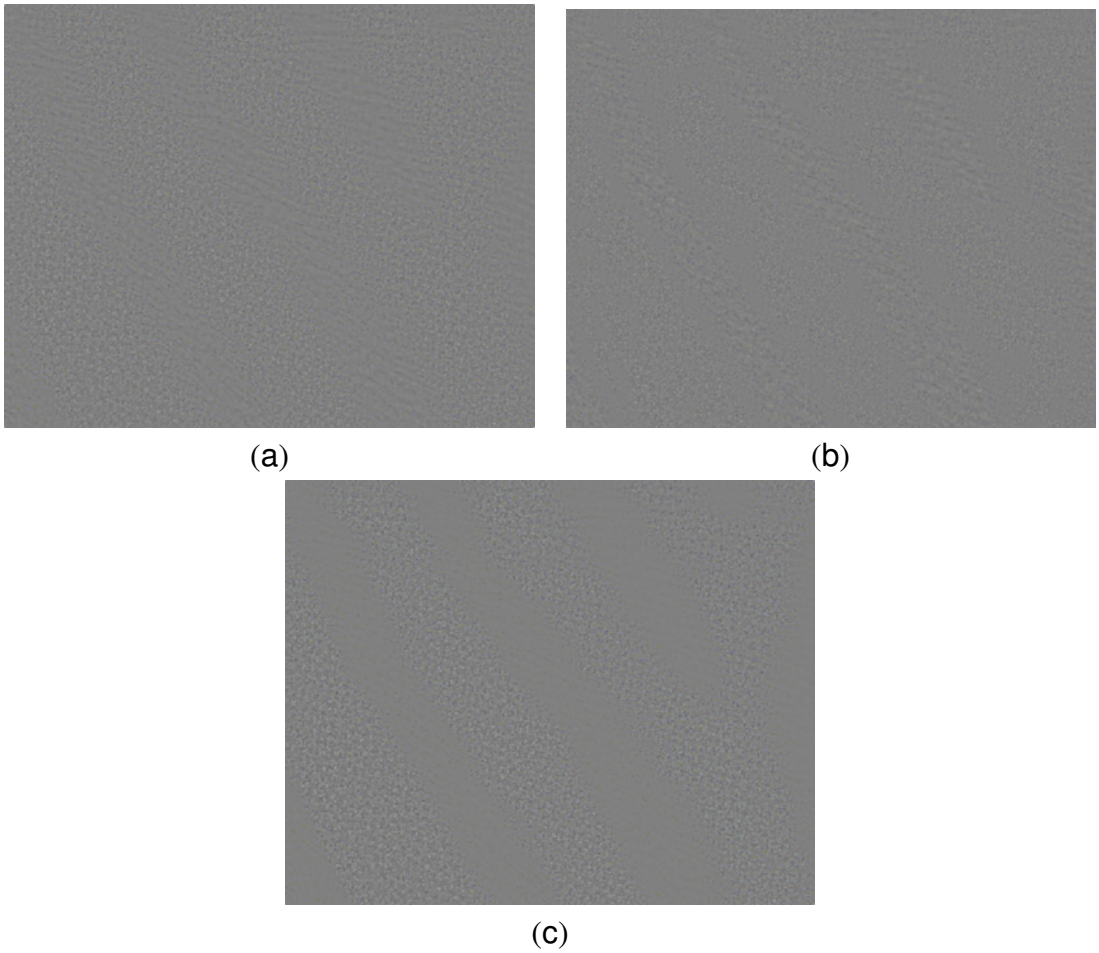


Fig. 6.6: Difference between the noisy image and the restored one using (a) our method with fixed noise variance (b) Image restored using the NLmean algorithm (c) Image restored using our method with variable noise variance

magnet strength, or a rather low signal-to-noise ratio. Therefore, signal reconstruction is crucial to obtain an appropriate estimate of the tensor field and for subsequent use of this estimate in applications like fiber tractography.

Several methods have been proposed to address diffusion tensor regularization. In [40], a two-step regularization was proposed consisting of the restoration of the principal diffusion directions using a total variation-model followed by the smoothing of the eigenvalues using an anisotropic tensor-driven formulation. In [17], the maximization of a log-posterior probability based on the Rician noise model is considered to smooth directly the diffusion-weighted images. A Bayesian model based on a Gaussian Markov Random Field was used in [105] to smooth the diffusion tensors. In [32], the authors consider the tensors as lying on a Riemannian manifold and use the corresponding distance to derive a local weighted averaging for DTI denoising. Tensors are assumed to be positive-definite matrices which was taken into account in [43] where an anisotropic filtering of the L^2 norm of the gradient of the diffusion tensor was considered and their proposed PDE scheme constrains the estimation to lie on this space. Such a concept was further developed in [144] where a variational method was proposed that aimed to minimize the L^p norm of the spatial gradient of the diffusion tensor under a constraint involving the non-linear form of Stejskal-Tanner equation. A non linear diffusion scheme is described in [147] where smoothing is made direction-dependent using a diffusion matrix in the PDE system. More recently, in [55] a joint reconstruction and regularization was proposed in the context of an energy minimization in a Log-Euclidean framework.
The existing variational methods and PDE based ones, involve regularization term that is a function of the norm of the gradient, which limits their ability to detect efficiently data structure and to preserve its variability. Besides, some of the methods above cited perform only regularization while it is interesting estimate and regularize tensors simultaneously. Furthermore, some of the proposed cost functional are non-convex which entails a preliminary initialization step.

In this section we propose an approach to jointly estimate and regularize diffusion tensor fields. We use a convex energy functional which combines the linearized form of Stejskal-Tanner equation as a data fidelity term. The regularization term is the one introduced in chapter (5) but we will provide the appropriate weights definition. The experimental results presented in this part on synthetic datasets as well as on real diffusion weighted images show the potential of the proposed method.

6.2.2 DTI Estimation and Regularization

Let us assume that n DTI acquisitions $(S_k)_{k=1...n}$ with respect to different magnetic gradient directions $(\mathbf{g}_k)_{k=1...n}$ are available. Ideally, the expected signal at a voxel $\mathbf{x}$ for the direction k as explained in [131] should respect the following condition

$$S_k(\mathbf{x}) = S_0(\mathbf{x}) \exp\big(- b\mathbf{g}_k^t \mathbf{D}(\mathbf{x}) \mathbf{g}_k \big) \tag{6.10}$$

with the tensor $\mathbf{D}$ being the unknown variable and b is a constant value that depends on the acquisition settings. The estimation of the tensors in the volume domain Ω can be done through direct inference (6 acquisitions are at least available), which is equivalent to minimizing:

$$E_{data}(\mathbf{D}) = \int_\Omega \sum_{k=1}^{n} \Big(\log\big(S_k(\mathbf{x})/S_0(\mathbf{x})\big) + b\mathbf{g}_k^t \mathbf{D}(\mathbf{x}) \mathbf{g}_k \Big)^2 d\mathbf{x} \tag{6.11}$$

This energy is based on the linearized diffusion tensor model which is reasonable for moderate values of SNR [124]. Such a direct estimation is quite sensitive to noise, on the other hand, it refers to a convex term, which is rather convenient when seeking its lowest potential. The most common approach to account for noise is through the use of an additional regularization term which constrains the estimation of $\mathbf{D}$ to be locally smooth. Like the natural image case we assume that, the tensor can be expressed as a linear combination of the tensors lying in its neighborhood since they are likely to represent the same population of fibers. This assumption still holds at the boundaries between different groups of fibers as long as the linear combination is thoroughly chosen to ensure that the contribution of tensors belonging to a different fiber population is negligible. This leads us to define the following regularization component:

$$E_{Reg}(\mathbf{D}) = \int_\Omega \left|\left| \mathbf{D}(\mathbf{x}) - \frac{1}{Z(\mathbf{x})} \int_{\mathbf{y}\in\mathcal{N}_\mathbf{x}} w_{\mathbf{xy}} \mathbf{D}(\mathbf{y}) d\mathbf{y} \right|\right|_F^2 d\mathbf{x} \tag{6.12}$$

where $w_{\mathbf{xy}}$ reflects the similarity between tensors $\mathbf{D}(\mathbf{x})$ and $\mathbf{D}(\mathbf{y})$, $||A||_F$ being the Frobenius norm $||A||_F = \sqrt{tr(A^t A)}$ and $Z(\mathbf{x})$ is a normalization factor, i.e $Z(\mathbf{x}) = \int_{\mathbf{y}\in\mathcal{N}_\mathbf{x}} w_{\mathbf{xy}} d\mathbf{y}$. The most critical aspect of such an approximation model is the definition of weights, measuring the similarity between tensors within the local neighborhood. The use of Gaussian weights is a common weight's selection, i.e $\left[w_{\mathbf{xy}} = e^{\frac{-d^2(\mathbf{D}(\mathbf{x}),\mathbf{D}(\mathbf{y}))}{2\sigma^2}} \right]$, where $d(.;.)$ is a distance between tensors and σ a scale factor.

Therefore, in order to perform a joint estimation and regularization of the tensors filed one has to

minimize the following energy function

$$E = E_{Reg} + \lambda E_{data} \tag{6.13}$$

In the context of direct estimation and regularization it is more appropriate to define similarities directly on the observation space rather than the estimation space. Such a choice will lead to a tractable estimation, while preserving the convexity of the cost function. Our distance definition as well as our minimization step are based on the representation of symmetric positive semi-definite matrices S_+^3 as a convex closed cone in the Hilbert space of symmetric matrices S^3, where the standard scalar product is defined by $\langle A, B\rangle_F = tr(A^t B)$ which induces the corresponding Frobenius norm.

Measuring Similarities from diffusion weighted images

We aim at simultaneously estimating and smoothing the tensor field, therefore the weights $w(\mathbf{x}, \mathbf{y})$ in E_{smooth} should be precalculated using the raw data. The most straightforward estimation of the distances can be done through the algebric distance between the $\log(S_k/S_0)$ for two neighborhood voxels in any direction

$$d\big(\mathbf{D}(\mathbf{x}), \mathbf{D}(\mathbf{y})\big) = \frac{1}{b}\sqrt{\sum_{k=1}^{N}\Big(\log\big(S_k(\mathbf{x})/S_0(\mathbf{x})\big) - \log\big(S_k(\mathbf{y})/S_0(\mathbf{y})\big)\Big)^2} \tag{6.14}$$

One can easily show that such an expression does not reflect similarity between tensors according to the norm $||.||_F$. In fact, this leads to

$$d\big(\mathbf{D}(\mathbf{x}), \mathbf{D}(\mathbf{y})\big) = \sqrt{\sum_{k=1}^{N}\Big(\mathbf{g}_k^t\big(\mathbf{D}(\mathbf{x}) - \mathbf{D}(\mathbf{y})\big)\mathbf{g}_k\Big)^2} = \sqrt{\sum_{k=1}^{N} < \mathbf{D}(\mathbf{x}) - \mathbf{D}(\mathbf{y}), \mathbf{G}_k >_F^2} \tag{6.15}$$

where $\{\mathbf{G}_k = \mathbf{g}_k\mathbf{g}_k^t\}_{1\leq k\leq N}$ do not form necessarily an orthonormal basis. We use a Gram-Schmidt orthogonalization scheme to calculate an orthonormal basis $\left\{\widetilde{\mathbf{G}_k}\right\}_{1\leq k\leq N}$ such that $\widetilde{\mathbf{G}_k} = \sum_l \alpha_{kl}\mathbf{G}_l$ (each new vector of the new basis is a linear combination of the vectors of the initial basis). This procedure allows us to have an approximation of $||\mathbf{D}(\mathbf{x}) - \mathbf{D}(\mathbf{y})||_F$ directly from the raw data S_k

and S_0 as follows

$$
\begin{aligned}
||\mathbf{D}(\mathbf{x}) - \mathbf{D}(\mathbf{y})||_F &= \sqrt{\sum_{k=1}^{N} < \mathbf{D}(\mathbf{x}) - \mathbf{D}(\mathbf{y}), \widetilde{\mathbf{G}_k} >_F^2} \qquad (6.16)\\
&= \frac{1}{b}\sqrt{\sum_{k=1}^{N} \Big(\sum_{l} \alpha_{kl} (\log\big(S_k(\mathbf{x})/S_0(\mathbf{x})\big) - \log\big(S_k(\mathbf{y})/S_0(\mathbf{y})\big)) \Big)^2}
\end{aligned}
$$

Semi-Definite Positive Gradient Descent

One now can seek the lowest potential of the cost function towards recovering the optimal solution on the tensor space. The present framework consists of a convex energy with a unique minimum which can be reached using a projected gradient descent on the space of semi-definitive positive matrices. The projection from S^3 onto S^3_+ denoted by $\Pi_{S^3_+}$ is well defined and has an explicit expression. Indeed, projecting M amounts to replacing the negative eigenvalues in its spectral decomposition by 0 [136, 67]. Note that we minimize over the set of semi-definite positive matrices because it is topologically closed, as opposed to the set of definite positive matrices. In the current setting, the problem is well posed and the projected gradient descent algorithm is convergent for a suitable choice of the time step dt. The gradient descent can be expressed as the following equation

$$
\begin{aligned}
\mathbf{D}^{t+1}(\mathbf{x}) &= \Pi_{S^3_+}\Big(\mathbf{D}^t(\mathbf{x}) - dt\frac{\partial E}{\partial \mathbf{D}(\mathbf{x})}(\mathbf{D}^t)\Big) \qquad (6.17)\\
&= \Pi_{S^3_+}\Big(\mathbf{D}^t(\mathbf{x}) - dt\frac{\partial E_{smooth}}{\partial \mathbf{D}(\mathbf{x})}(\mathbf{D}^t) - dt\lambda\frac{\partial E_{data}}{\partial \mathbf{D}(\mathbf{x})}(\mathbf{D}^t)\Big) \qquad (6.18)
\end{aligned}
$$

where

$$
\begin{aligned}
\frac{\partial E_{Reg}}{\partial \mathbf{D}(\mathbf{x})}(\mathbf{D}) &= 2\mathbf{D}(\mathbf{x}) - 2\int_{\mathbf{y}\in\mathcal{N}_\mathbf{x}} \frac{w_{\mathbf{xy}}}{Z(\mathbf{x})}\mathbf{D}(\mathbf{y})d\mathbf{y}\\
&\quad - 2\int_{\mathbf{y}\in\mathcal{N}_\mathbf{x}} \frac{w_{\mathbf{xy}}}{Z(\mathbf{y})}\Big(\mathbf{D}(\mathbf{y}) - \int_{\mathbf{z}\in\mathcal{N}_\mathbf{y}} \frac{w_{\mathbf{yz}}}{Z(\mathbf{y})}\mathbf{D}(\mathbf{z})d\mathbf{z}\Big)d\mathbf{y} \qquad (6.19)\\
\frac{\partial E_{data}}{\partial \mathbf{D}(\mathbf{x})}(\mathbf{D}) &= 2b\sum_{k=1}^{N}\Big(\log\big(S_k(\mathbf{x})/S_0(\mathbf{x})\big) + b\mathbf{g}_k^t\mathbf{D}(\mathbf{x})\mathbf{g}_k\Big)\mathbf{G}_k \qquad (6.20)
\end{aligned}
$$

Let us define the norm $||.||_{TF}$ over the whole tensor field $\mathbf{D}$ as $||\mathbf{D}||_{TF} = \int_\Omega ||\mathbf{D}(\mathbf{x})||_F d\mathbf{x}$. Considering two tensor fields $\mathbf{D}_1$ and $\mathbf{D}_2$, we show in the following that the gradient of our energy functional is L-Lipschitz. The constant L will allow us to choose automatically a time step that

insures the convergence of the algorithm.

$$\begin{aligned}\left\|\frac{\partial E_{data}}{\partial \mathbf{D}(\mathbf{x})}(\mathbf{D}_1)-\frac{\partial E_{data}}{\partial \mathbf{D}(\mathbf{x})}(\mathbf{D}_2)\right\|_F &= \left\|2b^2\sum_{k=1}^{N}\left[\mathbf{g}_k^t\left(\mathbf{D_1}(\mathbf{x})-\mathbf{D_2}(\mathbf{x})\right)\mathbf{g}_k\right]\mathbf{G}_k\right\|_F \\ &\leq 2b^2\sum_{k=1}^{N}\|<\mathbf{G}_k,\mathbf{D}_1(\mathbf{x})-\mathbf{D}_2(\mathbf{x})>_F\mathbf{G}_k\|_F \\ &\leq 2b^2\sum_{k=1}^{N}|<\mathbf{G}_k,\mathbf{D}_1(\mathbf{x})-\mathbf{D}_2(\mathbf{x})>_F|\,\|\mathbf{G}_k\|_F \\ &\leq 2b^2\sum_{k=1}^{N}||\mathbf{G}_k||_F^2||\mathbf{D}_1(\mathbf{x})-\mathbf{D}_2(\mathbf{x})||_F \qquad (6.21)\end{aligned}$$

Therefore $\|\nabla E_{data}(\mathbf{D}_1)-\nabla E_{data}(\mathbf{D}_2)\|_{TF}\leq 2b^2\sum_{k=1}^{N}||\mathbf{G}_k||_F^2||\mathbf{D}_1-\mathbf{D}_2||_{TF}$. Besides, we can easily show the following inequality

$$\|\nabla E_{Reg}(\mathbf{D}_1)-\nabla E_{Reg}(\mathbf{D}_2)\|_{TF}\leq 2(1+2|\mathcal{N}_\mathbf{x}|+|\mathcal{N}_\mathbf{x}|^2)||\mathbf{D}_1-\mathbf{D}_2||_{TF}$$

where $|\mathcal{N}_\mathbf{x}|$ is the number of the considered neighbors. Thus the gradient of the objective function is L-Lipschitz with $L=2\lambda b^2\sum_{k=1}^{N}||\mathbf{G}_k||_F^2+2(|\mathcal{N}_\mathbf{x}|+1)^2$. Choosing $0<dt<\frac{1}{\lambda b^2\sum_{k=1}^{N}||\mathbf{G}_k||_F^2+(|\mathcal{N}_\mathbf{x}|+1)^2}$ makes the projected gradient descent convergent [21].

We can give an interpretation of our regularization energy in terms of diffusion-weighted images smoothing. It can be easily verified that for each direction k

$$\begin{aligned}&\int_\Omega <\mathbf{D}(\mathbf{x})-\int_{\mathbf{y}\in\mathcal{N}_\mathbf{x}}\frac{w_{\mathbf{xy}}}{Z(\mathbf{x})}\mathbf{D}(\mathbf{y})d\mathbf{y},\mathbf{G}_k>_F^2 d\mathbf{x}= \\ &\frac{1}{b^2}\int_\Omega\left[\log\Big(\frac{S_k(\mathbf{x})}{S_0(\mathbf{x})}\Big)-\log\Big(\prod_{\mathbf{y}\in\mathcal{N}_\mathbf{x}}\big(\frac{S_k(\mathbf{y})}{S_0(\mathbf{y})}\big)^{\frac{w_{\mathbf{xy}}}{Z(\mathbf{x})}}\Big)\right]^2 d\mathbf{x}\end{aligned} \qquad (6.22)$$

Using Cauchy-Schwartz inequality we obtain :

$$\frac{1}{b^2}\int_\Omega\left[\log\Big(\frac{S_k(\mathbf{x})}{S_0(\mathbf{x})}\Big)-\log\Big(\prod_{\mathbf{y}\in\mathcal{N}_\mathbf{x}}\big(\frac{S_k(\mathbf{y})}{S_0(\mathbf{y})}\big)^{\frac{w_{\mathbf{xy}}}{Z(\mathbf{x})}}\Big)\right]^2 d\mathbf{x}\leq E_{Reg}||\mathbf{G}_k||_F^2$$

We can see that minimizing E_{smooth} has a direct implication on the normalized diffusion weighted images $\frac{S_k}{S_0}$. Reconstructing the tensors using a linear combination of the tensors in its neighborhood leads to the reconstruction of the normalized signals using a weighted geometric mean of the neighboring signals where the weights are not calculated only with a single volume S_k but also with the volumes obtained from the other magnetic gradient directions.

6.2.3 Experimental Validation

In order to validate the performance of the method we (i) have generated artificial tensor volumes corrupted with synthetic noise, (ii) used manual segmentation on T1 muscle images and tried to improve the separability of classes in the DTI space after regularization.

Artificially Corrupted Tensors

Let us consider two volumes. One is a $20 \times 20 \times 20$ lattice that consists of two homogeneous tensor fields. For this volume, the tensor fields for each region are

$$\mathbf{D_1} = 0.001 \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 0.5 \end{pmatrix} \quad \mathbf{D_2} = 0.001 \times \begin{pmatrix} 0.2 & 0 & 0 \\ 0 & 0.4 & 0 \\ 0 & 0 & 0.2 \end{pmatrix}$$

The second data set is a helix in which the internal voxels are anisotropic and the external ones are spherical [Fig.6.8-b]. We considered for both datasets a field strength $b = 700\, s.mm^{-2}$, a constant value $S_0 = 60$ for all volume voxels. The diffusion gradient directions, which are used to generate the DTI volume corresponding to the described tensors are the columns of the following matrix

$$\mathcal{G} = \begin{pmatrix} 1 & 1 & 1 & 1 & 0.41 & 0.41 & 0.41 & 0.41 & 0.41 & 0.41 & 0.41 & 0.41 \\ 0.41 & -0.41 & -0.41 & 0.41 & 0.41 & 1 & 1 & 0.41 & -0.41 & -1 & -1 & -0.41 \\ -0.41 & -0.41 & 0.41 & 0.41 & 1 & 0.41 & -0.41 & -1 & -1 & -0.41 & 0.41 & 1 \end{pmatrix}$$

The images were generated according to the Stejskal-Tanner equation (6.10) and then corrupted with a white zero-mean Gaussian noise forming a data set where ground-truth on the tensor are available.

For comparison we considered the regularization algorithm presented in [34] on noisy tensors that were estimated by minimizing E_{data} defined in (6.11). This approach is based on a variational formulation where a classical anisotropic Φ-function of the spatial gradient of the tensor field is used as a smoothing term. It is minimized using coupled and constrained PDE's with an adequate gradient flow that respects the geometry of the manifold of symmetric positive definite matrices. The flow considered for comparison is the rank preserving flow that imposes the definite positiveness on the tensors. The parameters of the method were tuned to obtain an optimal performance. As far as our method is concerned, the following parameters were used: $\lambda = 50$, $\mathcal{N}_{\mathbf{x}} = 3 \times 3 \times 3$, $dt = 10^{-7}$. To evaluate the performance of these methods, we considered the average sum of squared differences (**SSD**) between the regularized tensors and ground truth ones. Table (6.1)

demonstrates that our estimation and regularization approach achieves better results in terms of (**SSD**) than the approach presented in [34] mainly when the noise variance is important. The definition of the weights ensures a better interaction between tensors and preserves their variability. Indeed, the estimation of a noise free tensor involves only similar neighboring tensors. On the other hand, the anisotropic diffusion regularization relies on gradient information which is not robust when the noise variance is important with respect to the gradient magnitude.

In order to assess qualitatively our algorithm, we reported in [Fig.(6.7), Fig.(6.8)] the resulting

	Helix dataset			Homogeneous regions		
σ_n	0.5	1.2	3	1.5	4	9
Noisy tensor	1.08	6.24	39.54	9.82	71.25	393.38
Method in [34]	0.33	1.60	10.57	3.32	22.47	120.70
Our Method	0.41	1.38	3.78	0.44	4.23	18.30

Tab. 6.1: Average Sum of Square Differences (SSD)$\times 10^4$. Comparisons between our method and the one in [43]

tensors using our regularization method and the constrained anisotropic one. We can observe that our method achieves a better restoration, even in the presence of a high noise level.

DTI towards Understanding the Human Skeletal Muscle

In order to perform validation on real data, the following experiment was considered. DTI acquisitions of human skeletal muscle (calf) using 12 directions were carried out on a 1.5 T MRI scanner with $b = 700\ s.mm^{-2}$. In order to improve the signal-to-noise ratio, the acquisition was repeated thirteen times (one can use the average of the measurements) while a high resolution T1-weighted volume was also obtained and manually segmented [Fig.(6.9)].

In order to proceed with an evaluation of the proposed method, we consider the classification result on the regularized tensors and we compare it to the ground truth provided by the manual segmentation on the T1-weighted volume. The muscles that were considered in our experiments are the soleus (SOL), lateral gastrocnemius (LG) and the medial gastrocnemius (MG). The classification process is based on weak linear classifiers (in our case a multi-class linear SVM [73]) separating each class of muscle versus all others. Then, the success rate (percentage of voxels being attributed to the right class) from the classifier with respect to the ground truth was determined. In our validation process, we have performed the classification test thirteen times (for each volume) on the tensors computed using : (i) direct estimation without any regularization (**DE**), (ii) direct estimation and regularization (**DER**). An average correct classification rate was computed for the thirteen experiences for noisy tensors as well as regularized ones. Besides, we performed

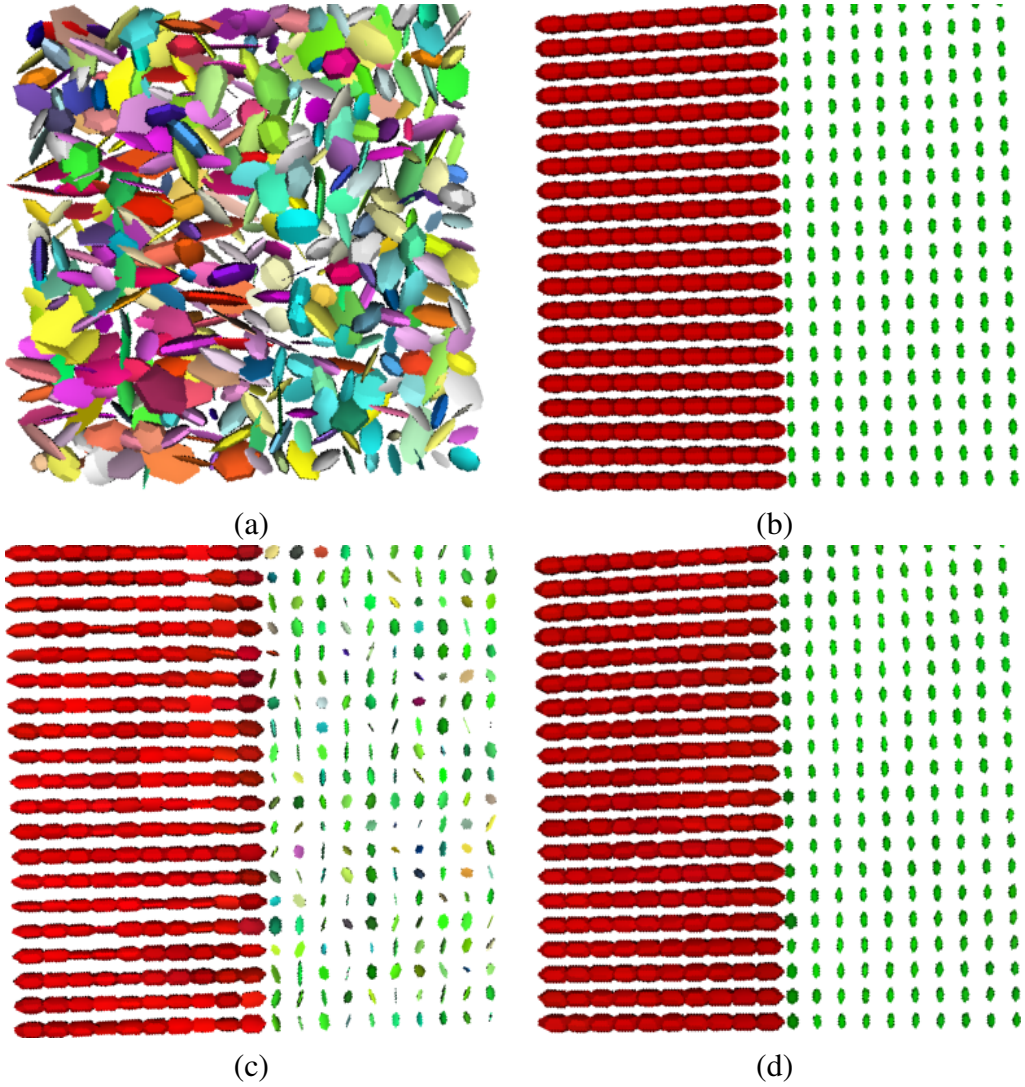


Fig. 6.7: Tensors on a volume slice (Homogeneous tensor field): (a) Noisy tensors (b) Ground-truth (c) Result obtained with [43] (d) Result obtained with our method

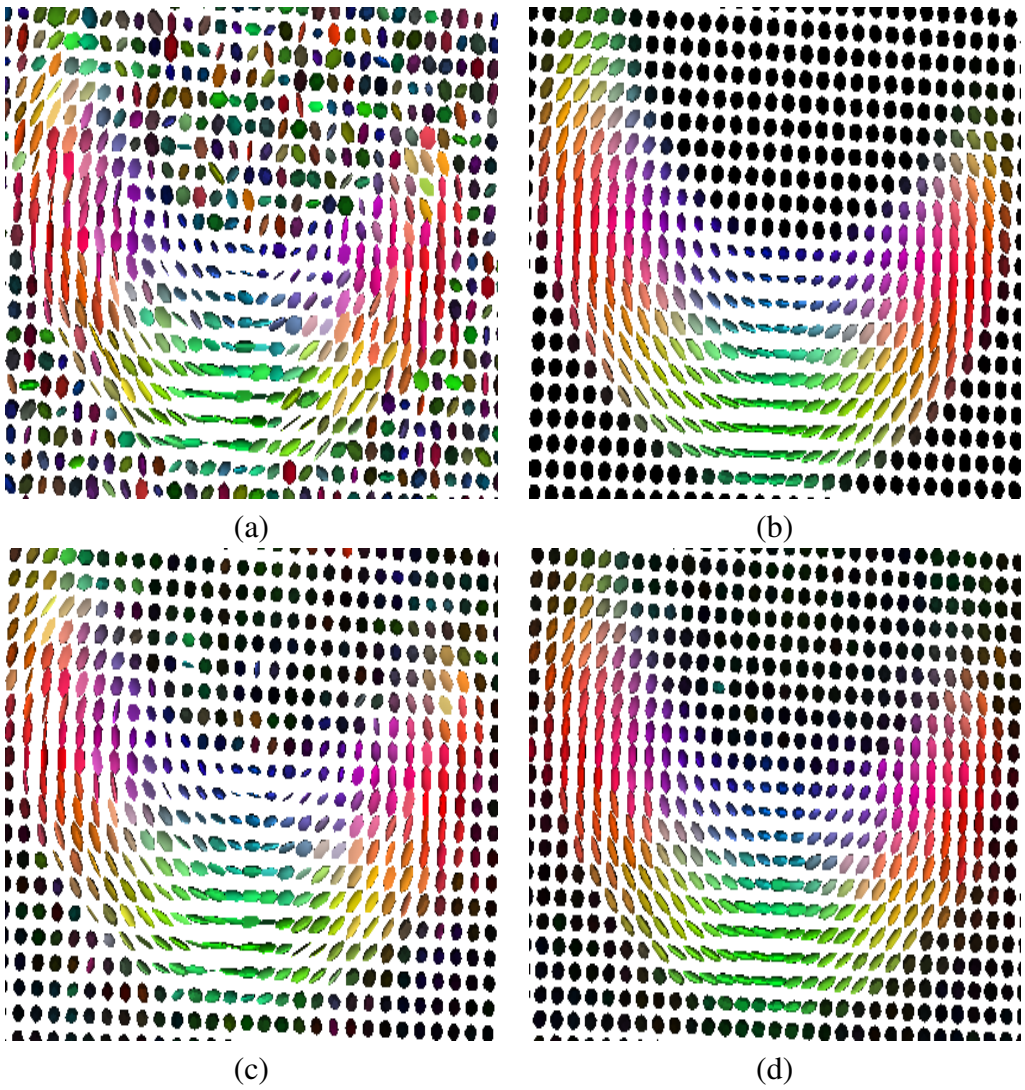


Fig. 6.8: Tensors on a volume slice (helix): (a) Noisy tensors (b) Ground-truth (c) Result obtained with [43] (d) Result obtained with our method

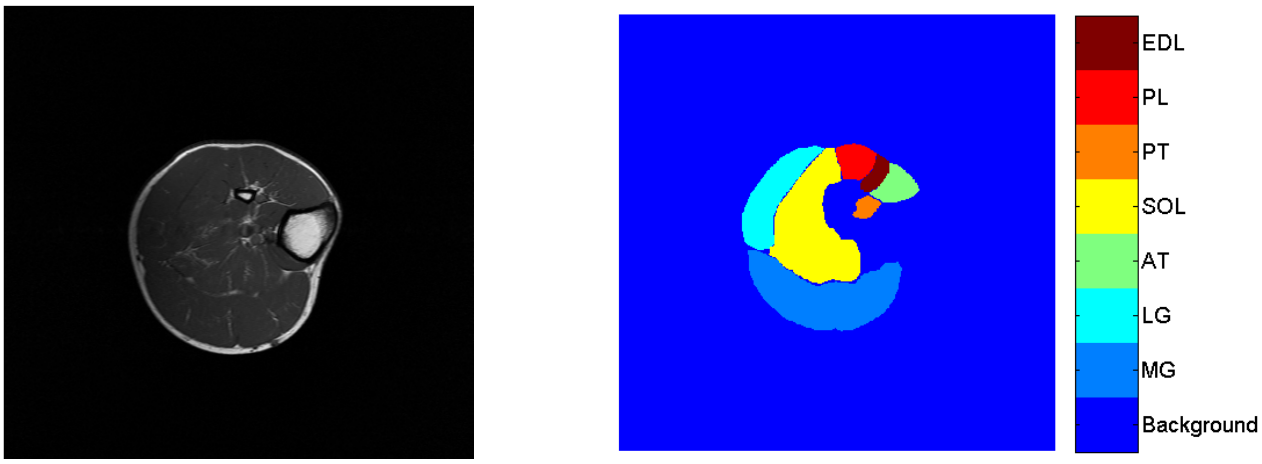


Fig. 6.9: A slice of the T1-weighted volume, different muscle groups segmented manually

classification using tensor field estimated on the average of the thirteen acquisitions (**ADE**). One would expect that since muscles consist of myo-fibers of the same nature, the classification should be improved if the estimation of the tensors is properly done, i.e. with appropriate regularization. In [Table 6.2], we present quantitative validation of the present framework for the considered muscles. One can see that our method improves in the correct classification rates with respect to a plain direct estimation. We also obtain better results when compared to the averaging+estimation method. We also show the result of our regularization on a slice of the volume in [Fig. (6.10)].

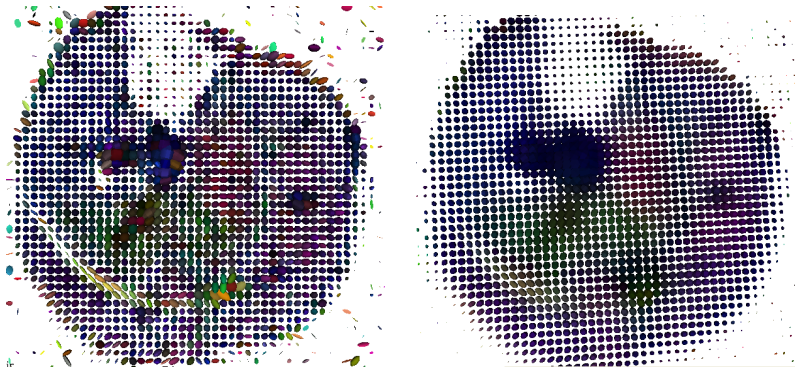


Fig. 6.10: Estimated tensors without regularization, tensors obtained with our method

6.2.4 Discussion

We proposed an approach to direct estimation and regularization of diffusion tensor images. The contribution was the regularization term that assumes linear approximation of neighborhood tensors as well as the definition of the similarity measure between tensors. The convex nature of

	Overall	MG	LG	SOL
DE	78.1%	86.16%	51.1 %	84.43%
ADE	84.46%	90.47%	65.72%	88.43%
DER	86.45%	91.82%	69.76%	89.97%

Tab. 6.2: Correct classification rates for the different methods and for each muscle group. The first and third row show the average correct classification rates for the set of 13 volumes

the proposed cost function which can be easily optimized is an attractive aspect of our method. Our algorithm was compared and outperformed the anisotropic constrained regularization using generated data with known noise model, and significantly improved human skeletal muscle segmentation/classification through DTI using real data. Another possible extension of this work is to replace the Frobenius norm in the energy functional by the Riemannian distance [32] or the Log-Euclidean distance [55]. However this will be done at the expense of the convexity of the function and the computational time.

6.3 Speckle suppression in ultrasound sequences

6.3.1 Introduction

Ultrasound imaging is a popular non invasive and low cost technique to observe the dynamical behavior of organs. Unfortunately the produced images are corrupted by a multiplicative noise (speckle) that affects their quality and complicates the diagnosis task. Therefore, speckle removal could provide better images while preserving image details is fundamental before exploiting the data.

In the literature, numerous methods for speckle suppression were proposed. Some of them transform the multiplicative noise problem into an additive one by considering the logarithm of the image and assume that the noise is Gaussian [72, 2]. The limitation of such a method lies in the fact that the logarithm function reduces the contrast in the image and makes the task of denoising more complex. Other techniques based on total variation minimization in the wavelet domain [111] fail to preserve image details since they rely on the piecewise smoothness assumption. The anisotropic diffusion [154] relies on the gradient information which is not robust when dealing with high noise levels. We want to point out also that these methods, consider only fixed images and not sequences. The temporal aspect was addressed in [63, 1] using temporal averaging. This technique is efficient tool of speckle removal but due to motion, fine details can be blurred unless

an appropriate weight definition is done. Techniques that aim to perform signal correction through motion correction were also investigated [104]. The main challenge of these methods refers to the reliable estimation of the motion fields, a difficult task due to presence of speckle.
In this section, we will introduce a speckle suppression approach for ultrasound sequences. The regularization model (5.13) will be extended to temporal sequences as well. The interactions between pixels are determined by weights that reflect the similarity in the temporal and the spatial domain with implicit constraints imposing motion consistency. As far as fidelity to data term is considered, we will adapt it to the noise model.

6.3.2 Problem Statement

Let us consider two image sequences U and I defined on $T \times \Omega$ where T is the length of the sequence and Ω refers to a frame spatial domain. These two sequences are related according to $I = U * n$ with n being a noise sequence that follows a Rayleigh distribution.

$$r(n) = \frac{n}{\sigma^2} \exp(-\frac{n^2}{2\sigma^2})$$

In order to recover the noise free sequence U we have to minimize a cost function that insures smoothness while imposing a fidelity constraint, to the observed sequence.
To define the fidelity to data term, we will consider the link between the Maximum a Posteriori (MAP) estimation framework and the total variation formulation. We recall The MAP estimator amounts to minimizing an energy composed of two terms : the image model that is the regularization term and the noise model that corresponds to the fidelity to data term (second part of (2.2.4)). Hence, the observed noise likelihood can be chosen as a data term. Explicitly for a frame U_t

$$E_{data}(U_t) = -log(P(I_t|U_t)) = -log(P(n_t))$$

Under the assumption that (i) the noise observations are independent and (ii) identically distributed, and (iii) U and I have non zero intensity for each pixel in the sequence, then we can write:

$$\begin{aligned} E_{data}(u_t) &= -\log\left(\prod_{i=1}^{|\Omega|} r(n_t(i))\right) = -\sum_{i=1}^{|\Omega|} \log\left(r(n_t(i))\right) \\ &= \sum_{i=1}^{|\Omega|}\left[-log\left(\frac{n_t(i)}{\sigma^2}\right) + \frac{n_t(i)^2}{2\sigma^2}\right] \\ &= \sum_{i=1}^{|\Omega|}\left[-\log\left(\frac{I_t(i)}{\sigma^2 U_t(i)}\right) + \frac{I_t^2(i)}{2\sigma^2 U_t^2(i)}\right] \end{aligned}$$

The continuous formulation for the data energy is

$$E_{data}(U_t) = \int_\Omega \left[-\log\left(\frac{I_t(\mathbf{x})}{\sigma^2 U_t(\mathbf{x})}\right) + \frac{I_t^2(\mathbf{x})}{2\sigma^2 U_t^2(\mathbf{x})}\right] d\mathbf{x} \tag{6.23}$$

$$= \int_\Omega \left[\log\left(U_t(\mathbf{x})\right) + \frac{I_t^2(\mathbf{x})}{2\sigma^2 U_t^2(\mathbf{x})} + Cte\right] d\mathbf{x} \tag{6.24}$$

Now, if we consider the function $f(x) = \log(x) + \frac{1}{2\sigma^2}\frac{I_t^2(\mathbf{x})}{x^2}$, it can be easily shown that the second derivative of f is positive for $x \leq \frac{\sqrt{6}I_t(\mathbf{x})}{2\sigma^2}$. This function is convex on the interval $]0, 255]$ if $\frac{\sqrt{6}I_t(\mathbf{x})}{2\sigma^2} > 255$ for any intensity value $I_t(\mathbf{x})$. For $\left(\sigma^2 \leq \frac{\sqrt{6}\times \text{Inf}(I_t)}{2\times 255}\right)$ the function f is convex. In practice it is verified when $(\sigma \leq 0.1)$ which is equivalent to small noise amount. We can then conclude that E_{data} is a convex function for low values of σ.

As far as the regularization term is concerned, we select the model introduced in the previous chapter that leads to a smooth sequence where each pixel is expressed as a weighted mean of the remaining pixels of the sequence. This model is more relevant for natural images than the piecewise constant one. For this application, we will consider the extension of this formulation to ultrasound sequences to account for the time dimension and to perform temporal filtering as well as spatial filtering. In this case E_{reg} is defined as

$$E_{reg}(U) = \iint_{T\times\Omega} \left(LU_t(\mathbf{x}) - U_t(\mathbf{x})\right)^2 d\mathbf{x}dt \tag{6.25}$$

$$LU_t(\mathbf{x}) = \frac{\iint_{T\times\Omega} w(\mathbf{x}, t, \mathbf{y}, t_1) U_{t_1}(y) d\mathbf{y}dt_1}{\iint_{T\times\Omega} w(\mathbf{x}, t, \mathbf{y}, t_1) d\mathbf{y}dt_1}$$

where $w(\mathbf{x}, t, \mathbf{y}, t_1)$ is a similarity measure between two pixels $\mathbf{x}$ and $\mathbf{y}$ observed on the frame t and t_1. Finally, we define the global cost function that is minimized to estimate the noise free image as

$$E(u) = E_{reg}(U) + \lambda \int_0^T E_{data}(U_t)dt \tag{6.26}$$

where λ is a trade-off parameter between the regularization and the fidelity to data term.
Finally the estimation of the noise free ultrasound sequence is obtained through minimizing the cost function given in equation (6.26). The minimization is done through a gradient descent scheme. Starting from the noisy observation $U^0 = I$, the restored image is obtained at the convergence of the sequence U_t^k

$$U_t^{k+1}(\mathbf{x}) = U_t^k(\mathbf{x}) - dt\left[\frac{\partial E_{reg}}{\partial U_t(\mathbf{x})} + \lambda\left(\frac{1}{U_t(\mathbf{x})} - \frac{1}{\sigma^2}\frac{I_t(\mathbf{x})}{U_t(\mathbf{x})^3}\right)\right] \tag{6.27}$$

In order to decrease the computation cost we will restrict the interaction between pixels to a local neighborhood domain and not the whole image. Therefore, we note : $\Pi_{\mathbf{x}}$ the spatial neighborhood centered on a pixel $\mathbf{x}$ and T_w the size of the temporal window (the number of frames that interact directly with a given frame). In this case, the derivative of the regularization term is equal to

$$\begin{aligned}\frac{\partial E_{reg}}{\partial U_t(\mathbf{x})} &= 2\int_{\Pi_{\mathbf{x}}}\int_{t-T_w}^{t+T_w}\left(U_{t_1}(\mathbf{z}) - \int_{\Pi_{\mathbf{z}}}\int_{t_1-T_w}^{t_1+T_w}\frac{w(\mathbf{z}, t_1, \mathbf{y}, t_2)}{Z(\mathbf{z}, t_1)}U_{t_2}(\mathbf{y})d\mathbf{y}dt_2\right)\frac{w(\mathbf{z}, t_1, \mathbf{x}, t)}{Z(\mathbf{z}, t_1)}d\mathbf{z}dt_1 \\ &+ 2\left(U_t(\mathbf{x}) - \int_{\Pi_{\mathbf{x}}}\int_{t-T_w}^{t+T_w}\frac{w(\mathbf{x}, t, \mathbf{y}, t_1)}{Z(\mathbf{x}, t)}U_{t_1}(\mathbf{y})d\mathbf{y}dt_1\right)\end{aligned} \tag{6.28}$$

with $Z(\mathbf{x}, t) = \int_{\Pi_{\mathbf{x}}}\int_{t-T_w}^{t+T_w} w(\mathbf{x}, t, \mathbf{y}, t_1)d\mathbf{y}dt_1$ the normalization coefficient. Note that such a derivative formulation assumes that the weights are independent from the image U which is a valid assumption because weights are calculated in the beginning of the process on the noisy observations.

6.3.3 Weights Computation

In our algorithm we combine spatial and temporal filtering. Therefore, a robust weight definition is crucial to avoid details blurring due to motion. More explicitly, with a good definition of weights there is no need to estimate the motion of the organ. In the context of video denoising corrupted by additive Gaussian noise [27], the authors provided a similarity measure to perform a weighted average based denoising without any motion compensation. In this work, we will suggest a suitable definition of the weights in the case of multiplicative Rayleigh noise. In particular, we will consider a distance based on the spatial information as well as the temporal one. Hence, we will consider patch based similarity toward weight computation. We recall that this measure relies on

the assumption that an image is stationary (at least at local scale), and redundant which means that a given patch has several copies in the entire ultrasound sequence.

Let us characterize a pixel $\mathbf{x}$ at the frame U_t by the set of neighboring pixels in the same frame. Let us consider now the L^2 distance between two patches of size $(2p+1) \times (2p+1)$ centered on $\mathbf{x}$ and $\mathbf{y}$ in two different frames U_{t_1} and U_{t_2} in the sequence.

$$\begin{aligned} d_s(\mathbf{x}, t_1, \mathbf{y}, t_2) &= \sum_{\mathbf{z} \in [-p,p]^2} (I_{t_1}(\mathbf{x}+\mathbf{z}) - I_{t_2}(\mathbf{y}+\mathbf{z}))^2 \\ &= \sum_{\mathbf{z} \in [-p,p]^2} (U_{t_1}(\mathbf{x}+\mathbf{z}) n_{t_1}(\mathbf{x}+\mathbf{z}) - U_{t_2}(\mathbf{y}+\mathbf{z}) n_{t_2}(\mathbf{y}+\mathbf{z}))^2 \end{aligned}$$

In case the observed patches are derived from the same speckle free patch, we can write:

$$d_s(\mathbf{x}, t_1, \mathbf{y}, t_2) = \sum_{\mathbf{z} \in [-p,p]^2} U_{t_1}^2(\mathbf{x}+\mathbf{z})(n_{t_1}(\mathbf{x}+\mathbf{z}) - n_{t_2}(\mathbf{y}+\mathbf{z}))^2$$

This equation shows that contrarily to additive noise, the L^2 distance is not only dependent on noise distribution but also on the noise image intensity. To overcome this dependence one can compute

$$d_s(\mathbf{x}, t_1, \mathbf{y}, t_2) = \sum_{\mathbf{z} \in [-p,p]^2} \frac{(I_{t_1}(\mathbf{x}+\mathbf{z}) - I_{t_2}(\mathbf{y}+\mathbf{z}))^2}{U_{t_1}^2(\mathbf{x}+\mathbf{z})}$$

To determine the distance $d_s(\mathbf{x}, t_1, \mathbf{y}, t_2)$, one needs an estimation of $U_{t_1}(\mathbf{x})$. A simple way to do that is to compute the average of I_{t_1} over the patch around $\mathbf{x}$. Furthermore, in order to obtain a symmetric distance with respect to t_1 and t_2 we will consider $U_{t_1}^2 = U_{t_1} * U_{t_2}$. Thus, we obtain the following distance definition

$$d_s(\mathbf{x}, t_1, \mathbf{y}, t_2) = \sum_{\mathbf{z} \in [-p,p]^2} \frac{(I_{t_1}(\mathbf{x}+\mathbf{z}) - I_{t_2}(\mathbf{y}+\mathbf{z}))^2}{\tilde{U}_{t_1}(\mathbf{x}+\mathbf{z})\tilde{U}_{t_2}(\mathbf{y}+\mathbf{z})} \tag{6.29}$$

$$\tilde{U}_{t_1}(\mathbf{x}) = \frac{1}{(2p+1)^2} \sum_{\mathbf{z} \in [-p,p]^2} I_{t_1}(\mathbf{x}+\mathbf{z}) \tag{6.30}$$

Now if we consider the temporal aspect of the sequence, we can assume that two similar pixels that belong to the same structure have similar displacement vectors. Thus, in addition to the distance between patches, the variation of pixel intensity over time should be taken into account. To this

end, we consider temporal windows of size $(2t_c+1)$ and similarly to the spatial distance, we define

$$d_t(\mathbf{x},\mathbf{y},t_1,t_2) = \sum_{k=-t_c}^{t_c} \frac{\left[I_{t_1+k}(\mathbf{x}) - I_{t_2+k}(\mathbf{y})\right]^2}{U_{t_1+k}(\mathbf{x})U_{t_2+k}(\mathbf{y})} \tag{6.31}$$

In this case we will also replace the actual intensity $U_t(\mathbf{x})$ for a given frame t by its approximation $\tilde{U}_t(\mathbf{x})$ defined in (6.30) so that

$$d_t(\mathbf{x},\mathbf{y},t_1,t_2) = \sum_{k=-t_c}^{t_c} \frac{\left[I_{t_1+k}(\mathbf{x}) - I_{t_2+k}(\mathbf{y})\right]^2}{\tilde{U}_{t_1+k}(\mathbf{x})\tilde{U}_{t_2+k}(\mathbf{y})} \tag{6.32}$$

Finally we define the similarity measure between $\mathbf{x}$ and $\mathbf{y}$ belonging to two different frames as a decreasing function of the distances $d_s(\mathbf{x},\mathbf{y},t_1,t_2)$ and $d_t(\mathbf{x},\mathbf{y},t_1,t_2)$. A possible weight definition is

$$w(\mathbf{x},t_1,\mathbf{y},t_2) = \exp\left(-\frac{d_s(\mathbf{x},t_1,\mathbf{y},t_2)}{2h_s^2}\right)\exp\left(-\frac{d_t(\mathbf{x},t_1,\mathbf{y},t_2)}{2h_t^2}\right) \tag{6.33}$$

h_s and h_t are parameters that have an impact on the selection degree of pixels that interact together. They have an influence on the smoothness of the final result.

To summarize, the proposed weights are designed to take into account the noise properties, mainly the fact that it is multiplicative. With such definition, we restrict interactions between pixels only to similar ones in order to preserve details in each frame. We take also into account motion consistency by computing similarity measure on temporal windows.

6.3.4 Experimental Results

In the present section we will focus on the experimental validation of the proposed method. To this end we used both synthetic and real data. In the first experiment, we used synthetic images with artificial speckle. The speckle is simulated by low pass filtering a complex Gaussian random field and computing its magnitude [2]. We compared our algorithm to the anisotropic diffusion algorithm proposed in [154] and a wavelet based technique proposed in [116]. We have to point out that we used the Matlab implementation provided by the authors, and the parameters were selected in order to achieve the optimal performance of each algorithm. Regarding our algorithm we considered 11×11 window size for Π_x while $\lambda = 0.05$. For weight computation the parameters setting is: $h_s = \frac{\sigma_n}{5}$ and a 3×3 patch size. We did not consider the temporal aspect for this experiment because the synthetic data are single images and not sequences. For an evaluation we considered the PSNR of the reconstructed images $\hat{U}$ that is a function of the mean square error with respect to the ground truth image.

	Sythetic1 [116]			Synthetic2		
σ_n	1	0.5	0.25	1	0.5	0.25
Corrupted image	22.20	27.93	33.95	20.44	26.47	32.62
OurMethod	29.16	37.32	38.01	33.88	38.89	43.53
SRAD[154]	24.88	33.09	40.95	31.55	38.36	46.15
wav [116]	29.14	35.48	41.69	31.06	35.85	39.75

Tab. 6.3: PSNR values for denoised image corrupted by Speckle

$$PSNR = 10log_{10}\frac{255^2}{MSE} \quad MSE = \sum_{t=0}^{T}\sum_{\mathbf{x}\in\Omega}(U_t(\mathbf{x}) - \hat{U}_t(\mathbf{x}))^2$$

In table (6.3) we reported the SNR value for the different methods and with different noise variance. This results show that our method achieve high performance when the noise variance is important. Such a behavior is justified by the fact that we consider large spatial neighborhood (Π_x) for filtering which reduces considerably the variance inside each region. The anisotropic diffusion algorithm (SRAD)[154] is more local and its numerical scheme is based on interactions between pixels at local scale when compared to our algorithm. Besides, the gradient information is not reliable in case of high level noise. The approach presented in [116] and the choice of the parameter of threshold used in this algorithm to compute the noise free wavelet coefficients is critical to insure the balance between edge preserving and noise suppression. In figure [Fig.(6.11)] one can see the restoration results for the synthetic image using the different methods and different variance. The method based on the wavelet transform provides images with sharp contours but the noise remains in homogeneous regions. The anisotropic diffusion results in blurry edges while smoothing the homogeneous areas. Our algorithm reaches an optimal balance where the obtained images are completely smoothed inside each region while the edges being sharp. As we stated before we consider large neighborhood size to remove the speckle, while encoding the image structure in our weight definition to preserve image discontinuities and details. This figure show that our method preserves the image discontinuities and contrast between regions better than the two other algorithms. As far as real data are concerned we considered an ultrasound sequence of the left ventricle on which we applied our speckle removal algorithm. The parameters used for this experiment are: a 7×7 window size for Π_x, $\lambda = 0.5$ and $T_w = 5$. The weight computation was performed using 3×3 patch size, a temporal window for comparison of size $t_c = 2$ and $h_s = h_t = 0.05$. In figure [Fig.(6.12)] we reported the restoration result on one frame of an ultrasound sequence. To evaluate the quality of restoration, we extracted the boundaries of the ventricle, using a level set based technique. The contours obtained using our method are smooth and the structure being present in the image were preserved. The contour extracted in the frame restored using our approach is similar to the one associated to the anisotropic diffusion, but this method yields piecewise smooth images. Regarding the wavelet based technique, it produces image with sharp details but the con-

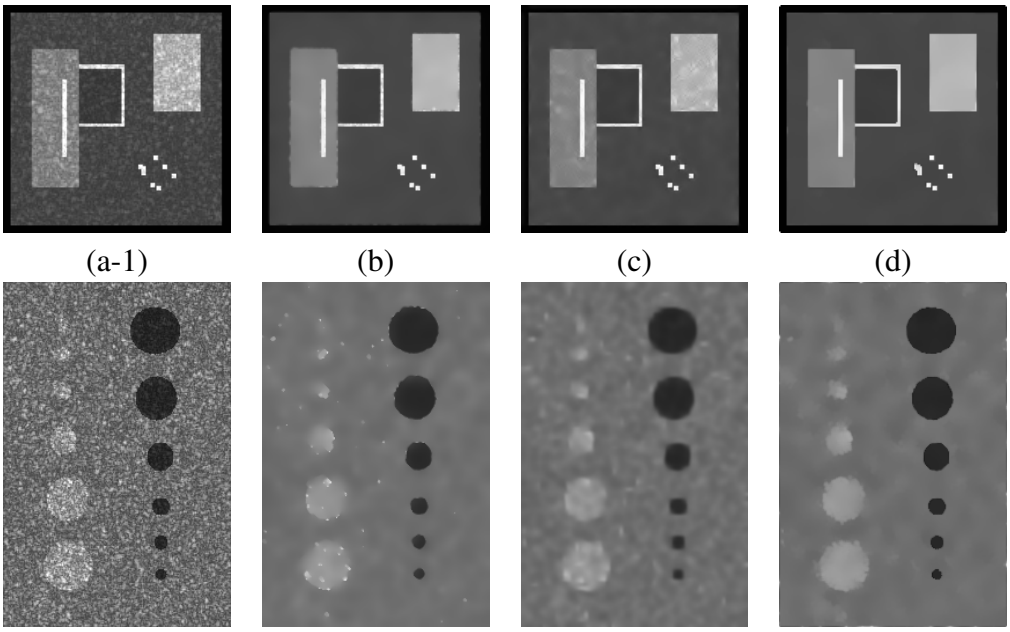
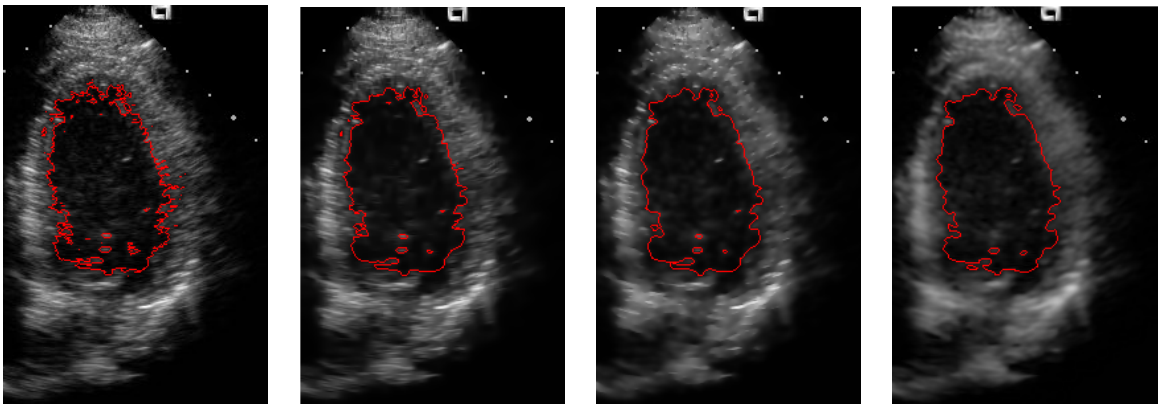


Fig. 6.11: (a-1) Image (Sythetic1) corrupted by the speckle $\sigma_n = 0.5$ (a-2) Image (Sythetic2) corrupted by the speckle $\sigma_n = 1$ (b) Result using the anisotropic diffusion [154] (c) Result using the wavelet based technique [116] (d) Result using our algorithm

tours are not smooth. To evaluate the impact of the time component we compared the algorithm performance by processing the whole sequence using $T_w = 2$ and by processing it frame by frame ($T_w = 0$). and the obtained results are shown in figure [Fig.(6.13)]. One can see that using a temporal window is more efficient for speckle reduction and this is illustrated by the residual images [Fig.(6.13-c),Fig.(6.13-e)](difference between the filtered image and the observed one). Besides, we notice that considering the time component don't introduce a blur on the images because our weight definition is robust enough to compensate the effect of motion. To conclude we can say that the experimental results show that our method can deal with correlated speckle even though we make the assumption of independence between pixels. We have to point out that the proposed speckle removal approach is flexible with the selection of the scale of interaction between pixels contrarily to the PDE based technique presented in [154]

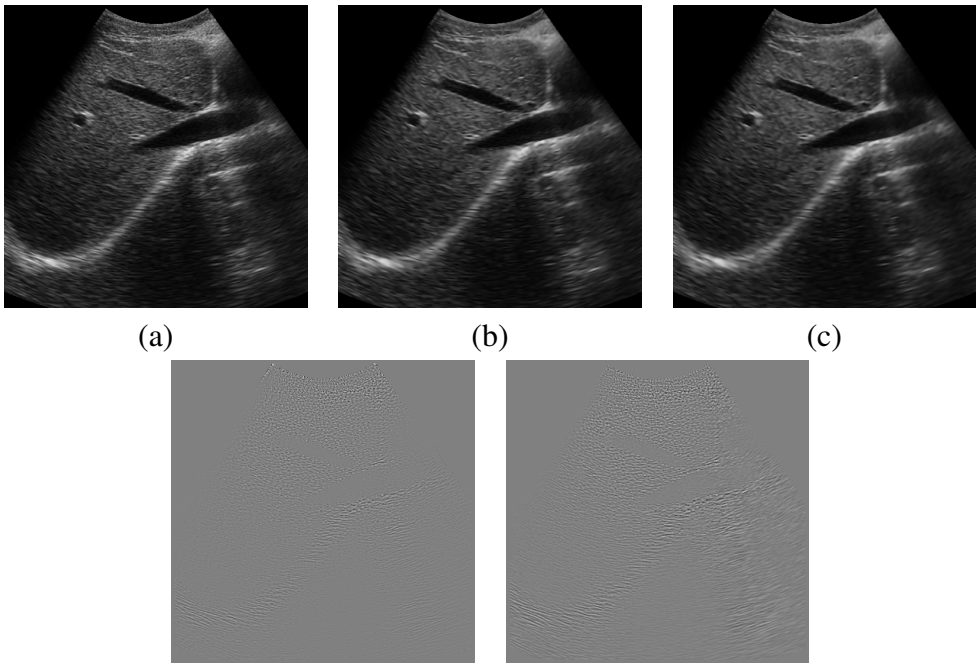
6.3.5 Discussion

In this section we addressed the problem of speckle removal in ultrasound sequence. It was shown that the underlying linear image model is a good assumption even in the case of multiplicative noise. The similarity measure between patches was modified to fit the noise model because the basic L^2 distance between patches has a distribution that depends on the actual intensity of the pixel.



(a) (b) (c) (d)

Fig. 6.12: Results of restoration on real ultrasound frames. (a) Observed image (b) Result using algorithm [116] (c) Anisotropic diffusion [154] (d) Results of our algorithm



(a) (b) (c)

(d) (e)

Fig. 6.13: Results of filtering real ultrasound frames (a) Observed image (b) Result of our algorithm without temporal component ($T_w = 0$) (frame by frame filtering) (c) Result of our algorithm using temporal filtering ($T_w = 2$) (d) Residual obtained with our algorithm without temporal component ($T_w = 0$) (frame by frame filtering)(e) Residual obtained with our algorithm using temporal filtering ($T_w = 2$)

Further improvement can be considered like selecting a more appropriate optimization technique to avoid local minima. A weight definition that is specific to the Rayleigh distribution can also be explored.

6.4 Conclusion

In this chapter, we demonstrated that the regularization model proposed in the previous chapter can be successfully exploited in order to tackle other signal reconstruction problems with different data and noise models. We demonstrated how we can adapt this model to address signal reconstruction for more complex scenarios. We demonstrated how one can incorporate more appropriate image-driven noise-models, as well as how to deal with complex signal perturbations. Last, but not least we show how such a model can be extended to address the case of constrained manifolds. Finally the experimental results demonstrated the flexibility and the accuracy of such model.

Chapter 7

Conclusion

In this thesis we studied the problem of image denoising and texture preservation, as well as their applications in a number of fields. We have introduced certain novel ideas that incorporate image structure in the denoising process. Towards better preservation of fine details and texture, we considered data driven image models. These ideas were explored using appropriate mathematical formulations and have been considered to address complex real scenarios. Promising qualitative results, as well as comparisons with the state of the art methods demonstrated the potentials of our contributions.

In this document, we focused on image enhancement with main aim being textured image restoration. An efficient regularization process must involve an appropriate definition of the interaction between pixels that have to be consistent with the image content. This is the issue that we addressed during the thesis and our main contributions are

- Addressing the limitation of conventional approaches, that often consider global parametric models without adapting them to local image context, through (i) introducing the notion of soft classification in the denoising process (ii) considering kernel based non parametric image model exploiting the available noisy observations. In order to reflect the data variability and mainly to account for the difference between patches distribution in textured and smooth regions, we considered variable bandwidth kernels. The result of the soft classification step was used for the bandwidth selection. This model was considered in an MPM estimation framework to determine the noise-free intensity. The novel proposed approach is able to better preserve texture and experimental results demonstrated its potential
- Introducing random walks and particle filtering technique that aims to adapt the filtering domain to image geometry. To this end, image structure is captured using local co-occurrence

statistics and is considered in the enhancement algorithm. The reconstruction process is modeled using a dynamical system where a number of particles are exploring the image domain to recover the most appropriate observations to be considered to estimate the intensity of a given noisy observation. Toward optimal exploration of the image (i) we defined the particle evolution using the prior learning of the co-occurrence between observations (ii) and considered likelihood function to evaluate the relevance of the state of a particle and its consistency with the observation to be recovered. Promising results using additive and multiplicative noise models demonstrate the efficiency of such a technique.

- Considering a global image model and a convex functional to be minimized. This functional relies on linear image model instead of the constant smooth one classically used in the TV based formulation. Within the proposed framework, the interaction between pixels describes better the image complexity. These interactions are expressed through weights that reflect the photometric distance between pixels as well as the spatial one. Regarding the photometric distance, we investigated new means of defining it (i) using an alternative features vector to describe pixels based on projection on subspaces of image patches (ii) taking into consideration the noise distribution and its impact on the distribution of the distance between patches. Finally, we introduced the notion of automatic spatial bandwidth selection to better capture the image structure. Such formulation provides a simple way to model natural image complexity. Comparison with state of the art methods shows interesting quantitative and qualitative results that reflect the potentials of the proposed framework.

- Investigating the use of the variational technique proposed in this thesis towards addressing applications with specific signal properties. First, we addressed the RGB image enhancement issue including more realistic noise properties. Next, we were concerned with speckle reduction in ultrasound images by defining appropriate data term consistent with the noise model. Last but not least, we extended our framework to Diffusion tensors estimation and regularization. The variety of fields, noise models and image properties demonstrate the flexibility and the efficiency of the proposed technique.

The different components introduced in this thesis carry on certain strengths as well as some limitations. It is natural, to asses a comparison between the three different methods. Like in the previous examples, the PSNR values are presented in table 7.1. They show that the functional minimization approach is the most efficient one. This confirms the observation that considering global cost function is better then the local ones since it enforces coherence of the denoising process for neighboring pixels. Furthermore, the underlying image model and the variable bandwidth kernel are more realistic. The images presented in figure [Fig. (7.1), Fig. (7.2)] proved that better image

	Barbara	Boat	Fingerprint	House	Lena	Baboon
MPM-var	28.9	29.11	26.68	31.02	31.25	25.39
Random-walk	29.29	28.72	26.4	31.53	31.07	25.51
Convex-Min	30.46	29.94	27.65	32.34	32.12	26.02

Tab. 7.1: PSNR values for denoised image (The PSNR of the image corrupted by gaussian noise of std=20 is equal to 22.15)

quality is obtained with the third approach.

Perspectives

In this thesis our objective was to provide some answers to the problem of natural images denoising and texture restoration. Nevertheless, many questions remain open. This concerns mainly the definition of capturing photometric correlation between pixels. The definition of new set of features is an interesting direction but the selection of the most appropriate features that have to be considered could bring further improvement. As we pointed out when we discussed the distribution of the distance between patches, the size of the patches considered for comparing pixels is important. Hence, more generally adapting the scale of the features considered when comparing pixels to the texture scale is a prominent issue. Regarding the spatial bandwidth of the filter additional constraints based on some prior knowledge on the image could be pertinent information to determine the optimal spatial bandwidth at each position. The automatic recognition of oscillations due to noise and other related to texture is a challenging question that was addressed in the second chapter of this thesis. Nevertheless, improving the classification outcome by considering other local descriptors is a direction to be explored. For instance such a classification could be integrated as prior information in different image processing tasks (demosaicing, inpainting, denoising ...). Extensive validation using a larger image database and a comparison with commercial denoising software could lead to better assessment of the performance of the proposed methods. Decreasing the computational complexity is also a task to be considered. As far as noise model is considered, we have shown in the previous chapter that the noise model is far from being white noise since we observe an inter-channel and intra-channel correlation between noise samples. Including more realistic noise models where the correlations between pixels are taken into account is an interesting direction to deal with.

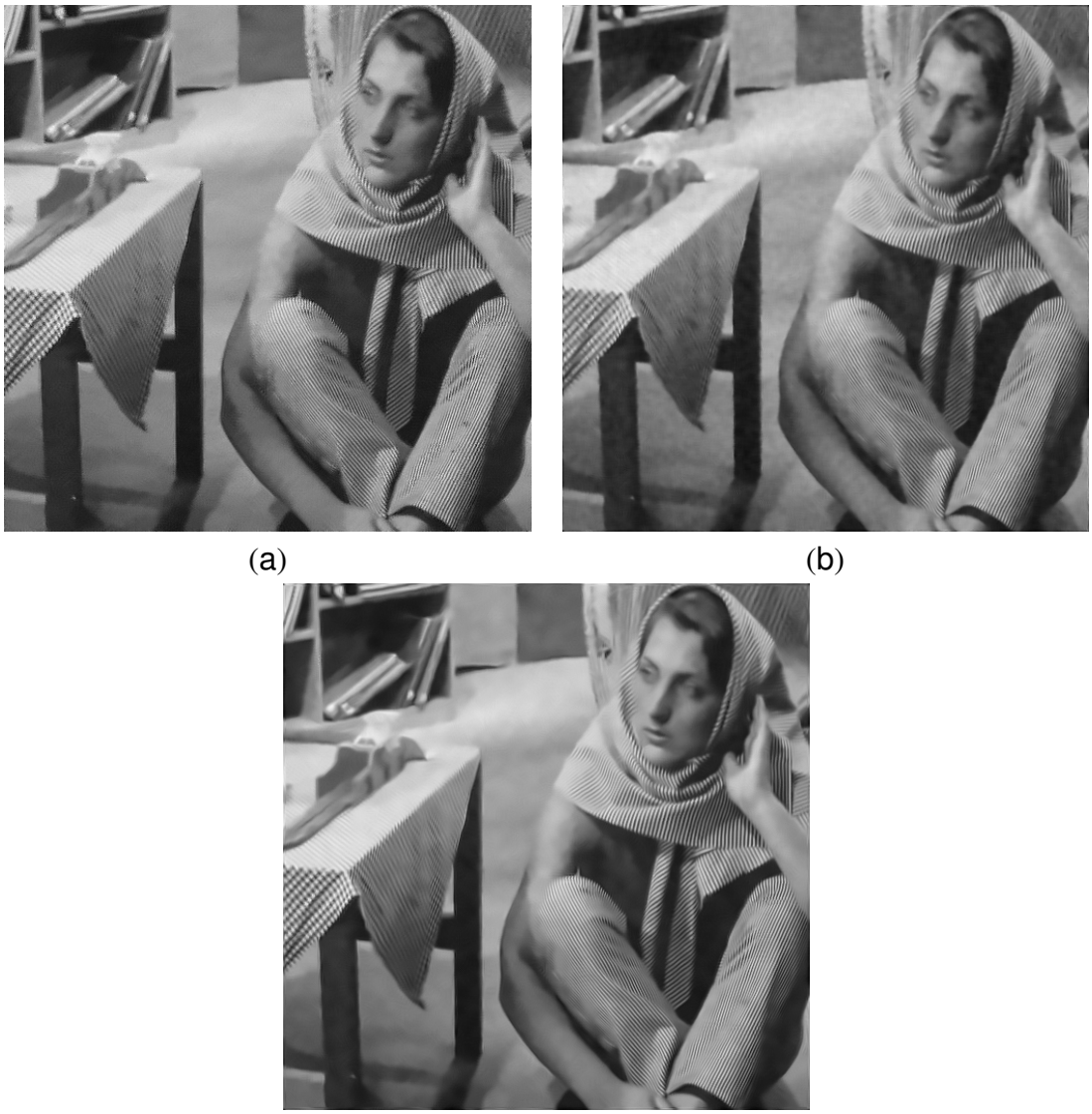


Fig. 7.1: Results of Barbara image restoration (a) using Random walks (b) using MPM estimation with variable kernels bandwidth (c) using functional minimization based approach with variable bandwidth spatial kernel

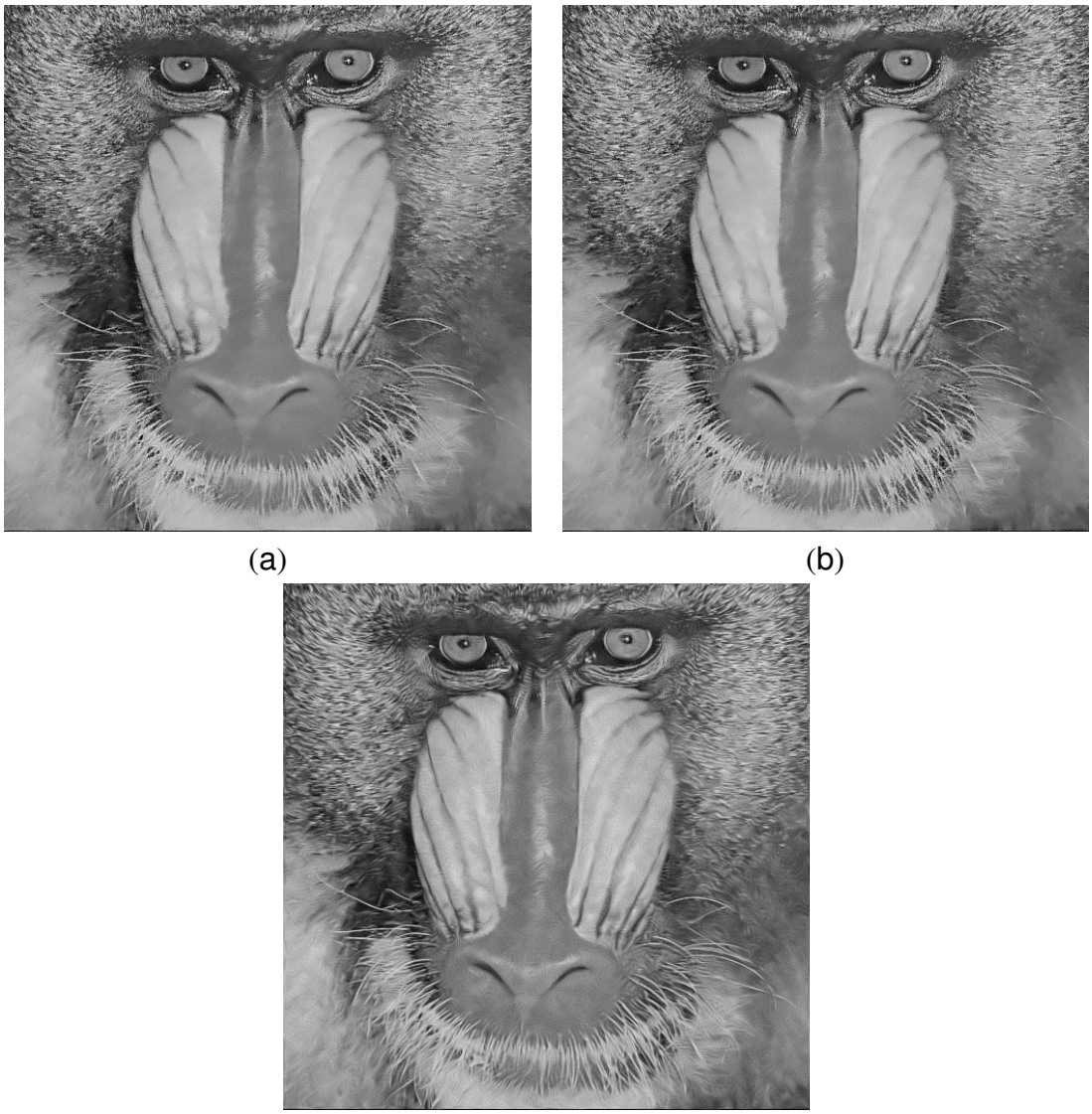


Fig. 7.2: Results of Baboon image restoration (a) using Random walks (b) using MPM estimation with variable kernels bandwidth (c) using functional minimization based approach with variable bandwidth spatial kernel

Chapter 8

Conclusion en Français

Au cours de cette thèse nous avons étudié les techniques de débruitage des images et la préservation de la texture. Nous avons introduit quelques idées nouvelles pour incorporer la structure de l'image dans le processus de filtrage. Pour une meilleure préservation des détails et de la texture, nous avons considéré des modèles construits à partir de l'observation bruitée. Cette idée a été mise en place à l'aide de formulation mathématique où on a essayé de prendre en considération des scénarios plus complexes. Des résultats qualitatifs prometteurs et des comparaisons avec l'état de l'art ont démontré le potentiel de nos contributions.

Dans ce document, nous nous somme intéressés à la restauration des images avec un objectif de préservation de texture. Un processus de régularisation efficace dépend de la définition des interactions entre les pixels qui doit être cohérente avec le contenu de l'image. Ceci est le problème que nous avons traité dans cette thèse et nos plus importantes contributions sont

- Proposer une alternative aux approches conventionnelles qui considèrent souvent des modèles globaux et paramétriques qui ne sont pas adaptés au contexte local du pixel et ceci à travers (i) l'introduction de la notion de classification dans le processus de débruitage (ii) considérer des modèles d'image non paramétriques en exploitant l'observation bruitée. Afin de refléter la variabilité des données et tenir compte de la différence dans la distribution des patches entre les zones texturées et les zones homogènes, nous avons considéré des noyaux de taille variable. Le résultat de l'étape de classification a été utilisé pour la sélection de la taille du noyau. Ce modèle a été considéré dans le cadre d'une estimation MPM afin de déterminer l'intensité sans bruit. Cette approche nouvelle permet une meilleure préservation de la texture et les résultats expérimentaux ont démontré son potentiel.
- Introduire les marches aléatoires et les filtres à particules dans le but d'adapter le domaine de

filtrage à la géométrie de l'image. Pour cette raison, nous allons proposer une caractérisation statistique de la structure de l'image (inspirée de la matrice de co-occurrence) et l'utiliser dans l'algorithme de débruitage. Le processus de reconstruction est modélisé à l'aide d'un système dynamique où un nombre de particules est entrain d'explorer le domaine de l'image. Ceci permettra de trouver les observations les plus pertinentes qui seront considérées pour estimer l'intensité d'une observation bruitée. Pour une exploration optimale du domaine de l'image (i) nous avons défini l'évolution des particules en utilisant une information à priori en considérant la co-occurrence entre les observations (ii) nous avons considéré une fonction de vraisemblance afin d'évaluer la pertinence de l'état d'une particule et sa cohérence avec l'observation qu'on veut reconstruire. Des résultats expérimentaux prometteurs pour le cas du bruit multiplicatif et additif démontrent l'efficacité de cette technique.

- Considérer un modèle d'image global et une fonctionnelle convexe à minimiser. Cette fonctionnelle s'appuie sur un modèle d'image linéaire contrairement aux formulations variationelles classiques qui supposent que l'image est constante par morceaux. Avec l'approche que nous avons adoptée les interactions entre les pixels décrivent mieux la complexité de l'image. Ces interactions sont exprimées à travers la fonction de poids qui reflètent la similarité photométrique et spatiale entre les pixels. Concernant la similarité photométrique, nous avons exploré d'autres moyens pour l'exprimer (i) en utilisant un nouveau vecteur caractéristique pour décrire les pixels par projection des patches de l'image dans un autre sous espace (ii) prendre en considération la distribution du bruit et son impact sur la distribution de la distance entre les patches. Finalement, nous avons introduit la notion de la sélection automatique de la taille du noyau spatial afin de mieux intégrer la structure de l'image dans le filtrage. Cette formulation donne un moyen simple pour modéliser la complexité des images naturelles. La comparaison avec l'état de l'art montre des résultats quantitatifs et qualitatifs intéressants qui reflètent les potentiels de cette technique.

- Appliquer la technique de minimisation d'une fonction d'énergie proposée dans cette thèse à d'autres problèmes et l'adapter aux spécifications de ces derniers. Au début, nous avons étudié le problème de la restauration des images RGB tout en considérant des propriétés de bruit plus réelles. Ensuite, nous avons étendu cette formulation au problème d'estimation et régularisation des tenseurs de diffusion. Finalement, nous nous sommes intéressés à la réduction du Speckle dans des séquences d'images ultrasons en définissant un terme d'attache aux données plus adapté au modèle de bruit.

Les différentes idées exposées au cours de cette thèse présentent des éléments de réponse au problème de filtrage mais ont aussi leurs limitations. Pour procéder à une comparaison entre les

	Barbara	Boat	Fingerprint	House	Lena	Baboon
MPM-var	28.9	29.11	26.68	31.02	31.25	25.39
Random-walk	29.29	28.72	26.4	31.53	31.07	25.51
Convex-Min	30.46	29.94	27.65	32.34	32.12	26.02

Tab. 8.1: PSNR des images débruitées par les différentes méthodes (le PSNR des images corrompue par un bruit additif Gaussien d'écart type std=20 est égal à 22.15)

différentes méthodes, nous avons présenté les valeurs de PSNR pour la série d'images test dans le tableau (8.1). Ces résultats montrent que l'approche basée sur la minimisation de fonctionnelles est l'approche la plus efficace. Ceci est un argument en faveur de l'usage des fonctions de coût globales puisque cela renforce la cohérence du processus de débruitage pour tous les pixels voisins. De plus, le modèle sous-jacent ainsi que le choix des noyaux de taille variable sont plus proches de la réalité. Les images présentées dans les figure [Fig. (8.1), Fig. (8.2)] prouvent que une meilleure restauration des images en utilisant la dernière approche.

8.1 Perspectives

Au cours de cette thèse notre objectif était de fournir quelques éléments de réponse au problème de la restauration des images et principalement la texture. Toutefois plusieurs questions restent ouvertes. Ceci concerne principalement la définition de la corrélation photométrique entre les pixels. La définition appropriée d'un vecteur caractéristique est une direction intéressante et peut apporter des améliorations. Comme nous l'avons signalé quand nous avons discuté de la distribution statistique de la distance entre patches, la taille du patch à considérer est importante. Ainsi, adapter la taille des patches et plus généralement l'échelle des descripteurs à l'échelle de la texture est une question qui mérite d'être étudiée. Considérant les interactions spatiale entre les pixel et la taille du noyau, l'ajout de contraintes supplémentaires ainsi qu'une connaissance à priori sur l'image pourrait être une information pertinente afin de déterminer la taille du noyau optimal pour chaque pixel. La reconnaissance des oscillations qui sont dûes au bruit et celles liées à la texture présente encore un défi et cette question a été étudiée dans le deuxième chapitre de cette thèse. Cependant, l'amélioration des résultats de la classification en considérant d'autres descripteurs locaux est une direction qui peut être explorée davantage. En effet, cette classification peut être intégrée en tant q'information à priori dans plusieurs applications de traitement des images (dématriçage, inpainting, débruitage,...). Une validation extensive en utilisant une base d'images plus large ainsi qu'une comparaison avec des logiciels commerciaux de filtrage d'image peut être effectuée afin de mieux évaluer les performances des méthodes proposées dans ce travail. Réduire le temps de calcul est

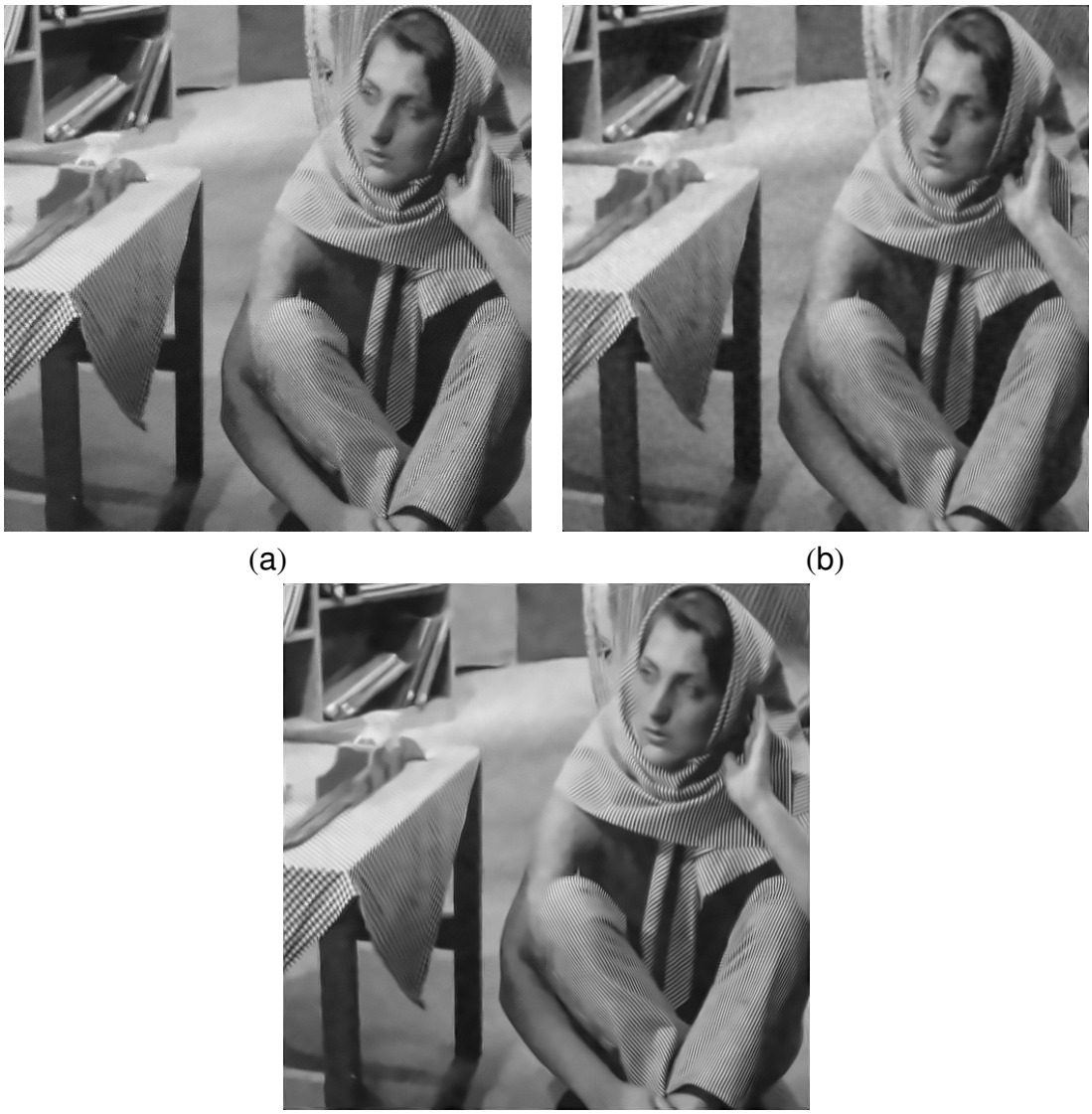


Fig. 8.1: Résultat de la restoration de l'image Barbara (a) En utilisant les marches aléatoires (b) En utilisant l'estimation MPM et les noyaux de taille variable (c) Minimisation de fonctionnelle en utilisant des noyaux de taille variable

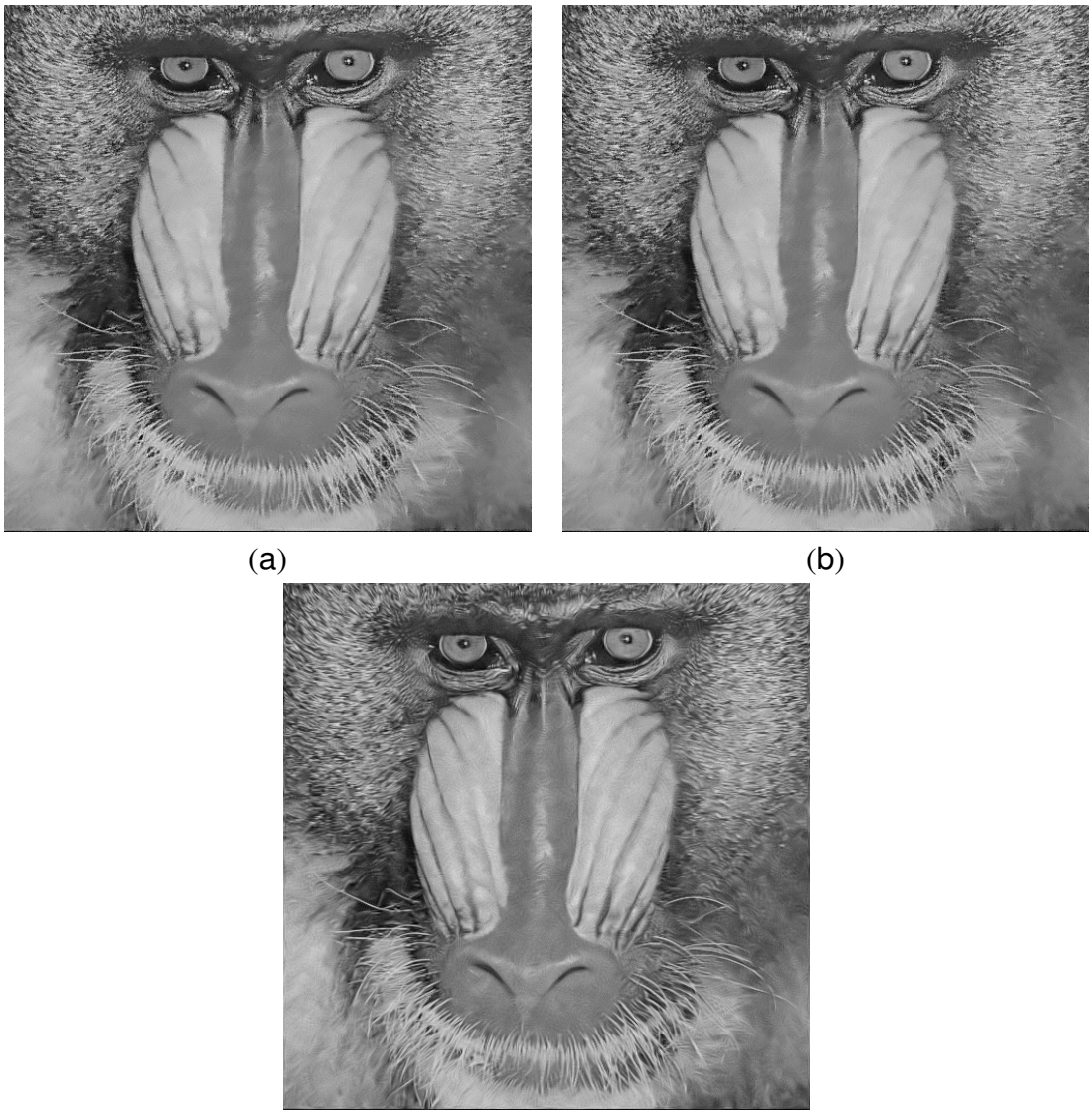


Fig. 8.2: Résultat de la restoration de l'image Baboon (a) En utilisant les marches aléatoires (b) En utilisant l'estimation MPM et les noyaux de taille variable (c) Minimisation de fonctionnelle en utilisant des noyaux de taille variable

aussi une tâche à accomplir dans le futur. En ce qui concerne le modèle du bruit nous avons montré dans le chapitre 4 que le bruit n'est pas blanc étant donné qu'on observe une corrélation aussi bien dans un canal que entre les canaux. Pour cette raison, inclure des modèles de bruit plus complexes où on tient compte de cette corrélation est une direction qui mérité d'être considérée.

Bibliography

[1] J. G. Abbott and F. L. Thurstone. Acoustic speckle: Theory and experimental analysis. *Ultrason. Imag*, 1:303–324, 1979.

[2] A. Achim, A. Bezerianos, and P. Tsakalides. Novel bayesian multiscale method for speckle removal in medical ultrasound images. *IEEE trans. on Medical Imaging*, 20:772–783, August 2001.

[3] M. Aharon, M. Elad, and A. Bruckstein. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE trans. on Signal Processing*, 54(11):4311 – 4322, 2006.

[4] L. Alvarez, F. Guichard, P.L. Lions, and J.M. Morel. Axioms and fundamental equations in image processing. *Archive for Rational Mechanics and Analysis*, 123:199–257, 1993.

[5] L. Alvarez, P.L. Lions, and J.M. Morel. Image selective smoothing and edge detection by nonlinear diffusion, ii. *Journal of Numerical Analysis*, 29(3):845–866, 1992.

[6] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online non-linear/non-gaussian bayesian tracking. *IEEE trans. on Signal Processing*, 50:174–188, 2002.

[7] G. Aubert and J.F. Aujol. A nonconvex model to remove multiplicative noise. In *Int. Conf. on Scale Space Methods and Variational Methods in Computer Vision*, pages 68–79, 2007.

[8] JF. Aujol and Chambolle A. Dual norms and image decomposition models. Technical Report 5130, INRIA, 2004.

[9] JF. Aujol and Chambolle A. Dual norms and image decomposition models. *Int. Journal of Computer Vision*, 63(1):85–104, 2005.

[10] S. Awate and R. Whitaker. Higher-order image statistics for unsupervised, information-theoretic, adaptive, image filtering. In *Proc. of the IEEE Int. Conf. on Computer Vision and Pattern Recognition*, pages 44–51, 2005.

[11] N. Azzabou and Paragios. Spatio-temporal speckle reduction in ultrasound sequences. In *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention*, 2008.

[12] N. Azzabou, N. Paragios, and F. Guichard. Random walks, constrained multiple hypothesis testing and image enhancement. In *Proc. of the Europ. Conf. on Computer Vision*, pages 379–390, 2006.

[13] N. Azzabou, N. Paragios, and F. Guichard. Image denoising based on adapted dictionary computation. In *Proc. of the IEEE Int. Conf. on Image Processing*, pages 109–112, 2007.

[14] N. Azzabou, N. Paragios, and F. Guichard. Uniform and textured regions separation in natural images towards mpm adaptive denoising. In *Int. Conf. on Scale Space Methods and Variational Methods in Computer Vision*, pages 418–429, 2007.

[15] N. Azzabou, N. Paragios, F. Guichard, and F. Cao. Variable bandwidth image denoising using image-based noise models. In *Proc. of the IEEE Int. Conf. on Computer Vision and Pattern Recognition*, pages 1–7, 2007.

[16] R. Bajcsy. Computer description of textured surfaces. In *International Joint Conference on Artificial Intelligence*, pages 572–579, 1973.

[17] S. Basu, P. Fletcher, and R. Whitaker. Rician noise removal in diffusion tensor mri. In *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention*, pages 117–125, 2006.

[18] E. Bennett, M. Uyttendaele, C.L. Zitnick2, R. Szeliski, and S.B. Kang. Video and image bayesian demosaicing with a two color image prior. In *Proc. of the Europ. Conf. on Computer Vision*, pages 508– 521, 2006.

[19] F. Bergeaud and S. Mallat. Matching pursuit of images. In *Proc. of the IEEE Int. Conf. on Image Processing*, pages 53–56, 1995.

[20] N. Bergman. *Recursive Bayesian estimation: Navigation and tracking applications*. PhD thesis, Linköping Univ., Linköping, Sweden, 1999.

[21] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, Belmont, MA, 1999.

[22] M. Black, G. Sapiro, D. Marimont, and D. Heeger. Robust anisotropic diffusion. *IEEE trans. on Image Processing*, 7:421–432, 1998.

[23] A.C. Bovik, M. Clark, and W.S. Geisler. Multichannel texture analysis using localized spatial filters. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 12:55–73, 1990.

[24] A.M. Bruckstein and M. Elad. On sparse signal representations. In *Proc. of the IEEE Int. Conf. on Image Processing*, pages 3–6, 2001.

[25] A. Buades. *Image and Film denoising by non-local means*. PhD thesis, Universitat de les Illes Balears, 2006.

[26] A. Buades, B. Coll, and J-M. Morel. A non-local algorithm for image denoising. In *Proc. of the IEEE Int. Conf. on Computer Vision and Pattern Recognition*, pages 60–65, 2005.

[27] A. Buades, B. Coll, and J.M Morel. Denoising image sequences does not require motion estimation. Technical Report 2005-18, CMLA Preprint, 2005.

[28] A. Buades, B. Coll, and J.M. Morel. Image enhancement by non-local reverse heat equation. Technical Report 2006-22, CMLA Preprint, 2006.

[29] A. Buades, B. Coll, J.M Morel, and C. Sbert. Non local demosaicing. Technical Report 2007-15, CMLA Preprint, 2007.

[30] A. Buades, JM. Morel, B. Coll, F. Cao, F. Guichard, and N. Azzabou. Conservation de la taille du grain en photographie numérique. French patent N: 07 05754, Aout 2007.

[31] V. Caselles, J.M. Morel, G. Sapiro, and A. Tannenbaum. Introduction to the special issue on partial-differential equations and geometry-driven diffusion in image-processing and analysis. *IEEE trans. on Image Processing*, 7(3):269–273, 1998.

[32] C. A. Castano-Moraga, C. Lenglet, R. Deriche, and J. Ruiz-Alzola. A Riemannian approach to anisotropic filtering of tensor fields. *Signal Processing, Special Issue on Tensor Signal Processing,*, 87(2):263–276, 2007.

[33] A. Chambolle and P.L. Lions. Image recovery via total variation minimization and related problems. *Numerische Mathematik*, 76:167–188, 1997.

[34] C. Chefd'hotel, D. Tschumperle, R. Deriche, and O.D. Faugeras. Regularizing flows for constrained matrix-valued images. *Journal of Mathematical Imaging and Vision*, 20(1-2):147–162, 2004.

[35] P.C. Chen and T. Pavlidis. Segmentation by texture using correlation. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 5(1):64–69, 1983.

[36] Y. Cheng. Mean shift, mode seeking, and clustering. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 17:790–799, 1995.

[37] R. Coifman, Y. Meyer, and V. Wickerhauser. Wavelet analysis and signal processing. *Wavelets and their applications*, pages 153–173, 1992.

[38] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 15(5):603–619, 2002.

[39] D. Comaniciu, V. Ramesh, and P Meer. The variable bandwidth mean shift and data-driven scale selection. In *Proc. of the IEEE Int. Conf. on Computer Vision and Pattern Recognition*, volume 1, pages 438 – 445, 2001.

[40] O. Coulon, D. C. Alexander, and S. Arridge. Diffusion tensor magnetic resonance image regularization. *Medical Image Analysis*, 8(1):47–67, 2004.

[41] I. Daubechies. The wavelet transform, time-frequency localization and signal analysis. *IEEE trans. on Information Theory*, 36:961–1005, 1990.

[42] A. Dempster, M. Laird, and D. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society*, 39:1–38, 1977.

[43] R. Deriche, D. Tschumperle, C. Lenglet, and M. Rousson. Variational approaches to the estimation, regularization and segmentation of diffusion tensor images. In Faugeras Paragios, Chen, editor, *Mathematical Models in Computer Vision: The Handbook*. Springer, 2005 edition, 2005.

[44] K. I. Diamantaras and S. Y. Kung. *Principal component neural networks: theory and applications*. John Wiley & Sons, Inc., 1996.

[45] D. L. Donoho and I. M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–455, 1994.

[46] A. Doucet. On sequential simulation-based methods for bayesian filtering. Technical Report CUED/F-INFENG/TR. 310, Cambridge University Department of Engineering, 1998.

[47] A. Doucet, J. de Freitas, and N. Gordon. *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, New York, 2001.

[48] D.Tschumperle and R.Deriche. Vector-valued image regularization with pdes: A common framework for different applications. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 27(4):506–517, 2005.

[49] R.O. Duda and P. Hart. *Pattern Classification And Scene Analysis*. WileyInterscience Publication, 1973.

[50] D. Dunn and W.E. Higgins. Optimal gabor filters for texture segmentation. *IEEE trans. on Image Processing*, 4:947–964, 1995.

[51] A. Efros and T. Leung. Texture synthesis by non-parametric sampling. In *Proc. of the Int. Conf. on Computer Vision*, pages 1033–1038, 1999.

[52] K. Egiazarian, V. Katkovnik, and J. Astola. Adaptive window size image denoising based on ici rule. In *Proc. of the IEEE Int. Conf. on Acoustic, Speech and Signal Processing*, pages 1869–1872, 2001.

[53] M. Elad and M. Aharon. Image denoising via learned dictionaries and sparse representation. In *Proc. of the IEEE Int. Conf. on Computer Vision and Pattern Recognition*, pages 895–900, 2006.

[54] I. Elfadel and R. Picard. Gibbs random fields, cooccurrences, and texture modeling. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 16(1):24–37, 1994.

[55] P. Fillard, V. Arsigny, X. Pennec, and N. Ayache. Clinical DT-MRI estimation, smoothing and fiber tracking with log-Euclidean metrics. In *IEEE Int. Symp. on Biomedical Imaging: From Nano to Macro*, pages 786–789, 2006.

[56] A. Foi, S. Alenius, V. Katkovnik, and K. Egiazarian. Image denoising using scale mixtures of gaussians in the wavelet domain. *IEEE Sensors Journal*, 7(10):1456–1461, 2007.

[57] D. Gabor. Theory of communication. *J. IEE (London)*, 93:429–457, 1946.

[58] S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.

[59] G. Gilboa, J. Darbon, S. Osher, and T. Chan. Nonlocal convex functionals for image regularization. Technical Report 06-57, UCLA CAM Report, 2006.

[60] G. Gilboa and S. Osher. Nonlocal linear image regularization and supervised segmentation. Technical Report 06-47, UCLA CAM Report, 2006.

[61] G. Gilboa and S. Osher. Nonlocal linear image regularization and supervised segmentation. *SIAM Multiscale Modeling and Simulation*, 6:595–630, 2007.

[62] G. Gilboa, N. Sochen, and Zeevi Y.Y. Variational denoising of partly-textured images by spatially varying constraints. *IEEE trans. on Image Processing*, 24(8):2281–2289, 2006.

[63] J.W. Goodman. Some fundamental properties of speckle. *J. Opt. Soc*, 66:1145–1150, 1976.

[64] N. Gordon, D. Salmond, and A. Smith. Novel approach to nonlinear/non-gaussian bayesian state estimation. *IEEE Proceedings on Radar and Signal Processing*, 140:107–113, 1993.

[65] R.M. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernatics*, 6:610–621, 1973.

[66] C. Herley and M. Vetterli. Wavelets and recursive filter banks. *IEEE trans. on Signal Processing*, 41:2536–2556, 1993.

[67] J.-B. Hiriart-Urruty and C. Lemaréchal. *Fundamentals of Convex Analysis*. Springer Verlag, Heidelberg, 2001.

[68] J. Huang and D. Mumford. Statistics of natural images and models. In *Proc. of the IEEE Int. Conf. on Computer Vision and Pattern Recognition*, pages 541–547, 1999.

[69] D. H. Hubel and T. N. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, 160:106–154, 1962.

[70] J. Hwang, S. Lay, and A. Lippman. Nonparametric multivariate density estimation: a comparative study. *IEEE trans. on Signal Processing*, 42:2795–2810, 1994.

[71] Besag. J. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society. Series B*, 36(2):192–236, 1974.

[72] A.K. Jain. *Fundamentals of digital image processing*. Prentice-Hall, Inc., 1989.

[73] T. Joachims. Making large-scale support vector machine learning practical. In A. Smola B. Schölkopf, C. Burges, editor, *Advances in Kernel Methods: Support Vector Machines*. MIT Press, Cambridge, MA, 1998.

[74] A. Katsaggelos and M. Kang. Simultaneous iterative image restoration and evaluation of the regularization parameter. *IEEE trans. on Signal Processing*, 40:2329–2334, 1992.

[75] C. Kervrann and J. Boulanger. Local adaptivity to variable smoothness for exemplar-based image denoising and representation. Technical Report 5624, INRIA, July 2005.

[76] Ch. Kervrann and J. Boulanger. Unsupervised patch-based image regularization and representation. In *Proc. of the Europ. Conf. on Computer Vision*, pages 555–567, 2006.

[77] Ch. Kervrann, J. Boulanger, and P. Coupé. Bayesian non-local means filter, image redundancy and adaptive dictionaries for noise removal. In *Int. Conf. on Scale Space Methods and Variational Methods in Computer Vision*, pages 520–532, 2007.

[78] B.B. Kimia, A. Tannenbaum, and S.W. Zucker. On the evolution of curves via a function of curvature. i. the classical case. *Journal Mathematical Analysis Applications*, 163:438–458, 1992.

[79] R. Kimmel. Demosaicing: Image reconstruction from color CCD samples. *IEEE trans. on Image Processing*, 8(9):1221–1228, 1999.

[80] R. Kimmel, R. Malladi, and N. Sochen. Image processing via the beltrami operator. In *Proc. of the Asian. Conf. on Computer Vision*, pages 574–581, 1998.

[81] R. Kimmel, N.A. Sochen, and R. Malladi. From high energy physics to low level vision. In *Proc. of the Int. Conf. on Scale-Space Theory in Computer Vision*, pages 236 – 247, 1997.

[82] R. Kimmel, N.A. Sochen, and R. Malladi. Images as embedding maps and minimal surfaces: A unified approach for image diffusion. In *Proc. of the IEEE Int. Conf. on Image Processing*, pages 368–371, 1997.

[83] S. Kindermann, S. Osher, and P. Jones. Deblurring and denoising of images by nonlocal functionals. *SIAM Multiscale Modeling and Simulation*, 4:1091 – 1115, 2005.

[84] G. Kitagawa. Monte carlo filter and smoother for non-gaussian nonlinear state space models. *Journal of Computational and Graphical Statistics*, 5:1–25, 1996.

[85] J.J. Koenderink. The structure of images. *BioCyber*, 50, 1984.

[86] P. Kornprobst and G. Aubert. *Mathematical Problems in Image Processing Partial Differential Equations and the Calculus of Variations*. Springer, 2002.

[87] P. Kornprobst, R. Deriche, and A. Gilles. Nonlinear operators in image restoration. In *Proc. of the IEEE Int. Conf. on Computer Vision and Pattern Recognition*, pages 325–330, 1997.

[88] H. Krim, D. Tucker, S. Mallat, and D. Donoho. On denoising and best signal representation. *IEEE trans. on Information Theory*, 45(7):2225–2238, 1999.

[89] D. Le Bihan, J.-F. Mangin, C. Poupon, C. A. Clark, S. Pappata, N. Molko, and H. Chabriat. Diffusion tensor imaging: concepts and applications. *Journal of Magnetic Resonance Imaging*, 13:534–546, 2001.

[90] E. Le Pennec and S. Mallat. Sparse geometric image representations with bandelets. *IEEE trans. on Image Processing*, pages 423– 438, 2005.

[91] A. Lee, K. Pedersen, and D. Mumford. The nonlinear statistics of high-contrast patches in natural images. *Int. Journal of Computer Vision*, pages 83 – 103, 2003.

[92] S. Lee. Digital image smoothing and the sigma filter. *CVGIP*, 24(2):255–269, November 1983.

[93] O. V. Lepski, E. Mammen, and V. G. Spokoiny. Optimal spatial adaptation to inhomogeneous smoothness: an approach based on kernel estimates with variable bandwidth selectors. *Annals of Statistics*, 25(3):929–947, 1997.

[94] M. S. Lewicki and T. J. Sejnowski. Learning overcomplete representations. *Neural Computation*, 12(2):337–365, 2000.

[95] S. Z. Li. *Markov Random Field Modeling in Computer Vision*. Springer, 1995.

[96] C. Liu, W.T. Freeman, R. Szeliski, and S.B. Kang. Noise estimation from a single image. In *Proc. of the IEEE Int. Conf. on Computer Vision and Pattern Recognition*, pages 901– 908, 2006.

[97] M.R. Luettgen, W.C. Karl, A.S. Willsky, and R. Tenney. Multiscale representations of markov random fields. *IEEE trans. on Signal Processing*, 41(12):3377–3396, 1993.

[98] M. Mahmoudi and G. Sapiro. Fast image and video denoising via nonlocal means of similar neighborhoods. *IEEE, Signal Processing Letters*, 12:839 – 842, 2005.

[99] J. Mairal, M. Elad, and G. Sapiro. Sparse representation for color image restoration. *IEEE trans. on Image Processing*, 17(1):53–69, 2008.

[100] J. Mairal, G. Sapiro, and M. Elad. Multiscale sparse image representation with learned dictionaries. In *Proc. of the IEEE Int. Conf. on Image Processing*, pages 105–108, 2007.

[101] R. Malladi and J.A. Sethian. Flows under min/max curvature flow and mean curvature: Applications in image processing. In *Proc. of the Europ. Conf. on Computer Vision*, pages 251–262, 1996.

[102] S. Mallat. A theory for multiresolution signal decomposition: the waveletrepresentation. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 11:674 – 693, 1989.

[103] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 1998.

[104] Lee Markosien. A motion-compensated filter for ultrasound image sequences. Technical report, Providence, RI, USA, 1996.

[105] M. Martín-Fernández, C.F. Westin, and C. Alberola-López. 3D Bayesian regularization of diffusion tensor MRI using multivariate Gaussian Markov random fields. In *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention*, pages 351–359, 2004.

[106] G. McLachlan, D. Peel, and P. Prado. Maximum likelihood clustering via normal mixture models. *Signal Processing Image Communication*, 8:105–111, 1996.

[107] Y. Meyer. *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations: The Fifteenth Dean Jacqueline B. Lewis Memorial Lectures*. American Mathematical Society, 2001.

[108] L. Moisan. Affine plane curve evolution: A fully consistent scheme. *IEEE trans. on Image Processing*, 7(3):411–420, 1998.

[109] N. Neji, R. Azzabou, G. Fleury, and N. Paragios. Diffusion tensor estimation, regularization and classification. In N. Ayache N.Paragios, J. Duncan, editor, *Biomedical Image Analysis: Methodologies and Applications*. Springer, 2008 edition, 2008.

[110] R. Neji, N. Azzabou, N. Paragios, and G. Fleury. A convex semi-definite positive framework for dti estimation and regularization. In *Int. Symp. in Visual Computing,*, pages 220–229, 2007.

[111] A. Ogier, P. Hellier, and C. Barillot. Speckle reduction on ultrasound image by variational methods and adaptive Lagrangian multipliers. In *IEEE Int. Symp. on Biomedical Imaging: From Nano to Macro*, pages 547–550, 2004.

[112] S. Paris and F. Durand. A fast approximation of the bilateral filter using a signal processing approach. In *Proc. of the Europ. Conf. on Computer Vision*, pages 568–580, 2006.

[113] E. Parzen. On the estimation of a probability density function and mode. *Annals of Mathematical Statistics*, 33:1065–1076, 1962.

[114] A. Perez. Markov random fields and images. Technical Report N1196, IRISA, 1998.

[115] P. Perona and J. Malik. Scale space and edge detection using anisotropic diffusion. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 12:629–639, 1990.

[116] A. Pizurica, W. Philips, I. Lemahieu, and M. Acheroy. A versatile wavelet domain noise filtration technique for medical imaging. *IEEE trans. on Medical Imaging*, 22(3):323–331, 2003.

[117] J. Polzehl and V. Spokoiny. Adaptive weights smoothing with applications to image restoration. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 62:335–354, 2000.

[118] K. Popat and R. W. Picard. Cluster based probability model and its application to image and texture processing. *IEEE trans. on Image Processing*, 6(2):268–284, 1997.

[119] J. Portilla, V. Strela, M.J. Wainwright, and E.P. Simoncelli. Image denoising using scale mixtures of gaussians in the wavelet domain. *IEEE trans. on Image Processing*, 12(11):1338–1351, 2003.

[120] S. Roth and M.J. Black. Fields of experts: A framework for learning image priors. In *Proc. of the IEEE Int. Conf. on Computer Vision and Pattern Recognition*, pages 860–867, 2005.

[121] L. Rudin and S. Osher. Total variation based image restoration with free local constraints. *Proc. of the IEEE Int. Conf. on Image Processing*, 1:31–35, 1994.

[122] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal. *Physica D*, 60:259–268, 1992.

[123] P. Sallee and B.A. Olshausen. Image denoising using learned overcomplete representations. In *Proc. of the IEEE Int. Conf. on Image Processing*, pages 381–384, 2003.

[124] R. Salvador, A. Peña, D. Menon, T. Carpenter, J. Pickard, and Ed. Bullmore. Formal characterization and extension of the linearized diffusion tensor model. *Human Brain Mapping*, 24(2):144–155, 2005.

[125] S.J Sheather and M. C. Jones. A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society. Series B (Methodological)*, 53:683–690, 1991.

[126] E. Simoncelli. Modeling the joint statistics of images in the wavelet domain. In *Proc SPIE, 44th Annual Meeting*, volume 3813, pages 188–195, 1999.

[127] S.M. Smith and J. M. Brady. SUSAN: A new approach to low level image processing. *Int. Journal of Computer Vision*, 23(1):45–78, 1997.

[128] B. Smolka and K. Wojciechowski. Random walk approach to image enhancement. *Signal Processing*, 81:465–482, 2001.

[129] N.A. Sochen, R. Kimmel, and R. Malladi. A general framework for low-level vision. *IEEE trans. on Image Processing*, 7(3):310–318, 1998.

[130] J.L. Starck, E.J. Candes, and D.L. Donoho. The curvelet transform for image denoising. *IEEE trans. on Image Processing*, 11(6):670–684, 2002.

[131] E.O. Stejskal and J.E. Tanner. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *Journal of Chemical Physics*, 42:288–292, 1965.

[132] A.D. Szlam. *Non-stationary analysis of datasets and applications*. PhD thesis, Yale, 2006.

[133] A.N. Tikhonov. The regularization of ill-posed problems. *Soviet. Math. Dokl. Akad. Nauk.*, 153(1):49–52, 1963.

[134] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *Proc. of the Int. Conf. on Computer Vision*, pages 839–846, 1998.

[135] D. Tschumperlé. *PDE's Based Regularization of Multivalued Images and Applications*. PhD thesis, Unversite de Nice Sophia Antipolis, 2002.

[136] D. Tschumperle and R. Deriche. Diffusion tensor regularization with constraints preservation. In *Proc. of the IEEE Int. Conf. on Computer Vision and Pattern Recognition*, pages 948–953, 2001.

[137] D. Tschumperle and R. Deriche. Vector-valued image regularization with pdes: A common framework for different applications. In *Proc. of the IEEE Int. Conf. on Computer Vision and Pattern Recognition*, pages 651–656, 2003.

[138] M. Tur, K.-C. Chin, and J.W. Goodman. When is speckle noise multiplicative? *Applied Optics*, 21:1157–1159, 1982.

[139] M. Unser. Local linear transforms for texture measurements. *Signal Processing.*, 11:61–79, 1986.

[140] L. Vese and S. Osher. Modeling textures with total variation minimization and oscillating patterns in image processing. *Journal of Scientific Computing*, 19:553–572, 2003.

[141] M.J. Wainwright, E.P. Simoncelli, and A.S. Willsky. Random cascades of gaussian scale mixtures and their use in modeling natural images with application to denoising. In *Proc. of the IEEE Int. Conf. on Image Processing*, pages 260–263, 2000.

[142] M. Wand and M. Jones. *Kernel Smoothing*. Chapman & Hall,, 1995.

[143] Ching-Chun Wang. *A Study of CMOS Technologies for Image Sensor application*. PhD thesis, MIT, 2001.

[144] Z. Wang, B. Vemuri, Y. Chen, and T. Mareci. A constrained variational principle for direct estimation and smoothing of the diffusion tensor field from complex DWI. *IEEE trans. on Medical Imaging*, 23(8):930–939, 2004.

[145] J. Weickert. A review of nonlinear diffusion filtering. In *Scale-Space Theory in Computer Vision*, pages 3–28, 1997.

[146] J. Weickert. Coherence-enhancing diffusion filtering. *Int. Journal of Computer Vision*, 31:111–127, 1999.

[147] J. Weickert, C. Feddern, M. Welk, B. Burgeth, and T. Brox. PDEs for tensor image processing. In J. Weickert and H. Hagen, editors, *Visualization and Processing of Tensor Fields*, pages 399–414. Springer, 2006.

[148] J. Weickert, B.M. ter Haar Romeny, and M. Viergever. Image recovery via total variation minimization and related problems. *Numerische Mathematik*, 76(2):167–188, 1997.

[149] J. Weickert, B.M. ter Haar Romeny, and M. Viergever. Efficient and reliable schemes for nonlinear diffusion filtering. *IEEE trans. on Image Processing*, 7(2):398, 1998.

[150] A.S. Willsky. Multiresolution markov models for signal and image processing. *Proceedings of the IEEE*, 90(8):1396– 1458, 2002.

[151] A.P. Witkin. Scale-space filtering. In *International Joint Conf. on Artificial Intelligence*, pages 1019–1022, 1983.

[152] L. P. Yaroslavsky. Local adaptive image restoration and enhancement with the use of dft and dct in a running window. *Proceedings of SPIE Wavelet Applications in Signal and Image Processing*, 2825:2–13, 1996.

[153] L.P. Yaroslavsky. *Digital Picture Processing*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1985.

[154] Y. Yu and S.T. Acton. Speckle reducing anisotropic diffusion. *IEEE trans. on Image Processing*, 11(11):1260–1270, 2002.

[155] C. Yung-Gang, G. Yoshikazu, and G. Shun'ichi. Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations. *J. Differ Geom*, 33:–, 1991.

[156] S.C. Zhu, Y. Wu, and M. Mumford. Frame: Filters, random fields, and minimax entropy–towards a unified theory for texture modeling. In *Proc. of the IEEE Int. Conf. on Computer Vision and Pattern Recognition*, page 686, 1996.

[157] S.C. Zhu, Y.N. Wu, and D. Mumford. Filters, random-fields and maximum-entropy (frame): Towards a unified theory for texture modeling. *Int. Journal of Computer Vision*, 27(2):107–126, 1998.