



École Doctorale
d'Informatique,
Télécommunications
et Électronique
de Paris

Thèse

présentée pour obtenir le grade de docteur

de Télécom Paris Tech

Spécialité : Signal et Images

Aymen SHABOU

**Minimisation multi-étiquette d'énergies
markoviennes par coupe-minimum sur
graphe : application à la reconstruction
de la phase interférométrique en imagerie
RSO**

Soutenue le 9 Novembre 2010 devant le jury composé de

Vito Pascazio	Professeur à l'Université de Napoli Parthenope (Italy)	Président
Emmanuel Trouvé	Professeur à Polytech'Annecy-Chambéry	Rapporteur
Xavier Descombes	Directeur de recherches à INRIA Sophia Antipolis	Rapporteur
Hugues Talbot	Professeur associé à l'ESIEE	Examineur
Marc Sigelle	Maître de conférence à TELECOM ParisTech	Examineur
Florence Tupin	Professeur à TELECOM ParisTech	Directeur de thèse
Jérôme Darbon	Chargé de recherche CNRS à l'ENS Cachan	Co-encadreur

Résumé

Les approches markoviennes en imagerie et vision par ordinateur offrent un cadre mathématique élégant pour résoudre certains problèmes complexes. Des modèles de formation des données observées ainsi que des modèles décrivant le contexte spatial des images à reconstruire sont alors nécessaires pour exploiter de telles approches. Le plus souvent, la fonction d'énergie globale obtenue demeure difficile à minimiser en raison de ses propriétés mathématiques et de la grande dimension des données. On est alors confronté d'une part à des problèmes combinatoires et d'autre part aux limites en puissance des machines actuelles.

C'est dans ce cadre que s'inscrit le premier objectif de ces travaux de thèse, où nous proposons des algorithmes d'optimisation efficaces d'une classe d'énergies markoviennes multi-étiquettes de premier ordre qui permet de modéliser un nombre important d'applications en imagerie. C'est la classe d'énergies ayant comme attache aux données une fonction quelconque, et comme *a priori* une fonction convexe. Les algorithmes proposés reposent d'une part sur une technique assez récente d'optimisation discrète en vision, à savoir la coupe-minimum sur graphe, et d'autre part sur une stratégie d'optimisation itérative par des mouvements de partitions. De nouveaux mouvements, que nous avons appelés mouvements de partitions larges et multi-étiquettes (MPLM), permettent alors d'avoir un compromis entre la qualité de l'optimum atteint et la complexité algorithmique. Appliqués et validés sur une série d'expériences traitant de la restauration d'images naturelles et radar, ils ont pu montrer de bonnes performances comparés aux algorithmes d'optimisation de l'état de l'art.

Le cadre applicatif principal de cette thèse est la reconstruction du relief par interférométrie radar à synthèse d'ouverture. Cette méthode de calcul de modèles numériques de terrain est confrontée le plus souvent à la nature très bruitée des données interférométriques radar et aussi à la complexité des scènes naturelles et urbaines à reconstruire. Les approches markoviennes trouvent alors leur justification pour s'appliquer à ce type de problèmes, grâce à des modèles bien adaptés au type de bruit sur ces données et au contexte spatial des scènes. Ainsi, le deuxième objectif de ces travaux de thèse consiste à proposer des modèles markoviens robustes face à la diversité des scènes à reconstruire dans le cas général. Un modèle a été alors proposé dans un premier temps qui repose d'une part sur un ensemble d'observations multicanaux de phase et d'autre part sur la variation totale comme *a priori*, pour parvenir à une reconstruction précise du relief. Ce modèle est validé tout d'abord sur des données de synthèse et testé ensuite sur des données réelles de scènes naturelles acquises par le capteur radar ERS-2, avec prise en compte de perturbations atmosphériques globales.

Confrontés à des scènes urbaines nécessitant une meilleure précision et localisation des structures non-naturelles, un deuxième modèle a été exploré, exploitant cette fois-ci conjointement l'information de la phase et de l'amplitude. Des algorithmes d'optimisation conjointe bien appropriés à cette nouvelle forme d'énergie vectorielle sont alors nécessaires. Appliquée à des données radar aériennes d'une scène urbaine, acquises par le capteur E-SAR, cette approche conjointe apporte une meilleure reconstruction en terme de précision et offre des perspectives pour exploiter conjointement d'autres données de natures différentes.

Abstract

The MRF in computer vision and image analysis is a powerful framework for solving complex problems using the MAP estimation. First of all an energy function that encodes both the noise affecting the observed data and the *a priori* on the estimated solution has to be fixed. Then an appropriate optimization algorithm of this energy is necessary in order to solve the problem efficiently. However, some image processing problems deal with high dimensional data and require complex MRF energy functions. Thus, optimization process becomes a hard task.

One of the main contributions of this thesis is developing new efficient optimization algorithms of a class of first order multi-label MRF energies : any likelihood terms and convex prior. This energy form characterizes a large set of image processing applications. The proposed algorithms rely on the combinatorial optimization technique, called graph-cut, and an iterative strategy to converge toward a local optimum, called partition move. We propose new moves, called large and multilabel partition moves (LMPM), that offer a trade-off between optimum quality and algorithm complexity. The graph-cut technique is adapted and used to compute the optimal moves in efficient way. These new algorithms are tested in the case of image denoising task, showing us good performances compared to the state of the art algorithms.

The main application of our work is the digital elevation model (DEM) estimation using interferometric synthetic aperture radar (InSAR) data. This problem is usually considered as a complex and an ill-posed one, because of the high noise rate affecting interferograms and the complex structures qualifying real natural and urban area. Therefore, MRF approaches are well adapted to deal with this task, since specific statistical and *a priori* energy functions could be proposed to well encode several problems affecting these data and also to well describe the complex knowledge on the estimated solution. So, a first model is proposed in this work that relies on a multichannel interferometric likelihood density function and the total variation regularization. Optimizing this model using the proposed LMPM based algorithms leads to a precise DEM. This approach is tested on several synthetic and real InSAR ERS-2 images, and also adapted to deal with multichannel data even if global atmospheric perturbations characterize them.

To reconstruct DEM in urban area, especially the buildings, more precise model is required. Therefore, a second MRF model is proposed that lies in both amplitude and phase images. This new joint model preserves well the discontinuities in the reconstructed profiles. However, a new class of vectorial energy minimization algorithms has to be studied, in order to minimize efficiently the joint model. The proposed approach is then applied to E-SAR interferometric and amplitude images of urban area, showing us a more precise reconstruction than the previous approach.

Remerciements

Je tiens à exprimer mes profondes gratitude et reconnaissances à toutes les personnes qui ont contribué à la réalisation et l'aboutissement de ce travail de recherche.

Je commence par remercier M. Emmanuel Trouvé (Professeur à Polytech'Annecy-Chambéry) et M. Xavier Descombes (Directeur de recherche à INRIA Sophia Antipolis Méditerranée), rapporteurs de ce travail ; ainsi que M. Vito Pascazio (Professeur à l'Université de Napoli Parthenope, Italy), président du Jury ; M. Marc Sigelle (Maître de conférence à TELECOM ParisTech) et M. Hugues Talbot (Professeur associé à l'ESIEE), examinateurs ; pour avoir accepté de participer au jury de thèse, pour leur minutieuse relecture du manuscrit, ainsi que leurs commentaires et remarques très riches.

Je remercie Mme. Florence Tupin (Professeur à TELECOM ParisTech) et M. Jérôme Darbon (Chargé de recherche CNRS à l'ENS Cachan) pour avoir accepté de diriger cette thèse. Leurs conseils et soutient m'ont toujours été très précieux pour me retrouver dans ce vaste domaine de la recherche en traitement d'image.

Une partie de ce travail est le fruit d'une collaboration avec Giampaolo Ferraioli (Post doctorant à l'université de Napoli Parthenope), durant son séjour doctoral à TSI. Je tiens à le remercier ainsi que son équipe pour la qualité de cette coopération.

J'adresse aussi mes remerciements à Mme. Marie-Pierre Béale (Chargée de recherche CNRS à l'ENS), pour m'avoir préparé et fournie des données InSAR sur la zone de Serre-Ponçon.

Je souhaite aussi remercier tous mes collègues du bureau pour la bonne ambiance et entente tout le long de ces trois années : Benoit, Guillaume, Charles, Gabrielle, Dina et Hélène ; sans oublier aussi les autres collègues de TSI : Nicolas, Sofiène, Vincent, Xia, Julien, ...et tout le monde.

Je remercie finalement ma famille, de l'autre côté de la méditerranée, de m'avoir encouragé à commencer et poursuivre jusqu'au bout cette aventure.

Table des matières

Résumé	2
Abstract	3
Remerciements	7
Table des matières	9
Introduction	18

I Minimiseurs multi-étiquettes d'énergies markoviennes de premier ordre 21

1	Approche markovienne et algorithmes d'optimisation	23
1	Problème d'étiquetage	23
1.1	L'estimation bayésienne	23
1.2	La modélisation markovienne	24
1.3	Résolution par minimisation d'énergie	27
2	Algorithmes de minimisation par coupe-minimum sur graphe	28
2.1	Coupe-minimum / flot-maximum	29
2.2	Champs markoviens binaires	30
2.2.1	Fonctions de $\mathbb{B}^2$	30
2.2.2	Fonctions de $\mathbb{B}^q$	33
2.3	Champs markoviens multi-étiquettes	34
2.3.1	Fonctions de $\mathbb{F}^2$	35
2.3.2	Fonctions de $\mathbb{F}^q$	38
2.4	Bilan	39
2.4.1	Limites de ces algorithmes	39
2.4.2	Quelques expériences	39
2.5	Conclusion	40
2	Optimisation par des mouvements de partitions multi-étiquettes	45
1	Optimisation multi-étiquette	45
2	Construction du graphe d'un MPLM optimal	48
3	Propriétés de convergence	52
3.1	Complexité algorithmique	52
3.2	Qualité de l'optimum	53
4	Expériences et résultats	56
4.1	Parcours des étiquettes	56

4.2	Restauration d'images bruitées par un bruit impulsif	59
5	Conclusion	73
3	Minimisation markovienne mixte : discrète et continue	75
1	Limites des approches de minimisation discrète ou continue	75
1.1	Minimisation discrète	75
1.2	Minimisation continue	76
2	Algorithme de minimisation mixte : discrète et continue	80
2.1	Phase 1 : Minimisation discrète	80
2.2	Phase 2 : Minimisation continue	80
3	Expériences et résultats	82
3.1	Illustration sur une fonction à deux variables	82
3.2	Restauration d'image radar à synthèse d'ouverture	85
4	Conclusion	90
II	Estimation markovienne de modèles numériques de terrain	93
4	Reconstruction du relief en imagerie RSO	95
1	Principes de l'imagerie radar	95
1.1	Capteurs radar	95
1.2	Résolution d'un capteur radar	96
1.3	Les systèmes RSO	97
1.4	L'image RSO	98
2	Approches de reconstruction du relief en imagerie RSO	100
2.1	Approches de reconstruction pixelique	101
2.2	Approches de reconstruction par régions et objets	102
3	Estimation des MNT par interférométrie	103
3.1	Principe de l'interférométrie	103
3.2	Contributions dans la phase mesurée	105
3.3	Produit interférométrique	106
3.4	Principales perturbations de la mesure interférométrique	108
3.5	Déroulement de phase InRSO et estimation de MNT	109
4	Approches bayésiennes de reconstruction de phase InRSO	111
4.1	Statistiques des données interférométriques	111
4.2	Ambiguïté de la phase InRSO	113
4.3	Reconstruction de la phase InRSO par l'approche multicanal	114
4.4	Algorithmes de reconstruction de phase InRSO	119
4.4.1	Algorithme de reconstruction de phase par une approche monocanal	119
4.4.2	Algorithme de reconstruction de phase par une approche multicanal	119
5	Conclusion	120
5	Reconstruction de MNT par une approche markovienne multicanal	123
1	Modèle MCPU+TV	124
2	Étude de l'influence de la cohérence sur l'estimation multicanal	125
3	Minimisation markovienne et estimation des hyperparamètres locaux	128
3.1	Étape 1 : optimisation discrète	128
3.2	Étape 2 : estimation des hyperparamètres locaux	130
3.3	Étape 3 : optimisation continue	131
4	Expériences sur des données de synthèse	132

4.1	Validation du modèle <i>a priori</i>	133
4.2	Étude de l'algorithme d'optimisation	137
5	Prise en compte des perturbations globales	141
5.1	Modèle de vraisemblance avec perturbations de phase globales	141
5.2	Algorithme d'estimation conjointe altitude et perturbations globales	142
5.3	Résultats de reconstruction	143
6	Résultats sur des données réelles	145
7	Conclusion	158
6	Reconstruction conjointe de la phase InRSO et de l'amplitude RSO	161
1	Introduction	161
2	Modèle markovien conjoint (MCPAR+TV)	162
2.1	Modèle de vraisemblance	163
2.2	Modèle <i>a priori</i>	164
3	Minimisation markovienne vectorielle	164
3.1	Mouvements binaires vectoriels	165
3.2	Mouvement vectoriel binaire optimal	166
3.2.1	Mouvement vectoriel binaire incomplet	167
3.2.2	Mouvement vectoriel binaire complet	168
3.2.3	Choix entre mouvement vectoriel binaire complet ou incomplet	170
3.3	Algorithme de minimisation vectorielle	171
3.3.1	Minimisation par des sauts vectoriels :	171
3.3.2	Minimisation par des extensions vectorielles :	171
3.3.3	Choix entre les stratégies de saut ou d'extension :	172
4	Évaluations expérimentales sur des données de synthèse	176
4.1	Validation du modèle conjoint	176
4.2	Comparaison entre les modèles MCPAR+TV et MCPU+TV	179
4.3	Comparaison entre les mouvements de saut et extension vectoriels	181
5	Résultats sur des données réelles	183
6	Conclusion	189
	Conclusions et perspectives	192
	Annexes	199
	A Construction de graphes binaires des algorithmes MPLB	199
	B Construction du graphe multi-étiquette d'Ishikawa	205
	C Flot-maximum dans un réseau	209
	D Géométrie de prise de vue en interférométrie Radar	215
	E Statistiques sur les données InRSO	217
	Liste des publications	221
	Bibliographie	221

Liste des abréviations

Pour des raisons de lisibilité, la signification d'une abréviation ou d'un acronyme n'est souvent rappelée qu'à sa première apparition dans le texte d'un chapitre. Nous donnons ainsi la liste complète des abréviations utilisées dans ce manuscrit de thèse. Certaines correspondent à des mots en anglais, leur traduction est aussi fournie.

FastDC	<i>Fast Discrete-Continuous</i> (Algorithme d'optimisation rapide mixte : discrète-continue)
ICM	<i>Iterated Conditional Mods</i> (Modes Conditionnels Itérés)
InRSO	<i>Interférométrie radar à synthèse d'ouverture</i>
LGMRF	<i>Local Gaussian Markov Random Fields</i> (Modèle markovien-gaussien local)
MAP	<i>Maximum a posteriori</i>
MCPAR	<i>Multi-Channel Phase and Amplitude Reconstruction</i> (Reconstruction multi-canal d'images de phase et d'amplitude)
MCPAR+TV	<i>Multi-Channel Phase and Amplitude Reconstruction with</i> <i>Total Variation prior on phase and amplitude</i> (Reconstruction multi-canal d'images de phase et d'amplitude avec un a priori de type variation totale sur la phase et l'amplitude)
MCPU	<i>Multi-Channel Phase Unwrapping</i> (Déroulement de phase multi-canal)
MCPU+TV	<i>Multi-Channel Phase Unwrapping with Total Variation prior</i> (Déroulement de phase multi-canal avec un a priori de type variation totale)
MNT	<i>Modèle Numérique de Terrain</i>
MPLB	<i>Mouvement de Partition Large et Binaire</i>
MPLM	<i>Mouvement de Partition Large et Multi-étiquette</i>
QPBO	<i>Quadratic Pseudo-boolean Optimization</i> (Optimisation de fonction quadratique pseudo-booléenne)
RSO	<i>Radar à Synthèse d'Ouverture</i>
TV	<i>Total Variation</i> (Variation totale)

Introduction générale

Contexte général et objectifs

De nombreux problèmes de traitement d'image peuvent être résolus par une approche markovienne, c'est à dire qu'ils s'appuient sur une information contextuelle. En effet, les champs de Markov, associés à la théorie bayésienne de décision, offrent un cadre mathématique cohérent et élégant pour résoudre ces problèmes. Le principe est de se ramener à un problème d'étiquetage, dont la résolution consiste à associer des étiquettes aux pixels de l'image de façon à minimiser un critère bien défini. Un tel formalisme nécessite tout d'abord le choix de la fonction de coût à minimiser en fonction des données traitées et des contraintes de l'application visée. Le plus souvent, cette fonction (communément appelée énergie) repose à la fois sur les statistiques du bruit qui affecte les données et sur le contexte spatial de l'image à reconstruire. Il est ensuite nécessaire de disposer d'un outil d'optimisation qui soit capable de calculer la configuration qui minimise cette fonction. L'espace des configurations étant extrêmement vaste, et augmentant en taille exponentiellement avec la taille des données, requiert des stratégies de recherche efficaces. C'est l'une des raisons qui a motivé le développement de techniques d'optimisation markovienne efficaces en imagerie et vision par ordinateur.

La minimisation de certaines classes de fonctions non-convexes et de grandes dimensions est désormais possible. Les algorithmes proposés dans la littérature varient entre algorithmes d'optimisation exacte et d'autres d'optimisation non-exacte ou approximative. L'avantage des premiers est la qualité de l'optimum vers lequel ils convergent, qui est un optimum global, mais au prix d'une complexité algorithmique qui explose avec la taille des données et même avec la nature de la fonction d'énergie qui doit par ailleurs satisfaire certaines conditions. En contrepartie, les algorithmes d'optimisation approximative offrent de bonnes performances de calcul, au prix de la qualité de l'optimum obtenu. Le compromis entre la qualité de l'optimum et la complexité algorithmique constitue alors le cœur des méthodes récentes d'optimisation markovienne appliquées aux problématiques d'imagerie traitant des données de grandes dimensions et des formes d'énergies non-convexes.

Nos travaux s'inscrivent dans le cadre applicatif de l'interférométrie satellitaire pour fournir des modèles numériques de terrain (MNT). Une telle problématique est d'un intérêt grandissant puisque ces modèles nourrissent d'autres problématiques en télédétection comme l'interférométrie différentielle pour la détection et le suivi de mouvement de terrain, ou la reconstruction de modèles numériques d'élévations et de structures 3D dans un milieu urbain. Plusieurs approches existent dans l'état de l'art pour parvenir à reconstruire le relief, telles que la stéréovision à partir d'images optiques et la radargrammétrie (ou stéréo radar) à partir d'images radar, ainsi que l'interférométrie radar. Dans cette dernière approche, le calcul du MNT se fait à partir de données de phase interférométriques, appelées interférogrammes. Ces données obtenues après une chaîne d'acquisition et de pré-traitement sont caractérisées par un fort taux de bruit qui rend leur exploitation complexe. De plus,

ces mesures de phases étant enroulées dans $[-\pi, \pi[$, illustrées sous forme de franges dans les interférogrammes, nécessitent un processus de déroulement pour retrouver la bonne altitude. Cette étape, connue sous le nom de déroulement de phase, se transforme en un problème difficile à résoudre en présence d'un fort bruit ou lorsqu'il existe de fortes discontinuités dans le milieu à reconstruire. Les méthodes markoviennes trouvent alors leur intérêt dans ce cadre, où il est possible grâce à des modèles de formation de ces données bruitées (ou observations) et des modèles apportant une information contextuelle (notamment les discontinuités et les structures) de bien reconstruire le relief d'une scène imagée. Ces approches permettent également de définir un cadre général supportant l'introduction de nombreuses informations. En revanche, elles sont confrontées aux problèmes qui sont liés à la nécessité d'une modélisation précise des phénomènes physiques contribuant à la formation des données et à l'optimisation numérique des fonctions d'énergie qui en découlent, qui demeurent le plus souvent difficiles à minimiser.

L'objectif de ces travaux de thèse est donc double :

1. Un premier objectif est d'étudier et proposer de nouveaux algorithmes d'optimisation non-convexe qui offrent un compromis entre la qualité de l'optimum atteint et la complexité algorithmique. Une telle étude est voulue générale et adressée à toute problématique en imagerie et vision par ordinateur qui nécessite un algorithme d'optimisation efficace et robuste à des modèles de fonctions d'énergie complexes. Les algorithmes proposés exploitent l'efficacité d'une technique assez récente en optimisation numérique dans le cadre de la vision par ordinateur, à savoir la coupe-minimum sur graphe. De plus, des méthodes d'optimisation continue sont aussi utilisées afin d'améliorer en précision l'estimation et en efficacité la convergence de l'algorithme.
2. Un second objectif plus applicatif consiste à proposer de nouveaux modèles markoviens pour la reconstruction de modèles numériques de terrains en imagerie radar. Ces modèles reposent sur les spécificités des données radar interférométriques (images de phase, amplitude et cohérence) et sur le contexte spatial des divers milieux à reconstruire : milieux naturels et milieux urbains. Ils s'inscrivent par ailleurs dans le cas où plusieurs acquisitions de la même scène sont disponibles, afin d'aboutir à des MNT précis. Disposant d'algorithmes efficaces d'optimisation non-convexe, l'optimisation de ces modèles est désormais possible.

Organisation du manuscrit

Le manuscrit de thèse est réparti sur deux axes selon les deux objectifs fixés.

La première partie traite de l’optimisation multi-étiquettes des énergies markoviennes de premier ordre.

Un bref aperçu des principaux algorithmes d’optimisation d’énergies markoviennes est proposé dans le premier chapitre. Nous nous focalisons sur les algorithmes reposant sur les techniques de coupe-minimum sur graphes et les mouvements de partitions. Nous illustrons brièvement au départ les différentes constructions de graphes selon les modèles étudiés. Ensuite, une analyse critique de ces méthodes est proposée à la fin du chapitre illustrant leurs limites et situant les problèmes auxquels nous nous attaquons.

Une nouvelle approche d’optimisation d’une classe de modèles markoviens non-convexes de premier ordre, où la fonction d’attache aux données est quelconque et celle de régularisation est convexe, est alors introduite dans le deuxième chapitre. Cette approche repose sur la technique de coupure minimale sur graphes et offre un ensemble d’algorithmes d’optimisation efficace grâce à des nouveaux mouvements de partitions, dits mouvements multi-étiquettes. Ces algorithmes sont beaucoup moins coûteux en complexité que les optimiseurs exacts et convergent vers des optima de meilleure qualité que ceux des optimiseurs approximatifs de l’état de l’art. Nous illustrons les contributions de ces nouveaux algorithmes dans le cadre de la restauration d’images bruitées, où le modèle de bruit est fortement non-convexe.

Pour améliorer d’avantage la qualité de l’optimum, nous proposons dans le troisième chapitre un algorithme mixte d’optimisation discrète et continue. Cet algorithme exploite les avantages des optimiseurs discrets déjà développés et ceux continus de l’état de l’art afin d’aboutir à une solution algorithmique qui permet à la fois de converger vers des optima de bonne précision tout en conservant un aspect rapide et non encombrant en mémoire. L’apport de cet algorithme est illustré à travers la restauration d’images radar en télédétection, où le besoin d’une bonne précision sur l’estimation de la réflectivité de la scène et la nécessité d’une efficacité dans le traitement se font de plus en plus ressentir pour des futurs traitements et interprétations de ces données.

La deuxième partie est consacrée à la reconstruction de modèles numériques de terrain en imagerie radar, et particulièrement en interférométrie.

Les études réalisées dans ce cadre d’application sont nombreuses, exploitant différents formalismes. Après un bref aperçu des principaux travaux présents dans la littérature, nous nous focalisons dans le quatrième chapitre sur les approches de reconstruction de relief par interférométrie. Une étude de telles approches est réalisée afin d’éclaircir les problèmes fondamentaux auxquels elles sont confrontées.

Un premier modèle markovien est alors proposé dans le cinquième chapitre. Ce modèle repose sur les données de phases interférométriques et exploite les points forts de certaines approches proposées dans la littérature, à savoir la technique multicanal en interférométrie pour réduire l’ambiguïté de la reconstruction du relief et la technique de régularisation par variation totale en imagerie pour conserver le mieux possible les structures du milieu reconstruit, tout en réduisant le bruit qui affecte les données. Le recours aux optimiseurs efficaces de la première partie du manuscrit de thèse nous a été d’une grande utilité pour mieux optimiser le modèle proposé. Des résultats sur des données synthétiques variées valident dans un premier temps le modèle et l’algorithme de résolution. Ensuite, des résultats sur des données interférométriques réelles sur une zone de Serre-Ponçon, jugées complexe

en raison de l'importance de la topographie et de la perte de corrélation, montrent la robustesse et l'efficacité de l'approche proposée.

Pour élargir notre étude à des scènes variées, urbaines ou non urbaines, nous proposons dans le sixième chapitre un deuxième modèle markovien. Ce dernier repose en plus des données de phase interférométrique sur les données d'amplitude pour apporter une information encore plus précise sur les structures à conserver dans le MNT reconstruit. Cependant, vue la nature conjointe de ce modèle, des algorithmes d'optimisation de nature différente de ceux utilisés jusqu'à ce stade du manuscrit de thèse sont nécessaires. Nous proposons ainsi une nouvelle famille d'algorithmes d'optimisation conjointe qui s'appuient sur les idées de la première partie de la thèse, profitant donc de l'efficacité des techniques de coupe-minimum sur graphe et des mouvements de partitions. La robustesse du modèle et l'efficacité de ces algorithmes nous ont permis de fournir des modèles numériques de terrains de scènes urbaines avec une bonne précision. Une zone urbaine de Dresde a alors été reconstruite par cette méthode, partant de données radar aériennes de phase et d'amplitude. Les résultats de reconstruction sont encourageants.

Le dernier chapitre conclut le manuscrit de thèse et présente un ensemble de perspectives et idées qui découlent du travail réalisé.

Première partie

Minimiseurs multi-étiquettes
d'énergies markoviennes de premier
ordre

Chapitre 1

Approche markovienne et algorithmes d'optimisation

Ce chapitre est consacré à l'étude bibliographique des approches d'optimisation des énergies markoviennes utilisées en traitement d'image et vision par ordinateur.

En premier lieu, la formulation de certains problèmes d'imagerie en un problème d'éti-quetage, résolu ensuite par une estimation bayésienne par maximum *a posteriori* (MAP), est décrite.

En deuxième lieu, l'approche markovienne est introduite pour modéliser l'information *a priori* de l'image.

En troisième lieu, des algorithmes d'optimisation du critère probabiliste MAP sont analysés. Plus particulièrement, nous étudions les algorithmes reposant sur la technique du flot-maximum/coupe-minimum sur un graphe. Dans des travaux récents, ces algorithmes ont montré leur efficacité et robustesse pour minimiser différentes formes d'énergies markoviennes utiles pour la vision par ordinateur.

1 Problème d'étiquetage

Différents problèmes en traitement d'image peuvent être reformulés sous la forme d'un problème d'étiquetage (*labelling*). C'est par exemple le cas pour la restauration d'images bruitées, la segmentation des images, la reconstruction 3D à partir d'observations 2D d'une même scène, etc.

Désignons par $\mathcal{P} = \{p_1, p_2, \dots, p_n\}$ l'ensemble des sites de l'image ou du graphe considéré et $\mathcal{L} = \{l_1, \dots, l_k\}$ l'ensemble des étiquettes qui leur sont associées. Le problème d'étiquetage consiste à mettre en correspondance une étiquette de l'ensemble $\mathcal{L}$ à chaque site de $\mathcal{P}$. Estimer une solution pour une application donnée revient à trouver le meilleur étiquetage selon un critère bien choisi.

1.1 L'estimation bayésienne

Il existe dans la littérature différentes approches d'estimation. Parmi celles-ci, nous trouvons la classe des approches probabilistes par inférence statistique bayésienne. Dans ce cadre, l'estimation consiste à inférer la solution du problème à partir d'un modèle probabiliste.

Partant d'une observation ou un ensemble d'observations, indiquées respectivement par $y = \{y_1, y_2, \dots, y_n\}$ ou $y = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n\}$ (avec $\mathbf{y}_p$ un vecteur d'observations du site p), la solution à estimer $x = \{x_1, x_2, \dots, x_n\}$ est obtenue en considérant x et y comme étant des

réalisations d'un champ aléatoire $z = \{z_1, z_2, \dots, z_n\}$ et en définissant une fonction de coût sur les erreurs de la solution.

Un des estimateurs largement utilisé en traitement d'image est l'estimateur bayésien MAP (Li, 2001). A cet estimateur correspond une fonction de coût en tout ou rien dans la solution globale. L'estimation consiste donc à maximiser pour x la densité de probabilité *a posteriori* $P(x|y)$:

$$\tilde{x} = \arg \max_{x \in \mathcal{X}} P(x|y), \quad (1.1)$$

avec $\tilde{x}$ la solution estimée et $\mathcal{X} = \mathcal{L} \times \mathcal{L} \times \dots \times \mathcal{L} = \mathcal{L}^n$ l'espace des solutions candidates.

Grâce à la règle de Bayes, ce problème se ramène à chercher la configuration x qui maximise le produit de la densité de probabilité *a priori* $P(x)$ par la fonction de vraisemblance $P(y|x)$:

$$\tilde{x} = \arg \max_{x \in \mathcal{X}} P(y|x)P(x). \quad (1.2)$$

Pour un problème donné, la fonction de vraisemblance correspond au modèle de formation des données (observations). En pratique, ce modèle est le plus souvent tiré d'une étude statistique sur ces observations tout en faisant une hypothèse d'indépendance conditionnelle entre les éléments y_p :

$$P(y|x) = \prod_{p=1}^n P(y_p|x_p), \quad (1.3)$$

avec $P(y_p|x_p)$ une loi connue (estimée statistiquement).

Par ailleurs, la distribution *a priori*, qui correspond au modèle qui décrit l'image, est déterminée à partir de la connaissance contextuelle *a priori* de l'estimation. En traitement d'image, un modèle permettant de formaliser cette information *a priori* dans une scène naturelle imagée est le modèle markovien. Ce dernier repose sur la localité des interactions entre observations.

1.2 La modélisation markovienne

Introduite en imagerie dans (Geman et Geman, 1984), cette approche consiste à modéliser la cohérence spatiale dans l'image, selon laquelle la probabilité de la configuration x_p d'un site donné p de l'image est déterminée à partir de la configuration des sites voisins. Un tel modèle est communément appelé modèle de régularisation permettant de lier la distribution globale $P(x)$ à des potentiels locaux.

Un système de voisinage, que l'on notera par $\mathcal{N}$, est alors défini de la façon suivante :

$$\mathcal{N}_p = \{q \in \mathcal{P}\} \text{ tels que } \begin{cases} p \notin \mathcal{N}_p, \\ q \in \mathcal{N}_p \Rightarrow p \in \mathcal{N}_q. \end{cases} \quad (1.4)$$

A titre d'exemple, en considérant les pixels d'une image, un système de voisinage possible est le système de *4-connexité*. Chaque pixel est voisin des pixels auxquels il est connecté dans la grille régulière.

Le système de voisinage est ensuite utilisé pour définir les sous-ensembles de sites en interaction spatiale, appelés cliques, que l'on notera par c . Une clique peut être par exemple un seul site $c = \{p\}$, une paire de sites voisins $c = (p, q)$, etc. L'ensemble que forment les cliques de même ordre m (c.-à-d. $m + 1$ sites par clique) sera indiqué par C_m . Par exemple,

$C_0 = \{p | p \in \mathcal{P}\}$, $C_1 = \{(p, q) | p \in \mathcal{P} ; q \in N_p\}$, etc. L'ensemble de toutes les cliques possibles sera noté par $\mathcal{C} = \bigcup_{i \in \mathbb{N}} C_i$.

En reprenant le formalisme bayésien, une image peut être alors considérée comme étant une réalisation d'un champ aléatoire bien particulier, à savoir un champ de Markov, défini par :

Définition 1.1. Champ de Markov

z est un champ de Markov sur $\mathcal{P}$ relativement au système de voisinage $\mathcal{N}$ si et seulement si z est un champ aléatoire qui satisfait les deux critères suivants :

- Positivité : $P(x) > 0 ; \forall x \in \mathcal{X}$
- Markoviannité : $P(x_p | x_{\mathcal{P}-\{p\}}) = P(x_p | x_{N_p})$. Avec, $x_{N_p} = \{x_q | q \in N_p\}$.

où x est une configuration du champ z .

Comme cette modélisation markovienne de l'image repose sur un ensemble d'interactions entre les sites voisins, il est alors nécessaire de les définir. La définition et le théorème suivants permettent donc de manipuler en pratique les modèles markoviens sur les images.

Définition 1.2. Champ de Gibbs

z est un champ de Gibbs sur $\mathcal{P}$ relativement au système de voisinage $\mathcal{N}$ si et seulement si toute configuration x de ce champ suit une distribution de Gibbs, c.-à-d. $P(x) = \frac{\exp(-U(x))}{Z}$, avec $Z = \sum_{x \in \mathcal{X}} \exp(-U(x))$ est une constante de normalisation appelée fonction de partition et $U(x) = \sum_{c \in \mathcal{C}} E_c(x)$ est une fonction d'énergie égale à la somme des potentiels E_c de toutes les cliques possibles dans $\mathcal{C}$. Le potentiel d'une clique dépend de la configuration locale associée à cette clique.

Théorème 1.1. Hammersley-Clifford (1971)

Un champ z est un champ de Markov défini sur un ensemble de sites $\mathcal{P}$ relativement à un système de voisinage $\mathcal{N}$ si et seulement si z est un champ de Gibbs sur $\mathcal{P}$ relativement au système de voisinage $\mathcal{N}$.

Cela nous amène à conclure que la distribution *a priori* de x selon un modèle de Markov est une distribution de Gibbs déterminée à partir de potentiels d'interactions définis sur des cliques selon la relation suivante :

$$P(x) = \frac{1}{Z} \exp \left(- \sum_{c \in \mathcal{C}} E_c(x) \right). \quad (1.5)$$

En injectant (1.3) et (1.5) dans (1.2), on voit que le problème de l'estimation MAP revient à résoudre le problème suivant :

$$\tilde{x} = \arg \max_{x \in \mathcal{X}} \left\{ \prod_{p=1}^n P(y_p | x_p) \exp \left(- \sum_{c \in \mathcal{C}} E_c(x) \right) \right\}. \quad (1.6)$$

Nous pouvons reformuler cette dernière expression par :

$$\tilde{x} = \arg \min_{x \in \mathcal{X}} E(x), \quad (1.7)$$

avec :

$$E(x) = \sum_{p=1}^n -\log(P(y_p | x_p)) + \sum_{c \in \mathcal{C}} E_c(x). \quad (1.8)$$

E est appelée fonction d'énergie markovienne associée au problème d'étiquetage.

Par la suite, nous considérons la notation suivante :

$$E(x_p) = E_p(y_p|x_p) = -\log(P(y_p|x_p)), \forall p \in \mathcal{P}. \quad (1.9)$$

Ces termes sont dits termes d'attache aux données ou de vraisemblance. Ils permettent de contraindre notre estimation à ressembler aux observations. Ainsi la fonction E_p doit pénaliser une variation importante entre l'estimation x_p et l'observation y_p .

Par ailleurs, le choix des potentiels des cliques E_c ainsi que l'ordre des cliques utilisées conditionnent directement l'*a priori* imposé à la solution. Souvent en imagerie, des cliques de premier ordre suffisent pour avoir une solution bien régularisée. On parle alors de *modèles markoviens de premier ordre*.

Dans certaines applications, telles que celles qui manipulent de la texture dans l'image, des cliques d'ordre supérieur deviennent nécessaires (Winkler, 2006). On utilise alors des *modèles markoviens d'ordre supérieur*.

Quant au choix des potentiels de cliques, il est souvent conditionné par la nature de la solution à estimer : homogène, lisse, lisse par morceaux, etc. Généralement, c'est ce dernier cas qui modélise une grande panoplie de scènes naturelles, où il est nécessaire de définir des potentiels qui permettent à la fois de régulariser la solution et de préserver les discontinuités. Les deux grandes familles de représentation des fonctions *a priori* modélisant ce type de connaissance sur l'image sont :

- La famille des potentiels faisant intervenir les discontinuités explicitement. Certains le font à travers des variables auxiliaires appelées processus de lignes booléens (Geman et Geman, 1984) marquant ces discontinuités. D'autres le font par des hyperparamètres locaux (Saqib *et al.*, 1998) estimés localement. Typiquement pour ce dernier cas, sur des cliques de premier ordre, on peut retrouver cette forme de potentiels :

$$E_{p,q}(x_p, x_q) = \beta_{p,q} \|x_p - x_q\|, \quad (1.10)$$

où $\beta_{p,q}$ est un hyperparamètre local et $\|\cdot\|$ est une distance pénalisant une variation importante entre les configurations de sites voisins.

- La famille des potentiels faisant intervenir les discontinuités implicitement par la fonction qui définit le potentiel lui même (Besag, 1986), (Hebert et Leahy, 1989), (Charbonnier, 1994). A titre d'exemple, nous trouvons les potentiels sur des cliques de premier ordre de type ϕ – *fonctions* telles que :

$$E_{p,q}(x_p, x_q) = \frac{(x_p - x_q)^2}{1 + (x_p - x_q)^2}, \quad (1.11)$$

ou de type quadratique tronquée telles que :

$$E_{p,q}(x_p, x_q) = \min \{(x_p - x_q)^2, b\}, \quad (1.12)$$

avec b une constante de troncature évitant de pénaliser les discontinuités à partir d'une valeur donnée.

A noter que quelques équivalences entre certains modèles des deux familles sont disponibles (Winkler, 2006).

Ayant défini les distributions de vraisemblance et *a priori* du modèle bayésien, le problème de l'estimation MAP se ramène donc à un problème de minimisation de la fonction d'énergie E .

1.3 Résolution par minimisation d'énergie

Minimiser la fonction E relève du domaine de l'optimisation numérique. Un algorithme de minimisation utilisé devrait répondre aux critères suivants :

- **Efficacité** : l'algorithme doit être de faible complexité en temps de calcul et en mémoire ;
- **Précision** : l'algorithme doit converger vers un optimum de bonne qualité, idéalement un optimum global de la fonction.

Dans la littérature, différentes méthodes d'optimisation existent. Ces méthodes peuvent être classées selon plusieurs critères. A titre d'exemple, nous distinguons les approches continues v.s. discrètes (combinatoires) ou celles linéaires v.s. non-linéaires, ou celles déterministes vs. stochastiques, ou encore celles d'optimisation de fonctions convexes vs. non-convexes, etc.

Intéressons nous aux méthodes qui ont été proposées pour maximiser le critère MAP, ou minimiser la fonction d'énergie markovienne, dans le cadre des applications en imagerie. Nous retrouvons parmi les approches les plus utilisées dans la littérature :

- Le recuit simulé, initialement exploité en traitement d'image dans (Geman et Geman, 1984), est une approche d'optimisation discrète, stochastique, qui en théorie permet d'optimiser exactement toute fonction objectif. Cependant, en pratique, cette méthode s'avère très lourde en temps de calcul pour converger vers un bon optimum, même si la convergence vers un optimum global est prouvée théoriquement.
- L'algorithme ICM (*Modes Conditionnels Itérés*), introduit en imagerie dans (Besag, 1986), est une approche d'optimisation discrète, déterministe, sous-optimale, c.-à-d. qui converge vers un optimum local de la fonction. La technique consiste à minimiser localement la fonction d'énergie en chaque site. La convergence est rapide et dans le cas d'énergies convexes, de bons optima sont obtenus (pas nécessairement globaux pour de bonnes initialisations). En revanche, dans le cas non-convexe, cette méthode n'est plus robuste. Elle est très sensible à l'ordre de balayage des sites ainsi qu'à l'initialisation et risque de converger vers un optimum de mauvaise qualité.
- L'algorithme GNC (*Graduated Non Convexity*), introduit en vision dans (Blake et Zisserman, 1987), est une approche d'optimisation déterministe qui consiste à définir une suite de problèmes d'estimation partant d'un modèle plus simple à optimiser, pour lequel la solution est unique, et convergeant vers le modèle désiré. Typiquement, dans le cas d'une optimisation non-convexe, la méthode part d'une fonction convexe à optimiser que l'on déforme progressivement. Cette technique est plus efficace en temps de calcul que l'approche stochastique par recuit et plus robuste que l'approche déterministe par ICM. Cependant, il n'est pas toujours possible de l'utiliser.
- La descente de gradient, utilisée dans le cadre des approches variationnelles de régularisation (Rudin *et al.*, 1992), est une approche d'optimisation continue, déterministe. Dans le cas de fonctions convexes, elle converge vers un optimum global. Cependant, dans le cas non-convexe, la méthode n'est plus intéressante en terme de qualité

d'optimum. Des versions stochastiques de cette technique existent en utilisant par exemple les échantillonneurs de Gibbs pour éviter les optima locaux (Younes, 1988).

- La propagation de croyances ou *Belief propagation*, introduite en imagerie dans (Weiss et Freeman, 2001), est une méthode récente d'optimisation discrète et déterministe. Elle est utile pour optimiser différentes formes de modèles de fonctions d'énergies. Elle repose sur une technique de propagation de messages entre les sites de l'image dans le but d'évaluer pour chaque site sa croyance en une étiquette donnée. En revanche, la convergence peut s'avérer lourde en temps de calcul dans la pratique.
- La coupe-minimum sur graphes ou *graph-cut*, introduite en vision dans (Greig *et al.*, 1989) pour le cas binaire et reprise et étendue à la fin des années 90, est aussi une technique assez récente d'optimisation discrète et déterministe. Elle permet d'optimiser exactement certaines formes d'énergies et approximativement d'autres. Elle repose sur la construction d'un graphe spécifique à la fonction d'énergie à optimiser et le calcul de la coupe de poids minimal sur ce graphe. Cette méthode est connue par son efficacité en temps de calcul et sa robustesse pour la majorité des formes d'énergies utilisées en traitement d'image (Szeliski *et al.*, 2006).

Toutes ces techniques d'optimisation, et bien d'autres moins utilisées, sont plus adaptées à certaines fonctions objectifs ou problèmes d'optimisation que d'autres et présentent dans certains cas des limites et dans d'autres cas de bonnes performances. Toutefois, nous nous focalisons dans la suite sur les techniques d'optimisation par coupe-minimum sur graphes pour les raisons suivantes :

- la possibilité d'adapter l'approche à différents modèles markoviens selon une construction de graphe spécifique ;
- le choix entre une optimisation exacte ou non-exacte (approximative) du modèle selon la construction choisie ;
- et les progrès réalisés récemment dans le domaine de l'optimisation combinatoire, apportant des algorithmes efficaces de coupe-minimum sur des graphes spécifiques aux problèmes de vision par ordinateur.

2 Algorithmes de minimisation par coupe-minimum sur graphe

La problématique d'étiquetage en imagerie, introduite au début de ce chapitre, peut être perçue aussi comme étant un problème d'étiquetage dans un graphe. L'étiquetage consiste alors à faire correspondre les sites $\mathcal{P}$ de l'image aux nœuds d'un graphe $\mathcal{G}$ orienté et à estimer ensuite la bonne étiquette pour chaque nœud de ce graphe de manière à avoir la meilleure configuration possible des nœuds au sens du critère établi.

En théorie des graphes, ce problème correspond à un problème de coupe multiterminal minimum sur ce graphe ; où un terminal correspond à une étiquette et la coupe regroupe les nœuds du graphe en sous-ensembles partageant une même étiquette. Étant donné alors un graphe $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ avec $\mathcal{V}$ l'ensemble de ses nœuds et $\mathcal{E}$ l'ensemble de ses arcs à poids positifs, cherchons la coupe optimale, c.-à-d. celle dont la somme des poids des arcs coupés est minimale. Ce problème, que l'on appellera par la suite (*k-coupe-minimum*), est connu en combinatoire (Hu, 1969), où k désigne la cardinalité de l'ensemble des étiquettes considérées. Sa résolution fait l'objet de plusieurs études dans les domaines de l'optimisation discrète et de la théorie des graphes. En général, il s'agit d'un problème NP-difficile. Tout de même, pour certains cas particuliers de graphes, sa résolution en temps polynômial

est possible, tel est le cas des graphes planaires (Yeh, 2001), (Chen et Wu, 2004). Pour d'autres, des auteurs proposent des approches de résolution approximative.

Dans le cas où deux nœuds terminaux sont définis, cela correspond à $k = \text{Card}\{\mathcal{L}\} = 2$, la k -coupe-minimum est connue sous le nom de s,t -coupe-minimum ou tout simplement *coupe-minimum*. Cette coupe répartit les nœuds du graphe en uniquement deux ensembles de nœuds. L'un correspond au terminal s , appelé *source*, et l'autre correspond au terminal t , appelé *puits*. Sous certaines hypothèses sur la construction du graphe, ce cas particulier de coupe multiterminal minimum peut être résolu en un temps polynômial en passant au problème dual, qui est le calcul de flot-maximum sur le graphe $\mathcal{G}$ (Ford et Fulkerson, 1956), (Elias *et al.*, 1956).

Dans la suite, nous présentons brièvement ce problème de coupe-minimum/flot-maximum sur un graphe et son adaptation à la problématique d'étiquetage en traitement d'image.

2.1 Coupe-minimum / flot-maximum

Pour introduire ce problème, on considère le graphe $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ ainsi que les deux nœuds terminaux s et t , respectivement la source et le puits. On définit une coupe $\mathcal{C}_{s,t} = \{\mathcal{S}, \mathcal{T}\}$ dans un graphe par une partition $(\mathcal{S}, \mathcal{T})$ de l'ensemble des sommets telle que :

$$\begin{cases} \mathcal{S} \cup \mathcal{T} &= \mathcal{V}, \\ \mathcal{S} \cap \mathcal{T} &= \emptyset, \\ s &\in \mathcal{S}, \\ t &\in \mathcal{T}. \end{cases} \quad (1.13)$$

Le poids d'une coupe $|\mathcal{C}_{s,t}|$ est défini par le réel positif de valeur :

$$|\mathcal{C}_{s,t}| = \sum_{\substack{u \in \mathcal{S} \\ v \in \mathcal{T} \\ (u,v) \in \mathcal{E}}} w(u,v), \quad (1.14)$$

où $w(u,v)$ est le poids associé à l'arc orienté reliant les deux nœuds u et v , considéré positif.

La coupe-minimum dans le graphe $\mathcal{G}$ est donc la coupe de poids minimal dans ce graphe parmi toutes les coupes possibles.

On définit ensuite un réseau par le couple $(\mathcal{G}, f)$, où $\mathcal{G}$ est le graphe précédemment défini et $f : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$ est une fonction appelée *flot* définie par les contraintes suivantes :

$$\begin{cases} \forall (u,v) \in \mathcal{V}^2, & f(u,v) \leq w(u,v), \\ \forall u \in \mathcal{V} - \{s,t\}, & \sum_{v \in \mathcal{V}} f(u,v) = 0, \\ \forall (u,v) \in \mathcal{V}^2, & f(u,v) = -f(v,u). \end{cases} \quad (1.15)$$

La première contrainte stipule qu'un arc ne peut contenir un flot qui dépasse sa capacité. La seconde est une contrainte de conservation de flot en chaque nœud (loi de Kirchoff). La dernière est dite contrainte de symétrie.

La valeur du flot $|f|$ est définie par :

$$|f| = \sum_{v \in \mathcal{V}} f(s,v) = \sum_{v \in \mathcal{V}} f(v,t), \quad (1.16)$$

et elle indique le flot total que l'on peut faire transiter de la source vers le puits.

Le flot maximum dans un réseau correspond alors au flot de valeur maximale que l'on peut faire passer de la source au puits.

La dualité entre les deux problèmes flot-maximum et coupe-minimum a été prouvée indépendamment dans (Ford et Fulkerson, 1956) et (Elias *et al.*, 1956).

Théorème 2.1. *Dans un graphe $\mathcal{G}$, comme défini précédemment, la valeur d'une coupe-minimum est égale à la valeur d'un flot-maximum dans un réseau associé à ce graphe. De plus, tout arc contenu dans la coupe est un arc saturé (c.-à-d. on ne peut pas y faire passer plus de flot).*

De nombreux travaux dans la littérature ont été consacrés à la recherche efficace du flot-maximum dans un réseau. Un aperçu sur les grandes familles de ces algorithmes est présenté dans l'annexe C.

Dans le cadre des problèmes de traitement d'image, si les sites de $\mathcal{P}$ sont les pixels de l'image, nous disposons alors de graphes $\mathcal{G}$ réguliers et bien structurés. D'où l'intérêt d'adapter ces algorithmes classiques de calcul de flot-maximum/coupe-minimum à ces graphes particuliers en vue d'améliorer leurs performances.

Dans la suite, nous présentons une variété d'algorithmes d'optimisation par coupe-minimum sur graphes en imagerie. Pour différents modèles markoviens associés à certains problèmes d'étiquetage en imagerie, nous décrivons les structures des graphes correspondants ainsi que les algorithmes de minimisation qui reposent sur ces constructions. Nous considérons pour la clarté du texte la répartition des champs markoviens entre champs de Markov binaires et champs de Markov multi-étiquettes. Pour chacune de ces deux classes, nous traitons séparément les cas des cliques de premier ordre et le cas des cliques d'ordre supérieur.

2.2 Champs markoviens binaires

Les champs markoviens binaires ou booléens sont utiles pour certaines applications en traitement d'image telles que la restauration d'images binaires (Greig *et al.*, 1989) et la segmentation d'images en objet/fond (Rother *et al.*, 2004). L'ensemble des étiquettes est dans ce cas de figure une paire d'étiquettes, que l'on notera par $\mathcal{L} = \{0, 1\}$.

Ainsi étiqueter l'ensemble des sites $\mathcal{P}$ revient à minimiser la fonction d'énergie markovienne pseudo-booléenne E définie sur des variables binaires $\{x_p\}_{p \in \{1, 2, \dots, n\}}$ par :

$$E(x) = \sum_{p=1}^n \underbrace{E_p(x_p)}_{=E_p(y_p|x_p)} + \sum_{c \in \mathcal{C}} E_c(x). \quad (1.17)$$

2.2.1 Fonctions de $\mathbb{B}^2$

Traisons en premier lieu les modèles markoviens de premier ordre. Nous noterons la classe de fonctions d'énergies pseudo-boléennes de tels modèles par $\mathbb{B}^2$. Elles s'écrivent sous la forme suivante :

$$E(x) = \sum_{p=1}^n E_p(x_p) + \sum_{(p,q)} E_{p,q}(x_p, x_q). \quad (1.18)$$

La minimisation de cette forme de fonction pseudo-boléenne a été traitée initialement dans (Picard et Ratliff, 1975) ensuite dans (Greig *et al.*, 1989). Elle a été reprise en détails dans (Kolmogorov et Zabih, 2004).

La résolution par coupe-minimum paraît immédiate dans ce cas de figure puisque l'objectif est d'étiqueter les sites de $\mathcal{P}$ par 0 ou 1, de façon à avoir une configuration binaire qui minimise E .

La construction d'un tel graphe est néanmoins restreinte à une classe de fonctions de $\mathbb{B}^2$ dites fonctions *sous-modulaires*. Pour les autres fonctions de $\mathbb{B}^2$, la recherche de la coupe-minimum devient un problème NP-difficile (Kolmogorov et Zabih, 2004).

Les termes de vraisemblance (à une variable) sont toujours sous-modulaires, tandis que les termes d'interaction de paires de variables ne sont sous-modulaires que si la relation suivante est vérifiée :

$$E_{p,q}(0,0) + E_{p,q}(1,1) \leq E_{p,q}(0,1) + E_{p,q}(1,0). \quad (1.19)$$

Le cas de fonctions sous-modulaires de $\mathbb{B}^2$

Si nous disposons de fonctions d'énergies markoviennes sous-modulaires de $\mathbb{B}^2$, la minimisation est désormais possible grâce aux algorithmes de calcul de flot-maximum appliqués sur un graphe particulier. De plus, le minimum obtenu est un minimum global. Pour construire le graphe, considérons la reformulation suivante de la fonction d'énergie :

$$\begin{aligned} E(x) = & \sum_{p=1}^n \left(E_p(1)x_p + E_p(0)(1-x_p) \right) \\ & + \sum_{(p,q)} \left(E_{p,q}(0,1) + E_{p,q}(1,0) - E_{p,q}(0,0) - E_{p,q}(1,1) \right) (1-x_p)x_q \\ & + \sum_{(p,q)} \left(E_{p,q}(1,0) - E_{p,q}(0,0) \right) x_p + \left(E_{p,q}(1,1) - E_{p,q}(1,0) \right) x_q, \end{aligned} \quad (1.20)$$

avec $E_p(1) = E_p(y_p|x_p = 1)$ et $E_p(0) = E_p(y_p|x_p = 0)$. On peut ainsi construire le graphe directement à partir de ces termes de l'énergie de la façon suivante :

- chaque nœud v_p du graphe, correspondant à un site p , est relié à la source par un arc orienté $\overrightarrow{(s, v_p)}$ de poids $E_p(0)$ et au puits par un arc orienté $\overrightarrow{(v_p, t)}$ de poids $E_p(1)$. Ces arcs sont appelés *arcs de données* (cf. FIG. 1.1). Une construction équivalente présentée dans la même figure est proposée dans (Kolmogorov et Zabih, 2004)

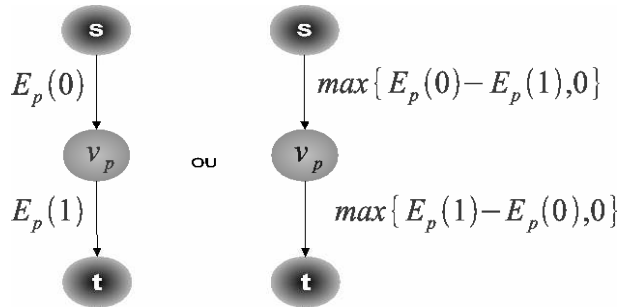


FIGURE 1.1 – Structure répétitive du graphe sur tous les sites de $\mathcal{P}$. Exemple des deux constructions possibles.

- pour chaque couple de nœuds voisins (v_p, v_q) , correspondant à un couple de sites voisins (p, q) , nous créons un arc orienté $\overrightarrow{(v_p, v_q)}$ de poids $\left(E_{p,q}(0,1) + E_{p,q}(1,0) - E_{p,q}(0,0) - E_{p,q}(1,1) \right)$. Cet arc est appelé *arc d'interaction*. Ensuite, un arc ori-

enté $\overrightarrow{(s, v_p)}$ de poids $(E_{p,q}(1, 0) - E_{p,q}(0, 0))$ si ce poids est positif, sinon $\overrightarrow{(v_p, t)}$ de poids $-(E_{p,q}(1, 0) - E_{p,q}(0, 0))$, est construit. On ajoute ensuite un arc orienté $\overrightarrow{(v_q, t)}$ de poids $(E_{p,q}(1, 1) - E_{p,q}(1, 0))$ si ce poids est positif, sinon $\overrightarrow{(s, v_q)}$ de poids $-(E_{p,q}(1, 1) - E_{p,q}(1, 0))$. Ces deux arcs sont appelés *arcs de bord* et sont souvent fusionnés avec les arcs de données (cf. FIG. 1.2) .

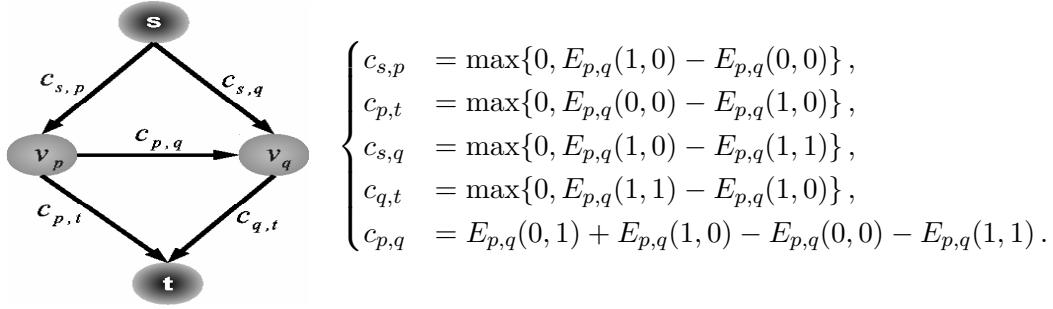


FIGURE 1.2 – Structure répétitive du graphe pour chaque couple de sites voisins de $\mathcal{P}$ telle que $p < q$.

On remarque la condition de sous-modularité qui intervient dans cette construction pour assurer un arc à poids positif $c_{p,q}$ entre les deux nœuds voisins (v_p, v_q) .

La figure 1.3 illustre le graphe total correspondant à une image de taille 3×3 avec un système de voisinage 4-connexité.

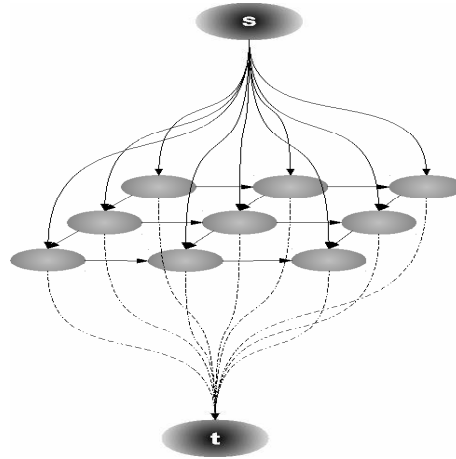


FIGURE 1.3 – Construction du graphe total correspondant à une image de taille 3×3 avec un système de voisinage en 4-connexité.

La taille du graphe est donc de l'ordre de $O(n)$ et la complexité de l'algorithme de calcul du flot-maximum est polynômiale.

La résolution d'un problème de vision par une telle approche, en supposant que les potentiels des cliques de premier ordre sont sous-modulaires, est donc très efficace en temps de calcul. De plus, un minimum global est obtenu. En revanche, si les potentiels ne sont pas tous sous-modulaires, la construction proposée n'est plus appropriée. Dans ce cas, des approches de minimisation non-exacte ou approximative ont été proposées.

Le cas de fonctions non sous-modulaires de $\mathbb{B}^2$

Une première méthode d'optimisation approximative a été présentée dans (Kwatra *et al.*, 2003), (Agarwala *et al.*, 2004) et (Rother *et al.*, 2005). Elle consiste à ignorer les termes non sous-modulaires ou les approximer par d'autres sous-modulaires pour pouvoir faire appel à la construction de graphe décrite précédemment et minimiser approximativement l'énergie globale par coupe-minimum sur ce graphe. Néanmoins, cette méthode n'est valable qu'en présence d'un nombre restreint de termes d'interactions non sous-modulaires. Le cas échéant, l'approximation de la fonction d'énergie risque de détruire le vrai modèle.

Une deuxième méthode qui a montré son efficacité et qui reste à ce jour la plus utilisée est celle proposée dans (Kolmogorov et Rother, 2007). Cette méthode repose sur la technique de minimisation de fonctions quadratiques pseudo-bouliennes par flot-maximum parue dans (Hammer *et al.*, 1984), appelée *roof duality*. Cette technique consiste à proposer un étiquetage partiel des nœuds du graphe. Dans le cas où les termes de l'énergie sont sous-modulaires, un étiquetage booléen est proposé aux nœuds qui leurs correspondent, et les autres nœuds restent sans étiquetage (ou portent l'étiquette $\emptyset$). Un algorithme de construction de graphe approprié et d'étiquetage par calcul de flot-maximum, appelé *QPBO*, est initialement proposé dans (Boros *et al.*, 1991). Ensuite, les auteurs de (Kolmogorov et Rother, 2007) l'ont adapté aux problèmes de vision par une construction de graphe plus efficace. Cependant, cet algorithme reste encore non fiable quand plusieurs termes de l'énergie ne vérifient pas la condition de sous-modularité. Une extension a été alors proposée dans (Rother *et al.*, 2007) pour étiqueter le maximum possible de nœuds correspondant aux termes non sous-modulaires. L'idée repose sur la technique proposée dans (Boros *et al.*, 2006) et qui consiste à partir de l'étiquetage partiel donné par QPBO et à étiqueter itérativement les nœuds qui restent, un par un, de façon à minimiser l'énergie globale. L'algorithme reste toujours rapide puisqu'il s'agit de minimisations binaires et converge vers un optimum local de meilleure qualité.

Ainsi minimiser efficacement une fonction d'énergie markovienne booléenne de premier ordre quelconque est désormais possible par coupe-minimum sur graphe, dans la mesure où un nombre faible de termes de l'énergie sont non sous-modulaires.

2.2.2 Fonctions de $\mathbb{B}^q$

Considérons maintenant la famille des modèles markoviens d'ordre supérieur. Nous noterons la classe de fonctions d'énergie pseudo-boulienne de tels modèles par $\mathbb{B}^q$, avec q l'ordre de la clique de cardinalité maximale du modèle. Elles s'écrivent sous la forme :

$$E(x) = \sum_{p=1}^n E_p(x_p) + \sum_{c \in \mathcal{C}} E_c(x), \quad (1.21)$$

avec $\mathcal{C} = \bigcup_{i \leq q} C_i$.

Cette classe de fonctions d'énergie s'avère utile pour mieux modéliser des images de scènes naturelles texturées. Cependant, leur optimisation est complexe. Des travaux récents en optimisation par coupe-minimum ont adressé cette problématique. L'idée fondamentale consiste à réduire les cliques d'ordre supérieur en cliques d'ordre 0 et 1 et à optimiser le nouveau modèle réduit par les optimiseurs décrits précédemment.

Une première approche de réduction a été proposée initialement dans (Kolmogorov et Zabih, 2004) pour des cliques d'ordre 2. Ensuite, elle a été généralisée dans (Freedman et Drineas, 2005) pour tout ordre. La réduction consiste à transformer les termes à 3 variables booléennes ou plus en termes à une ou deux variables.

Exemple : Les termes d'ordre 2 de l'énergie de la forme $f(x_1, x_2, x_3) = ax_1x_2x_3$, avec a un réel et x_1, x_2 et x_3 des variables binaires peuvent s'écrire sous la forme :

$$\begin{aligned} f(x_1, x_2, x_3) &= \tilde{f}(x_1, x_2, x_3, y) \\ &= \begin{cases} ay(x_1 + x_2 + x_3 - 2) & \text{si } a < 0, \\ a(y(x_1 + x_2 + x_3 - 1) + (x_1x_2 + x_2x_3 + x_3x_1) - (x_1 + x_2 + x_3) + 1) & \text{si } a > 0, \end{cases} \end{aligned} \quad (1.22)$$

avec y une variable binaire auxiliaire à ajouter à l'ensemble des variables x_p sur lesquelles la fonction E est définie.

En revanche, une telle transformation de la fonction d'énergie E n'est généralisable que si les coefficients a des termes d'ordre supérieur sont négatifs, autrement une telle réduction n'est possible que pour certains ordres. Pour les autres, la transformation est encore plus compliquée pour pouvoir minimiser efficacement l'énergie totale. Dans un travail récent (Ishikawa, 2009a), l'auteur propose une extension à cette approche pour tout ordre, quelque soient les signes des coefficients des termes d'ordre supérieur. En revanche, la méthode introduit un nombre important de variables auxiliaires. Son utilisation reste en pratique restreinte aux modèles markoviens d'ordre supérieur assez faible.

Une deuxième approche par substitution a été proposée initialement dans (Rosenberg, 1975), ensuite reprise dans (Boros et Hammer, 2002) et appliquée en vision dans (Ali *et al.*, 2008). Elle procède par transformation de toute fonction pseudo-booléenne en fonction pseudo-booléenne quadratique. La réduction consiste à substituer tout produit de deux variables binaires x_1x_2 par une nouvelle variable binaire y ayant la même valeur que le produit en minimisant la fonction globale.

Exemple : Les termes d'ordre 2 de l'énergie de la forme $f(x_1, x_2, x_3) = x_1x_2x_3$, avec x_1, x_2 et x_3 des variables binaires peuvent s'écrire sous la forme :

$$f(x_1, x_2, x_3) = \tilde{f}(x_1, x_2, x_3, y) = yx_3 + bD(x_1, x_2, y), \quad (1.23)$$

avec

$$D(x_1, x_2, y) = x_1x_2 - 2x_1y - 2x_2y + 3y, \quad (1.24)$$

et b un réel positif. On remarque que si $y = x_1x_2$ alors D est minimum ($= 0$), sinon $D > 0$. Ainsi minimiser la nouvelle forme réduite de l'énergie permet de minimiser aussi la forme d'énergie d'origine, à condition que b soit fixé très grand.

Le problème de cette réduction est que le plus souvent le modèle obtenu présente un nombre important de termes non sous-modulaires ; ce qui laisse un grand nombre de sites non-étiquetés après recours aux approches d'optimisation non sous-modulaire (QPBO).

En conclusion, optimiser efficacement les énergies markoviennes pseudo-booléennes d'ordre supérieur reste à ce jour un problème ouvert et d'une grande importance pour résoudre certaines applications de haut niveau en vision. Ces applications font intervenir des primitives à la place des pixels de l'image, tel est le cas pour la reconstruction des réseaux routiers en imagerie satellitaire à partir des primitives routes (Stoica, 2001) et la reconnaissance automatique des sillons corticaux en imagerie médicale à partir des primitives linéiques (Rivière, 2000). Dans ce cadre, la modélisation de *l'a priori* nécessite le recours à des cliques d'ordre supérieur dont les potentiels ne sont pas forcément sous-modulaires.

2.3 Champs markoviens multi-étiquettes

Les modèles markoviens multi-étiquettes modélisent un nombre plus important d'applications en imagerie, comparés aux modèle binaires. L'ensemble des étiquettes est alors de

cardinalité supérieure à 2. Optimiser par coupe-minimum les modèles d'énergie associés devient une tâche plus complexe et dans le cas général, il s'agit d'un problème d'optimisation combinatoire NP-difficile.

2.3.1 Fonctions de $\mathbb{F}^2$

Similairement à l'étude des fonctions booléennes, nous commençons par traiter le cas des modèles markoviens de premier ordre. La classe des fonctions d'énergie markovienne associées à de tels modèles sera notée par $\mathbb{F}^2$. La forme de l'énergie est la même que celle indiquée dans (1.18), mais en considérant cette fois-ci $\mathcal{L} = \{l_1, l_2, \dots, l_k\}$.

L'optimisation exacte par coupe-minimum d'une fonction de la classe $\mathbb{F}^2$ peut être effectuée par deux approches. La première a été initialement proposée dans (Roy et Cox, 1998; Roy, 1999) ensuite étudiée en détails dans (Ishikawa et Geiger, 1998a,b; Ishikawa, 2000, 2003). La deuxième a été proposée dans (Darbon, 2005; Darbon et Sigelle, 2006a,b; Darbon, 2008) exploitant la technique de décomposition d'une variable sur ses ensembles de niveaux (Guichard et Morel, 2002). Une approche similaire a été aussi présentée indépendamment dans (Zalesky, 2003a).

Ces approches proposent différentes constructions de graphe pour minimiser exactement la famille de fonctions sous-modulaires de $\mathbb{F}^2$. Dans cette classe de fonctions, si l'*a priori* $E_{p,q}$ peut s'écrire sous la forme de $E_{p,q}(x_p, x_q) = g(x_p - x_q)$, $\forall (p, q)$, g est alors une fonction convexe (Ishikawa, 2003; Darbon, 2008). La taille des graphes construits est de l'ordre de $O(nk)$ et le calcul de flot-maximum est de complexité pseudo-polynomiale. Dans l'annexe B, le cas de la construction du graphe multi-étiquette d'Ishikawa est présenté.

Pour les autres fonctions de $\mathbb{F}^2$, c.-à-d. ayant un *a priori* non-convexe, différentes stratégies d'optimisation ont été proposées dans la littérature. Les approches développées reposent sur la technique de *mouvement de partitions*. Cette technique est initialement introduite dans (Otten et Van Ginneken, 1989) pour décrire l'algorithme de minimisation par recuit simulé. Ensuite, elle a été reprise dans (Veksler, 1999) en proposant une nouvelle classe de mouvements de partitions dits *larges*. L'idée consiste en fait à minimiser itérativement la fonction d'énergie E , de façon à proposer à chaque itération i une configuration $x^{(i)}$ des sites de $\mathcal{P}$ plus optimale (c.-à-d. ayant une énergie plus faible) par rapport à la configuration courante $x^{(i-1)}$. L'algorithme converge quand aucune configuration possible par le mouvement considéré n'est optimale par rapport à la configuration courante.

Un mouvement de partition est dit large si plusieurs sites sont concernés par le mouvement, en une itération donnée de l'algorithme. Ces sites sont appelés des sites *actifs*.

Les mouvements larges souvent exploités dans la littérature sont des mouvements *binaires* (Boykov *et al.*, 2001), que l'on notera par MPLB (Mouvements de Partitions Larges et Binaires). A chaque itération, le mouvement de partition consiste à ne proposer à chacun des sites actifs que deux étiquettes, pour pouvoir exploiter les techniques efficaces d'optimisation par s,t-coupe-minimum présentées dans le paragraphe 2.2.

Trois classes de mouvements binaires sont le plus souvent utilisées dans la littérature :

Échange : Un échange d'étiquettes est un mouvement de partitions large permettant de passer d'une configuration à une autre en proposant aux sites actifs d'échanger leurs étiquettes. Le mouvement binaire correspondant est appelé $\alpha\beta$ -échange et consiste à faire échanger les étiquettes courantes des sites actifs (c.-à-d. ceux dont les étiquettes sont dans $\{\alpha, \beta\}$, avec $\alpha, \beta \in \mathcal{L}$) de façon à minimiser l'énergie de la configuration courante. Un mouvement optimal peut être obtenu grâce à la technique de coupe-minimum en considérant l'étiquetage par α ou β comme étant un étiquetage binaire. Ainsi un graphe similaire

au graphe binaire présenté dans le paragraphe 2.2 est construit sur l'ensemble des sites actifs. Les poids des arcs orientés sont définis de façon à faire correspondre une coupe-minimum sur le graphe à une configuration d'énergie E minimale (voir l'annexe A pour les détails de la construction). Comme la construction de ce graphe nécessite la condition de sous-modularité, une classe de fonctions *a priori* vérifiant cette contrainte est la classe des semi-métriques, c.-à-d.

$$\begin{cases} E_{p,q}(x_p, x_q) > 0 & , \forall (x_p, x_q) \in \mathcal{L}^2, \text{ and } x_p \neq x_q, \\ E_{p,q}(x_p, x_p) = 0 & , \forall x_p \in \mathcal{L}, \\ E_{p,q}(x_p, x_q) = E_{q,p}(x_q, x_p) & , \forall (x_p, x_q) \in \mathcal{L}^2. \end{cases} \quad (1.25)$$

L'algorithme de minimisation de E par $\alpha\beta$ -échange est un algorithme itératif. En partant d'une configuration initiale $x^{(0)}$, un couple d'étiquettes $(\alpha, \beta) \in \mathcal{L}^2$ est choisi à chaque itération. Étant donné ce couple, le graphe est construit et le calcul de flot-maximum donne le mouvement optimal des étiquettes. L'algorithme itère sur tous les couples d'étiquettes possibles pour finir un cycle. Les cycles reprennent jusqu'à ce que la configuration finale des sites, obtenue à la fin d'un cycle, ne change plus. L'algorithme converge alors vers un optimum local de E .

Extension : Une extension est un mouvement de partitions large permettant de passer d'une configuration à une autre en proposant à tous les sites de nouvelles étiquettes. Le mouvement binaire correspondant est appelé α -extension et consiste à proposer à chacun des sites de $\mathcal{P}$ soit de garder l'étiquette courante, soit de la remplacer par l'étiquette $\alpha \in \mathcal{L}$, de façon à minimiser l'énergie totale de la configuration. Un mouvement optimal peut être obtenu grâce à la technique de coupe-minimum en considérant l'étiquetage par α ou $\bar{\alpha}$ (c.-à-d. étiquette courante) comme étant un étiquetage binaire. Ainsi un graphe similaire au graphe binaire est construit sur tous les sites de $\mathcal{P}$. Les poids des arcs orientés sont définis de façon à faire correspondre une coupe-minimum sur le graphe à une configuration d'énergie E minimale (voir l'annexe A pour les détails de la construction). De la même façon que pour le mouvement d'échange, la construction du graphe binaire nécessite la condition de sous-modularité. Une classe de fonctions *a priori* vérifiant cette contrainte est la classe des métriques, c.-à-d.

$$\begin{cases} E_{p,q}(x_p, x_q) \geq 0 & , \forall (x_p, x_q) \in \mathcal{L}^2, \\ E_{p,q}(x_p, x_p) = 0 & , \forall x_p \in \mathcal{L}, \\ E_{p,q}(x_p, x_q) = E_{q,p}(x_q, x_p) & , \forall (x_p, x_q) \in \mathcal{L}^2, \\ E_{p,q}(\gamma, \beta) \leq E_{p,q}(\gamma, \alpha) + E_{p,q}(\alpha, \beta) & , \forall (\gamma, \beta, \alpha) \in \mathcal{L}^3. \end{cases} \quad (1.26)$$

L'algorithme de minimisation de E par α -extension est un algorithme itératif. En partant d'une configuration initiale $x^{(0)}$, une étiquette $\alpha \in \mathcal{L}$ est choisie à chaque itération et le graphe est construit pour le calcul de flot-maximum donnant le mouvement optimal des partitions. L'algorithme itère sur toutes les étiquettes possibles pour finir un cycle. Les cycles reprennent jusqu'à ce que la configuration finale des sites, obtenue à la fin d'un cycle, ne change plus. L'algorithme converge alors vers un optimum local de E .

Saut : Un saut est un mouvement de partitions large permettant de passer d'une configuration à une autre en proposant aux sites actifs d'ajouter une quantité à leurs étiquettes courantes. Le mouvement binaire correspondant est appelé δ -saut et consiste à proposer à chacun des sites de $\mathcal{P}$ soit de garder l'étiquette courante soit de l'augmenter ou de la

diminuer d'une quantité δ , de façon à minimiser l'énergie totale de la configuration. La nouvelle étiquette doit être toujours dans $\mathcal{L}$. Puisqu'il s'agit d'un mouvement binaire, le mouvement optimal est obtenu en deux étapes : la première consiste à proposer un saut positif (augmenter l'étiquette de δ) et la deuxième est un saut négatif (diminuer la nouvelle étiquette de δ). A chaque étape, un mouvement optimal est obtenu grâce à la technique de coupe-minimum en considérant l'étiquetage binaire par $\pm\delta$ ou $+0$ (c.-à-d. étiquette courante, puisque la quantité ajoutée est nulle). Ainsi un graphe similaire au graphe binaire est construit sur les sites actifs de $\mathcal{P}$. Les poids des arcs orientés sont définis de façon à faire correspondre une coupe-minimum sur le graphe à une configuration d'énergie E minimale (voir l'annexe A pour les détails de la construction). De même, comme la construction de ce graphe nécessite la condition de sous-modularité, une classe de fonctions *a priori* vérifiant cette contrainte est la classe des fonctions convexes, c.-à-d.

$$E_{p,q}(x_p, x_q) + E_{p,q}(x_p + \delta, x_q + \delta) \leq E_{p,q}(x_p + \delta, x_q) + E_{p,q}(x_p, x_q + \delta), \forall (x_p, x_q, x_p + \delta, x_q + \delta) \in \mathcal{L}^4. \quad (1.27)$$

L'algorithme de minimisation de E par δ -saut est un algorithme itératif. En partant d'une configuration initiale $x^{(0)}$, une quantité δ est choisie à chaque itération et le graphe est construit pour le calcul de flot-maximum donnant le mouvement optimal des partitions. L'algorithme itère sur toutes les quantités possibles (en supposant un pas de quantification fixe défini) pour finir un cycle. Les cycles reprennent jusqu'à ce que la configuration finale des sites, obtenue à la fin d'un cycle, ne change plus. L'algorithme converge alors vers un optimum local de E .

Ces différentes stratégies de mouvements de partitions permettent de minimiser différentes classes d'énergies de $\mathbb{F}^2$ en ayant recours à une même structure d'algorithme de minimisation. En revanche, ces algorithmes ne garantissent pas la convergence vers un optimum global. En pratique, avec des modèles d'énergies E non fortement convexes, l'optimisation est efficace et converge en un nombre fini d'itérations. Souvent, un optimum local de bonne qualité est obtenu (voir même global).

A noter que des modifications et améliorations apportées à ces mouvements ont été adressées dans des travaux récents, dans le but soit d'accélérer encore la minimisation soit d'adapter la stratégie de mouvement de partitions à la nature des données et de l'ensemble des étiquettes. Nous citons parmi ces travaux :

- le mouvement de saut dichotomique, initialement introduit dans (Darbon, 2005) pour minimiser des modèles markoviens convexes d'une manière très efficace. Au cours d'un cycle, des sauts binaires sont effectués en partant d'un pas de valeur maximale et en le divisant par 2 itérativement, jusqu'à atteindre un pas de quantification fixé. Les cycles s'enchaînent jusqu'à converger vers une configuration optimale. Ce mouvement a aussi montré de bonnes performances pour minimiser certains modèles markoviens non-convexes. C'est le cas pour le modèle markovien utilisé pour la restauration d'images radar en amplitude dans (Denis *et al.*, 2009b), où la fonction de vraisemblance est une fonction quasi-convexe et l'*a priori* est choisie comme étant une fonction convexe.
- le mouvement de saut multi-précision (Valadão et Bioucas-Dias, 2009) qui consiste à effectuer, au cours d'un cycle, des sauts binaires de pas $\pm\delta$, où δ est considéré

fixe dans un cycle, jusqu'à ce que la configuration ne change plus. Un nouveau cycle enchaîne avec un pas de saut égal à la moitié du précédent, ainsi de suite. L'algorithme s'arrête à un pas fixé initialement qui correspond à la précision désirée sur l'étiquetage. Cette technique est utile pour estimer des données à étiquettes dans un ensemble continu, dont la quantification entraîne une perte d'information.

Les mouvements par échange, extension et saut nécessitent des potentiels d'interaction respectivement semi-métrique, métrique et convexe. Pour les autres classes de potentiels, l'optimisation est aussi possible grâce à la technique d'optimisation partielle décrite dans le paragraphe 2.2.1, où à chaque itération, le mouvement binaire optimal est obtenu en utilisant la technique QPBO.

Une généralisation de ces différents algorithmes d'optimisation par MPLB est présentée dans (Lempitsky *et al.*, 2009). Un mouvement qui les généralise, appelé mouvement par *fusion de partitions*, est introduit. À chaque itération, la configuration courante est fusionnée avec l'étiquetage proposé selon une stratégie de fusion (extension, échange, saut, ou autre). Pour traiter le cas des fonctions *a priori* non sous-modulaires relativement au mouvement proposé, le mouvement binaire optimal à chaque itération est obtenu par la technique de QPBO.

2.3.2 Fonctions de $\mathbb{F}^q$

Considérons maintenant la famille des modèles markoviens multi-étiquettes d'ordre supérieur. Nous noterons la classe de ces fonctions d'énergies par $\mathbb{F}^q$, avec q l'ordre de la clique de cardinalité maximale du modèle. Elles s'écrivent sous la même forme que (1.21) mais en considérant cette fois-ci $\mathcal{L} = \{l_1, l_2, \dots, l_k\}$.

Optimiser cette famille de fonctions d'énergie est une tâche beaucoup plus complexe. Quelques travaux lui ont été consacrés et qui reposent d'une part sur la technique de projection des cliques d'ordre supérieur en cliques d'ordre 0 ou 1 et de l'autre part sur la technique d'optimisation par des mouvements de partition binaires de l'énergie markovienne de premier ordre obtenue. Nous citons parmi ces travaux :

- dans (Kohli *et al.*, 2007, 2009), les auteurs étendent les algorithmes d'optimisation par α -extension et $\alpha\beta$ -échange pour minimiser efficacement une classe de fonctions de $\mathbb{F}^q$. Cette classe de fonctions est définie à partir de potentiels de cliques généralisant le modèle de Potts à des cliques d'ordre supérieur :

$$E_c(x) = \begin{cases} \gamma_k & \text{si } x_p = l_k, \forall p \in c \\ \gamma_{max} & \text{sinon,} \end{cases} \quad (1.28)$$

avec $\gamma_{max} \geq \gamma_k, \forall l_k \in \mathcal{L}$.

- dans (Ishikawa, 2009b), l'algorithme de minimisation procède par des minimisations successives avec des mouvements binaires. À chaque itération, le mouvement optimal est obtenu en réduisant les cliques d'ordre supérieur en cliques d'ordre 0 ou 1 et en utilisant la technique par fusion de partitions. L'algorithme QPBO est utilisé à chaque itération puisque l'étiquetage proposé ne garantit pas la sous-modularité de la fonction *a priori*. En effet, le nouveau mouvement binaire par fusion introduit consiste à proposer aux sites soit de garder leur étiquetage courant soit de prendre l'étiquetage obtenu par une descente de gradient à partir de la configuration courante.

2.4 Bilan

Minimiser une fonction d'énergie markovienne par coupe-minimum est désormais possible grâce aux différentes approches et algorithmes proposés dans la littérature. Toutefois, l'optimisation exacte n'est pas toujours possible, des optimiseurs non-exacts sont dès lors utilisés. En pratique, ces algorithmes convergent vers un optimum local de faible énergie (proche du global).

Dans le tableau 1.1, un récapitulatif de ces méthodes est proposé présentant différents modèles d'énergie avec des optimiseurs possibles. A chaque fois, le ou les algorithmes par coupe-minimum sur graphe appropriés et le(s) plus utilisé(s) en vision sont indiqués tout en précisant si l'optimisation est exacte ou non-exacte et si possible la complexité algorithmique. Focalisons nous sur les modèles markoviens de premier ordre et les algorithmes de minimisation par coupe-minimum qui leur sont dédiés. Les approches que nous venons de présenter le long de ce chapitre ont montré leur utilité, robustesse et efficacité dans différents travaux en vision et traitement d'image. Cependant, elles ont un certain nombre de limites, notamment pour les fonctions d'énergies non-convexes.

2.4.1 Limites de ces algorithmes

Un cas courant de fonctions d'énergie non-convexes en pratique est celui où l'énergie d'attache aux données est non-convexe et l'énergie de régularisation est convexe. Ce cas de figure caractérise différentes applications en imagerie. Nous citons à titre d'exemple, le débruitage des images corrompues par un bruit de type multiplicatif (tel est le cas du chatoiement en imagerie radar ou d'échographie, le bruit impulsif, etc.) et le déroulement de phase interférométrique où le modèle de vraisemblance est multimodal, c.-à-d. plusieurs optima locaux caractérisent la fonction d'énergie. Pour ces problèmes, d'un côté les algorithmes de minimisation exacte permettent d'atteindre le minimum global de l'énergie au prix d'une importante consommation de mémoire, nécessaire pour la construction des graphes. De l'autre côté les algorithmes de minimisation par MPLB, compte tenu de leur aspect itératif, convergent vers des solutions non satisfaisantes. La qualité du minimum obtenu dépend à la fois de la fonction de l'énergie et du déroulement de l'algorithme (initialisation de la solution à estimer, parcours des étiquettes, etc.). Il est donc souvent nécessaire de recourir aux optimiseurs exacts pour garantir une bonne solution au problème traité. A l'heure actuelle, la seule alternative au problème de complexité en mémoire des approches exactes est la décomposition de l'ensemble de variables de la fonction (l'ensemble des sites $\mathcal{P}$) en sous-ensembles et la résolution du problème par minimisation exacte sur chacun d'eux. A la fin, la solution au problème est obtenue en regroupant les solutions des sous-problèmes résolus. Cette approche est aussi en pratique peu appréciée compte tenu de la perte de l'information contextuelle globale qu'elle engendre sur les données.

Pour illustrer plus clairement ces différentes problématiques, nous présentons dans la suite deux expériences montrant les difficultés à minimiser certains modèles markoviens par les algorithmes proposés dans la littérature.

2.4.2 Quelques expériences

La première expérience (cf. FIG. 1.4) illustre la difficulté que rencontrent les optimiseurs par MPLB à converger vers un optimum de bonne qualité. Le recours dans ce cas aux optimiseurs exacts est nécessaire pour obtenir une solution satisfaisante, mais au prix d'une construction de graphe coûteuse en mémoire. L'application en question est la reconstruction de la phase interférométrique en radar à synthèse d'ouverture. Les mesures de phase

observées sont bruitées et enroulées dans $[-\pi, \pi[$. Ce problème sera présenté clairement plus tard dans le manuscrit. Nous nous contentons ici de présenter quelques résultats illustrant la difficulté à converger vers un bon optimum par un algorithme à MPLB. La fonction de vraisemblance associée à ce problème est une fonction non-convexe et multimodale. Cette particularité engendre la présence de plusieurs optima locaux. Ainsi un algorithme de minimisation itératif risque de stagner dans l'un de ces optima. Ce que confirme le résultat dans la figure 1.4(d).

La deuxième expérience (cf. FIG. 1.5) illustre la difficulté que rencontre les optimiseurs exacts en présence de données de grandes dimensions pour minimiser la fonction d'énergie et le recours à une résolution du problème en décomposant les données d'origine en d'autres de tailles plus petites. Ensuite, la solution au problème de départ est reconstruite à partir des solutions aux sous-problèmes. Elle montre bien l'effet de blocs qui apparaît, dû à la perte du contexte global décrivant les données d'origine. L'application considérée est le débruitage des images radar à synthèse d'ouverture. Le modèle de vraisemblance considéré est non-convexe, tiré des statistiques du chatoiement pleinement développé caractérisant ces données radar. Le recours à l'optimisation exacte est aussi nécessaire pour converger vers un optimum global de l'énergie.

2.5 Conclusion

Compte tenu des limites en terme de qualité de l'optimum et de complexité algorithmique en minimisant certaines formes d'énergies markoviennes (vraisemblance quelconque et *a priori* convexe) par les approches de coupe-minimum sur graphes, nous proposons dans ce travail de thèse de relever ce défi.

Dans le chapitre suivant, nous introduisons de nouveaux algorithmes de minimisation non-exacte permettant à la fois de converger vers de meilleurs minima locaux de ces formes d'énergie, comparés à ceux atteints par les MPLB, tout en consommant de faibles espaces mémoires, nettement inférieurs à ceux nécessaires pour les optimiseurs exacts. Cette contribution est d'un intérêt double. D'une part, il est désormais possible d'obtenir une bonne solution (voir même l'optimale) à certains problèmes inverses en imagerie, modélisés par une fonction markovienne avec une régularisation convexe. D'autre part, il est aussi possible de contrôler conjointement la complexité (en mémoire et en temps de calcul) de l'algorithme de minimisation ainsi que la qualité de la solution que l'on cherche. En pratique, ce contrôle est d'autant plus nécessaire en présence d'applications traitant des données de grandes dimensions, comme c'est le cas en télédétection.

Remarque : Dans la suite du manuscrit, nous désignons par une énergie markovienne non-convexe tout modèle d'énergie ayant un attache aux données quelconque et un *a priori* convexe. Le cas général de fonctions markoviennes non-convexes où l'*a priori* est aussi non-convexe dépasse notre étude.

	Énergies markoviennes			
	Pseudo-booléennes : $\mathbb{B}$		Multi-étiquettes : $\mathbb{F}$	
	1 ^{er} ordre : $\mathbb{B}^2$	q^{eme} ordre : $\mathbb{B}^q$	1 ^{er} ordre : $\mathbb{F}^2$	q^{eme} ordre : $\mathbb{F}^q$
Sous-modulaires	Graphe binaire (exacte, polynômial) (Kolmogorov et Zabih, 2004)	Réduction + Graphe binaire (non-exacte) (Freedman et Drineas, 2005)	MPLB (non-exacte, pseudo-polynômial) (Boykov <i>et al.</i> , 2001) Ishikawa, Darbon (exacte, pseudo-polynômial) (Ishikawa, 2003)(Darbon, 2008)	Projection + MPLB (non-exacte, pseudo-polynômial) (Kohli <i>et al.</i> , 2009)
Non sous-modulaires	QPBO (non-exacte, polynômial) (Kolmogorov et Rother, 2007)	Réduction + QPBO (non-exacte) (Ishikawa, 2009a)	Fusion (non-exacte) (Lempitsky <i>et al.</i> , 2009)	Réduction + Fusion (non-exacte) (Ishikawa, 2009b)

TABLE 1.1 – Algorithmes d’optimisation de différentes énergies markoviennes en vision.

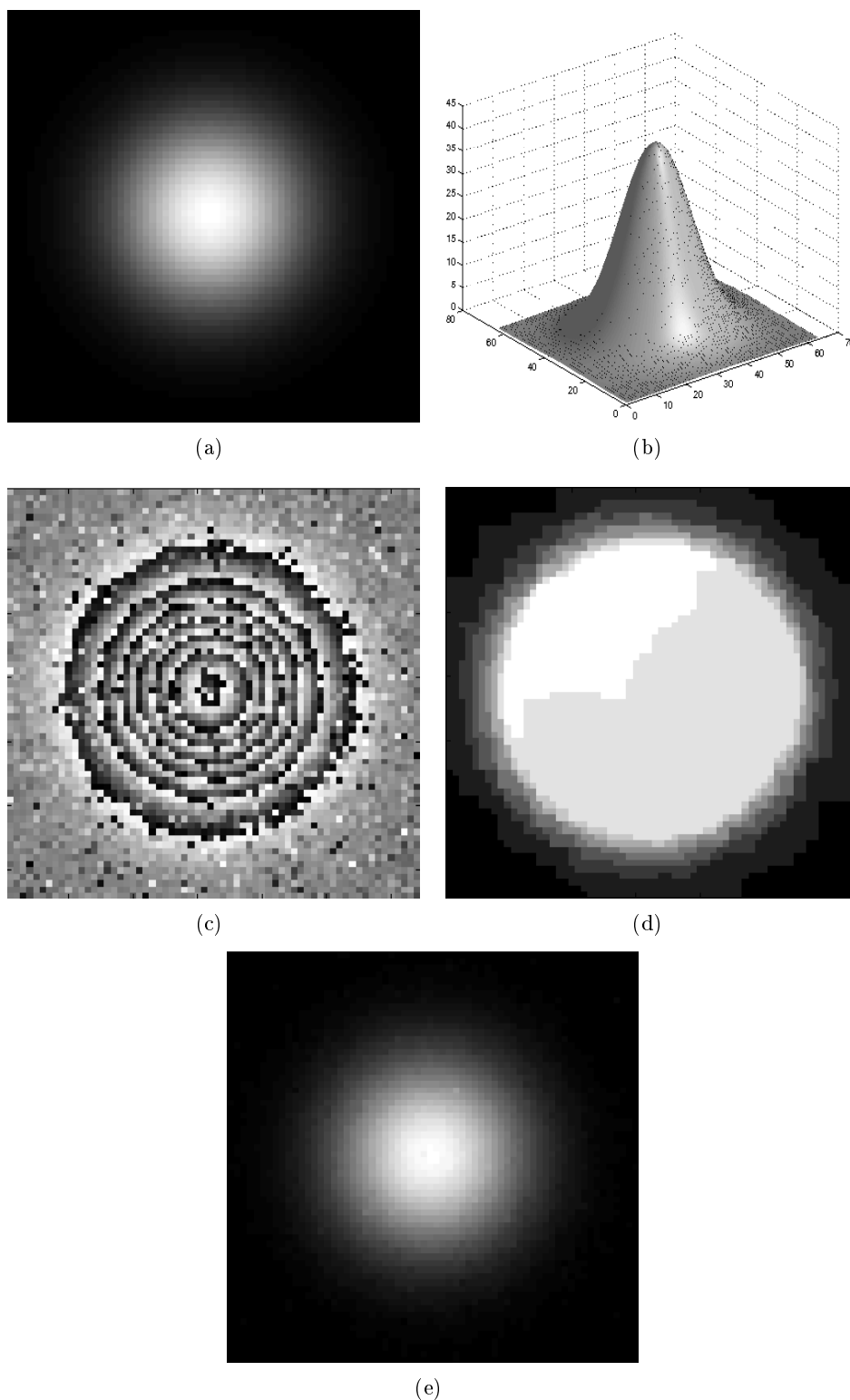


FIGURE 1.4 – Déroulement et débruitage de la phase interférométrique en imagerie radar. Le modèle markovien est défini par une attache aux données multimodal et une régularisation convexe par variation totale. (a) Image de synthèse d'un profil gaussien ($n = 64 \times 64$, $k = 113$), (b) sa visualisation en 3D, (c) image de phases bruitée et enroulée, résultats de la minimisation de l'énergie par (d) α -extension, et (e) par un optimiseur exact.

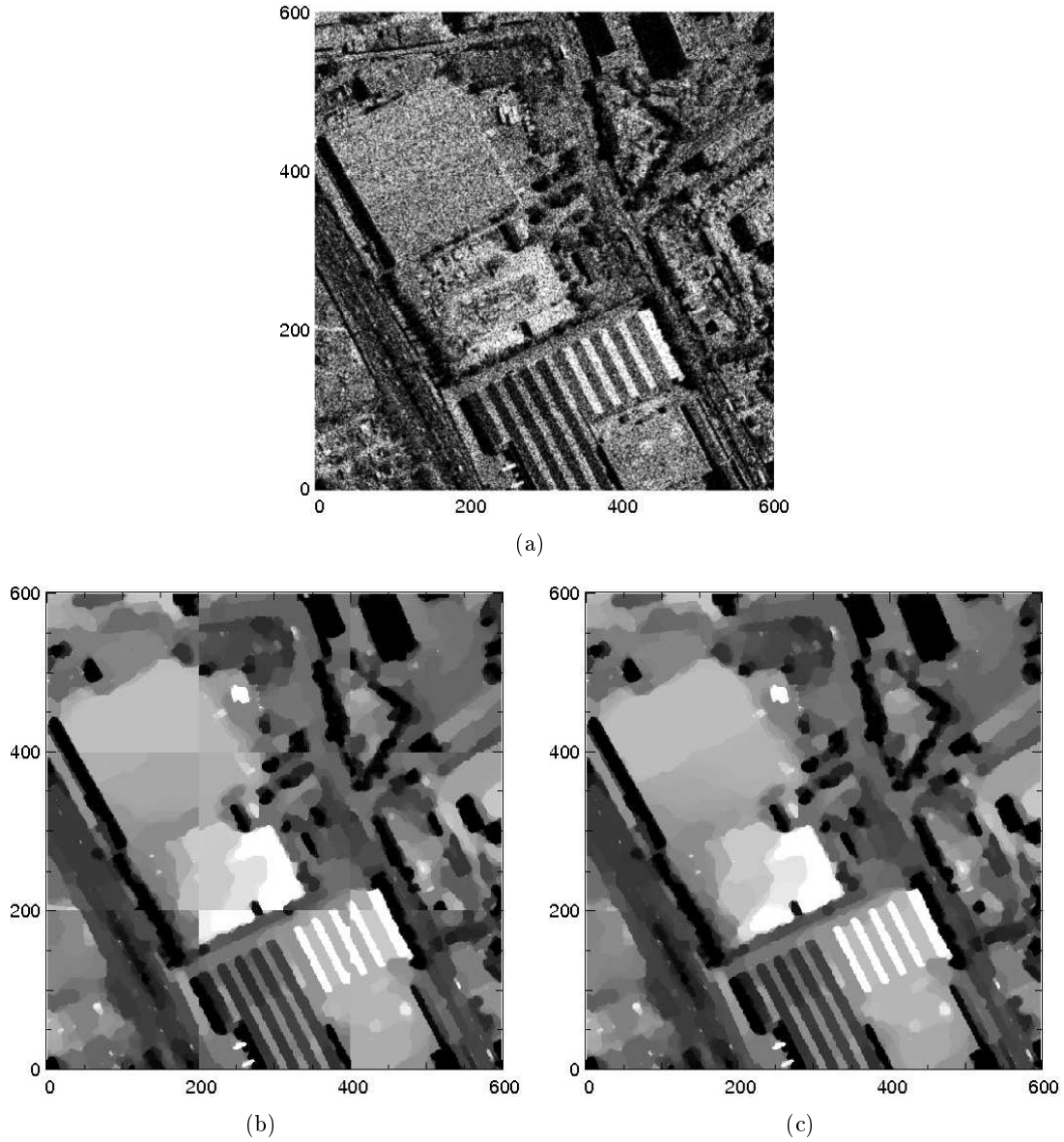


FIGURE 1.5 – Débruitage d'une image radar à synthèse d'ouverture. Le modèle markovien est défini par une attache aux données non-convexe et une régularisation convexe par variation totale. (a) Image radar ($n = 600 \times 600$, $k = 256$), (b) image régularisée par blocs de 200×200 (une minimisation exacte est effectuée sur chaque bloc), (c) image régularisée en entier (un optimiseur multi-étiquette qui sera introduit dans le chapitre suivant est utilisé. La minimisation exacte sur une telle image nécessite une mémoire physique et un temps de calcul prohibitifs).

Chapitre 2

Optimisation par des mouvements de partitions multi-étiquettes

Malgré les avancées en minimisation d'énergies markoviennes multi-étiquettes de premier ordre, présentées dans le chapitre précédent, la classe des algorithmes d'optimisation non-exacte demeure encore limitée pour converger vers des optima d'énergies markoviennes non-convexes à énergies proches d'un optimum global. Le recours aux optimiseurs exacts est alors nécessaire pour obtenir des résultats satisfaisants (optimum global de l'énergie), au prix d'une complexité algorithmique qui explose avec la taille des données.

Dans ce chapitre, nous proposons de nouveaux algorithmes d'optimisation de ces énergies dont l'*a priori* est convexe et l'attache aux données est quelconque. Les nouveaux optimiseurs sont plus robustes que les optimiseurs approximatifs qui existent dans l'état de l'art et plus efficaces que ceux exacts en termes de mémoire requise. Ils permettent de converger vers des optima locaux de bonne qualité (voir même globaux). La robustesse est obtenue grâce à de nouveaux mouvements de partitions que l'on appellera *mouvements de partitions larges et multi-étiquettes (MPLM)*. Par ailleurs, l'efficacité est due au recours à la technique de coupe-minimum sur des graphes de taille réduite comparée à celle des graphes des optimiseurs exacts.

Des travaux ont été réalisés récemment, indépendamment du nôtre, pour explorer des mouvements multi-étiquettes (Veksler, 2007, 2009) (Carr et Hartley, 2009) dans un contexte différent. Les auteurs cherchent en fait à définir ces mouvements pour la minimisation des modèles markoviens ayant un *a priori* quasi-convexe (par exemple : quadratique tronquée). Les mouvements proposés généralisent la technique d'échange d'étiquettes. Tandis que les mouvements que nous proposons dans ce chapitre généralisent tous les mouvements MPLB afin de mieux optimiser différentes formes d'énergie markoviennes sous-modulaires.

Ce chapitre présente une des contributions importantes de cette thèse. Une partie de ces travaux a été publiée dans les actes de la conférence internationale *IEEE International Conference on Image Processing (ICIP'09)* (Shabou *et al.*, 2009). La totalité de ce travail a été soumise à la revue *Image and Vision Computing*.

1 Optimisation multi-étiquette

Nous considérons dans ce chapitre les modèles markoviens multi-étiquettes de premier ordre définis par :

$$E(x) = \sum_{p=1}^n E_p(x_p) + \sum_{(p,q)} E_{p,q}(x_p, x_q), \quad (2.1)$$

tout en gardant les mêmes notations qu'au chapitre précédent.

Définition 1.1. Optimisation MPLM

Optimiser une fonction d'énergie markovienne multi-étiquette par des mouvements de partitions larges et multi-étiquettes (MPLM) consiste à proposer aux sites actifs, à chaque itération de l'algorithme d'optimisation, soit de conserver leur étiquette courante soit d'en prendre une nouvelle dans un sous-ensemble $\mathcal{L}_m$ de $\mathcal{L}$, avec m un entier dénotant la taille de ce sous-ensemble ($2 \leq m \leq k$).

Étant donnée cette définition, nous proposons une classification (cf. FIG. 2.1) des algorithmes d'optimisation par des mouvements de partitions selon les deux paramètres suivants :

- la largeur du mouvement ou son domaine spatial qui définit l'ensemble des sites actifs selon le mouvement,
- la taille du sous-ensemble d'étiquettes $\mathcal{L}_m$ ou son domaine radiométrique.

Domaine spatial du mouvement Domaine radiométrique du mouvement	Non-large	<i>largeur</i>	► Large	
Binaire		$\alpha\beta$ -échange	δ -saut	α -extension
Étiquettes		Algorithmes proposés: larges et multi-étiquettes		
▼ Card(L)	ICM			Exacte (Ishikawa, Darbon)

FIGURE 2.1 – Classification des algorithmes d'optimisation par mouvement de partitions.

Nous retrouvons dans ce tableau les MPLB (α -extension, $\alpha\beta$ -échange et δ -saut), les algorithmes d'optimisation exacte et un exemple d'algorithme classique à mouvement de partitions non-large et multi-étiquettes, l'ICM. La partie rayée de ce tableau correspond aux mouvements de partitions non encore explorés dans la littérature pour l'optimisation markovienne par coupure minimale sur graphes. C'est l'objet d'étude de ce chapitre.

Parmi les mouvements de partitions multi-étiquettes possibles, nous considérons les mouvements particuliers ci-après comme extension aux définitions des mouvements binaires MPLB introduits dans le chapitre précédent.

Définition 1.2. α_m -extension

Le mouvement de partition α_m -extension consiste à proposer, à une itération donnée, à

```

Initialisation de  $\tilde{x}$ ;
Pour chaque cycle  $c$  faire
    fin := faux;
    Pour chaque itération  $i$  faire
        1. Création de l'ensemble d'étiquettes de taille  $m$  :  $\mathcal{L}_m^{(i)}$ ;
        2. Recherche du mouvement de partition optimal :
            
$$x^{(i)} = \arg \min_{x \in \mathcal{X}} E(x),$$

            Si ( $E(x^{(i)}) < E(\tilde{x})$ ) Alors
                 $\tilde{x} := x^{(i)}$ ;
                fin := vrai;
            Fin Si
        Fin Pour
    Si (fin == faux) Alors
        retourner  $\tilde{x}$ ;
    Fin Si
Fin Pour

```

FIGURE 2.2 – Algorithme d'optimisation par MPLM.

tous les sites soit de garder leurs étiquettes courantes, soit de prendre une étiquette de l'ensemble $\mathcal{L}_m = \{\alpha_1, \alpha_2, \dots, \alpha_m\} \subset \mathcal{L}$.

Définition 1.3. β_m -échange

Le mouvement de partition β_m -échange consiste à proposer aux sites actifs, ceux dont les étiquettes sont dans $\mathcal{L}_m = \{\beta_1, \beta_2, \dots, \beta_m\} \subset \mathcal{L}$, d'échanger leurs étiquettes.

Définition 1.4. δ_m -saut

Le mouvement de partition δ_m -saut consiste à proposer aux sites actifs de diminuer/augmenter leurs étiquettes courantes par les quantités

$\mathcal{L}_m = \{-\delta_{\lfloor m/2 \rfloor}, -\delta_{\lfloor m/2 \rfloor - 1}, \dots, -\delta_1, 0, +\delta_1, \dots, +\delta_{\lfloor m/2 \rfloor - 1}, +\delta_{\lfloor m/2 \rfloor}\}$, avec $\lfloor \cdot \rfloor$ l'opération partie entière. Ces sites actifs sont les sites ayant des nouvelles étiquettes toujours dans $\mathcal{L}$.

Pour chacun de ces mouvements, le mouvement optimal permet d'obtenir une configuration ayant une énergie minimale par rapport à celle courante.

L'algorithme général d'optimisation par ces mouvements MPLM, illustré dans la figure 2.2, consiste alors à effectuer une série de mouvements optimaux partant d'une configuration initiale. Chaque mouvement diminue progressivement l'énergie globale de la configuration x , jusqu'à converger vers un optimum $\tilde{x}$ de la fonction d'énergie E .

En effet, à chaque itération (i), un ensemble d'étiquettes de taille m , noté par $\mathcal{L}_m$, est construit à partir de $\mathcal{L}$ selon la nature du mouvement (extension, échange ou saut). Ensuite, un mouvement de partition optimal est obtenu selon le mouvement choisi et l'ensemble d'étiquettes construit, qui décroît nécessairement l'énergie. Un cycle s'achève si toutes les étiquettes de $\mathcal{L}$ ont été proposées aux sites. Un nouveau cycle commence si

l'énergie a diminué au cours du dernier cycle. Autrement, l'algorithme a convergé vers un optimum de E .

Le problème du choix des étiquettes de $\mathcal{L}_m$ selon le type de mouvement sera discuté plus loin dans le chapitre. La détermination du mouvement optimal est détaillée dans le prochain paragraphe.

2 Construction du graphe d'un MPLM optimal

Pour réaliser un mouvement optimal, à une itération donnée de l'algorithme 2.2, nous nous reposons sur la technique de coupe-minimum sur graphe. La réalisation d'un mouvement multi-étiquette nécessitera ainsi la construction d'un graphe particulier que l'on appellera *graphe multi-étiquette* et que l'on notera par $\mathcal{G}_m$.

En nous inspirant de la construction du graphe multi-étiquette d'Ishikawa (cf. annexe B) pour la minimisation exacte des énergies markoviennes multi-étiquettes de premier ordre (Ishikawa, 2003), nous généralisons une telle construction pour des mouvements multi-étiquettes faisant intervenir des graphes ayant les particularités suivantes :

- Les nœuds étiquettes associés à un site donné, à une itération donnée, tiennent compte de l'étiquette courante du site pour pouvoir optimiser itérativement l'énergie globale.
- A chaque site est associé un sous-ensemble d'étiquettes de $\mathcal{L}$ propre à lui, différemment de la construction d'Ishikawa, où le même ensemble $\mathcal{L}$ est proposé à tous les sites.
- Le graphe n'est construit que sur les sites actifs.

Pour la suite, on notera par $\mathcal{L}_m(p)$ l'ensemble d'étiquettes associé au site p de $\mathcal{P}$ pour un MPLM de taille m . Cet ensemble sera considéré ordonné et de cardinalité k_p . Ainsi on a :

$$\mathcal{L}_m(p) = \{l_1^p, l_2^p, \dots, l_{k_p}^p\}, \quad (2.2)$$

avec

$$\begin{cases} l_1^p < l_2^p < \dots < l_{k_p}^p, \forall p \in \mathcal{P}, \\ l_i^p \in \mathcal{L}, \forall (i, p) \in \mathcal{L} \times \mathcal{P}. \end{cases}$$

Le graphe $\mathcal{G}_m = (\mathcal{V}_m, \mathcal{U}_m)$ est un graphe orienté, où $\mathcal{V}_m$ et $\mathcal{U}_m$ représentent respectivement l'ensemble des nœuds et l'ensemble des arcs du graphe. A ce graphe on ajoute deux nœuds spéciaux s et t , respectivement la source et le puits, afin de pouvoir y calculer le flot-maximum. La construction du graphe est réalisée de façon à ce qu'une coupe sur le graphe corresponde à une configuration possible de x et que la coupe minimum corresponde alors à la configuration optimale obtenue par le mouvement MPLM considéré, en partant d'une configuration courante.

Dans un premier temps, nous définissons l'ensemble des nœuds $\mathcal{V}_m$ comme suit :

$$\mathcal{V}_m = \{\mathcal{V}_m(p) | p \in \mathcal{P}\}, \quad (2.3)$$

avec

$$\mathcal{V}_m(p) = \{v_{p,i} | i \in \{1, 2, \dots, k_p\}\}. \quad (2.4)$$

Ainsi à chaque site actif p correspond k_p nœuds correspondants aux étiquettes de $\mathcal{L}_m$ et celle courante.

Ensuite, nous définissons l'ensemble des arcs $\mathcal{E}_m$ comme suit :

$$\mathcal{E}_m = \mathcal{E}_m^D \cup \mathcal{E}_m^C \cup \mathcal{E}_m^P, \quad (2.5)$$

où $\mathcal{E}_m^D$ dénote l'ensemble des arcs d'attache aux données (ou tout simplement arcs de données), $\mathcal{E}_m^C$ correspond à l'ensemble des arcs de contraintes, et $\mathcal{E}_m^P$ correspond à l'ensemble des arcs de pénalité ou d'interaction.

Les arcs de données $\mathcal{E}_m^D$ correspondent aux termes de vraisemblance de l'énergie et se décomposent ainsi :

$$\mathcal{E}_m^D = \bigcup_{p \in \mathcal{P}} \mathcal{E}_m^D(p), \quad (2.6)$$

avec

$$\mathcal{E}_m^D(p) = \{(s, v_{p,1})\} \cup \{(v_{p,i}, v_{p,i+1}) | i \in \{1, 2, \dots, k_p - 1\}\} \cup \{(v_{p,k_p}, t)\}, \quad (2.7)$$

et leurs capacités sont définies comme suit :

$$\begin{cases} c(s, v_{p,1}) & = +\infty, \\ c(v_{p,i}, v_{p,i+1}) & = E_p(x_p = l_i^p); i \in \{1, 2, \dots, k_p - 1\}, \\ c(v_{p,k_p}, t) & = E_p(x_p = l_{k_p}^p). \end{cases} \quad (2.8)$$

Afin de garantir la correspondance exacte entre une coupe sur le graphe et une configuration x (c.-à-d. pour chaque site p , uniquement un arc d'attache aux données est coupé), nous avons besoin de définir des arcs de contraintes, $\mathcal{E}_m^C$. Ces arcs sont répartis comme suit :

$$\mathcal{E}_m^C = \bigcup_{p \in \mathcal{P}} \mathcal{E}_m^C(p), \quad (2.9)$$

avec

$$\mathcal{E}_m^C(p) = \{(v_{p,i+1}, v_{p,i}) | i \in \{1, 2, \dots, k_p - 1\}\}, \quad (2.10)$$

et leurs capacités sont fixées à une valeur infinie, c.-à-d.

$$c(v_{p,i+1}, v_{p,i}) = +\infty; \forall i \in \{1, 2, \dots, k_p\}. \quad (2.11)$$

Enfin, les arcs d'interaction $\mathcal{E}_m^P$ correspondent aux termes de régularisation de l'énergie et sont répartis comme suit :

$$\mathcal{E}_m^P = \{(v_{p,i}, v_{q,j}) | (p, q) \in \mathcal{N}; i \in \{1, 2, \dots, k_p\}; j \in \{1, 2, \dots, k_q\}\}. \quad (2.12)$$

Pour définir les poids sur ces arcs, nous commençons par étudier la valeur de la coupe dans $\mathcal{G}_m$ entre deux sites voisins.

Considérons deux pixels voisins p et q en interaction. Une coupe quelconque sur le graphe $\mathcal{G}_m$ fera intervenir en plus des arcs de données associés à ces deux pixels les arcs d'interaction, telle que la valeur de la coupe sur ces derniers arcs vaut :

$$C(i, j) = \sum_{a=1}^{a=i} \sum_{b=j}^{b=k_q} c(v_{p,a}, v_{q,b}) + \sum_{a=i}^{a=k_p} \sum_{b=1}^{b=j} c(v_{q,b}, v_{p,a}), \quad (2.13)$$

où $C(i, j)$ correspond à la pénalité de l'interaction entre les deux pixels, $c(v_{p,a}, v_{q,b})$ est la capacité sur l'arc orienté $e = (v_{p,a}, v_{q,b})$ avec a et b deux indices d'étiquettes qui vivent toutes dans $\mathcal{L}_m(p)$ et $\mathcal{L}_m(q)$, et (i, j) est le couple d'indices d'étiquettes associées au couple de sites voisins (p, q) résultant de la coupe. En d'autres termes, la valeur $C(i, j)$ est la somme des poids sur les arcs que rencontre la coupe entre deux sites voisins sans compter les poids des arcs de données.

Similairement au raisonnement adopté dans (Ishikawa, 2003), on se propose dans la suite de définir l'expression analytique des poids sur ces arcs d'interaction en fonction des termes de régularisation de l'énergie. Dans un premier temps, nous montrons que si une telle construction existe alors la fonction de régularisation est nécessairement *convexe*. Dans un second temps, en considérant l'hypothèse de convexité de cette fonction, nous proposons une expression analytique possible des poids d'interaction dans le graphe. Cette expression généralisera le cas particulier de poids proposé dans (Ishikawa, 2003).

A. condition nécessaire de convexité :

Proposition 2.1. *Si la capacité de la coupe $C(i, j)$ obtenue en sommant toutes les capacités des arcs coupés entre deux sites voisins p et q peut s'écrire sous la forme d'une fonction g de la différence entre les étiquettes l_i^p et l_j^q , associées respectivement aux sites p et q après la coupe, alors la fonction g de cette différence est une fonction à une variable convexe sur $\mathbb{R}$.*

Avant de donner la preuve de cette proposition, nous mentionnons une propriété d'une fonction convexe discrète, à valeurs réelles, qui nous intéressera dans la preuve.

Lemme 2.1. *Soit f une fonction discrète unidimensionnelle : $f : \mathbb{R} \rightarrow \mathbb{R}$. On a l'équivalence entre ces deux propositions :*

1. f est convexe,
2. $\forall x, \forall y \geq x, \forall d \leq (y - x), f(x) + f(y) \geq f(x + d) + f(y - d)$.

La preuve de la proposition 2.1 est comme suit :

Preuve. Soit $a = d_i^p = l_i^p - l_{i-1}^p \geq 0$ représentant la différence entre les deux étiquettes successives l_{i-1}^p et l_i^p associé au site p . La positivité est justifiée par l'ordonnancement choisi ascendant des étiquettes dans $\mathcal{L}_m(p)$. De même notons par $b = d_j^q = l_j^q - l_{j-1}^q \geq 0$ qui représente la différence entre les deux étiquettes successives l_{j-1}^q et l_j^q associées au site q .

On a bien : $a + b \geq 0$.

Soit $x = l_i^p - l_j^q$ un réel. On a toujours : $x + b \geq x - a$.

Notons par $X = x + b$ et $Y = x - a$.

Soit g une fonction discrète unidimensionnelle : $(g : \mathbb{R} \rightarrow \mathbb{R})$. On a :

$$\begin{aligned}
 & g(X) + g(Y) - g(X - b) - g(Y + b) \\
 = & g(x + b) + g(x - a) - g(x) - g(x - a + b) \\
 = & (g(x + b) - g(x)) - (g(x - a + b) - g(x - a)) \\
 = & (g(l_i^p - l_j^q + d_j^q) - g(l_i^p - l_j^q)) - (g(l_i^p - l_j^q - d_i^p + d_j^q) - g(l_i^p - l_j^q - d_i^p)) \\
 = & (f(l_i^p, l_j^q - d_j^q) - f(l_i^p, l_j^q)) - (f(l_i^p - d_i^p, l_j^q - d_j^q) - f(l_i^p - d_i^p, l_j^q)) \\
 = & (f(l_i^p, l_{j-1}^q) - f(l_i^p, l_j^q)) - (f(l_{i-1}^p, l_{j-1}^q) - f(l_{i-1}^p, l_j^q)) \\
 = & (C(i, j - 1) - C(i, j)) - (C(i - 1, j - 1) - C(i - 1, j)).
 \end{aligned}$$

En développant l'expression finale de cette égalité, utilisant l'équation (2.13), nous obtenons :

$$\left(C(i, j-1) - C(i, j) \right) - \left(C(i-1, j-1) - C(i-1, j) \right) = c(v_{p,i}, v_{q,j}) + c(v_{q,j}, v_{p,i}) \geq 0.$$

La positivité est due au fait que les poids des arcs du graphe sont nécessairement positifs. Ainsi on a :

$$\begin{aligned} g(X) + g(Y) - g(X-b) - g(Y+b) &\geq 0 \\ \Rightarrow g(X) + g(Y) &\geq g(Y+b) + g(X-b), \\ \text{où } X \geq Y \text{ et } (X-Y) &= (b+a) \geq b \text{ (puisque } a \geq 0). \end{aligned}$$

Cela conclut la convexité de la fonction g selon le lemme 2.1. $\square$

B. Expression analytique des poids d'interaction :

Considérons maintenant une fonction $g(x)$ convexe et paire (dans le cas non paire, nous pouvons toujours prendre la fonction définie par $\frac{g(x)+g(-x)}{2}$ qui est aussi convexe et paire). On définit les poids sur les arcs d'interaction du graphe G_m comme suit :

$$c(v_{p,a}, v_{q,b}) = \frac{1}{2} \left(g(l_a^p - l_{b-1}^q) + g(l_{a-1}^p - l_b^q) - g(l_{a-1}^p - l_{b-1}^q) - g(l_a^p - l_b^q) \right). \quad (2.14)$$

En reprenant la propriété de convexité citée dans le lemme 2.1 et la preuve précédente, nous pouvons conclure que ces capacités sont toujours positives grâce à la convexité de la fonction g .

En injectant cette expression de capacités dans les deux sommes de l'équation (2.13), nous obtenons l'égalité suivante :

$$C(i, j) = g(l_i^p - l_j^q) - D_{q,p}(j) - D_{p,q}(i) + Cst, \quad (2.15)$$

avec :

$$\begin{cases} D_{p,q}(i) &= \frac{1}{2} \left(g(l_i^p - l_1^q) + g(l_i^p - l_{k_q}^q) \right), \\ D_{q,p}(j) &= \frac{1}{2} \left(g(l_j^q - l_1^p) + g(l_j^q - l_{k_p}^p) \right), \\ Cst &= \frac{1}{2} \left(g(l_1^p - l_{k_q}^q) + g(l_1^q - l_{k_p}^p) \right). \end{cases} \quad (2.16)$$

Ainsi la capacité d'un arc d'interaction est bien définie par l'expression (2.14) puisque la somme des capacités des arcs d'interaction coupées, entre deux sites voisins, est égale au terme d'interaction de la fonction markovienne définie sur ces deux sites, à une constante près. Les termes $D_{p,q}(i)$ et $D_{q,p}(j)$ peuvent être ajoutés respectivement aux arcs de données associés aux pixels p et q , puisqu'ils ne dépendent que de leurs étiquettes séparément. On les appellera par la suite des *termes de bord*.

Ayant défini les bonnes capacités sur les différents arcs du graphe $\mathcal{G}_m$, nous obtenons le théorème suivant :

Théorème 2.1. *Soit $\mathcal{G}_m$ un graphe multi-étiquette construit comme il a été indiqué précédemment. Étant donnée une configuration x^C des sites p et un ensemble de labels $\mathcal{L}_m$, il existe une correspondance terme à terme entre une coupe $\mathcal{C}$ sur ce graphe et un étiquetage des sites dans l'ensemble $\mathcal{L}_m$. De plus, nous avons l'égalité suivante : $|\mathcal{C}| = E(x^C) + Cst$, où $|\mathcal{C}|$ est le poids de la coupe $\mathcal{C}$ et Cst est une constante.*

Preuve. Soit $\mathcal{C}$ une coupe dans le graphe $\mathcal{G}_m$ qui lui correspond un étiquetage unique x^C . L'étiquette x_p^C associée au site p par cette coupe indique quel arc de données est coupé (unique arc par site p). Le poids sur cet arc est égal à un terme d'attache aux données et un terme de bord. Pour deux sites voisins (p, q) , la coupe $\mathcal{C}$ passe par des arcs de poids définis précédemment dont la somme est égale au potentiel d'interaction associé aux deux étiquettes (x_p, x_q) à une constante près.

Notons par $\{d_i^p ; i \in \{1, 2, \dots, k_p\} ; p \in \mathcal{P}\}$ l'ensemble des arcs de données et par $\{e_{i,j}^{p,q} ; i \in \{1, 2, \dots, k_p\} , j \in \{1, 2, \dots, k_q\} ; (p, q) \in \mathcal{N}\}$ l'ensemble des arcs d'interactions entre deux sites voisins (p, q) .

On a :

$$|\mathcal{C}| = \sum_{p \in \mathcal{P}} |\mathcal{C} \cap \{d_i^p ; i \in \{1, 2, \dots, m\}\}| + \sum_{(p,q) \in \mathcal{N}} |\mathcal{C} \cap \{e_{i,j}^{p,q} ; i \in \{1, 2, \dots, k_p\} , i \in \{1, 2, \dots, k_q\}\}|, \quad (2.17)$$

tel que :

$$\begin{aligned} |\mathcal{C} \cap \{d_i^p ; i \in \{1, 2, \dots, k_p\}\}| &= |d_a^p| ; \text{ si } d_a^p \in \mathcal{C} \\ &= E^p(x_p = l_a^p) - \sum_{(p,q) \in \mathcal{N}} D_{p,q}(a) \\ &= E^p(x_p^C) - Cst_1, \end{aligned}$$

et

$$\begin{aligned} |\mathcal{C} \cap \{e_{i,j}^{p,q} ; i \in \{1, 2, \dots, k_p\} , j \in \{1, 2, \dots, k_q\}\}| &= \sum_{a=1}^{a=i} \sum_{b=j+1}^{b=k_q} |e_{a,b}^{p,q}| + \sum_{a=i+1}^{a=k_p} \sum_{b=1}^{b=j} |e_{b,a}^{q,p}| ; \text{ si } d_i^p \in \mathcal{C} \text{ et } d_j^q \in \mathcal{C} \\ &= E^{p,q}(x_p = l_i^p, x_q = l_j^q) + D_{p,q}(i) + D_{q,p}(j) + Cst \\ &= E^{p,q}(x_p^C, x_q^C) + Cst_2. \end{aligned}$$

Ainsi on a :

$$\begin{aligned} |\mathcal{C}| &= \sum_p (E^p(x_p^C) + Cst_1) + \sum_{(p,q) \in \mathcal{N}} (E^{p,q}(x_p^C, x_q^C) + Cst_2) \\ &= \sum_p E^p(x_p^C) + \sum_{p \sim q} E^{p,q}(x_p^C, x_q^C) + Cst \\ &= E(x^C) + Cst. \end{aligned}$$

□

D'où le corollaire suivant :

Corollaire 2.1. *Un MPLM optimal partant d'une configuration x donne une configuration x^C , où $\mathcal{C}$ est la coupe-minimum sur le graphe $\mathcal{G}_m$.*

Dans la figure 2.3, la construction des structures répétitives du graphe $\mathcal{G}_m$ est présentée. Une illustration 1D du graphe total est présentée dans la figure 2.4.

3 Propriétés de convergence

Pour étudier les propriétés de convergence d'un algorithme d'optimisation, on se ramène à étudier les caractéristiques suivantes :

1. la complexité algorithmique en temps de convergence et occupation en mémoire,
2. la qualité de l'optimum vers lequel il converge.

3.1 Complexité algorithmique

Il est clair que l'algorithme proposé effectue un nombre fini d'itérations par cycle. Ce nombre est strictement lié à la taille m des sous-ensembles d'étiquettes ainsi qu'à la stratégie de parcours des étiquettes. A chaque itération, nous cherchons la coupe-minimum sur le graphe $\mathcal{G}_m$ de taille nm en nœuds. Elle pourrait être estimée à une complexité pseudo-polynomiale, en utilisant l'algorithme de flot-maximum proposé dans (Boykov et Kolmogorov, 2004)¹.

1. Code source disponible sur [http : //vision.csd.uwo.ca/code/](http://vision.csd.uwo.ca/code/)

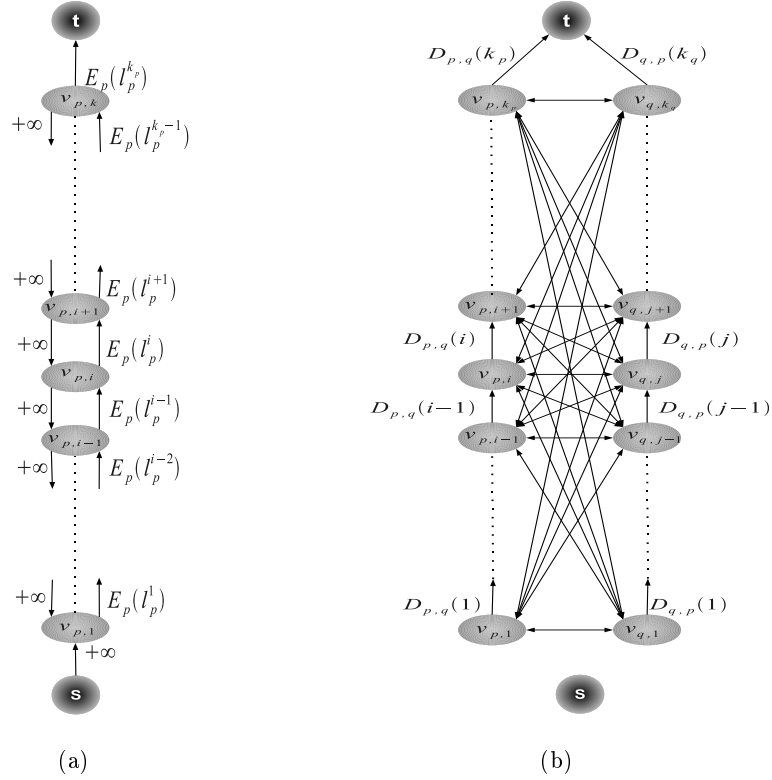


FIGURE 2.3 – Illustration des structures répétitives du graphe $\mathcal{G}_m$. (a) Structure répétitive par colonne des arcs de données et de contraintes associés à chaque site actif p , (b) structure répétitive par couples de colonnes d'arcs d'interaction associés aux sites voisins actifs (p, q) .

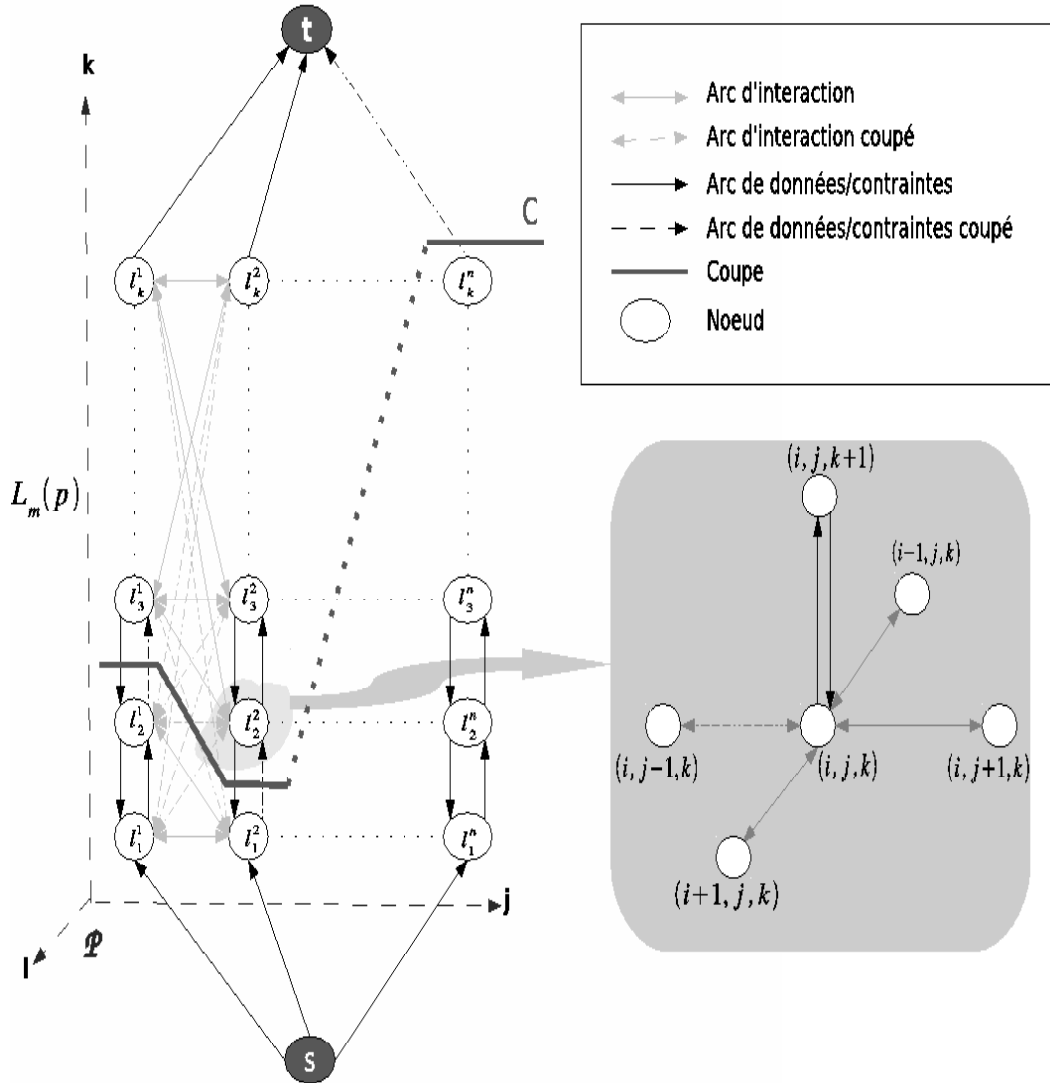
Le nombre de cycles nécessaires pour converger est délicat à prédire puisqu'il dépend de la convergence à un optimum de la fonction, c.-à-d. de la forme de la fonction d'énergie, du parcours effectué sur les étiquettes et aussi du mouvement considéré pour la recherche de la solution dans tout l'espace des solutions possibles. Compte tenu de tous ces facteurs qui varient d'un cas d'étude à un autre et des choix algorithmiques, il est très difficile de déterminer des bornes sur la convergence finale des algorithmes proposés. Néanmoins, nous pouvons indiquer une borne limite supérieure en complexité, en se plaçant dans les pires des cas. Elle est la même pour tous les optimiseurs à mouvements de partition, en particulier l' α -extension. Pour ce dernier, il a été montré dans (Veksler, 1999) que l'algorithme converge en $O(n)$ cycles.

Au final, l'algorithme d'optimisation par MPLM est de complexité pseudo-polynomiale. En pratique, la convergence est obtenue après un nombre notablement réduit de cycles comparé à n (typiquement une dizaine de cycles).

3.2 Qualité de l'optimum

Comme il a été mentionné au début de ce chapitre, les mouvements multi-étiquettes sont une alternative par rapport aux autres mouvements de partition afin de pouvoir mieux optimiser une classe de fonctions d'énergies non-convexes. Ainsi la qualité des optima vers lesquels ils convergent serait meilleure et leur valeur en énergie est plus proche d'un minimum global de la fonction.

D'un autre côté, face à une exploration complexe de l'espace des configurations possi-


 FIGURE 2.4 – Illustration 1D du graphe $\mathcal{G}_m$ total avec une coupe C .

bles par un algorithme MPLM, relier analytiquement la qualité de l'optimum obtenu au paramètre principal (m) de l'algorithme est une tâche complexe. Hormis pour des cas simples et restreints à des stratégies de parcours que nous étudions dans la suite, il nous a été impossible de généraliser les propriétés trouvées.

Proposition 3.1. Soit $\mathcal{L}_{m-1}$ et $\mathcal{L}_m$ deux ensembles d'étiquettes de $\mathcal{L}$ et x un étiquetage courant. Notons par x^m l'étiquetage optimal obtenu après une seule itération d'un mouvement MPLM sur $\mathcal{L}_m$ à partir de x et par x^{m-1} l'étiquetage optimal obtenu après une seule itération de ce même mouvement sur $\mathcal{L}_{m-1}$ à partir de x . On a alors l'implication suivante :

$$\mathcal{L}_{m-1} \subset \mathcal{L}_m \implies E(x^{m-1}) \geq E(x^m).$$

Preuve. La preuve est directe puisqu'à une itération donnée, et partant d'une même configuration x , en proposant un ensemble d'étiquettes $\mathcal{L}_m$ qui soit plus large que $\mathcal{L}_{m-1}$ et le contenant en même temps, l'ensemble de configurations possibles par un MPLM sur $\mathcal{L}_{m-1}$ sera forcément incluse dans celui par un MPLM sur $\mathcal{L}_m$. Ainsi l'énergie de la configuration optimale x^m sera inférieure ou égale à celle de x^{m-1} .

Il est aussi trivial de montrer que le sens opposé de la proposition n'est pas toujours vraie. Il est possible d'obtenir une configuration optimale d'énergie plus faible tout en proposant un ensemble d'étiquettes moins large. Par exemple, un mouvement par un ensemble d'étiquettes formé des étiquettes de la configuration de l'optimum global donne une meilleure configuration qu'un mouvement par un ensemble plus large ne contenant aucune étiquette de cet optimum. $\square$

Dans un cas général, on obtient la proposition suivante :

Proposition 3.2. *Partons d'une même configuration initiale x à l'itération (0) pour deux optimiseurs MPLM de paramètres m et $m - 1$. Soient $\mathcal{L}_{m-1}^{(i)}$ et $\mathcal{L}_m^{(i)}$ deux ensembles d'étiquettes de $\mathcal{L}$ à l'itération (i) , $x^{m-1(i)}$ un étiquetage courant à la i ème itération de l'algorithme d'optimisation par MPLM sur des sous-ensembles d'étiquettes de tailles $m - 1$ de $\mathcal{L}$, et $x^{m(i)}$ un étiquetage courant à la i ème itération de l'algorithme d'optimisation par MPLM sur des sous-ensembles d'étiquettes de tailles m de $\mathcal{L}$. Notons par x^{m-1} l'étiquetage optimal obtenu après une seule itération d'un mouvement MPLM sur $\mathcal{L}_{m-1}^{(i)}$ à partir de $x^{m-1(i)}$ et par x^m l'étiquetage optimal obtenu après une seule itération de ce même mouvement sur $\mathcal{L}_m^{(i)}$ à partir de $x^{m(i)}$. On a alors l'implication suivante :*

$$\mathcal{L}_{m-1}^{(k)} \subset \mathcal{L}_m^{(k)}, \forall 0 \leq k \leq i \implies E(x^{m-1}) \geq E(x^m).$$

Preuve. La preuve repose sur la même idée que précédemment. En effet, en arrivant à l'itération (i) de l'algorithme d'optimisation, l'ensemble des configurations possibles par le mouvement MPLM sur les sous-ensembles $\{\mathcal{L}_{m-1}^{(k)}, \forall 0 \leq k \leq i\}$ est toujours incluse dans l'ensemble des configurations possibles par le même mouvement sur les sous-ensembles $\{\mathcal{L}_m^{(k)}, \forall 0 \leq k \leq i\}$. Ce qui conclut que l'énergie de la configuration optimale x^m sera toujours inférieure ou égale à celle de x^{m-1} . $\square$

D'où le corollaire suivant sur la qualité de l'optimum obtenu par des MPLM en augmentant m :

Corollaire 3.1. *La configuration obtenue à la convergence d'un algorithme MPLM sur des sous-ensembles $\mathcal{L}_m^{(i)}$ de $\mathcal{L}$ de taille m est d'énergie inférieure à celle obtenue à la convergence du même algorithme sur des sous-ensembles $\mathcal{L}_{m-1}^{(i)}$ de $\mathcal{L}$ de taille $m - 1$, si $\mathcal{L}_{m-1}^{(i)} \subset \mathcal{L}_m^{(i)}, \forall i$ et en partant d'un même point de départ.*

Cette dernière condition nous l'appellerons par la suite *condition d'inclusion*. Elle permet d'obtenir un meilleur (ou égal) optimum en augmentant la taille des sous-ensembles d'étiquettes.

Dans un cas particulier du corollaire précédent, on obtient le corollaire suivant :

Corollaire 3.2. *L'algorithme de minimisation par un mouvement MPLM de paramètre m converge vers un meilleur optimum qu'un algorithme MPLB, si la condition d'inclusion est vérifiée et en partant d'une même configuration initiale.*

Cela se déduit immédiatement du corollaire précédent puisqu'un mouvement MPLB correspond à un mouvement MPLM à un paramètre minimum $m = 1$.

Ainsi nous avons pu prouver dans ce paragraphe la propriété de convergence vers un meilleur optimum en augmentant le paramètre m de l'algorithme MPLM, mais dans un cas bien particulier de parcours d'étiquettes.

Dans le cas général, il est *a priori* impossible de prouver de telles propriétés. Il nous a été impossible de comparer les ensembles de configurations par des mouvements MPLM quelconques entre eux. Cependant nous montrerons empiriquement qu'augmenter le nombre d'étiquettes m permet une convergence vers de meilleurs optima locaux, même avec des parcours aléatoires quelconques.

4 Expériences et résultats

4.1 Parcours des étiquettes

Le parcours des étiquettes dans le cadre des optimiseurs itératifs agit directement sur les mouvements des partitions et donc sur la convergence de l'algorithme. Cela se ressent plus dans le cas de fonction d'énergie non-convexe, où selon le choix de ce parcours, différents optima locaux sont obtenus. Pour une fonction convexe, comme il s'agit d'un unique optimum, le parcours des étiquettes agit uniquement sur la vitesse de convergence de l'algorithme. Ainsi pour les différentes fonctions et modèles markoviens que nous optimisons, il est judicieux de bien définir une stratégie de parcours des étiquettes la plus adaptée possible. Cela apportera à l'algorithme un gain en temps de calcul d'une part et une convergence vers un optimum de bonne qualité de l'autre part. La stratégie de parcours est directement liée à la nature du mouvement de partitions choisi. Ainsi pour les différents mouvements introduits dans ce chapitre, nous présentons des stratégies qui leurs sont appropriées.

α_m -extension

Dans ce cas de mouvement, des sous-ensembles d'étiquettes de $\mathcal{L}$ sont proposés aux sites itérativement. A une itération donnée, le sous-ensemble d'étiquettes est défini indépendamment des étiquettes des sites de la configuration courante. Deux stratégies de parcours d'étiquettes, illustrées dans la figure 2.5, sont alors envisageables :

1. Parcours ordonné, tel que le sous-ensemble d'étiquettes à l'itération i succède directement au sous-ensemble des étiquettes de l'itération $i - 1$, dans l'ensemble $\mathcal{L}$.
2. Parcours aléatoire obtenu en effectuant une permutation de l'ensemble $\mathcal{L}$ à chaque cycle. On choisit ensuite, itérativement, des sous-ensembles de même taille m ayant une intersection vide. A chaque itération, le sous-ensemble d'étiquettes sera ordonné pour pouvoir construire le graphe $\mathcal{G}_m$.

En terme de qualité d'optimum, un parcours aléatoire pourrait décroître mieux l'énergie compte tenu de la recherche aléatoire des solutions dans l'espace des configurations possibles. Seulement cela pourrait être non avantageux en terme de temps de calcul, où un parcours déterministe ordonné s'avère plus rapide, puisque les mêmes sous-ensembles d'étiquettes seront alors proposés à chaque cycle. La convergence est alors plus rapide sans pour autant pénaliser la qualité de l'optimum obtenu (qui serait plutôt liée à la taille m).

Dans la suite, nous choisissons un parcours ordonné pour un mouvement d'extension.

β_m -échange

Pour ce mouvement, le choix des sous-ensembles d'étiquettes de $\mathcal{L}$, à une itération donnée, détermine l'ensemble des sites actifs et donc les partitions qui changeront par ce mouvement. Ainsi le choix du parcours des étiquettes est en relation avec les configurations courantes des solutions. Deux stratégies illustrées dans la figure 2.6 sont alors envisageables :

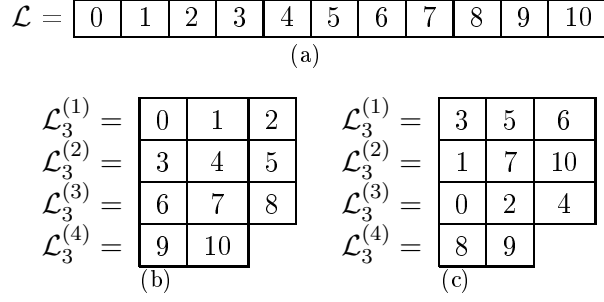


FIGURE 2.5 – Exemples de parcours d’étiquettes par un mouvement type α_3 -extension. (a) Ensemble des étiquettes possibles, (b) parcours ordonné des étiquettes par des paquets de 3, (c) parcours aléatoire des étiquettes par des paquets de 3 après avoir effectué une permutation sur $\mathcal{L}$ donnant $\mathcal{L} = \{3, 5, 6, 1, 7, 10, 0, 2, 4, 8, 9\}$.

1. Parcours par énumération de toutes les combinaisons possibles de sous-ensembles $\mathcal{L}_m$ de $\mathcal{L}$. Ceci permettra de réduire la dépendance entre parcours d’étiquettes et configurations courantes, puisque toutes les possibilités d’échange d’étiquettes seront proposées. Cependant, l’optimisation nécessitera un temps de calcul énorme. A chaque cycle, il faudrait effectuer $\binom{k}{m} = \frac{k!}{m!(k-m)!}$ itérations !
2. Parcours aléatoire, le même que celui décrit dans α_m -extension, mais avec chevauchement des paquets d’étiquettes entre deux itérations successives pour garantir un ensemble large de sites actifs durant un mouvement de partition.

En terme de qualité d’optimum, il est clair que le choix du premier parcours est plus approprié, mais au prix d’un temps de calcul non-intéressant en pratique. On préfère le deuxième parcours, qui par l’aspect aléatoire pourrait converger vers des optima locaux de qualité proche tout en ayant un temps de convergence raisonnable. C’est cette stratégie qui sera utilisée par la suite.

A noter que le parcours proposé dans (Veksler, 2009), pour les mouvements d’échange multi-étiquettes, est un parcours de type ordonné par paquets d’étiquettes adjacentes, comme celui proposé dans α_m -extension. Cette stratégie de parcours est adoptée pour contraindre l’*a priori* non-convexe à demeurer convexe sur ces sous-ensembles d’étiquettes. Le parcours, testé expérimentalement, ne permet pas de bien explorer l’ensemble de configurations possibles par un mouvement de type échange, ce qui pourrait limiter la qualité de l’optimum obtenu à la convergence.

δ_m -saut

Le mouvement de saut est un mouvement directement lié à la configuration courante. A chaque itération, les étiquettes courantes des sites sont augmentées/diminuées par des quantités de $\mathcal{L}_m$. Pour éviter une optimisation qui soit locale par rapport à la configuration courante, qui conduira l’algorithme à stagner dans un optimum local proche de l’initialisation, un parcours d’étiquettes non-local est alors nécessaire. En d’autres termes, durant un cycle, de faibles ainsi que de forts sauts radiométriques $\mathcal{L}_m$ sont proposés aux sites. Pour ce type de mouvement, l’ensemble d’étiquettes $\mathcal{L}$ n’est pas défini à l’avance, il est plutôt défini itérativement. L’algorithme d’optimisation a besoin uniquement de la valeur

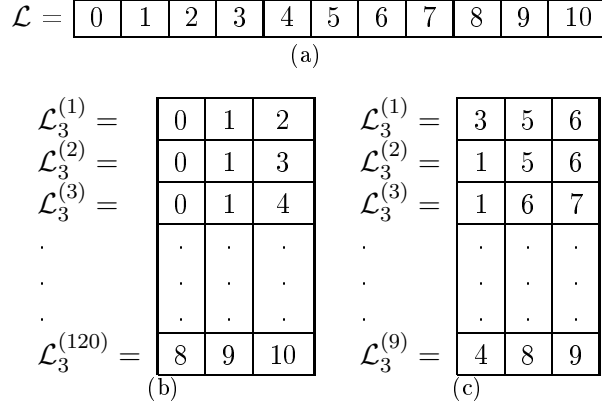


FIGURE 2.6 – Exemples de parcours d’étiquettes par un mouvement type β_3 -échange. (a) Ensemble des étiquettes possibles, (b) parcours ordonné des étiquettes par des paquets de 3 de toutes les combinaisons possibles de 3 étiquettes parmi les 11 étiquettes, (c) parcours aléatoire des étiquettes en effectuant une permutation sur $\mathcal{L}$ donnant $\mathcal{L} = \{3, 5, 6, 1, 7, 10, 0, 2, 4, 8, 9\}$ et en sélectionnant des paquets d’étiquettes de taille 3, en avançant d’un pas d’une seule étiquette dans le nouveau $\mathcal{L}$ et en ordonnant les paquets obtenus à chaque fois.

minimale ainsi que de celle maximale de $\mathcal{L}$ pour définir les sites actifs à chaque itération. Ce mouvement est donc intéressant dans le cas des applications où la quantification de l’ensemble des étiquettes possibles dans un unique ensemble $\mathcal{L}$ induit une perte d’information dans la solution estimée. Le recours à des étiquettes de manière adaptative réduit ce problème.

Nous proposons deux stratégies de parcours illustrées dans la figure 2.7 :

1. Saut ordonné, tel que les sous-ensembles de sauts $\mathcal{L}_m$, d’une itération à une autre, se succèdent dans un intervalle de saut large $[\delta_{min}, \delta_{max}]$. Ainsi les mouvements de partitions sont initialement locaux, par rapport aux configurations courantes. Plus on avance dans les itérations, plus des forts changements sont possibles, permettant de ne pas stagner dans un optimum local.
2. Saut dichotomique, tel qu’à une itération donnée i , le sous-ensemble d’étiquettes $\{\delta_{min}^{(i)}, \dots, \delta_{max}^{(i)}\}$, est décomposé par dichotomie, avec $\delta_{min}^{(i)} = \frac{\delta_{min}^{(i-1)}}{2}$ et $\delta_{max}^{(i)} = \frac{\delta_{max}^{(i-1)}}{2}$. Le nombre d’itérations dans un cycle est fixé initialement et correspond à la précision sur la solution estimée.

En terme de qualité de l’optimum, les deux techniques peuvent converger vers des optima différents, comme les étiquettes proposées par chacune sont complètement différentes. En terme de temps de convergence, le saut dichotomique est plus rapide à une même précision donnée par rapport au saut ordonné. Dans ce sens, il sera considéré par la suite pour effectuer les mouvements de type saut.

Après avoir défini les stratégies de choix des étiquettes, nous passons à l’évaluation des mouvements proposés pour optimiser certaines fonctions d’énergie markoviennes utilisées

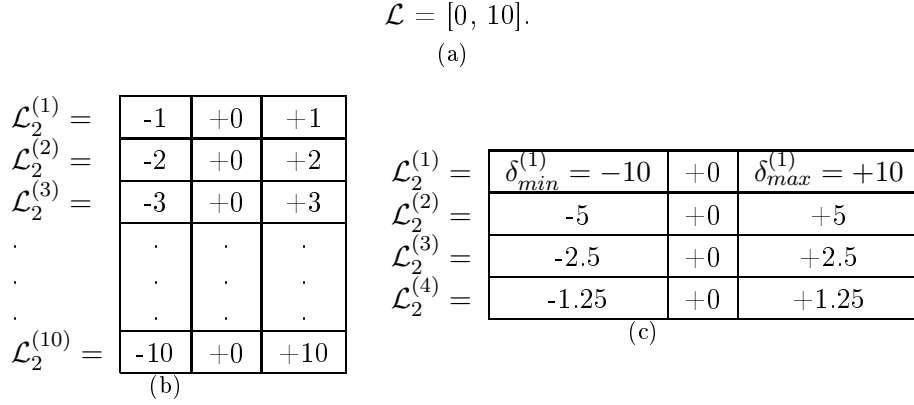


FIGURE 2.7 – Exemples de parcours d’étiquettes par un mouvement type δ_2 -saut. (a) Ensemble continu des étiquettes possibles, (b) saut ordonné par des paquets de 2, (c) saut dichotomique, au premier cycle, par des paquets de 2, à une précision de 1.

en traitement d’image.

4.2 Restauration d’images bruitées par un bruit impulsif

Dans ce paragraphe, nous expérimentons les optimiseurs multi-étiquettes par MPLM dans la cadre de la restauration d’image corrompue par un bruit de type impulsif. Le choix de ce type de bruit est intéressant dans notre cas, étant donnée la non-convexité de la fonction d’énergie de vraisemblance définie pour chaque site p par :

$$E_p(y_p|x_p) = \begin{cases} -\log\left((1-P) + \frac{P}{k}\right) & \text{si } y_p = x_p, \\ -\log\left(\frac{P}{k}\right) & \text{sinon.} \end{cases} \quad (2.18)$$

P est le paramètre de ce type de bruit, indiquant la probabilité pour qu’un site p soit bruité. Le pixel est bruité en prenant une étiquette aléatoirement dans l’ensemble des étiquettes $\mathcal{L}$ de taille k .

Pour régulariser l’image, le modèle discret de la variation totale (TV) défini par :

$$\sum_{(p,q)} E_{p,q}(x_p, x_q) = \beta \sum_{(p,q)} |x_p - x_q|, \quad (2.19)$$

avec β un coefficient global de régularisation, est considéré.

Sans perdre de généralité, des palettes d’étiquettes quantifiées entre l_{min} et l_{max} , d’un pas de 1 (par exemple $\mathcal{L} = \{0, 1, \dots, 254, 255\}$) sont considérées dans les expériences. Ainsi les optimiseurs MPLM par extension et échange sont plus appropriés dans ce cas d’application. Les résultats de l’optimiseur par sauts seront illustrés dans un autre cadre d’application en imagerie, présentée dans le chapitre suivant, où la quantification de l’ensemble des étiquettes n’est plus nécessaire, voir même peu appropriée.

Dans la suite, uniquement les résultats de la restauration par ces deux types de mouvements (extension et échange) seront illustrés. Les résultats seront comparés par rapport à ceux

- des optimiseurs MPLB,
 - d’un optimiseur exact,
- et par rapport à ce que donnent d’autres approches de type :
- ICM

- recuit simulé

Nous illustrons aussi quantitativement le comportement de ces différents algorithmes de minimisation en terme de minimum d'énergie atteint, de temps de calcul et d'occupation en mémoire. Par ailleurs, les propriétés de convergence des optimiseurs multi-étiquettes, discutées dans la section 3, seront aussi étudiées empiriquement.

Exemple 1 : Image du Pingouin

Cette première expérience illustre les performances des algorithmes d'optimisation par MPLM comparées aux autres algorithmes de l'état de l'art. L'image bruitée en impulsif est présentée dans la figure 2.8(b). La palette des étiquettes considérée est dans l'ensemble $\mathcal{L} = \{0, 1, \dots, 255\}$.

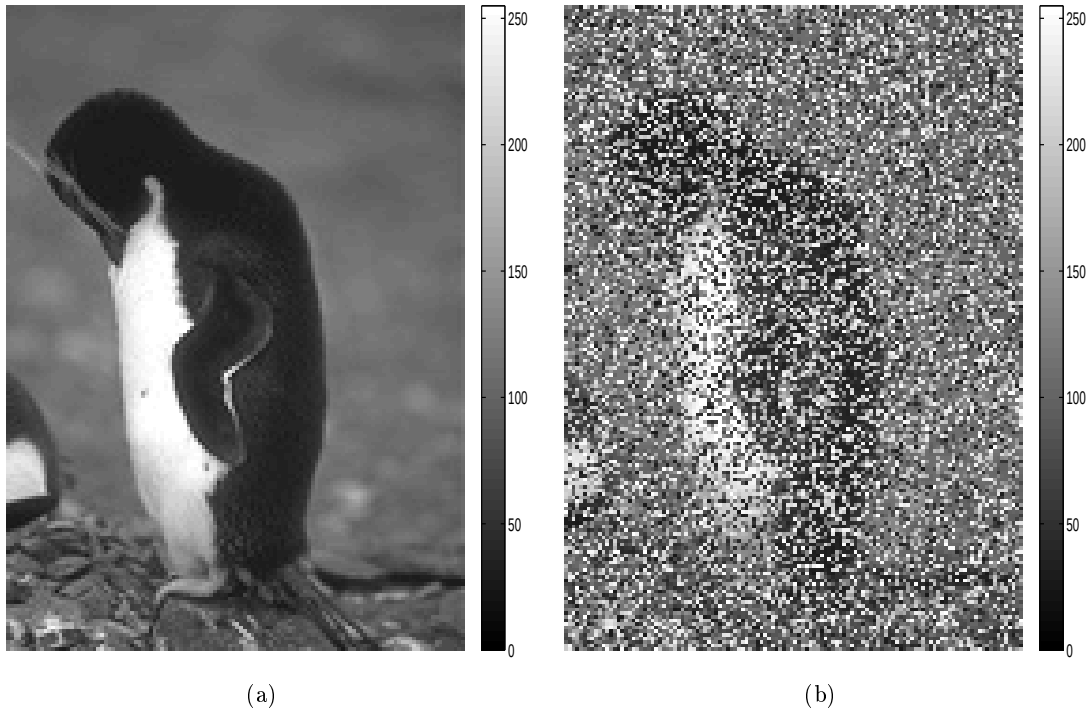
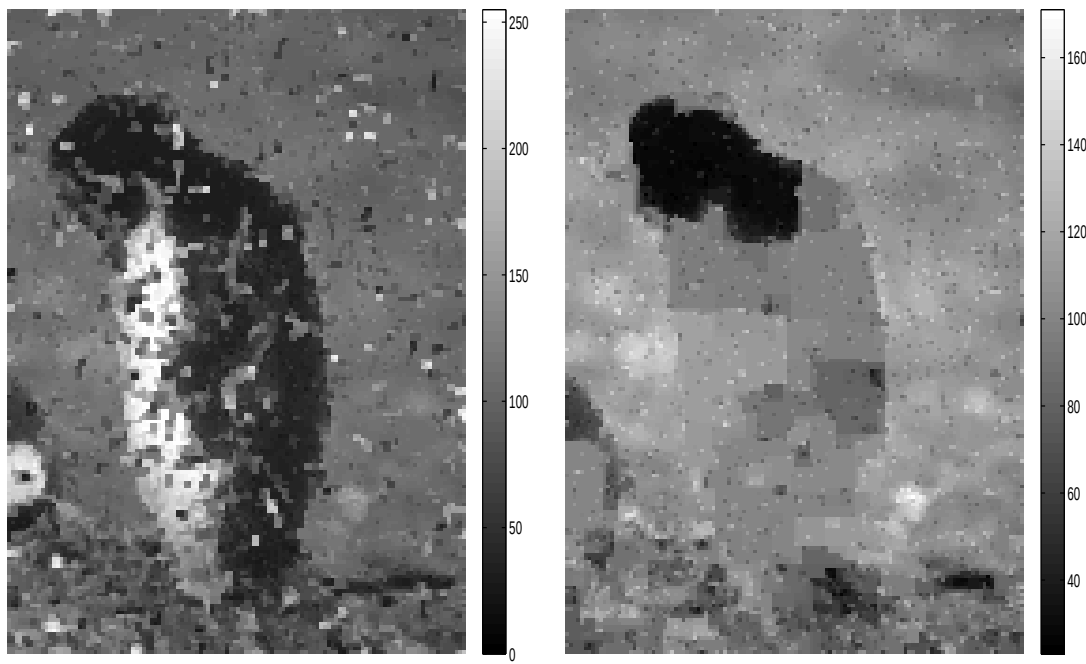


FIGURE 2.8 – Images d'origine et bruitée par un bruit impulsif avec $P = 0.6$. (f) Image d'origine de taille 179×122 pixels, (b) image bruitée.

Les résultats dans les figures 2.9(a), 2.9(b) et 2.9(c) montrent bien la faible qualité de l'image restaurée relativement au modèle markovien établi, comparé à ce que donne un optimiseur exact dans les figures 2.9(d) et 2.9(e).



(a)

(b)



(c)

(d)

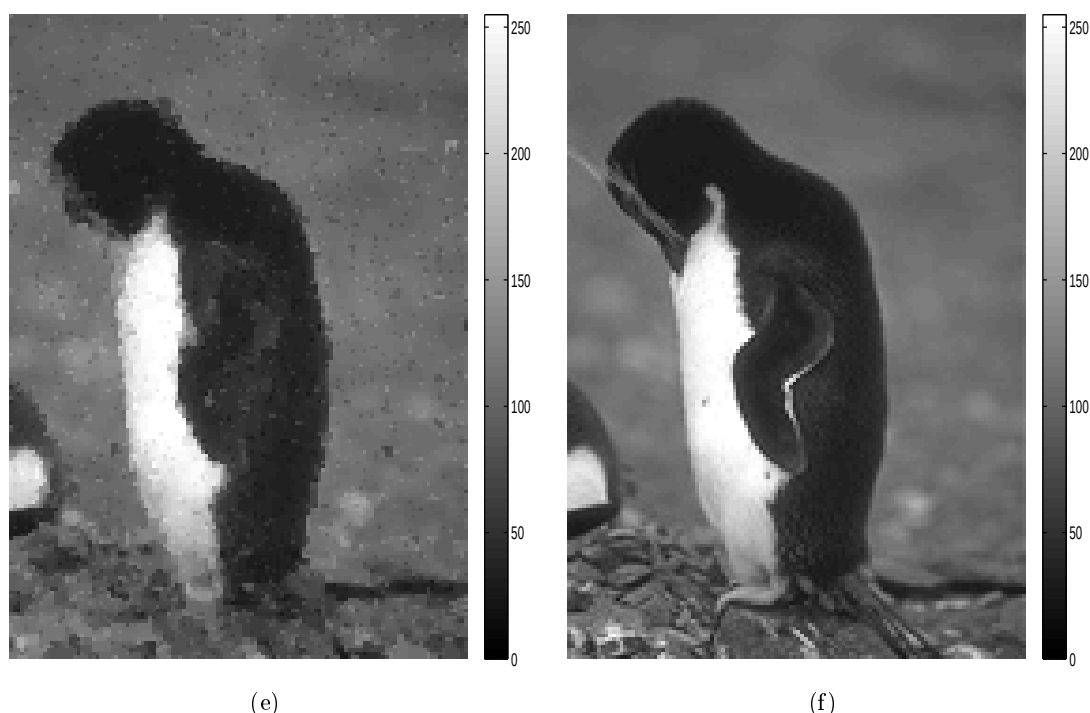


FIGURE 2.9 – Restauration par différents optimiseurs avec un paramètre de régularisation global $\beta = 0.87$ (normalisé dans $[0,1]$). Optima obtenus par (a) ICM, (b) α -extension, (c) $\alpha\beta$ -échange, (d) recuit simulé, (e) la construction d'Ishikawa. (f) Image d'origine.

Les résultats quantitatifs le prouvent aussi dans le tableau 2.1. En effet :

- pour l'ICM, en partant de l'image bruitée qui est considérée proche de la solution et en effectuant un parcours aléatoire des sites à chaque cycle, l'optimum obtenu reste très bruité à cause de l'aspect local de l'optimiseur. L'optimisation est très rapide mais non appropriée à ce type de bruit.
- pour l' α -extension, en partant d'une image uniforme nulle et en effectuant à chaque cycle un parcours aléatoire des étiquettes, l'optimum obtenu est bien régularisé grâce à l'aspect global de l'optimisation mais reste loin de la solution qu'on cherche. Plusieurs stratégies ont été testées : parcours non-aléatoire, partir de l'image bruitée, etc., mais aucune ne parvient à fournir une configuration meilleure.
- pour l' $\alpha\beta$ -échange, en partant d'une image uniforme nulle, l'optimum obtenu est aussi loin d'être satisfaisant.
- pour le recuit simulé, partant de l'image bruitée, l'optimisation s'effectue itérativement et localement en chaque pixel de l'image. La mémoire nécessaire est très réduite, mais un temps de calcul très lourd est nécessaire pour garantir une convergence vers un optimum local, non loin du global (ou le global en théorie, si les paramètres de l'algorithme sont fixés pour une convergence à temps infini).
- pour l'optimiseur exact d'Ishikawa, le minimum exact obtenu correspond à une image bien restaurée. Le temps de calcul est aussi raisonnable par rapport à la taille de l'image considérée. Cependant un espace mémoire gigantesque a été nécessaire pour stocker le graphe.

TABLE 2.1 – Comparaison entre les performances des algorithmes testés.

Algorithme	Nombre d'itérations	Minimum de l'énergie	Temps (mn :sec)	Mémoire (noeuds)
ICM	12	$1.2931e + 06$	01 : 21	1
α -extension	3840	$1.2892e + 06$	01 : 11	$0.2184e + 05$
$\alpha\beta$ -échange	460545	$1.2953e + 06$	50 : 45	$1 \rightarrow 0.2184e + 05$
Recuit simulé	244821	$1.1179e + 06$	60 : 32	1
Ishikawa	1	$1.1776e + 06$	01 : 16	$55.9104e + 05$

TABLE 2.2 – Optima obtenus par des α_m -extensions à différentes valeurs de m .

m	Nombre d'itérations	Minimum de l'énergie	Temps (mn :sec)	Mémoire (noeuds)
8	384	$1.1981e + 06$	3 : 15	$1.7472e + 05$
16	128	$1.1975e + 06$	2 : 59	$3.4944e + 05$
32	48	$1.1930e + 06$	4 : 15	$6.9888e + 05$
64	20	$1.1787e + 06$	5 : 40	$13.9776e + 05$
128	12	$1.1778e + 06$	12 : 45	$27.9552e + 05$

TABLE 2.3 – Optima obtenus par des β_m -échanges à différentes valeurs de m .

m	Nombre d'itérations	Minimum de l'énergie	Temps (mn :sec)	Mémoire (noeuds)
8	12400	$1.3374e + 06$	3 : 34	$1.7472e + 05$
16	12000	$1.3049e + 06$	9 : 19	$3.4944e + 05$
32	6720	$1.2933e + 06$	16 : 58	$6.9888e + 05$
64	5760	$1.2437e + 06$	59 : 43	$13.9776e + 05$
128	3840	$1.1817e + 06$	51 : 14	$27.9552e + 05$

Malgré leur efficacité en temps de calcul (sauf pour l' $\alpha\beta$ -échange qui est nettement le plus lent) et en occupation de mémoire, les optimiseurs approximatifs de l'état de l'art sont peu appropriés à ce type de bruit, dont le modèle de vraisemblance est non-convexe.

En revanche les résultats obtenus par le recuit simulé et Ishikawa sont nettement meilleurs autant visuellement que quantitativement. Cependant, nous retrouvons bien leur problèmes en temps de calcul ou occupation de mémoire.

Ainsi, un compromis entre le temps de calcul, l'occupation de mémoire et la qualité de l'optimum est nécessaire pour ce type d'application. Ce compromis est vérifié par les optimiseurs MPLM comme le montrent les résultats expérimentaux, visuellement dans la figure 2.10 pour des mouvements de type α_m -extension et dans la figure 2.11 pour des mouvements de type β_m -échange ; ainsi que quantitativement dans les tableaux 2.2 et 2.3 qui leur correspondent respectivement.

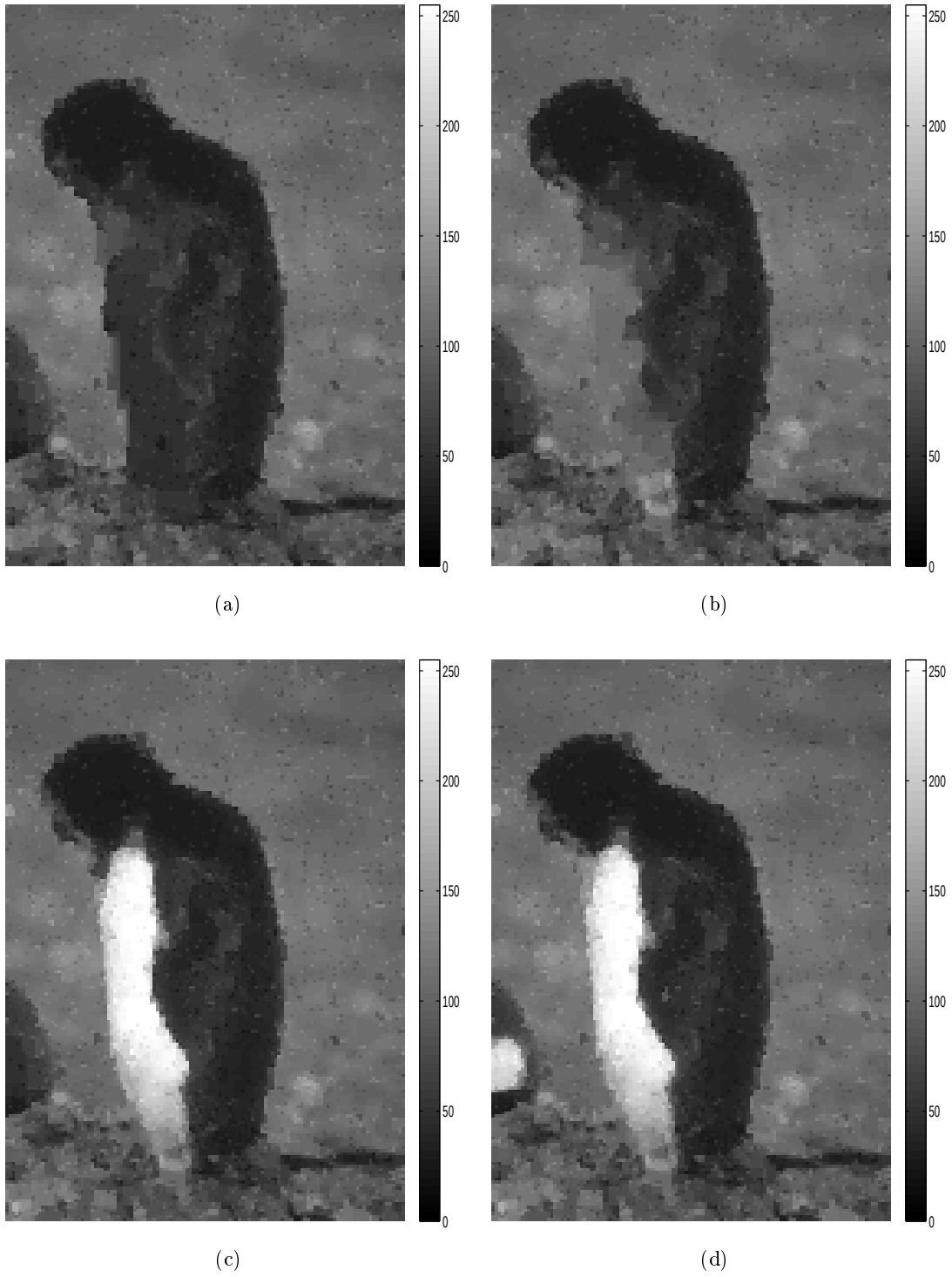


FIGURE 2.10 – Restauration par des α_m -extensions. Optima obtenus pour (a) $m = 16$, (b) $m = 32$, (c) $m = 64$, et (d) $m = 128$.

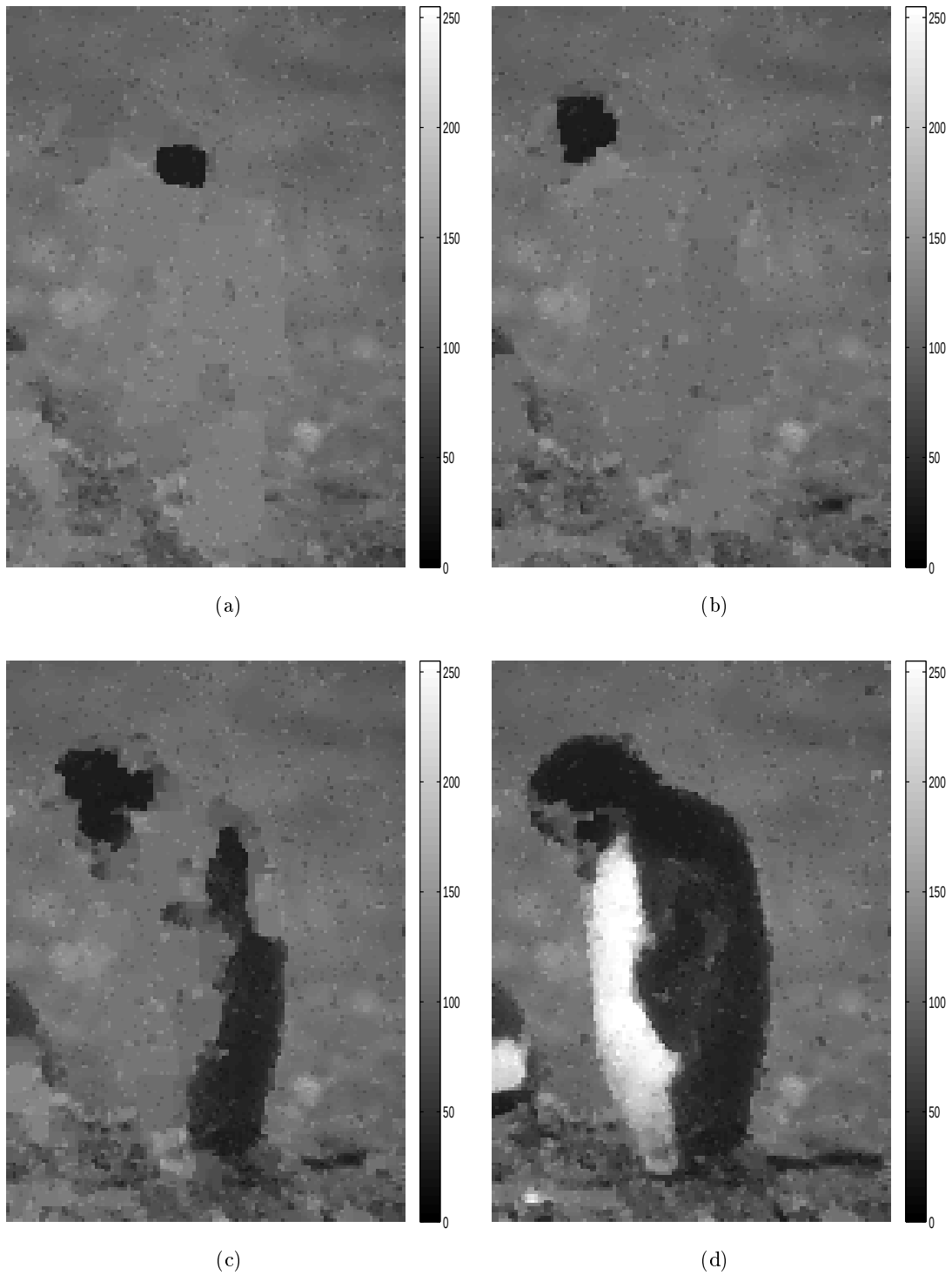


FIGURE 2.11 – Restauration par des β_m -échanges. Optima obtenus pour (a) $m = 16$, (b) $m = 32$, (c) $m = 64$, et (d) $m = 128$.

Les résultats avec les MPLM par extension ou échange montrent bien de meilleures performances par rapport aux algorithmes précédents :

- en terme de minimum de l'énergie, les optima locaux obtenus sont d'énergies plus

faibles et aussi plus proches d'un optimum global. Ces optima sont d'autant meilleurs qu'on augmente m ;

- en terme de qualité de restauration, les images restaurées sont moins bruitées et proches des solutions exactes par rapport au modèle markovien défini ;
- en terme de mémoire, la construction des graphes est nettement moins coûteuse que celle d'un optimiseur exact. La taille du graphe construit est réduite par 4 pour obtenir des résultats proches d'un optimum global ;
- en terme de temps de calcul, comparé à celui d'un optimiseur MPLB, ce temps peut être supérieur puisque les graphes sont plus volumineux et nécessitent donc un temps de calcul de flot-maximum plus important. Comparé à un optimiseur global, le temps de calcul peut être aussi supérieur puisque l'algorithme est itératif, ainsi le calcul de flot-maximum est effectué plusieurs fois. Mais, nous remarquons que pour des valeurs de m choisies suffisantes pour avoir un bon optimum local, le temps de calcul reste raisonnable et du même ordre que ceux des autres algorithmes (MPLB ou exact).

Par ailleurs, nous remarquons que le mouvement d'extension présente un meilleur compromis entre le temps de calcul, le nombre d'étiquettes et la qualité de l'optimum atteint comparé à l'échange. Cela s'explique par le fait que l'extension est un mouvement plus large, il offre donc une meilleure régularisation en tenant compte de l'aspect global de toute l'image.

Pour conclure sur cette première expérience, les optimiseurs approchés disponibles dans la littérature ne sont pas suffisamment robustes face à ce type de bruit. Les optimiseurs exacts sont alors le seul moyen qui existe pour pouvoir bien minimiser de tels modèles d'énergies. En revanche, ces optimiseurs restent limités à des données de tailles moyennes, en nombre de sites ou en nombre d'étiquettes.

Pour la plupart des applications, que nous rencontrons dans la réalité, les données varient énormément et peuvent atteindre de très grandes tailles. De plus, la formation de l'ensemble d'étiquettes nécessite souvent une étape de quantification précise, donnant des ensembles d'étiquettes de grandes tailles, afin de préserver l'information pertinente sur ces données. Les optimiseurs non-exacts par MPLM permettent alors de retrouver de bonnes solutions dans ce cas de figure, dont l'énergie est proche d'un minimum global, sans autant être limités par les problèmes de taille des données.

Exemple 2 : Image de synthèse

Cette seconde expérience illustre les propriétés de convergence proposées dans le paragraphe 3, dans le cas de stratégie de parcours d'étiquettes par inclusions successives.

L'image de synthèse bruitée par un bruit impulsif est présentée dans la figure 2.12(b). La palette des étiquettes considérée est dans l'ensemble $\mathcal{L} = \{0, 1, \dots, 63\}$.

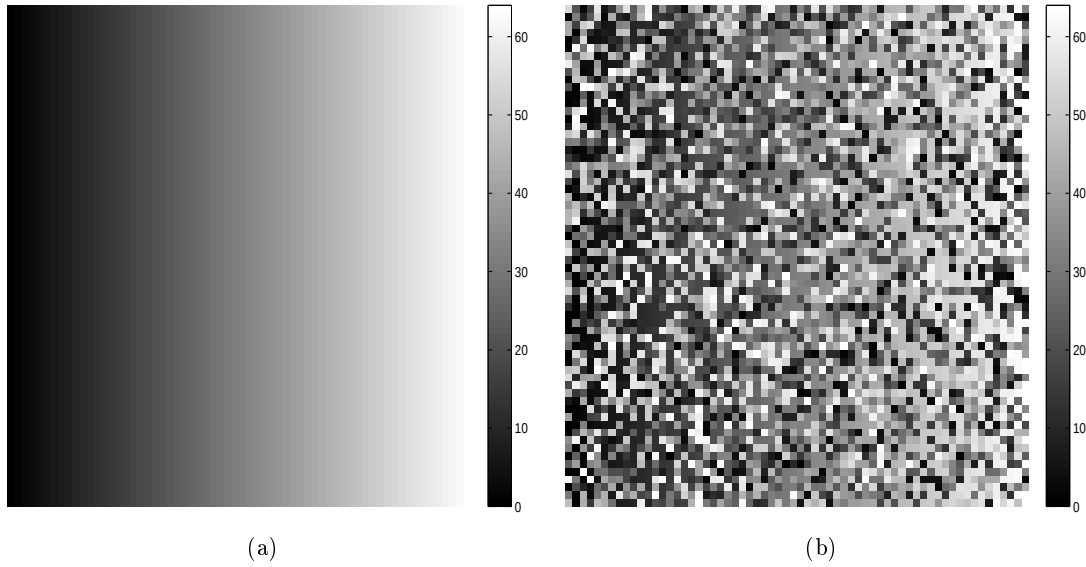


FIGURE 2.12 – Images d'origine et bruitée par un bruit impulsif avec $P = 0.7$. (a) Image d'origine de taille 64×64 pixels, (b) image bruitée.

L'image restaurée qui correspond à l'optimum global obtenu par l'approche d'Ishikawa est présentée dans la figure 2.13.

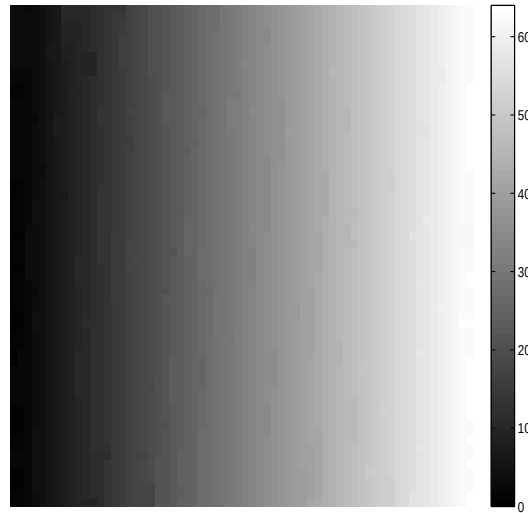


FIGURE 2.13 – Optimum global obtenu par la construction d'Ishikawa (le paramètre de régularisation normalisé est $\beta = 0.97$).

Les minima d'énergies obtenus à la convergence des MPLMs par extension et par échange, pour différentes valeurs de m dans $\{1, 2, \dots, |\mathcal{L}|\}$, sont présentés dans la figure 2.14.

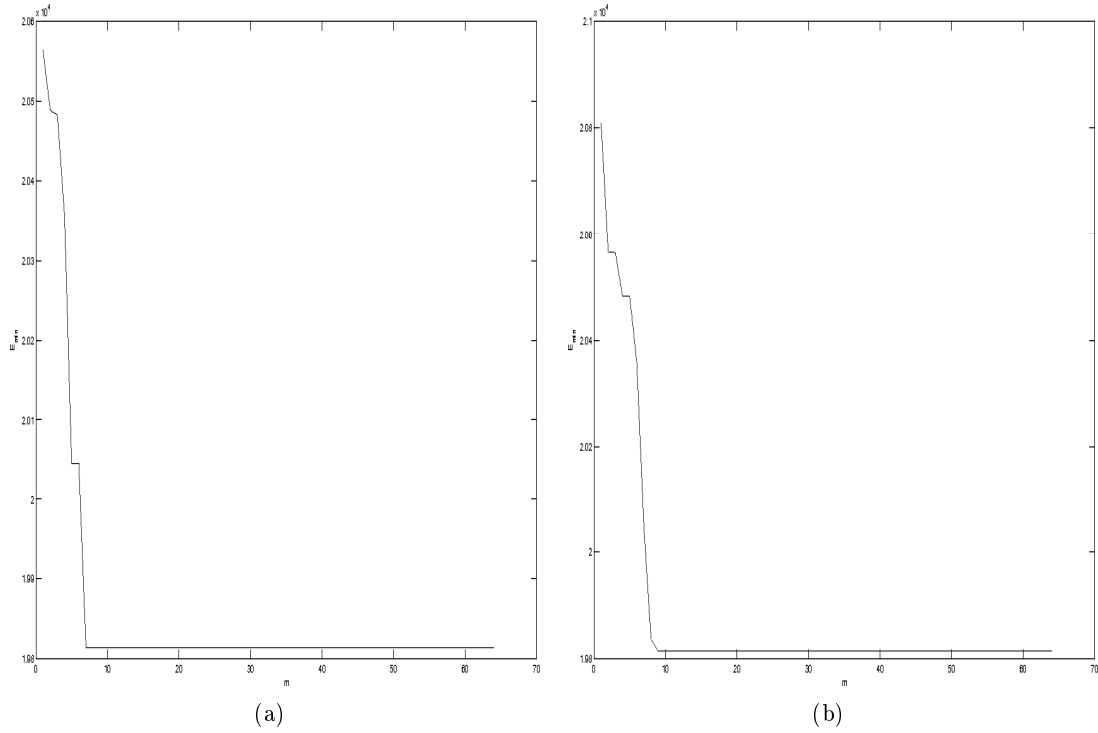


FIGURE 2.14 – Optima obtenus à la convergence des MPLMs par (a) α_m -extensions et (b) β_m -échange, à valeurs croissantes de m .

Nous retrouvons les propriétés de convergence en fonction du paramètre m , dans le cas d'un parcours à inclusions successives. Nous remarquons aussi, que pour $m = 1$, ce qui correspond aux algorithmes MPLB, l'optimum obtenu est loin d'un optimum global. Tandis qu'à partir d'un certain m , proche de 10, l'algorithme MPLM converge vers l'optimum global de l'énergie. Cela confirme encore l'utilité de ces mouvements multi-étiquettes pour mieux minimiser l'énergie markovienne, et même atteindre les optima globaux de la fonction sans recours à des constructions de graphes volumineux.

Dans la figure 2.15, d'autres propriétés des algorithmes MPLM en terme de temps de convergence et de mémoire allouée en fonction du paramètre m , sont présentés.

- En terme de temps de calcul, comme il a été précédemment mentionné, plus m augmente plus le temps de calcul augmente jusqu'à une certaine valeur de m où le temps de convergence décroît puisque le nombre d'itérations nécessaires pour la convergence diminue.
- En terme d'erreur de restauration RMSE, nous remarquons que plus m est grand, plus l'erreur diminue, jusqu'à un certain m , où l'erreur devient fixe (minimale) puisque l'algorithme a convergé vers l'optimum global.
- En terme de mémoire, la taille des graphes augmente linéairement avec m .

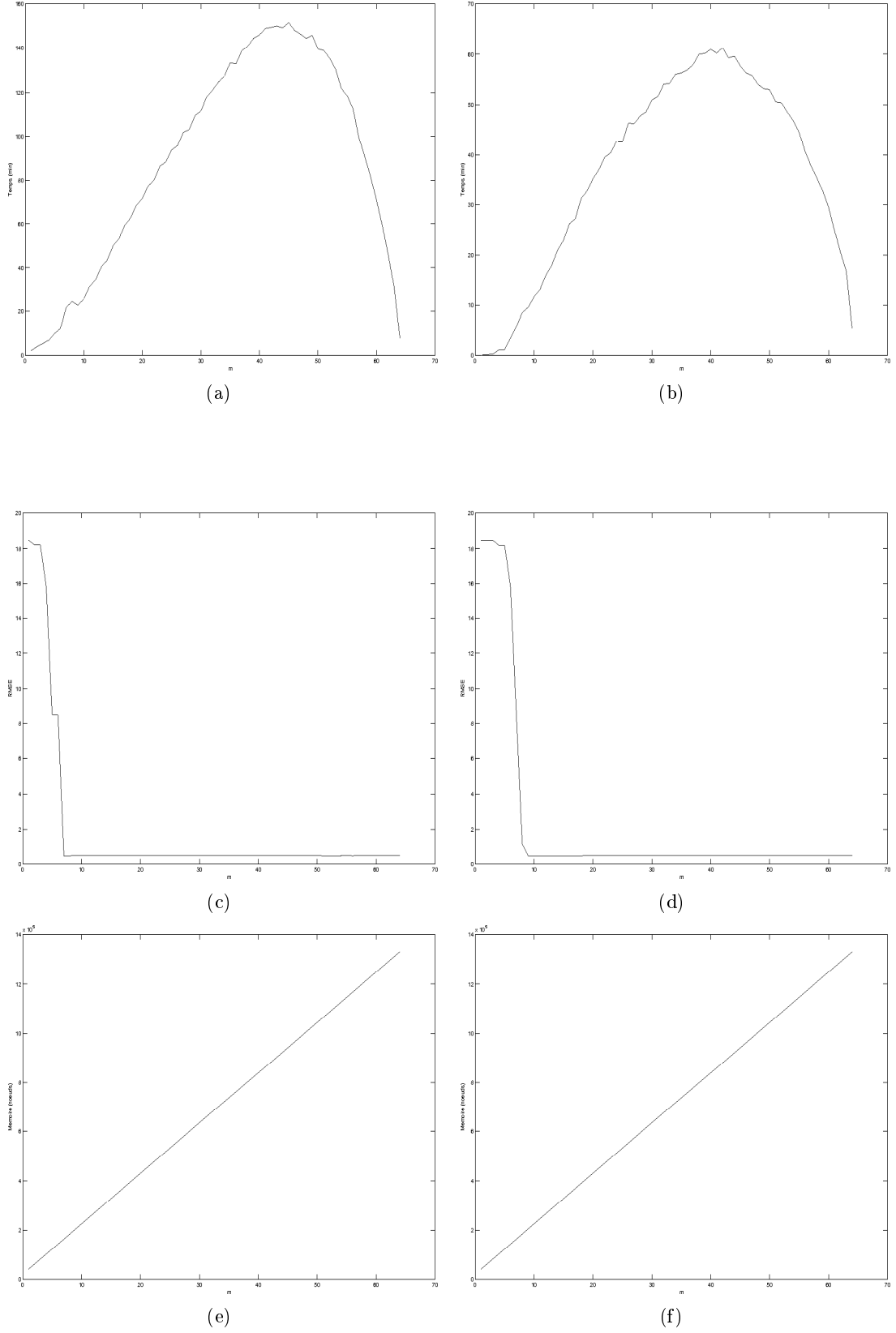


FIGURE 2.15 – Propriétés des MPLM en fonction de m , en terme de temps de convergence avec (a) α_m -extension et (b) β_m -échange ; erreur de restauration avec (c) α_m -extension et (d) β_m -échange ; et mémoire allouée avec (e) α_m -extension et (f) β_m -échange.

Exemple 3 : Image de l'Écureuil

Cette dernière expérience valide empiriquement les propriétés de convergence des algorithmes d'optimisation par MPLM, listées dans la section 3, dans le cas des stratégies de parcours quelconques (aléatoires) proposées dans le paragraphe 4.1.

L'image bruitée par un bruit impulsif est présentée dans la figure 2.16(b). La palette des étiquettes considérée est dans l'ensemble $\mathcal{L} = \{0, 1, \dots, 255\}$.

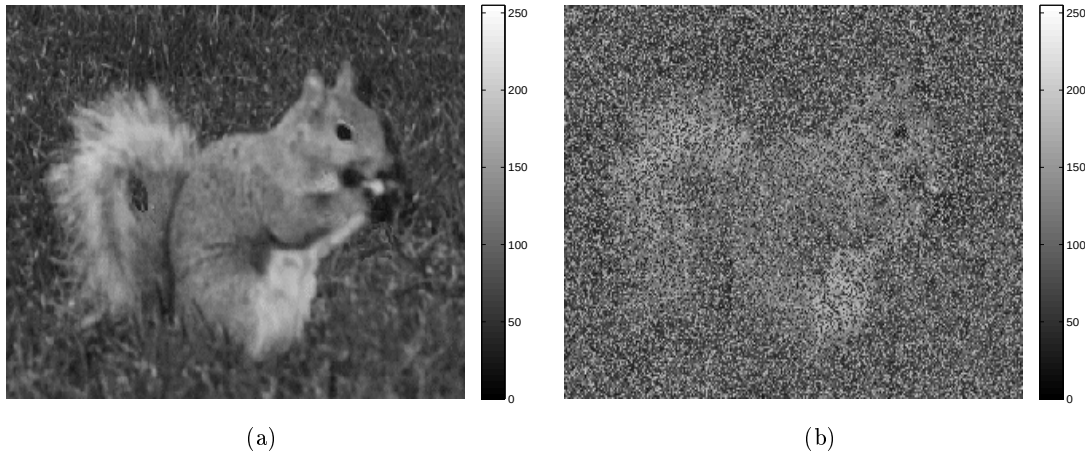


FIGURE 2.16 – Images d'origine et bruitée par un bruit impulsif avec $P = 0.7$. (a) Image d'origine de taille 209×288 pixels, (b) Image bruitée.

La figure 2.17 illustre les résultats de l'optimiseur non-exact α -extension et de l'optimiseur exact d'Ishikawa. Il est clair que le minimum local obtenu par l'algorithme d'optimisation approximative est loin d'un optimum global. L'image correspondante est mal reconstruite. En contre partie, l'optimum global obtenu montre une bonne reconstruction de l'image par le modèle markovien défini.

Nous précisons que pour cette expérience, le parcours est réalisé d'une manière aléatoire. Ainsi, de meilleurs optima sont obtenus par rapport à ce que donnerait un parcours non-aléatoire, au prix d'un temps de calcul plus important. Autrement, l' α -extension par un parcours ordonné donne des résultats pour une telle image bruitée loin d'être satisfaisants.

Dans la figure 2.18, quelques résultats des optimiseurs MPLM par extension sont présentés illustrant l'intérêt d'élargir la taille des sous-ensembles d'étiquettes $\mathcal{L}_m$ afin de mieux minimiser l'énergie. Le tableau 2.4 présente un ensemble de résultats quantitatifs obtenus par différents optimiseurs α_m -extensions, avec m qui varie. Les valeurs confirment bien les propriétés de convergence listées dans le paragraphe 3 malgré un parcours complètement aléatoire, où la condition d'inclusion n'est plus vérifiée. Le tableau est aussi illustré sous la forme graphique dans la figure 2.19 pour montrer le comportement d'un optimiseur MPLM en fonction du paramètre multi-étiquette m .

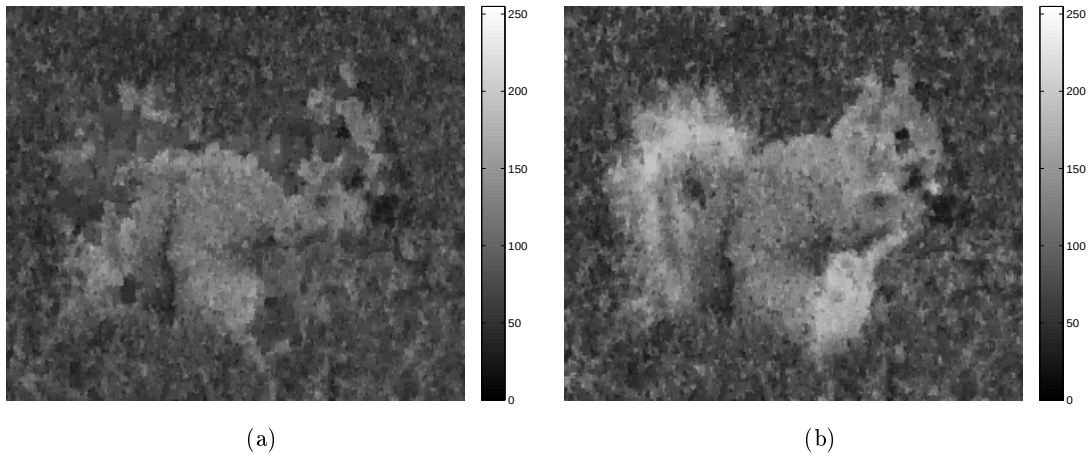


FIGURE 2.17 – Restauration par MPLB et un optimiseur exact. Optima obtenus par (a) α -extension, (b) un algorithme exact.

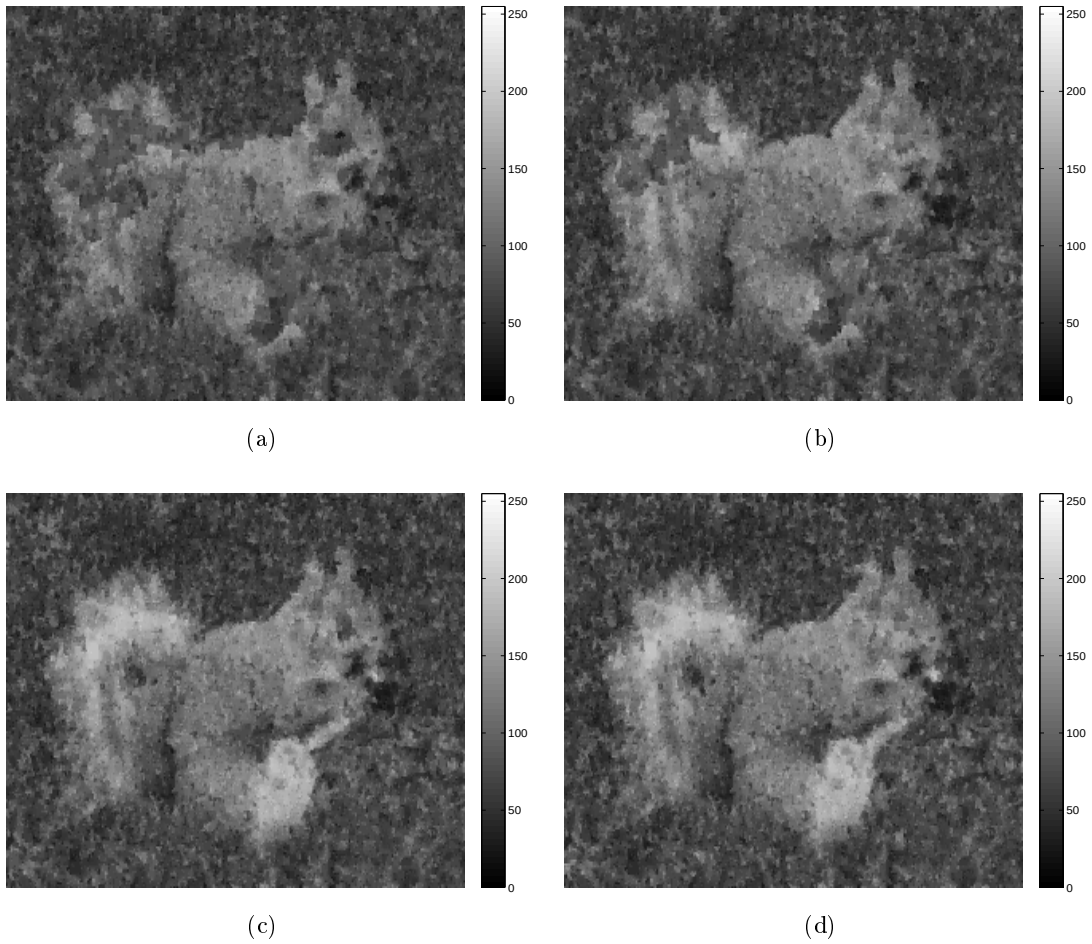


FIGURE 2.18 – Restauration par MPLM. Optima obtenus par (a) α_8 -extension, (b) α_{32} -extension, (c) α_{64} -extension, (d) α_{128} -extension.

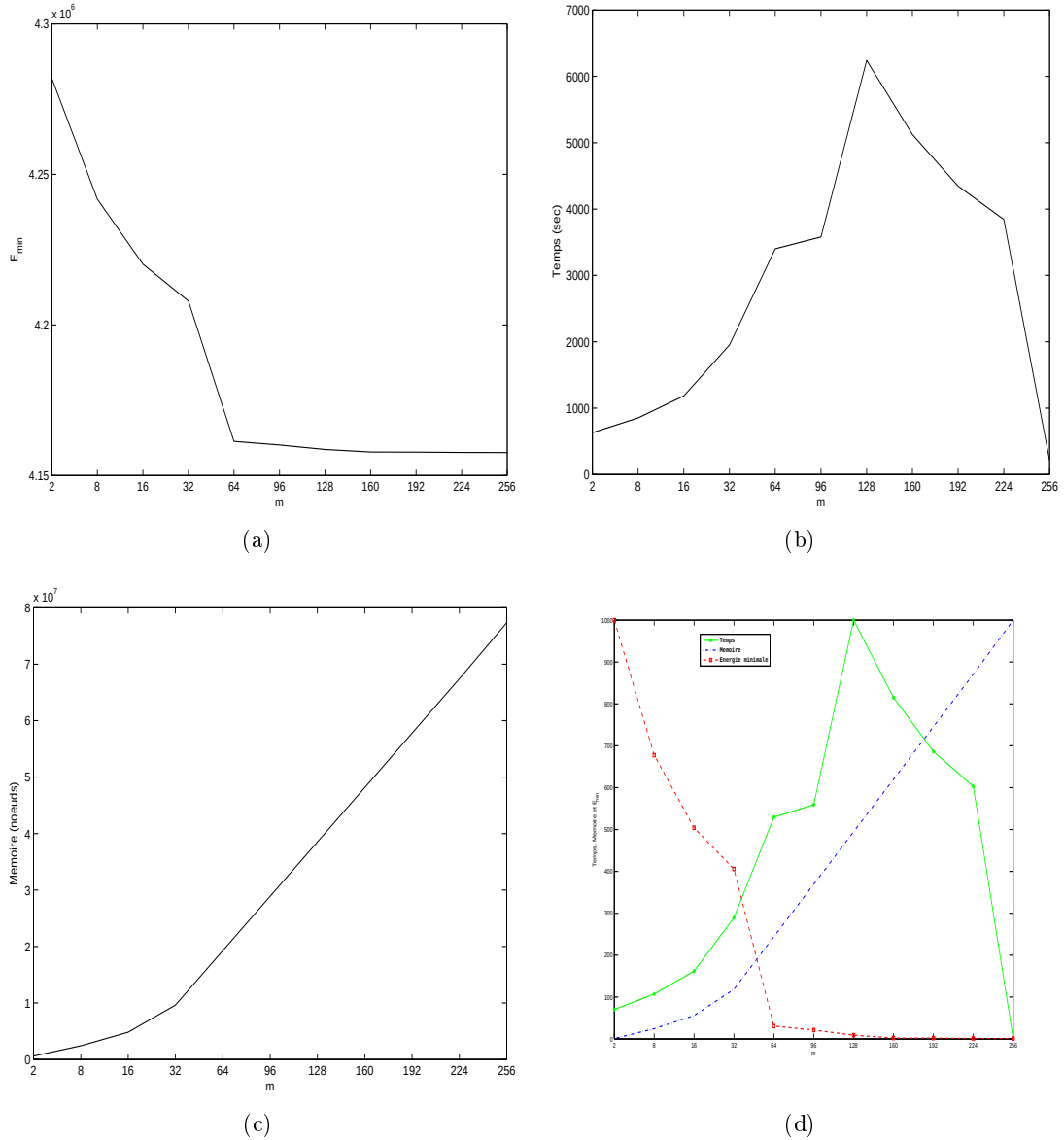


FIGURE 2.19 – Courbes d'évolution des minima, temps de calcul et mémoire occupée par les graphes obtenues par l'optimiseur α_m -extension avec différentes valeurs de m . (a) Courbe des minima de l'énergie, (b) courbe des temps de convergence, (c) courbe des espaces mémoires occupées par les graphes, (d) superposition des trois courbes (a), (b), et (c) avec une normalisation dans le même intervalle de valeurs $[1,1000]$ pour conserver la même grille d'affichage.

Remarque : Les temps de calcul enregistrés dans la courbe 2.19(b) semblent très importants même pour les faibles valeurs de m . Ceci est obtenu en raison du parcours aléatoire des étiquettes. Ainsi plusieurs cycles sont nécessaires pour converger, où l'amélioration apportée par les derniers cycles semble très faible. Un parcours ordonné permet de converger beaucoup plus rapidement dans le cas pratique, et que nous considérons par la suite dans le manuscrit. Nous nous sommes ici contentés de montrer uniquement les propriétés de convergence en fonction du paramètre m dans un cas extrême de parcours, où aucun contrôle sur les étiquettes n'est possible.

TABLE 2.4 – Optima obtenus avec différentes valeurs de m pour le débruitage de l'image de la figure 2.16(b).

m	Nombre d'itérations	Minimum de l'énergie	Temps de convergence (mn :sec)	Taille du graphe en nœuds
2	5120	$4.2819e + 06$	10 : 29	$006.02e + 05$
8	640	$4.2418e + 06$	14 : 12	$024.08e + 05$
16	320	$4.2202e + 06$	19 : 43	$048.16e + 05$
32	160	$4.2079e + 06$	32 : 32	$096.31e + 05$
64	80	$4.1613e + 06$	56 : 40	$192.61e + 05$
96	42	$4.1601e + 06$	59 : 40	$288.96e + 05$
128	40	$4.1586e + 06$	104 : 00	$385.23e + 05$
160	30	$4.1578e + 06$	85 : 23	$481.54e + 05$
192	15	$4.1578e + 06$	72 : 27	$577.84e + 05$
224	9	$4.1577e + 06$	64 : 04	$674.24e + 05$
256	1	$4.1576e + 06$	03 : 31	$773.47e + 05$

5 Conclusion

Pour conclure, nous avons pu relever dans le cadre des problèmes d'optimisation d'énergies markoviennes de premier ordre ayant un *a priori* convexe un défi important, à savoir optimiser efficacement des modèles markoviens avec une attache aux données fortement non-convexe, tout en convergeant vers des optima locaux de bonne qualité. Grâce aux mouvements multi-étiquettes et à une construction de graphes qui leur est appropriée, il est désormais possible de contrôler à la fois la complexité des algorithmes d'optimisation et la qualité de l'optimum obtenu par un paramètre multi-étiquette m . Comparés aux optimiseurs de l'état de l'art, les expériences montrent des résultats nettement meilleurs en terme de compromis qualité de l'optimum et complexité algorithmique.

Nous avons aussi souligné la limite de cette approche en terme de temps de calcul si le paramètre m est choisi assez grand. En effet, compte tenu de la taille du graphe $\mathcal{G}_m$ qui en résulte et de l'aspect itératif de l'algorithme d'optimisation, un temps de convergence important est enregistré pour atteindre l'optimum. Nous nous proposons dans le chapitre suivant de pallier à ce problème par un nouvel algorithme. Ce dernier étendra l'optimisation MPLM en offrant la possibilité de réduire l'ensemble d'étiquettes de départ par une sous-quantification de $\mathcal{L}$ afin de gagner en temps de convergence. Ensuite, grâce à une nouvelle étape d'optimisation continue, la qualité de l'optimum est améliorée. Ce nouvel algorithme montrera son utilité dans le cadre des applications nécessitant une grande précision sur l'ensemble d'étiquettes de la solution recherchée.

Chapitre 3

Minimisation markovienne mixte : discrète et continue

Dans ce chapitre, nous apportons des améliorations aux résultats de l'optimisation par les algorithmes MPLM en terme de temps de calcul et de précision de la solution obtenue. Une nouvelle approche d'optimisation est proposée. L'approche est mixte exploitant d'une part la robustesse des optimiseurs discrets par coupe-minimum sur graphe et d'autre part la précision des optimiseurs continus. Ces derniers convergent vers des solutions à étiquettes quelconques dans $\mathcal{R}$.

L'algorithme proposé apporte aussi un gain en temps de calcul par rapport aux algorithmes MPLM. Le recours à des grandes valeurs de m si besoin est désormais possible sans pénaliser le temps de convergence.

Dans un premier temps, un bref aperçu sur les optimiseurs continus est proposé. Développer un nouvel algorithme dans ce cadre d'optimisation continue ne constitue pas l'objectif de ce chapitre, ni de nos travaux de thèse. Nous nous sommes contentés de rappeler les approches élémentaires utilisées dans la littérature.

Dans un deuxième temps, l'algorithme mixte d'optimisation d'énergies markoviennes de premier ordre est décrit en détail.

Enfin, des résultats sont présentés afin d'illustrer d'une part les étapes de l'algorithme et d'autre part ses propriétés intéressantes. Le domaine d'application exploré dans ce chapitre est la restauration d'image RSO. Pour ce type de données, des optimiseurs robustes et efficaces sont nécessaires pour réduire le bruit de chatoiement sur des données généralement de grandes tailles et à forte dynamique radiométrique. Pour conserver l'information physique apportée par ces images, une bonne précision sur les solutions reconstruites est nécessaire.

Le travail proposé dans ce chapitre a été soumis à la revue *Geoscience and Remote Sensing Letters* (Shabou *et al.*, 2010a), où d'autres expériences ont été présentées dans le cadre du déroulement de la phase interférométrique, qui seront illustrées dans la deuxième partie de ce manuscrit.

1 Limites des approches de minimisation discrète ou continue

1.1 Minimisation discrète

Quoique les approches de minimisation discrète, présentées dans le premier chapitre, permettent d'optimiser des fonctions d'énergie complexes et non-convexes, nous avons montré certaines limites en terme de qualité de l'optimum et de complexité algorithmique.

Les algorithmes proposés dans le second chapitre ont permis d'apporter des améliorations notables, proposant un compromis entre complexité et qualité de l'optimum par le contrôle d'un paramètre de l'algorithme, noté par m , qui définit la taille des sous-ensembles d'étiquettes. Cependant, le recours à des grandes valeurs de m pourrait alourdir le temps de convergence, compte tenu de l'aspect itératif de l'algorithme. En effet, à chaque itération le calcul du flot-maximum est effectué sur un graphe $\mathcal{G}_m$ qui augmente de taille linéairement avec m .

Le recours à ces fortes valeurs de m pourrait se produire quand il s'agit de fonctions markoviennes fortement non-convexes définies sur des données à forte dynamique radiométrique. En effet, la non-convexité nécessitera le recours à des valeurs de m assez grandes pour garantir la convergence vers des optima de bonne qualité. De plus, avec une quantification très précise de l'ensemble d'étiquettes $[l_{min}, l_{max}]$, il serait alors nécessaire d'itérer l'optimisation MPLM plusieurs fois pour pouvoir couvrir toutes les étiquettes possibles. Vu l'efficacité du calcul du flot-maximum, le temps global de convergence est principalement lié au nombre d'itérations nécessaires pour atteindre un optimum. Ainsi réduire ce nombre d'itérations sans affecter la qualité de l'optimum est la meilleure alternative pour s'affranchir de ce temps de calcul pour les importantes valeurs de m .

1.2 Minimisation continue

D'autres approches de minimisation, non encore discutées dans ce manuscrit, existent dans la littérature. Il s'agit des approches d'optimisation continue. Pour les introduire, nous considérons le problème de minimisation sans contraintes dans le domaine continu suivant :

$$\mathbb{P} \begin{cases} \min E(x), \\ x \in \mathbb{R}^n. \end{cases} \quad (3.1)$$

Il a été prouvé que pour toute fonction $E : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ propre (c.-à-d. finie en un point au moins), continue et coercive, ce problème admet au moins une solution. Nous rappelons qu'une fonction $E : \mathbb{R}^n \rightarrow \mathbb{R}$ est dite coercive si $\lim_{\|x\| \rightarrow +\infty} E(x) = +\infty$, où $\|\cdot\|$ désigne la norme euclidienne de l'espace $\mathbb{R}^n$.

Toutefois, l'unicité de la solution n'est pas toujours vérifiée, sauf si la fonction est coercive et strictement convexe.

Les algorithmes de résolution d'un tel problème de minimisation reposent généralement sur la résolution de certaines équations qui en découlent en partant des conditions d'optimalité de la configuration $\tilde{x}$ (solution au problème $\mathbb{P}$).

Dans le cas des fonctions différentiables, la condition nécessaire d'optimalité de premier ordre est l'équation d'Euler définie par :

$$\nabla E(\tilde{x}) = 0. \quad (3.2)$$

Ce point $\tilde{x}$ est alors un point critique, appelé aussi point stationnaire qui pourrait être un optimum local ou global de E , comme aussi un maximum local ou global. Pour cette raison, cette condition d'optimalité est nécessaire mais non suffisante pour obtenir une solution au problème $\mathbb{P}$. Sauf dans le cas convexe, c.-à-d. E est en plus une fonction convexe sur $\mathbb{R}^n$, il s'agit bien d'un minimum global de la fonction.

La condition d'optimalité de second ordre qui permet d'affirmer l'existence d'une solution dans le cas non-convexe est la positivité de la matrice hessienne de E , si elle existe, calculée en un point critique.

Reposant sur ces deux conditions d'optimalité, différents algorithmes ont été proposés pour minimiser une fonction E coercive. Sans présenter de façon exhaustive l'état de l'art

Initialisation de $x^{(0)}$ et de $\rho^{(0)} > 0$;
Répéter
 | $x^{(i+1)} = x^{(i)} + \rho^{(i)} d^{(i)}$;
jusqu'à ce que ($\|x^{(i+1)} - x^{(i)}\| < \varepsilon$) ;
 $\tilde{x} = x^{(i)}$;

FIGURE 3.1 – Algorithme de minimisation par descente du gradient.

de ces approches, nous nous limitons à des algorithmes classiques élémentaires. Ces algorithmes cherchent à déterminer un optimum local de E en supposant sa coercivité ainsi que sa différentiabilité. Nous renvoyons le lecteur à l'ouvrage (Bonnans *et al.*, 1997) pour une étude plus générale, portant aussi sur le cas non-différentiable.

Ces algorithmes sont de nature itérative. En partant d'une initialisation et en fixant un critère d'arrêt, ils effectuent une série d'itérations pour converger vers un optimum de E .

Une des classes de ces algorithmes est la classe des algorithmes de descente. Le schéma général de ces algorithmes est la recherche itérative de la configuration $x^{(i+1)}$, à l'itération $i + 1$, vérifiant :

$$\begin{cases} x^{(0)} \in \mathfrak{R}^n \text{ donnée ,} \\ x^{(i+1)} = x^{(i)} + \rho^{(i)} d^{(i)}, \text{ où } E(x^{(i)} + \rho^{(i)} d^{(i)}) \leq E(x^{(i)}) , \end{cases} \quad (3.3)$$

avec $x^{(0)}$ le point de départ de la descente, appelé initialisation, $x^{(i)}$ est la configuration obtenue à l'itération précédente, $d^{(i)}$ un vecteur non-nul de $\mathfrak{R}^n$, appelé direction de descente et $\rho^{(i)}$ un réel strictement positif appelé pas de descente.

Une idée naturelle pour trouver une direction de descente est de passer par un développement de Taylor à l'ordre 2 de la fonction E , donné par :

$$E(x^{(i)} + \rho^{(i)} d^{(i)}) = E(x^{(i)}) + \rho^{(i)} (\nabla E(x^{(i)}), d^{(i)}) + o(\rho^{(i)} d^{(i)}) , \quad (3.4)$$

où ∇ désigne l'opérateur gradient et $(., .)$ désigne le produit scalaire sur $\mathfrak{R}$. En choisissant comme direction $d^{(i)} = -\nabla E(x^{(i)})$, on vérifie bien la condition : $E(x^{(i)} + \rho^{(i)} d^{(i)}) \leq E(x^{(i)})$. Le choix d'une telle direction de descente nous ramène à la méthode de minimisation par *descente du gradient*. L'algorithme est alors présenté dans la figure 3.1.

On se donne ε un réel positif donné qui représente la précision désirée. Le pas de descente peut être fixé tout le long de l'algorithme comme il peut être variable. Une variation de cette méthode de descente, appelée méthode du gradient à pas optimal, propose un choix judicieux de pas de descente pour accélérer la convergence. A chaque itération, le pas $\rho^{(i)}$ est choisi de façon à minimiser sur $\mathfrak{R}^+$ la fonction ϕ définie par :

$$\phi(\rho) = E(x^{(i)} - \rho \nabla E(x^{(i)})) . \quad (3.5)$$

En pratique, au lieu de chercher le minimum de cette fonction, on effectue à la place une recherche linéaire du pas optimal selon le schéma de l'algorithme 3.2.

Cette méthode d'optimisation, très simple à mettre en œuvre, est malheureusement lente en convergence si les paramètres $\rho^{(i)}$ et ε sont mal choisis.

```

Initialisation de  $\rho = 1$ ,  $\rho_- = \rho_+ = 0$  et  $0 < \beta_1 < \beta_2 < 1$ ;
Tant que (  $(\phi(\rho) > \phi(0) + \beta_1 \rho \phi'(0))$  OU  $(\phi(\rho) < \beta_2 \phi'(0))$  ) faire
    Si  $(\phi(\rho) > \phi(0) + \beta_1 \rho \phi'(0))$  Alors
         $\rho_+ = \rho$ ;
    Fin Si
    Si  $((\phi(\rho) \leq \phi(0) + \beta_1 \rho \phi'(0))$  ET  $(\phi(\rho) < \beta_2 \phi'(0))$  ) Alors
         $\rho_- = \rho$ ;
    Fin Si
    Si  $(\rho_+ == 0)$  Alors
         $\rho = 2 * \rho_-$ ; /*ou bien tout  $\rho > \rho_-^*$  /
    Sinon
         $\rho = \frac{\rho_- + \rho_+}{2}$ ; /*ou bien tout  $\rho \in ]\rho_-, \rho_+[^*$  /
    Fin Si
Fait
retourner  $\rho$ ;

```

FIGURE 3.2 – Algorithme de recherche linéaire du pas optimal.

```

Initialisation de  $x^{(0)}$ ; /*(proche de  $\tilde{x}$ )* /
Répéter
     $x^{(i+1)} = x^{(i)} + H(x^{(i)})^{-1} \nabla E(x^{(i)})$ ;
jusqu'à ce que ( $\|x^{(i+1)} - x^{(i)}\| < \varepsilon$ ;)
 $\tilde{x} = x^{(i)}$ ;

```

FIGURE 3.3 – Méthode de résolution de Newton.

Une autre approche classique utilisée pour minimiser E est la méthode de Newton. Cette méthode permet de chercher les points critiques de la fonction à partir de la résolution du système d'équations non-linéaires dans $\mathbb{R}^n$: $\nabla E(x) = 0$. En notant par $F = \nabla E$, une fonction de $\mathbb{R}^n$ dans $\mathbb{R}^n$ et par H la dérivée de F , c.-à-d. la matrice hessienne de E , la méthode de résolution de Newton s'effectue selon l'algorithme de la figure 3.3. Toutefois, nous n'obtenons que les points critiques de la fonction E . Une étape de vérification des minima est alors nécessaire en étudiant le signe de la hessienne aux points critiques retrouvés.

L'étape de calcul de H est la plus coûteuse dans cette méthode. Une variante possible pour remédier à cela est la méthode quasi-newton qui permet d'approximer H grâce à des formules algébriques. Cette approche est très efficace. La convergence est quadratique et donc beaucoup plus rapide que la descente du gradient. Néanmoins, elle est très sensible à l'initialisation. En pratique, une descente du gradient est au départ effectuée pour s'approcher d'une solution optimale, ensuite la méthode de Newton est appliquée pour

```

Initialisation de  $x^{(0)}$  ;
Répéter
    Pour  $i$  varie de 1 à  $n$  faire
         $\arg \min_{x_i} E(x_1^{(k+1)}, x_2^{(k+1)}, \dots, x_{i-1}^{(k+1)}, x_i, x_{i+1}^{(k)}, \dots, x_n^{(k)})$  ;
    Fin Pour
jusqu'à ce que  $(\|x^{(k+1)} - x^{(k)}\| < \varepsilon ;)$ 
 $\tilde{x} = x^{(k)}$  ;

```

FIGURE 3.4 – Méthode de relaxation.

converger rapidement.

Une autre famille d'approches d'optimisation continue est celles par relaxation. Le problème de minimisation dans $\mathfrak{R}^n$ se ramène à la résolution successive de n problèmes de minimisation dans $\mathfrak{R}$. Ainsi à chaque cycle (ensemble d'itérations), étant donnée une configuration courante de x , toutes les composantes sont fixées sauf une sur laquelle la minimisation sur $\mathfrak{R}$ est effectuée (par exemple par la méthode du gradient). On recommence ensuite pour minimiser sur toutes les autres composantes, itérativement. Ensuite, le cycle reprend avec la nouvelle configuration obtenue jusqu'à vérifier une condition d'arrêt. La figure 3.4 illustre cet algorithme.

Bien d'autres approches d'optimisation en continu existent, qui soit étendent les approches que nous venons de décrire, soit proposent d'autres schémas de minimisation pour explorer plus d'optima locaux et éviter de stagner dans un optimum local. En effet, les méthodes décrites jusqu'à maintenant sont des méthodes de recherche locale d'un minimum de E . Dans le cas de fonctions non-convexes, une initialisation proche de la solution désirée est nécessaire, autrement l'algorithme peut diverger ou stagner dans un optimum loin de ce que nous cherchons.

Dans ces travaux de thèse, comme nous nous intéressons essentiellement à une classe de fonctions non-convexes sur $\mathfrak{R}^n$ (les énergies markoviennes de premier ordre ayant un *a priori* convexe et une vraisemblance quelconque), les approches de minimisation continue que nous venons de décrire ne permettent pas de bien la minimiser. En revanche, en proposant une première étape d'optimisation discrète approximative permettant de s'approcher d'un bon optimum de la fonction E , et en considérant l'optimum obtenu comme étant un point de départ d'un algorithme de minimisation continue, l'optimum final ainsi obtenu serait de bonne qualité. Cette optimisation mixte semble être intéressante pour les deux raisons suivantes :

- accélérer la convergence localement, proche de l'optimum obtenu par minimisation discrète ;
- obtenir une solution dans le domaine continu, c.-à-d. avec des étiquettes dans $\mathfrak{R}^n$ et non plus dans $\mathcal{L}^n$.

Dans la section suivante, l'algorithme exploitant cette approche mixte discrète et continue est présenté.

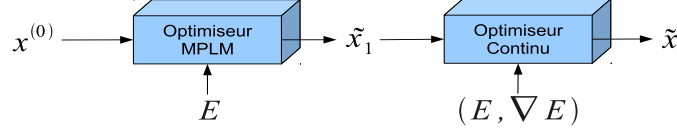


FIGURE 3.5 – Illustration des deux phases de l’algorithme d’optimisation.

2 Algorithme de minimisation mixte : discrète et continue

L’algorithme proposé s’effectue en deux phases comme l’indique la figure 3.5.

2.1 Phase 1 : Minimisation discrète

La première phase de l’algorithme est une optimisation discrète reposant sur l’approche d’optimisation multi-étiquette proposée dans le chapitre précédent.

Afin de pallier au problème du temps de calcul pour des valeurs importantes de m , l’ensemble des étiquettes $\mathcal{L}$ est obtenu à partir de l’ensemble d’étiquettes continu $[l_{min}, l_{max}]$ avec une quantification à pas grossier. Ainsi $\mathcal{L}$ est de taille réduite. Cette sous-quantification réduit le nombre d’itérations pour parcourir toutes les étiquettes possibles. Comme indiqué précédemment, réduire le nombre d’itérations semble être une bonne alternative pour accélérer la convergence d’un optimiseur MPLM.

La configuration $\tilde{x}_1$ obtenue par MPLM, même si elle est optimale par rapport à cette sous-quantification, peut être loin d’un optimum global de E et aussi de très faible précision.

2.2 Phase 2 : Minimisation continue

Partant de la configuration $\tilde{x}_1$, une minimisation locale continue est réalisée pour obtenir l’optimum local le plus proche, qui correspond à une configuration $\tilde{x}$ à étiquettes dans l’intervalle continu $[l_{min}, l_{max}]$. Cette nouvelle configuration est de meilleure précision, comparée à $\tilde{x}_1$.

L’algorithme le plus simple à mettre en place dans ce cadre est la descente du gradient. Ce dernier nécessite le calcul du gradient de E pour retrouver itérativement la direction de la descente. Le calcul du gradient requiert la différentiabilité des termes de E . Dans les différentes applications que nous traitons dans ce manuscrit de thèse, les modèles de vraisemblance sont tous des fonctions différentiables. Le modèle *a priori* toujours utilisé est la variation totale. Il présente néanmoins des problèmes de dérivabilité qu’il est nécessaire d’enlever pour garantir la convergence du minimiseur continu et éviter les problèmes numériques.

Initialement introduite dans (Rudin *et al.*, 1992) pour le débruitage des images, la variation totale (TV) est désormais un des modèles de régularisation largement utilisés en traitement d’image, permettant de régulariser l’image tout en préservant les discontinuités. Il s’agit d’un *a priori* bien adapté aux images de type constantes par morceaux. Le passage au modèle discret de l’énergie pour régulariser les images numériques nécessite une étape de discrétisation des différents termes. L’expression continue est donnée par :

$$TV(x) = \int_{\mathbb{R}^2} |\nabla x|, \quad (3.6)$$

où x dénote l'image à régulariser, définie sur $\mathbb{R}^2$.

La discrétisation de cette forme s'est avérée importante pour le déroulement de l'algorithme numérique de minimisation de E . Ainsi un choix judicieux de la forme discrète de TV est nécessaire.

Les formes classiques de discrétisation de TV reposent sur l'approximation du gradient par éléments finis. Nous trouvons à titre d'exemple le modèle discret de TV avec la norme l_1 et un voisinage en 4-connexité donnant l'expression :

$$\begin{cases} TV_d(x) = \sum_{0 \leq k, l < N} |\nabla x|_d(k, l), \\ |\nabla x|_d(k, l) = |x_{k+1, l} - x_{k, l}| + |x_{k, l+1} - x_{k, l}|, \end{cases} \quad (3.7)$$

avec $N \times N$ la taille de l'image numérique x et (k, l) les coordonnées d'un site quelconque de l'image.

Avec la norme l_2 et un voisinage en 4-connexité, nous trouvons le modèle discret suivant :

$$\begin{cases} TV_d(x) = \sum_{0 \leq k, l < N} |\nabla x|_d(k, l), \\ |\nabla x|_d(k, l) = \sqrt{(x_{k+1, l} - x_{k, l})^2 + (x_{k, l+1} - x_{k, l})^2}, \end{cases} \quad (3.8)$$

qui a l'avantage d'être euclidien, condition nécessaire pour l'algorithme rapide de minimisation proposé dans (Chambolle, 2004).

D'autres formes de discrétisation de TV existent. Nous citons à titre d'exemple la forme spectrale proposée dans (Moisan, 2007) permettant de discrétiser TV tout en respectant le théorème d'échantillonnage de Shannon pour éviter le problème d'aliasing sur la solution finale.

Puisque les algorithmes de minimisation continue nécessitent le calcul du gradient de l'énergie, on se retrouve face au problème de non différentiabilité des formes discrètes de la TV. Pour s'en affranchir, des approximations numériques du modèle TV discret sont nécessaires. Une première approximation de (Vogel et Oman, 1996) est la suivante :

$$\begin{cases} TV_\eta(x) = \sum_{0 \leq k, l < N} |\nabla x|_\eta(k, l), \\ |\nabla x|_\eta(k, l) = \sqrt{|\nabla x|_\eta(k, l)^2 + \eta^2}, \end{cases} \quad (3.9)$$

avec η une constante positive de précision, fixée très faible pour ne pas biaiser la valeur de TV sur x . La dérivation de ce modèle est alors possible, donnée par :

$$(\nabla TV_\eta(x))_{(p, q)} = -div \left(\frac{(\nabla x)_{(p, q)}}{\sqrt{|\nabla x|_{(p, q)}^2 + \eta^2}} \right), \forall (p, q) \in \mathcal{P}^2, \quad (3.10)$$

avec div l'opérateur divergence défini comme l'adjoint du gradient ∇ .

Une deuxième approximation de (Huber, 1973) est la fonction de Huber, une forme lisse de TV discrète qui est la suivante :

$$\begin{cases} TV_\mu(x) = \sum_{0 \leq k, l < N} \psi_\mu(|\nabla x|_\mu(k, l)), \\ \psi_\mu(x) = \begin{cases} |x| & \text{si } |x| \geq \mu, \\ \frac{1}{2\mu}(x^2 + \mu^2) & \text{sinon,} \end{cases} \end{cases} \quad (3.11)$$

avec μ une constante positive et faible, lissant la partie non-différentiable de TV. La dérivation de ce modèle est donnée alors par :

$$\begin{cases} (\nabla TV_\mu(x))_{(p,q)} = -div(\Psi_{(p,q)}) , \\ \Psi_{(p,q)} = \begin{cases} \frac{(\nabla x)_{(p,q)}}{|(\nabla x)_{(p,q)}|} & \text{si } |(\nabla x)_{(p,q)}| \geq \mu , \\ \frac{(\nabla x)_{(p,q)}}{\mu} & \text{sinon} . \end{cases} \end{cases} \quad (3.12)$$

Nous choisirons dans les expériences la première approximation pour conserver au mieux la forme discrète de la TV utilisée à l'étape d'optimisation discrète MPLM.

Étant donnée donc la fonction d'énergie markovienne à minimiser, ainsi que son gradient bien défini, il est alors possible d'utiliser l'approche d'optimisation mixte proposée pour minimiser la fonction. L'algorithme proposé sera appelé *FastDC*.

Dans la suite, des expériences sont présentées pour valider les performances et l'intérêt de cet algorithme.

3 Expériences et résultats

Deux expériences sont réalisées dans cette partie. La première illustre le fonctionnement de l'algorithme sur un cas de fonctionnelle simple, à deux dimensions. La deuxième valide l'algorithme FastDC dans un cas d'application : la restauration d'image RSO.

3.1 Illustration sur une fonction à deux variables

La fonction à deux variables utilisée est définie sur $[0, 10\pi[$ par :

$$E(x, y) = - \left(\cos(x - 5) + \cos\left(\frac{2}{5}x - 5\right) \right) - \left(\cos(x - 2) + \cos\left(\frac{2}{5}x - 2\right) \right) . \quad (3.13)$$

Cette fonction est périodique sur $\mathbb{R}$, de période égale à 10π , différentiable et présentant différents optima locaux au sein d'une même période. Ceci rend la tâche de minimisation difficile pour un algorithme local tel que la descente du gradient. Dans la figure 3.6, le graphe de la fonction E est illustré.

La figure 3.7 montre l'étape de minimisation discrète de cette fonction, en quantifiant l'intervalle de valeurs $[0, 10\pi[$ sur 80 éléments. La minimisation est effectuée par un mouvement d'extension. On remarque la convergence de l'algorithme vers l'optimum global de la fonction, à cette quantification près. Cependant, il ne s'agit pas d'un vrai optimum global de la fonction E définie sur $\mathbb{R}$.

La figure 3.8 illustre l'étape de minimisation continue de cette fonction, partant de l'optimum obtenu à l'étape précédente et en effectuant une descente du gradient. On remarque la convergence directe vers l'optimum global de la fonction E sur l'intervalle $[0, 10\pi[$.

Dans la figure 3.9, la décroissance de l'énergie aux étapes 1 et 2 est présentée montrant la convergence de l'algorithme mixte vers l'optimum global.

Ainsi étant donnée une fonction markovienne non-convexe, avec un *a priori* convexe, il est alors intéressant de l'optimiser par cette approche mixte pour s'approcher d'un bon optimum (idéalement le global) reposant sur l'optimisation MPLM sur un ensemble d'étiquettes quantifié, et ensuite de converger vers cet optimum par une descente du gradient. Les deux étapes sont efficaces en temps de calcul et en occupation de mémoire. L'optimum obtenu est de bonne précision comparé à ce que donnent des approches d'optimisation discrète. L'expérience suivante illustre ces propriétés dans un cas d'application réel.

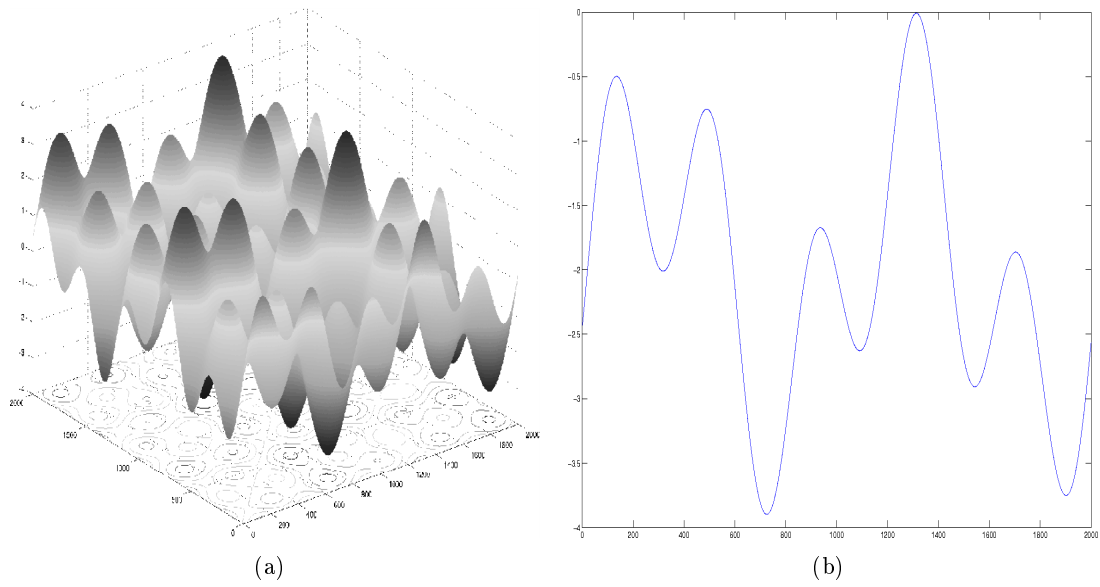


FIGURE 3.6 – Illustration (a) 3D et (b) d'une coupe 2D du graphe de la fonction E .

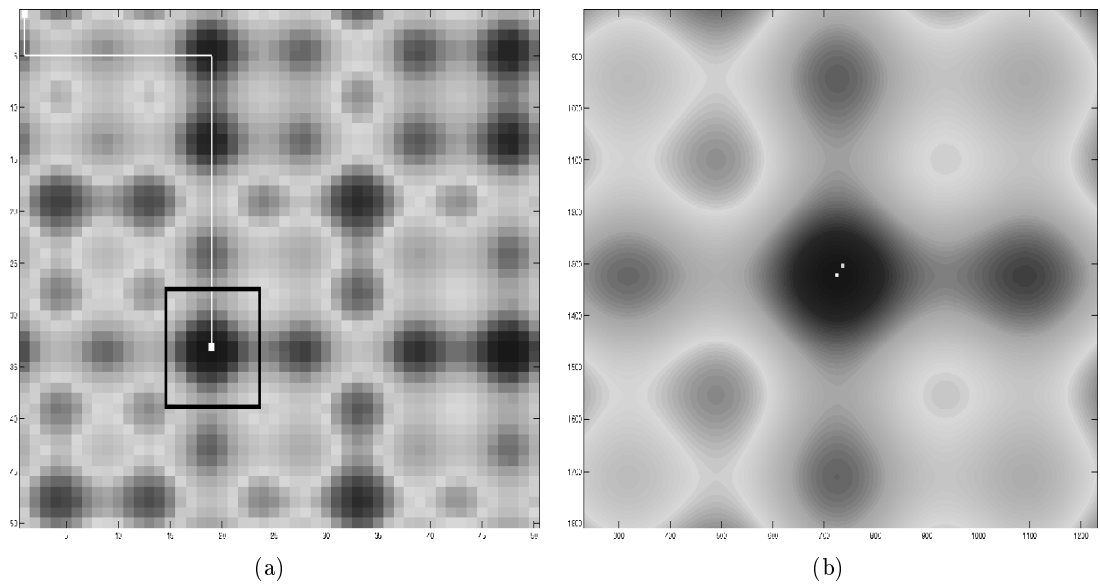


FIGURE 3.7 – Étape de minimisation discrète. (a) Convergence de l'algorithme de minimisation par extension, (b) optimum global obtenu versus l'optimum global de E (zoom à la fenêtre autour de l'optimum obtenu dans la figure (a)).

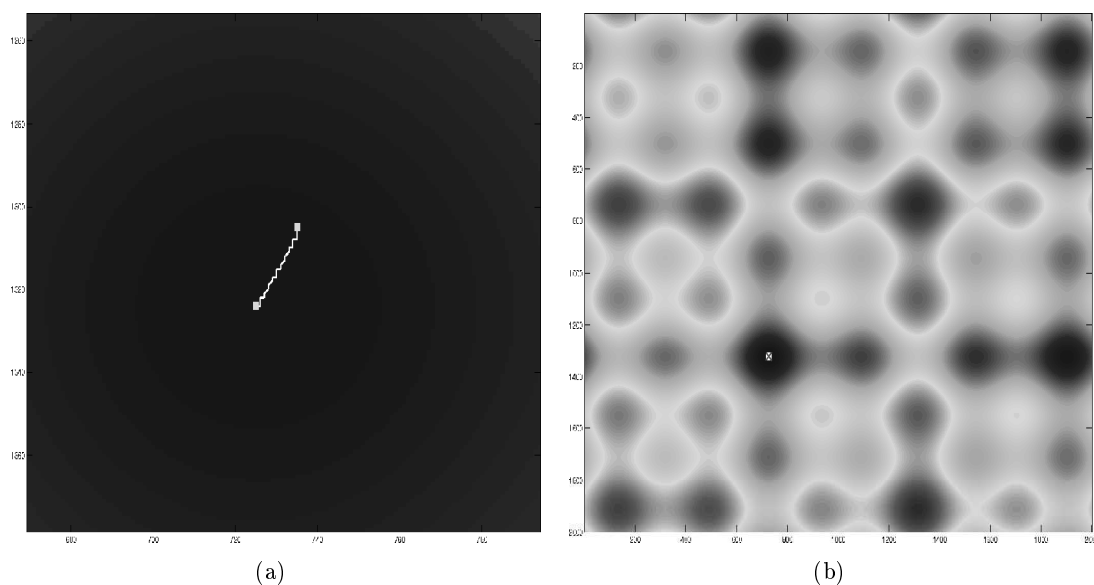


FIGURE 3.8 – Étape de minimisation continue. (a) Convergence de l'algorithme de minimisation par descente du gradient (zoom de la descente autour du précédent optimum), (b) optimum global obtenu vs l'optimum global de E (superposés).

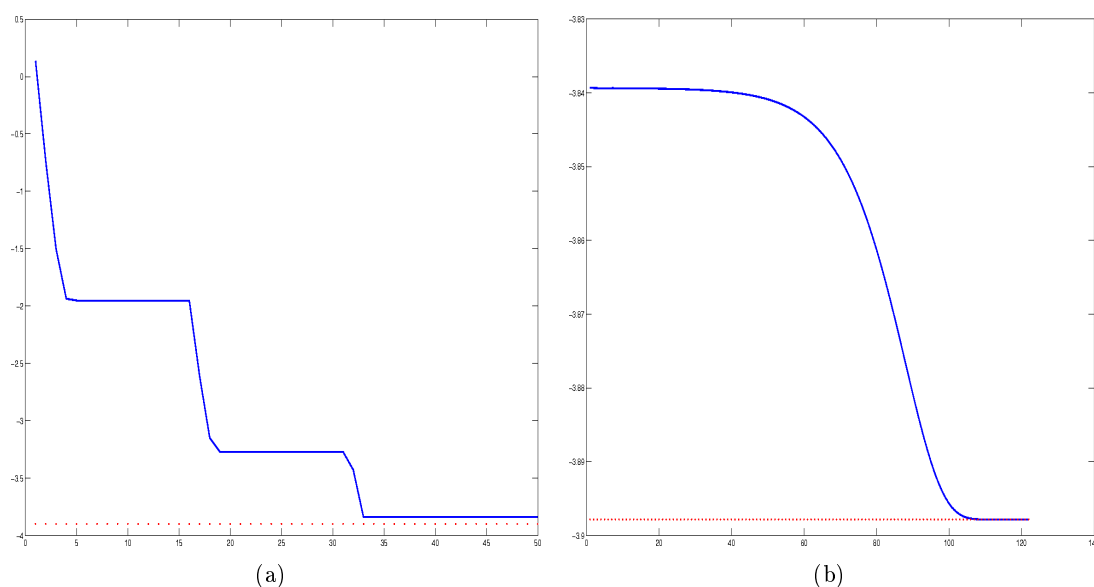


FIGURE 3.9 – Décroissance de l'énergie à (a) l'étape de minimisation discrète, (b) et à l'étape de minimisation continue. Le trait en points indique la valeur du minimum global de E .

3.2 Restauration d'image radar à synthèse d'ouverture

Les images RSO sont caractérisées par un bruit multiplicatif, souvent modélisé par du chatolement pleinement développé, rendant l'interprétation directe de la scène imagée difficile. Une image monovue RSO est une image de complexes dont le bruit sur les données en amplitude est modélisé statistiquement par la loi Nakagami (voir annexe E).

A partir de cette loi, les termes d'attache aux données du modèle markovien multi-étiquette peuvent s'écrire sous la forme :

$$E_p(y_p|x_p) = \frac{y_p^2}{x_p^2} + 2\log(x_p), \forall p \in \mathcal{P}. \quad (3.14)$$

Dans la figure 3.10, la courbe de la fonction de vraisemblance E_p est présentée en fonction de $(x_p$ et $y_p)$ et x_p toute seule (en fixant y_p) dans le but de montrer la non-convexité de cette fonction.

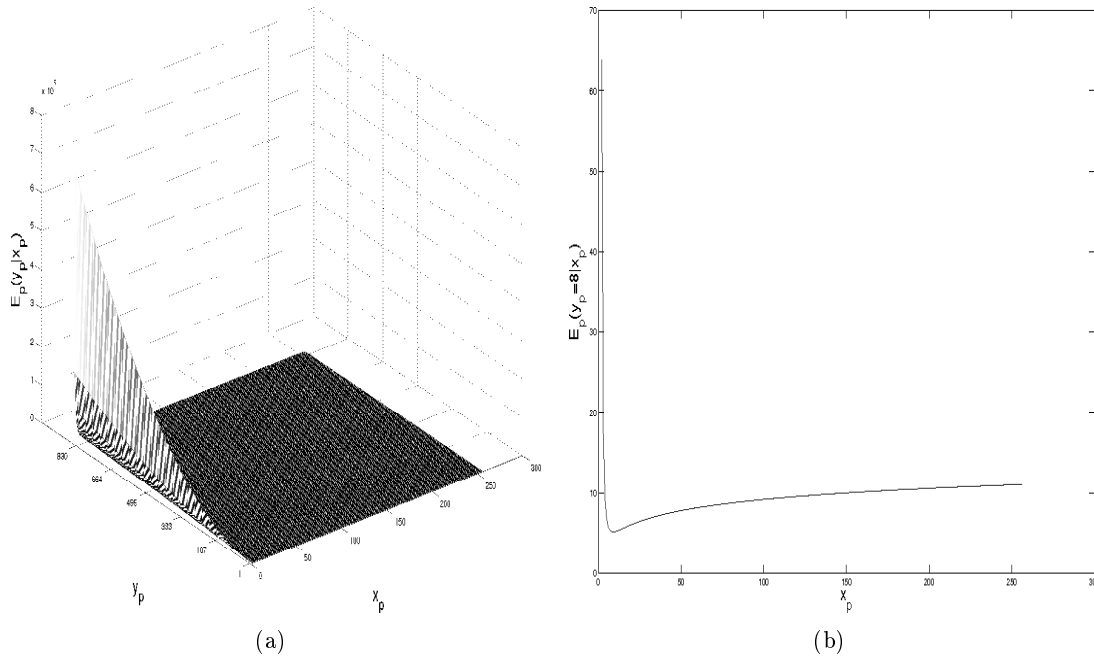


FIGURE 3.10 – Courbe de la fonction de vraisemblance associée au modèle de bruit : chatolement pleinement développé. (a) Courbe de E_p en fonction de l'estimation $x_p \in \{0, 1, \dots, 255\}$ et l'observation $y_p \in \{0, 1, \dots, 847\}$, (b) courbe de E_p en fonction de x_p , fixant y_p .

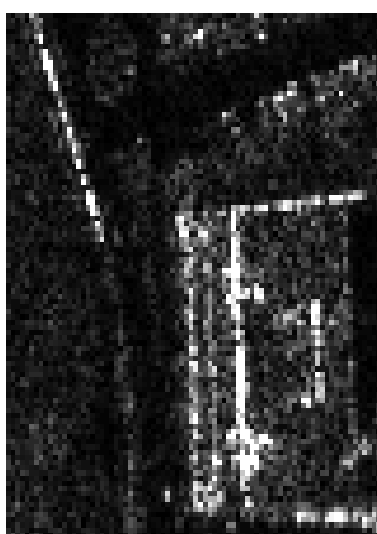
Il est à noter que sur les images RSO, la dynamique des niveaux de gris est très grande compte tenu du bruit multiplicatif. D'où la grande différence entre les ensembles d'étiquettes de x_p et y_p . Par ailleurs, considérer l'ensemble des étiquettes $\mathcal{L} = \{0, 1, \dots, 255\}$ (c.-à-d. l'ensemble discret de valeurs entre 0 et 255 avec un pas de quantification 1) semble peu approprié à la restauration de ce type d'images. Une telle quantification brute induira une perte de l'information physique apportée par ces images. Pour ne pas restreindre les étiquettes à rester dans cet ensemble, nous utilisons les mouvements d'optimisation par saut. Ces derniers exploitent des pas de quantification différents d'une itération à une autre, ce qui laisse plus de liberté de choix des étiquettes ayant des valeurs non forcément entières dans $[l_{min}, l_{max}]$.

La figure 3.11 illustre une scène de Toulouse imagée par un capteur RSO, avec deux extraits de zones différentes : bâtiment (avec des discontinuités) et croisement de routes avec un champ autour.

Les restaurations de l'image entière par les deux optimiseurs : MPLM par saut et FastDC, sont illustrées dans la figure 3.12.



(a)

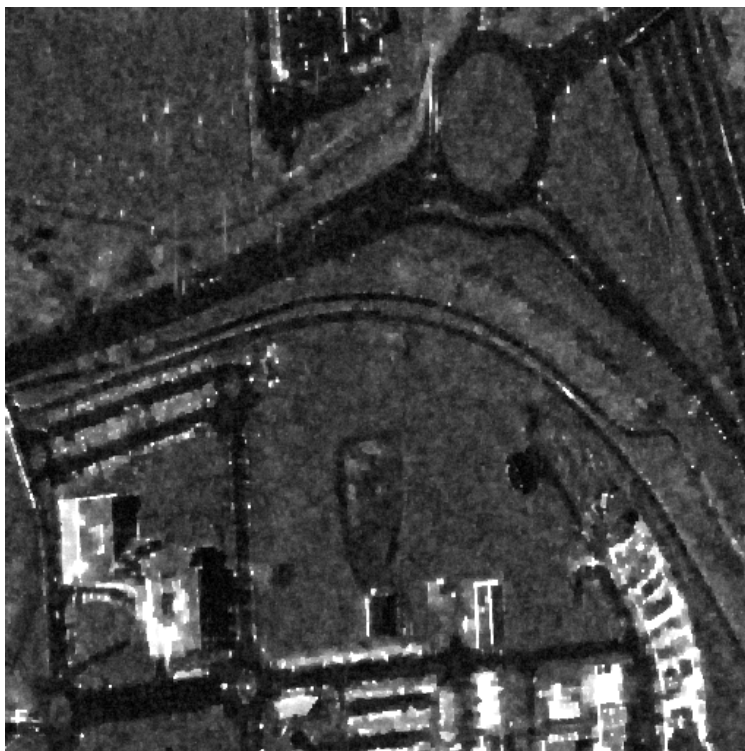


(b)



(c)

FIGURE 3.11 – Image RSO d'origine sur la zone de Toulouse de taille (512×512) . (a) Image complète à restaurer, (b) première zone d'intérêt, (c) deuxième zone d'intérêt.



(a)



(b)

FIGURE 3.12 – Résultats de la restauration par : (a) δ_{16} -saut, (b) FastDC(δ_4 -saut + Descente du gradient).

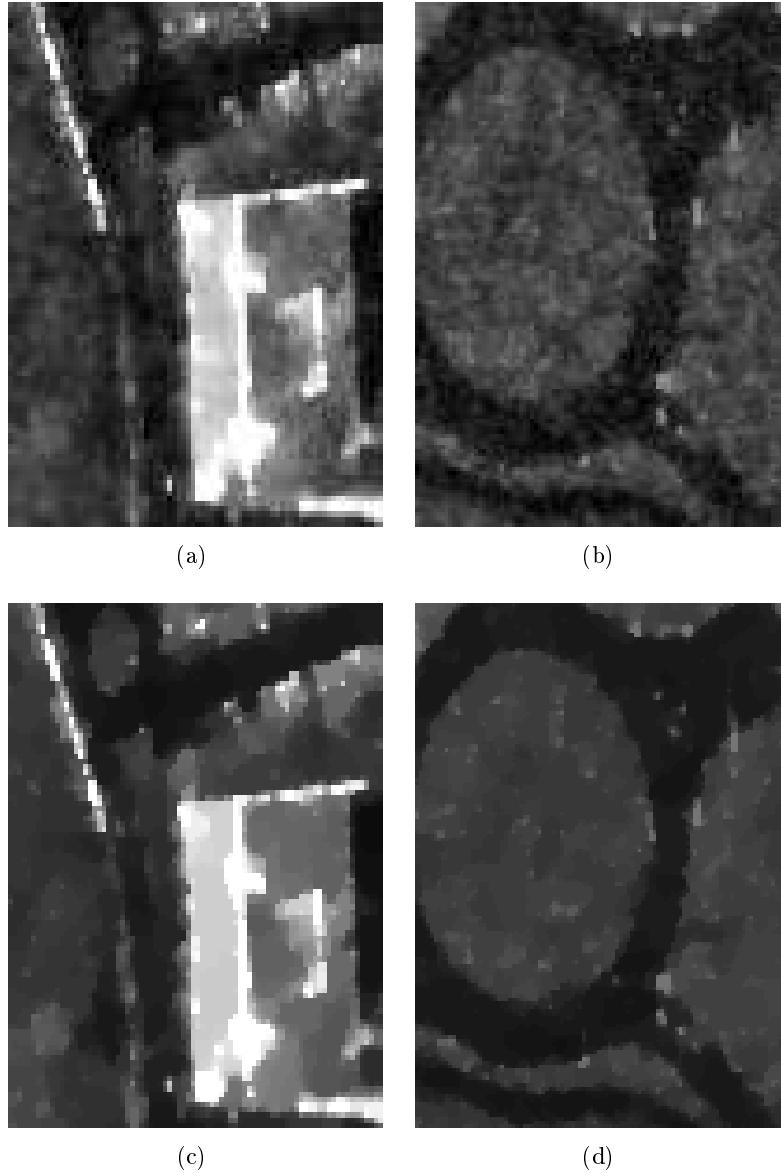


FIGURE 3.13 – Résultats de la restauration sur des zones d'intérêt de l'image de Toulouse. Restauration par δ_{16} -saut (a) de la zone 1, et (b) de la zone 2. Restauration par FastDC (δ_4 -saut + descente du gradient) (c) de la zone1, et (d) de la zone2.

On remarque l'avantage de l'approche mixte en terme de précision et temps de calcul (cf. TAB.3.1), malgré l'utilisation de plus d'étiquettes ($m = 16$) par le premier optimiseur MPLM par rapport à $m = 4$ par le deuxième. La précision de l'optimisation par saut sur les étiquettes est fixée à 0.1.

Ainsi réduire l'ensemble des étiquettes $\mathcal{L}_m$ pour une optimisation MPLM suivie d'une optimisation continue permet d'obtenir des minima de meilleure qualité en moins de temps de calcul. La précision sur la configuration obtenue à la convergence est aussi meilleure, vu la réduction de l'effet de la quantification des étiquettes de $\mathcal{L}$.

TABLE 3.1 – Résultats quantitatifs (image de taille 512×512 pixels).

Approche	Temps de convergence (mn :sec)	Minimum de l'énergie
MPLM (δ_{16} -saut)	03 : 08	$1.7613e + 06$
FastDC (δ_4 -saut + Descente du gradient)	02 : 09	$1.3818e + 06$

Les résultats visuels sur les deux zones extraites de l'image de Toulouse sont illustrés dans la figure 3.13, montrant plus clairement l'amélioration apportée par l'approche d'optimisation mixte sur la qualité des images restaurées : bords des bâtiments et routes plus nets et meilleur lissage des zones homogènes.

La figure 3.14 montre la décroissance de l'énergie itérativement par l'algorithme mixte de minimisation, ainsi que les optima obtenus respectivement par la première étape d'optimisation discrète et la deuxième étape d'optimisation continue.

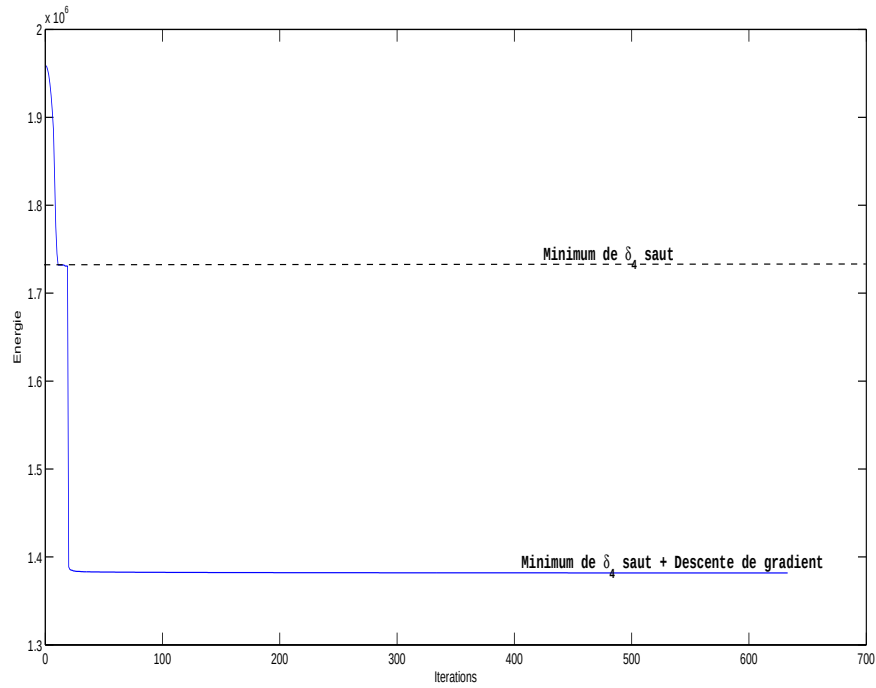


FIGURE 3.14 – Évolution de l'énergie itérativement par l'algorithme de minimisation FastDC.

4 Conclusion

Pour conclure, nous avons proposé dans ce chapitre un algorithme d'optimisation qui étend l'optimiseur MPLM proposé précédemment, dans le but d'améliorer ses performances en temps de calcul et précision de reconstruction. Ces performances ont été obtenues grâce à l'efficacité de la technique d'optimisation discrète par coupe-minimum suivie d'une optimisation convexe continue rapide. Les performances ont été illustrées sur un cas d'application de reconstruction d'images RSO présentant un challenge entre complexité algorithmique (temps de calcul et occupation en mémoire) et qualité/précision de la reconstruction. D'autres expériences dans un autre cadre applicatif seront illustrées dans la deuxième partie de ce manuscrit.

Au final, l'algorithme mixte d'optimisation FastDC nous a permis de nous affranchir des limites des optimiseurs présents dans l'état de l'art en traitement d'image. Les limites en mémoire occupée par les graphes ont été levées grâce aux mouvements larges et multi-étiquettes qui par ailleurs permettent de converger vers des optima locaux de bonne qualité. En contrôlant le paramètre m , il est dès lors possible d'optimiser la classe des modèles markoviens avec un *a priori* convexe et une attache aux données quelconque, tout en proposant un compromis entre complexité algorithmique et qualité de l'optimum. Il a été vérifié expérimentalement que la qualité de l'optimum est d'autant meilleure que m est grand. En pratique, avec un m fixé large et nettement inférieur à la taille de l'ensemble des étiquettes $\mathcal{L}$, la convergence vers un optimum global de l'énergie est possible. Pour pallier au problème de temps de calcul pour des valeurs élevées de m , dans le cas de données à forte dynamique radiométrique, un algorithme d'optimisation mixte a été proposé. L'ensemble des étiquettes est tout d'abord réduit pour accélérer le temps de convergence des optimiseurs MPLM. Une perte de précision sur les données est alors engendrée par une telle sous-quantification de l'intervalle des étiquettes possibles. Ainsi enchaîner l'optimisation discrète par une optimisation continue, type descente de gradient, est une alternative possible pour améliorer en précision l'optimum discret obtenu.

Dans la partie suivante de ce manuscrit, nous aborderons une problématique de grande importance en imagerie radar à synthèse d'ouverture, qui est la reconstruction de la phase interférométrique. Une bonne reconstruction ouvre les perspectives pour différentes applications en imagerie RSO nécessitant un modèle numérique de terrain précis. Pour cela, en nous reposant sur les approches markoviennes, des modèles seront élaborés et des algorithmes d'optimisation efficaces vont être nécessaires pour obtenir de bons résultats, en un temps raisonnable, sur des données de grandes dimensions, variées et complexes. Nous montrerons alors l'utilité des optimiseurs développés dans cette première partie du manuscrit de thèse pour bien résoudre ce problème d'imagerie.

Deuxième partie

Estimation markovienne de modèles numériques de terrain

Chapitre 4

Reconstruction du relief en imagerie RSO

L'imagerie radar est un domaine d'étude très actif en raison des multiples avantages que nous offrent ses systèmes imageurs et des nombreuses applications exploitant les données qu'ils génèrent. En effet, ces systèmes assurent une couverture continue du globe terrestre à tout moment de la journée et quelles que soient les conditions météorologiques. De plus, l'information apportée par les données acquises présente une complémentarité par rapport à l'information que fournissent les images optiques, ainsi qu'une importante spécificité à travers la phase du champ électromagnétique rétrodiffusé.

Parmi les applications en télédétection qui suscitent depuis des années un intérêt grandissant, la mesure du relief occupe une place très importante. Dans ce cadre, l'imagerie radar apporte une solution qui a ouvert de nouveaux horizons à cette thématique grâce à l'*interférométrie*.

Nous nous proposons d'étudier dans ce chapitre le problème de la reconstruction 3D de surfaces complexes à partir des données radar à synthèse d'ouverture (RSO) par interférométrie.

Dans un premier temps, un bref rappel des principales notions associées à l'imagerie radar, et plus particulièrement l'imagerie radar à synthèse d'ouverture est présenté. Nous renvoyons le lecteur cherchant une étude plus approfondie sur ce thème aux ouvrages suivants (Curlander et McDonough, 1991) et (Maître, 2001).

Dans un second temps, un bref aperçu des principales approches de reconstruction du relief à partir de ces images radar est proposé. Nous nous focalisons sur les techniques fondées sur l'interférométrie. Les points forts et faibles de telles approches seront aussi analysés, nous permettant de situer et justifier les méthodes que nous proposons dans les chapitres suivants.

1 Principes de l'imagerie radar

L'imagerie radar est une alternative à l'imagerie optique dans le sens où elle utilise un capteur actif dans le spectre des ondes radio (ondes centimétriques).

1.1 Capteurs radar

Les capteurs RADAR (RADio Detection And Ranging) sont apparus au cours de la deuxième guerre mondiale comme des systèmes militaires très utiles pour la détection d'intrusion d'ennemis ou la localisation des cibles par exemple. Puis, ces systèmes ont pris

un nouvel essor avec l'apparition des techniques de synthèse radar, améliorant notablement la résolution et donc élargissant le champ des applications tant militaires que civiles.

Différents capteurs radar ont alors été mis en service. On distingue deux familles :

- les capteurs de basse résolution (de l'ordre d'une dizaine de mètres) qui ont été initialement construits pour l'observation de la terre, de l'atmosphère et de la glace et pour l'étude des reliefs. Nous citons à titre d'exemple, SIR-A, SIR-B, RADARSAT, ERS1 et ENVISAT ;
- les capteurs de haute résolution (de l'ordre de 20cm à quelques mètres) et de très haute résolution (moins de 20cm) qui sont essentiellement militaires à l'origine et récemment mis en place pour l'étude précise de cibles et de bâtiments. Nous citons à titre d'exemple les capteurs aéroportés RAMSES et PAMIR (militaires), et les satellites TerraSAR-X (civil), RADARSAT-2 (civil) et Cosmo-Skymed (mixte).

Le spectre des fréquences des ondes radar varie de 30MHz à 300GHz et est décomposé en différentes bandes dont certaines sont réservées pour un usage militaire, d'autres sont absorbées par l'atmosphère et celles restant sont utilisées pour l'imagerie radar. A titre d'exemple, on trouve comme bandes utilisables : la bande X (de 5.75GHz à 10,90GHz), la bande L (de 0.39GHz à 1.55GHz) et la bande C (de 4.29GHz à 5.75GHz).

Composé d'une ou plusieurs antennes fonctionnant en émission et/ou réception, le système radar émet un train d'ondes cohérentes électromagnétiques et reçoit leurs réflexions par des cibles. En mesurant le temps entre l'émission t_e et la réception t_r de ce train d'ondes, on déduit la distance R entre la cible et le capteur par la relation suivante :

$$R = \frac{t_r - t_e}{2C}, \quad (4.1)$$

avec C la vitesse de la lumière supposée constante.

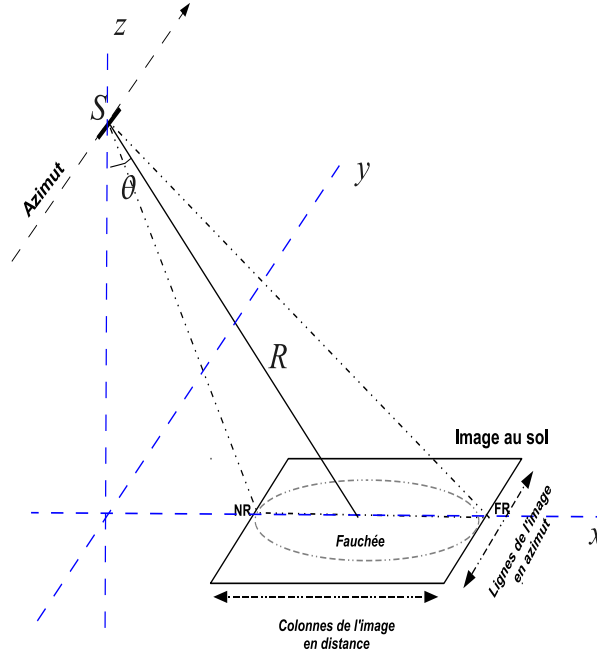
Les capteurs radar imageurs sont mobiles, aéroportés ou sur satellites. Leur trajectoire, parallèle à la zone imagée, définit l'axe azimut ou *along track* de l'image. L'acquisition est en visée *latérale* avec un angle d'émission θ (cf. FIG. 4.1). Cet axe de visée définit l'axe des distances ou *cross track*. La zone illuminée entre la distance minimale dite distance proximale ou *near range* (NR) et la distance maximale dite portée distale ou *far range* (FR) est appelée fauchée. Le temps de retour de l'impulsion électromagnétique fournit la position en distance des différents éléments réflecteurs de la surface.

1.2 Résolution d'un capteur radar

Une notion importante en imagerie radar est la résolution (en azimut et en distance). Elle se définit par la distance minimale entre deux sources identiques placées telles qu'un observateur puisse distinguer les deux sources.

En distance, la résolution caractérise la capacité à séparer deux objets selon l'axe des distances. Cette résolution est contrainte par la durée de l'impulsion radar émise. Les signaux réfléchis par des cibles sont séparables si la distance qui sépare ces cibles est suffisamment grande pour que le capteur ait le temps de recevoir deux impulsions distinctes de durée chacune τ . La résolution en distance est définie ainsi par :

$$r_{distance} = \frac{C\tau}{2}. \quad (4.2)$$

FIGURE 4.1 – Visée du radar d'antenne S .

En azimut, la résolution caractérise la capacité du capteur à séparer deux objets selon l'axe des azimuts. Cela est lié directement à la directivité de l'antenne qui est d'autant plus forte que la longueur d'onde est petite ou que la taille de l'antenne est plus grande. La résolution en azimut à la distance R de l'antenne est définie par :

$$r_{azimut} = 0.88 \frac{\lambda}{L} R, \quad (4.3)$$

avec L la longueur de l'antenne et λ la longueur d'onde du signal.

Cependant, en utilisant les expressions 4.2 et 4.3 pour obtenir une résolution de 50cm en azimut et en distance, à titre d'exemple, il faudrait une antenne de 220m émettant un train d'onde de durée $\tau = 3.3\text{nm}$. Un tel système est en pratique non concevable, puisque les antennes sont en réalité de l'ordre de quelques mètres, émettant des longueurs d'ondes généralement de l'ordre du centimètre.

Vu ces limitations de résolution en pratique, des techniques de traitement de signal sont alors utilisées. Il s'agit de techniques de compression d'impulsion pour la direction en distance et de la synthèse d'ouverture pour la direction azimutale. Cela nous amène aux systèmes Radar à Synthèse d'Ouverture (RSO).

1.3 Les systèmes RSO

La technique de compression d'impulsion consiste à conserver une impulsion longue afin de conserver une bonne énergie à la réception sans pour autant avoir une mauvaise résolution. L'idée est d'envoyer une impulsion modulée linéairement en fréquence autour d'une fréquence centrale. Ce type de signal, appelé *chirp*, est ensuite corrélé au signal retourné. L'effet résultant est qu'une cible ponctuelle produit un sinus cardinal d'autant

plus étroit que la largeur de bande B du chirp est grande. Grâce au filtrage adapté, la résolution peut être excellente même en utilisant des impulsions relativement longues. Ce procédé permet donc de produire des trains d'onde de forte énergie avec des puissances plus modestes. La résolution en distance de ce signal démodulé est désormais fonction de la bande passante du chirp comme suit :

$$r_{distance}^{RSO} = \frac{C}{2B}. \quad (4.4)$$

La résolution en distance sera donc d'autant meilleure que la bande passante sera grande.

La technique de compression en azimuth consiste à synthétiser artificiellement une antenne très grande en utilisant une antenne beaucoup plus petite (cf. FIG. 4.2). Cela est réalisé en exploitant le déplacement du porteur de l'antenne dans la direction de l'azimut et le recouvrement des zones illuminées. Ainsi l'information concernant une cible se retrouve répartie sur différentes acquisitions, avec une phase variable puisque la distance capteur/cible varie au cours du déplacement. En isolant chaque cible, nous nous retrouvons avec une quantité d'information sur la cible équivalente à celle que pourrait donner une antenne de grande taille et donc extrêmement directive. Grâce à ce procédé, la résolution en azimuth devient :

$$r_{azimut}^{RSO} = \frac{L}{2}. \quad (4.5)$$

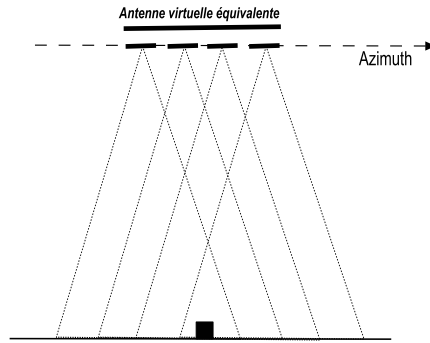


FIGURE 4.2 – Schéma de la synthèse en azimuth.

1.4 L'image RSO

Dans la figure 4.1, le capteur à bord de l'avion ou d'un satellite émet une onde et enregistre l'onde rétrodiffusée à partir de l'instant t_0 pendant une durée Δt . L'ensemble de ces mesures forme une ligne de l'image. Cette ligne correspond donc à la zone au sol dont la distance au capteur est comprise entre Ct_0 et $C(t_0 + \Delta t)$. Le porteur avance d'une distance $\frac{v}{f}$, avec f la fréquence d'émission des impulsions. Le capteur envoie une deuxième onde, une nouvelle ligne est formée dans l'image ; ainsi de suite jusqu'à balayer toute la

scène en azimut pour former l'image entière de la scène. Sur ces données, appelées *brutes*, sont alors appliqués des algorithmes de traitement de signal (compression d'impulsion et synthèse d'ouverture) pour obtenir l'image proprement dite.

L'image RSO obtenue est une image complexe (amplitude et phase). Son amplitude et sa phase sont caractérisées par des déformations géométriques et radiométriques spécifiques.

Les déformations géométriques, illustrées dans la figure 4.3, sont essentiellement dues à la vision latérale du capteur radar. Nous distinguons :

- La contraction ou *forshortening* : rétrécissement des pentes orientées vers le capteur et étirement des pentes orientées en sens opposé.
- L'avancée de relief : l'écho renvoyé par une zone de relief de pente forte, comme par exemple une montagne, précède celui de la base. Ce problème survient généralement pour des géométries de visée à petits angles d'incidence comparés à la pente du relief imagé.
- Le repliement ou recouvrement ou *layover* : des points de zones de relief distinctes sont confondus.
- L'ombre ou *shadow* : comme la visée est latérale, l'onde n'éclaire pas une partie de la scène masquée par une pente élevée. Ainsi nous perdons de l'information sur les signaux au sol dans cette zone.

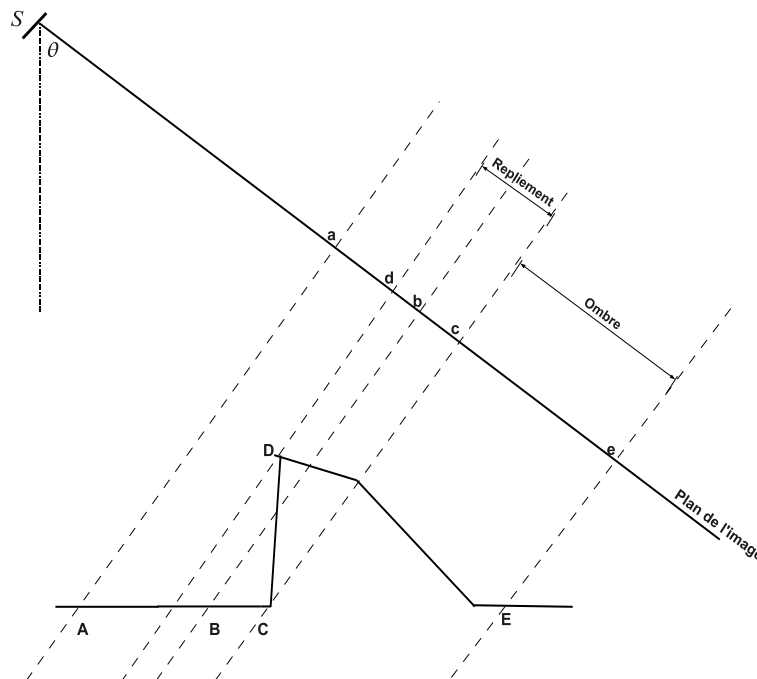


FIGURE 4.3 – Problèmes géométriques de la visée latérale du radar.

Les problèmes radiométriques sont principalement :

- Les coins réflecteurs et les angles droits parallèlement à la trajectoire, compte tenu de la visée latérale du radar, constituent des objets extrêmement rétrodiffusant. Ils entraînent une grande hétérogénéité radiométrique des zones urbaines, rendant difficile leur interprétation.
- Le bruit qui est dû à la somme cohérente des réponses des diffuseurs élémentaires dans une cellule de résolution. Cette addition cohérente crée un bruit multiplicatif appelé chatoiement ou *speckle* caractéristique d'une image d'amplitude RSO. Ce bruit gêne la plupart des traitements automatiques sur des images radar. Plusieurs travaux ont été consacrés à modéliser statistiquement ce bruit en vue de l'éliminer. Le plus simple des modèles proposés est celui du chatoiement pleinement développé proposé par Goodman (Goodman, 1985), modélisé par une loi *Rayleigh-Nakagami*. Ce dernier suppose que la surface est composée de réflecteurs élémentaires dont la distribution spatiale est complètement aléatoire et avec une forte rugosité comparée à la longueur d'onde. Il est plutôt adapté aux scènes naturelles à basse résolution. Dans le cas des scènes à haute résolution, les textures tendent à disparaître et les échos forts dominent. Les tailles et natures des objets changent clairement, ce qui rend l'hypothèse de zone homogène rugueuse non vérifiée. Nous trouvons d'autres modèles prenant en compte les textures ou les forts rétrodiffuseurs comme les lois **K** (Jakeman et Pusey, 1976), Beta (Lopes *et al.*, 1990), Log-Normale (Szajnowski, 1977), Nakagami-Rice (Dana et Knepp, 1986), Fisher (Nicolas, 2002), etc.

En ce qui concerne l'image de phase, l'information est non exploitable directement en raison de son caractère aléatoire d'un pixel à un autre dépendant des propriétés électromagnétiques des objets (cf. FIG. 4.4). Cependant, d'après la théorie de propagation de l'onde électromagnétique, elle contient également une information de distance relatant le trajet parcouru par l'onde. Ainsi en utilisant deux phases de deux images prises de la même scène, dans des conditions géométriques proches, et en les retranchant, on peut remonter à l'information en distance et estimer la hauteur d'un point observé. C'est le principe de l'interférométrie RSO qu'on détaillera par la suite.

Disposant de ces données RSO, différents traitements sont alors nécessaires dans le but de l'interprétation automatique ou semi-automatique de la scène imagée. Une des applications possibles est la reconstruction de modèles numériques de terrain pour estimer les altitudes des cibles imagées. Dans la suite, un bref aperçu sur les principales approches de reconstruction du relief proposées dans la littérature est présenté.

2 Approches de reconstruction du relief en imagerie RSO

La représentation d'une scène en relief peut se faire selon différents modèles numériques :

- le modèle numérique de terrain (MNT) qui est une reconstruction de l'information altimétrique du sol, souvent réalisé sur des données de moyenne résolution.
- le modèle numérique de surface (MNS) qui est une reconstruction indifféremment de la nature du relief : sol, bâtiment, etc., souvent obtenu à partir de données de haute résolution.
- le modèle numérique d'élévation (MNE) qui est généralement obtenu par une approche objet identifiant séparément les altitudes des éléments constituant la scène,

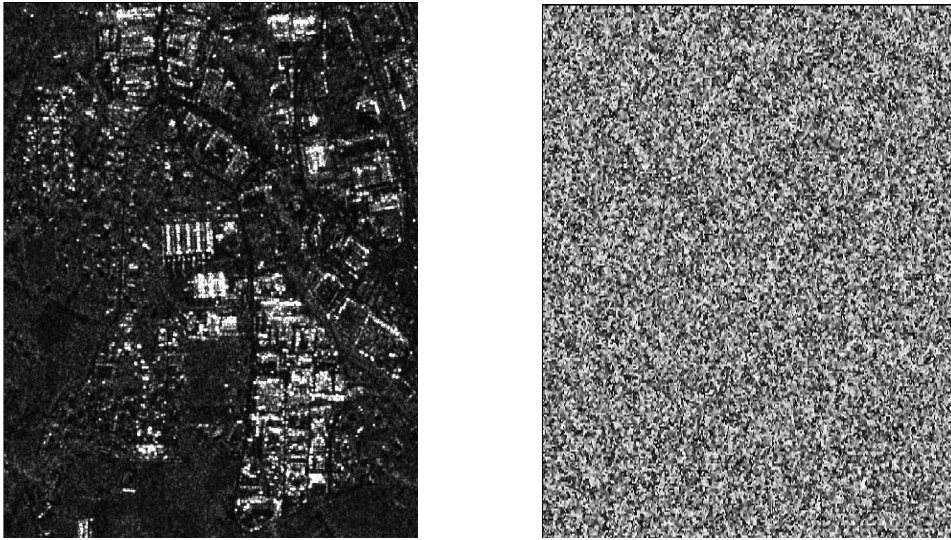


FIGURE 4.4 – Images d’amplitude et de phase d’une même zone.

utile pour interpréter une scène urbaine à partir de données de très haute résolution.

Ces différents modèles numériques sont construits selon les besoins de l’application. Nous répartissons alors les approches de reconstruction du relief en deux familles. La première famille regroupe les méthodes pixeliques fournissant généralement des MNS ou MNT et la deuxième regroupe les méthodes par objets ou régions donnant au final des MNE.

2.1 Approches de reconstruction pixelique

Reconstruire l’information de hauteur d’une scène donnée peut être réalisé à partir de l’image d’amplitude RSO, en exploitant les statistiques associées aux différentes textures de la scène et selon des approches se rapportant au *shape from shading* (forme à partir de l’ombrage). Une approche dans ce cadre a été au départ étudiée dans (Kirscht et Rinke, 1998) pour fournir des MNS. Elle exploite les moments du premier et deuxième ordre de ces statistiques pour discerner entre les trois classes : sol, ombres et zones élevées. Ensuite, l’information des hauteurs des zones élevées, donnant un MNT, est extraite à partir de la taille des ombres (Bolter, 2000). En revanche, cette technique est souvent non robuste aux problèmes radiométriques et géométriques qui caractérisent les images RSO.

D’autres approches ont été développées exploitant plutôt un couple d’images RSO. Ces approches reposent essentiellement sur deux techniques : l’interférométrie et la radargrammétrie.

L’interférométrie est une technique permettant de remonter à l’information altimétrique de la scène à partir des données de phase, appelées interférogrammes. Un interférogramme est construit à partir des différences des phases de couple d’images RSO d’une même scène prises de points de vue différents mais avec une faible variation de l’angle d’incidence. Pour en extraire l’information d’élévation, différents prétraitements sont indispensables. Au départ, l’information géométrique provenant de l’éloignement en distance le long de la fauchée est éliminée (franges orbitales). Cela nécessite la connaissance des paramètres du

système d'acquisition. Ensuite, les franges topographiques sont localisées ou déroulées pour retrouver le relief. Cette étape est délicate compte tenu du bruit qui caractérise ces interférogrammes. Enfin, des post-traitements sont effectués en vue d'améliorer les résultats de reconstruction et tenir compte des problèmes géométriques en imagerie RSO (Petit, 2004).

La radargrammétrie reprend le principe de la stéréoscopie optique, où l'information 3D d'une cible est obtenue à partir de la différence de ses localisations sur un couple d'images prises sur la même scène à des angles d'incidence différents (Leberl, 1990). Cette information 3D est obtenue à partir d'appariement (ou mise en correspondance) des deux images. C'est cette étape qui est la plus délicate dans l'approche, compte tenu des différences radiométriques des images prises avec un écart (même faible) entre les angles d'incidence. L'appariement peut être réalisé directement sur les pixels des images, cela nécessitera des critères de corrélation assez puissants et appropriés à ces différences de radiométrie (Simonetto, 2002). Il peut aussi être réalisé entre primitives extraites sur les deux images (Soergel *et al.*, 2006).

En réalité, pour raffiner les résultats de reconstruction et aboutir à des MNT plus précis et moins bruités, les différentes familles d'approches décrites précédemment sont parfois exploitées conjointement :

- Interférométrie et amplitude dans (Soergel *et al.*, 2000), (Dell'Acqua *et al.*, 2001), (Stilla *et al.*, 2001), où les auteurs procèdent par filtrage et segmentation de l'image en amplitude. Les segments obtenus sont combinés avec l'information altimétrique donnée par l'interférogramme, aboutissant à une meilleure reconstruction 3D.
- Interférométrie et radargrammétrie dans (Schubert *et al.*, 2002), (Schubert, 2004), où en premier lieu, une approche stéréoscopique multirésolution est adoptée pour la reconstruction des reliefs. Ensuite, les résultats sont corrigés en exploitant l'information complémentaire apportée par un MNT généré à partir de l'interférogramme disponible sur la même scène.

2.2 Approches de reconstruction par régions et objets

Pour représenter la scène sous la forme d'un ensemble d'objets 3D, des modèles sont alors nécessaires pour décrire les éléments de cette scène comme les structures urbaines et les structures naturelles. Ces modèles sont souvent décrits à partir des primitives extraites sur les structures et d'un ensemble d'hypothèses ou contraintes géométriques mettant en relation ces primitives.

Dans la littérature, les travaux développant une telle approche ont surtout été orientés vers la reconstruction 3D d'objets non-naturels (*man made objects*), comme les bâtiments, compte tenu de la connaissance *a priori* des structures géométriques qui les caractérisent le plus souvent. Un bâtiment est modélisé par exemple par une structure ayant une forme rectangulaire, cylindrique ou plus complexe (polygone), et ayant un toit plat, en pente ou double pente et une orientation globale.

Pour aboutir à un modèle, des primitives sont alors extraites tout d'abord sur l'image RSO et ensuite gérées par des contraintes imposant le modèle géométrique défini. Nous trouvons comme primitives utilisées :

- les contours à partir des ombres des bâtiments extraits sur l'image d'amplitude.

Différentes méthodes d'extraction de l'ombre existent (Bolter, 2000), (Soergel *et al.*, 2000), (Tison, 2004) ;

- les points d'intérêts à partir des échos doubles et points brillants sur l'image d'amplitude (Michaelsen *et al.*, 2006), (Petit, 2004), (Thiele *et al.*, 2006) ;
- les régions plates à partir d'une segmentation de l'interférogramme déroulé (Gamba *et al.*, 2000).

D'autres approches de modélisation, dites par géométrie stochastique, procèdent plutôt par paramétrisation des modèles géométriques à reconstruire (exemple : centroïdes des facettes de bâtiments, leurs dimensions, leurs réflectivités, etc.) (Quartulli et Datcu, 2003). Ensuite, par une approche probabiliste d'estimation bayésienne du critère MAP, les pixels de l'image d'amplitude RSO sont classifiés selon ces paramètres afin de reconstruire les hauteurs des bâtiments selon les classes définies.

Dans la suite, les approches de reconstruction de relief par interférométrie sont étudiées.

3 Estimation des MNT par interférométrie

La reconstruction de la topographie de la terre par interférométrie est la plus ancienne des approches pour estimer le MNT d'une scène imagée par un radar (Graham, 1974), (Zebker et Goldstein, 1986). L'interférogramme est généré à partir de deux signaux (couple interférométrique) provenant de deux antennes séparées spatialement. La distance qui les sépare est dite *base*. Elle ne doit être ni trop petite pour avoir des variations de phase suffisantes, ni trop grande pour éviter une importante décorrélation spatiale (qui génère plus de bruit interférométrique). Au delà d'une certaine base, les pixels sont décorrélés et l'interférométrie n'est plus possible.

Deux types de systèmes interférométriques existent :

- Système à une seule antenne dont le porteur effectue deux passages assez proches l'un de l'autre (Goldstein *et al.*, 1988) pour fournir les deux acquisitions interférométriques. Il s'agit de l'approche multipasse généralement utilisée pour les capteurs satellitaires.
- Système à deux antennes qui éclairent simultanément le même terrain (Zebker *et al.*, 1992) pour éviter le problème de décorrélation temporelle des surfaces provoqué par exemple lors d'une variation du sol imagé entre les deux passages de l'antenne ou par différentes perturbations atmosphériques. Il s'agit de l'approche monopasse qui a été utilisée dans la mission SRTM de la navette spatiale américaine et souvent par des capteurs aéroportés.

Pour ces deux types de systèmes, la reconstruction du relief est réalisée en se basant sur la même théorie (Li et Goldstein, 1987) que nous présentons dans la suite.

3.1 Principe de l'interférométrie

Considérons le cas d'un système interférométrique monopasse, l'altitude h d'une cible est calculée à partir de l'interférogramme obtenu par la différence de phase des deux sig-

naux qu'elle a retrodiffusés aux deux antennes S_1 et S_2 , appelées respectivement antenne maîtresse et antenne esclave (cf. FIG. 4.5).

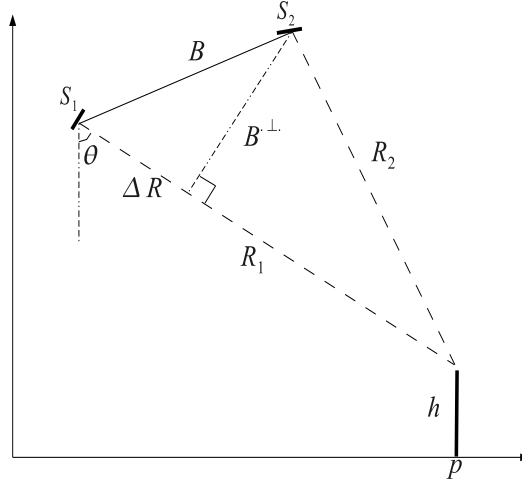


FIGURE 4.5 – Prise de vue en interférométrie.

En effet, la phase d'un signal reçu mesure le décalage temporel entre l'émission du train d'ondes et le retour de son écho. Cette phase mesurée est en réalité la combinaison de plusieurs phénomènes physiques qui interceptent l'acheminement de l'onde entre le capteur et la cible. Ces phénomènes sont divers, nous citons à titre d'exemple les conditions atmosphériques, les propriétés des matériaux constituant la cible et les procédés d'acquisition du capteur (Petit, 2004). Ces derniers ajoutent au déphasage ϕ_g entre le signal émis et son écho par une cible, dû uniquement à la propagation du signal, des quantités d'information non désirable que l'on notera par ϕ_r . La phase mesurée s'écrit alors comme étant :

$$\phi = \phi_g + \phi_r. \quad (4.6)$$

En se plaçant dans l'hypothèse où l'indice du milieu traversé par les trains d'ondes est le même que celui du vide, la phase due uniquement au trajet s'écrit comme étant :

$$\phi_g = \frac{2\pi}{\lambda} \times 2R, \quad (4.7)$$

où R désigne la distance séparant le centre de la cible du centre de l'antenne.

En se plaçant dans le cas de deux acquisitions d'une même cible, avec l'hypothèse de conditions d'acquisition presque identiques, nous avons :

$$\phi_1 = \frac{2\pi}{\lambda} 2R_1 + \phi_r, \quad (4.8)$$

$$\begin{aligned} \phi_2 &= \frac{2\pi}{\lambda} 2R_2 + \phi_r \\ &= \frac{2\pi}{\lambda} 2(R_1 - \Delta R) + \phi_r. \end{aligned} \quad (4.9)$$

Ainsi la différence de phases $\Delta\phi$ pour une cible donnée calculée à partir de ces deux acquisitions vaut :

$$\begin{aligned}\Delta\phi &= \phi_1 - \phi_2 \\ &= \frac{4\pi}{\lambda}\Delta R.\end{aligned}\tag{4.10}$$

En négligeant la courbure de la terre, la géométrie de prise de vue permet de relier ΔR avec l'altitude h de la cible imagée par l'expression suivante :

$$\Delta R = \frac{B^\perp h}{R_1 \sin(\theta)},\tag{4.11}$$

où $B^\perp$ est la composante perpendiculaire de la base et θ est l'angle d'incidence (voir annexe D pour les détails du calcul).

Connaissant les deux expressions (4.10) et (4.11), il est possible de lier l'information altimétrique d'une cible avec la différence de phases $\Delta\phi$ par l'expression suivante :

$$\begin{aligned}\Delta\phi &= \frac{4\pi B^\perp h}{\lambda R_1 \sin(\theta)} \\ &= 2\pi \frac{h}{h^{amb}},\end{aligned}\tag{4.12}$$

où h^{amb} est une quantité calculée à partir des paramètres d'acquisition des données InRSO, appelée *hauteur d'ambiguïté*, et dont l'expression est donnée par :

$$h^{amb} = \frac{\lambda R_1 \sin(\theta)}{2 B^\perp}.\tag{4.13}$$

Notons par ψ la mesure de phase interférométrique, appelée aussi phase empirique. Cette mesure est conventionnellement calculée modulo 2π , c.-à-d.

$$\psi = \langle \Delta\phi \rangle_{2\pi}.\tag{4.14}$$

Ainsi la hauteur d'ambiguïté est en fait l'altitude pour laquelle la phase interférométrique mesurée aura fait un tour de phase complet de 2π .

En interférométrie, les difficultés de reconstruction d'altitudes à partir des mesures de déphasage sont le plus souvent dues à cette opération d'enroulement de la vraie phase, en plus du bruit interférométrique qui s'ajoute aux mesures.

3.2 Contributions dans la phase mesurée

Comme le déphasage ϕ dans l'équation (4.6) fait intervenir en réalité différentes quantités dues aux phénomènes physiques interceptant l'onde en propagation, qui peuvent changer d'une acquisition à une autre, la phase interférométrique mesurée ψ dans (4.14) est en réalité composée en plus de la phase relative à l'altitude de la scène de différentes autres quantités qu'on peut résumer par l'expression suivante :

$$\psi = \psi_{topo} + \psi_{orb} + \psi_{def} + \psi_{atm} + \psi_{bruit},\tag{4.15}$$

où ψ_{topo} est la phase topographique relative au délai de phase dû au relief, ψ_{orb} , appelée phase orbitale, est le délai de phase associé à la contribution de la terre plate, ψ_{def} est le délai de phase représentant le déplacement du terrain sur la ligne de visée du radar, ψ_{atm} , appelée contribution atmosphérique, résulte de la variation des délais de phase apportés

par les couches de l'atmosphère dans le cas des acquisitions multi-passes, enfin ψ_{bruit} est un bruit qui pourrait être d'origine diverse telle que le capteur, la décorrélation temporelle ou géométrique, etc.

Ces quantités qui s'ajoutent à la mesure de phase, en plus de l'opération de l'enroulement, rendent la tâche de la reconstruction de la topographie d'une scène difficile et ambiguë.

3.3 Produit interférométrique

En pratique, pour disposer de cette information de déphasage ψ , une chaîne de traitements est indispensable à la suite de l'acquisition des données images RSO, livrant à la fin un ensemble de données interférométriques qu'on appellera *produit interférométrique*.

Initialement, nous disposons de données images complexes prises par les deux antennes S_1 et S_2 , notées respectivement par z_1 et z_2 :

$$z_1 = a_1 e^{j\psi_1}, \quad (4.16)$$

$$z_2 = a_2 e^{j\psi_2}, \quad (4.17)$$

où a_1 et a_2 correspondent respectivement aux amplitudes des images complexes z_1 et z_2 et ψ_1 et ψ_2 correspondent respectivement à leurs phases.

Pour un calcul précis de ces mesures de déphasage entre ψ_1 et ψ_2 , c.-à-d. de l'interférogramme, une étape de recalage est nécessaire pour bien faire se correspondre les pixels.

Différentes techniques de recalage d'images RSO existent dans la littérature. Certaines reposent sur la corrélation statistique (Li et Goldstein, 1987), d'autres sur la transformée de Fourier rapide (Gabriel et Goldstein, 1988), d'autres recourent à une minimisation d'une fonctionnelle dite de fluctuation moyenne (Lin *et al.*, 1991), etc. L'efficacité et la fiabilité d'une technique dépend de son adéquation avec les caractéristiques des images à recalculer telles que la texture et la radiométrie. Il est à noter que les méthodes de recalage permettent d'avoir un recalage subpixelique, de l'ordre du dixième du pixel. Cela montre la grande précision des interférogrammes que nous pouvons obtenir.

Disposant d'images parfaitement recalées, nous pouvons dès lors calculer la quantité dite *cohérence complexe* pour en déduire le produit interférométrique. Cette quantité est obtenue à partir du produit hermitien moyenné entre z_1 et z_2 (recalées), noté par ζ , dont chaque pixel p est obtenu par l'opération suivante :

$$\zeta(p) = \frac{\sum_{q=1}^L z_1(q) \overline{z_2(q)}}{\sqrt{\sum_{q=1}^L |z_1(q)|^2 \sum_{q=1}^L |z_2(q)|^2}} \quad (4.18)$$

$$= \gamma(p) e^{j\psi(p)} \quad (4.19)$$

où $\overline{z_2(q)}$ est la valeur complexe conjuguée au pixel q dans l'image z_2 et L est le nombre de pixels utilisés pour moyenner la valeur au pixel p à partir des valeurs aux pixels q , à la résolution originale, dans une fenêtre de taille L . Ce moyennage permet en fait de réduire le bruit de speckle.

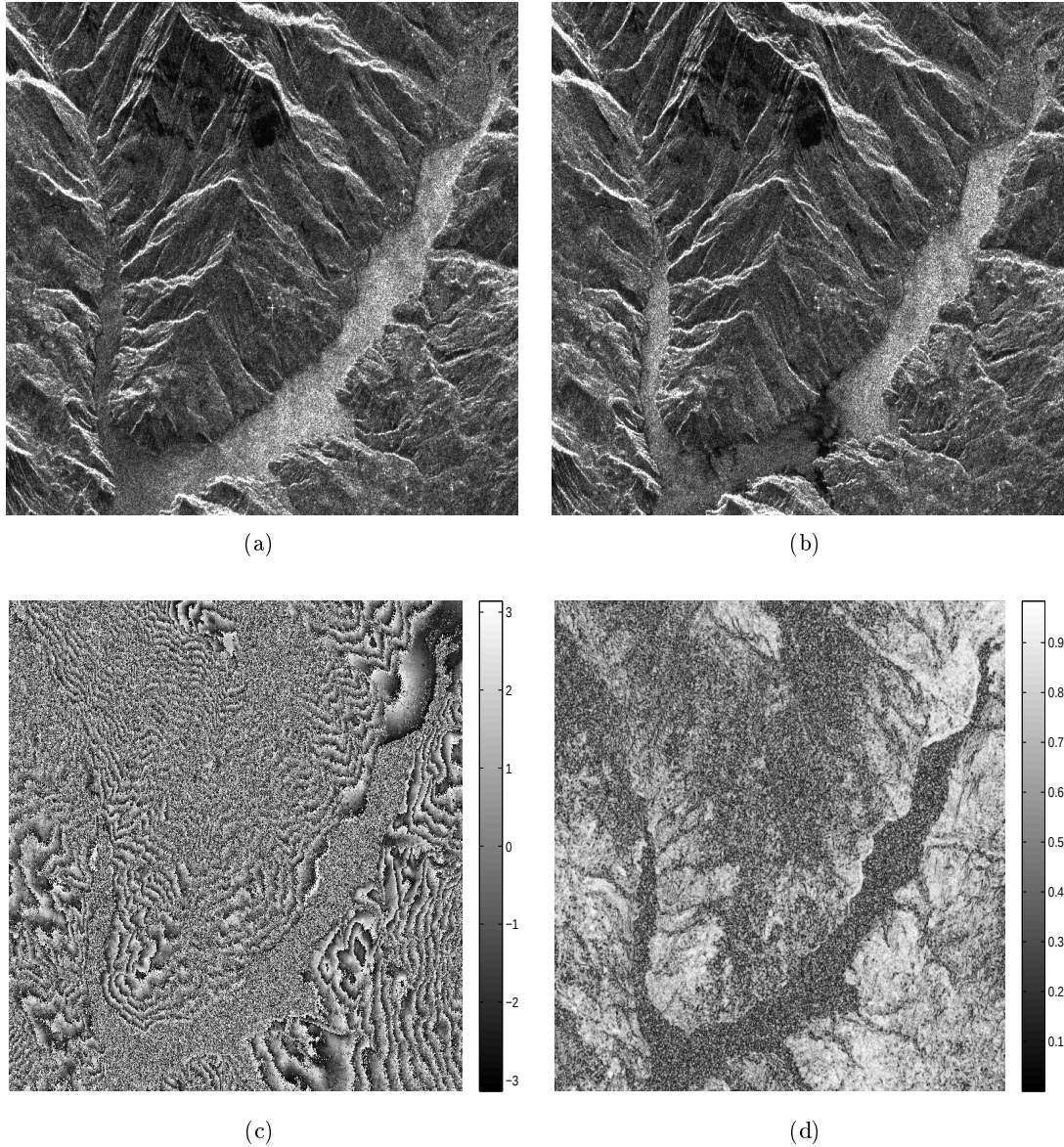


FIGURE 4.6 – Exemple de produit interférométrique (données RSO fournies par le capteur ERS-2 sur une zone de Serre-Ponçon). (a) image d'amplitude a_1 , (b) image d'amplitude a_2 recalée par rapport à a_1 , (c) image de différence de phases ou interférogramme ψ , et (d) image de cohérence γ .

De ce produit hermitien, le produit interférométrique fourni est :

- L'image des cohérences, notée par γ , qui n'est autre que le module de l'image complexe ζ . Elle peut être interprétée comme étant l'image de la corrélation entre les deux images RSO ou mesure de confiance dans l'interférogramme. La cohérence serait faible à cause d'un changement important de l'angle d'incidence, changement de la nature du sol dans le cas multipasse, présence de zones de très faibles rétrodiffusion où le bruit thermique domine, etc.
- L'interférogramme ou image des différences de phases, notée par ψ , qui n'est autre que la phase de l'image complexe ζ enroulée dans $[-\pi, \pi[$. Comme indiqué dans

le paragraphe 3.2, on observe sur cette image des franges (résultant de l'opération modulo) de plusieurs natures :

- franges orbitales dues à l'éloignement par rapport au capteur en distance,
- franges topographiques dues au relief,
- franges atmosphériques dues au changement des conditions atmosphériques entre les deux acquisitions,
- franges dues au déplacement ou changement de terrain de la zone imagée,
- franges dues à d'autres sources de bruit : chaînes d'acquisition et de formation des données, imprécision des mesures, etc.

Ces différentes franges sont en fait combinées ensemble.

Dans la figure 4.6, un exemple de produit interférométrique est présenté. Partant de ces données fournies, estimer la topographie d'une scène directement des mesures de l'interférogramme est donc bien une tâche complexe. Des prétraitements sont au préalable indispensables pour réduire la quantité d'information non-utile dans ces mesures pour ne garder à la fin que les mesures en relation avec le relief de la scène.

Les franges orbitales peuvent être enlevées dans un premier temps en considérant la géométrie de l'image et connaissant les paramètres d'acquisition du capteur. Les autres franges, dues à des phénomènes divers et imprévisibles, sont beaucoup plus difficiles à enlever. Leur identification et modélisation sont nécessaires pour mettre en place des approches permettant de les estimer.

Dans le paragraphe suivant, une description plus détaillée de ces artefacts de phase est alors proposée.

3.4 Principales perturbations de la mesure interférométrique

Nous répartissons ces perturbations en deux catégories :

A. Décorrélations :

La décorrélation entre les deux images RSO du couple interférométrique est l'une des sources majeures d'artefacts et bruits parasites sur les interférogrammes. Cette décorrélation est d'origine variée :

- bruit thermique additif lié au capteur,
- imprécision du recalage qui introduit des réponses de cellules des deux images RSO ne faisant pas intervenir les mêmes cibles élémentaires,
- changement de l'état de la scène entre deux acquisitions dans le cas de l'interférométrie multipasse, introduisant ainsi une décorrélation temporelle entre les cellules de résolution,
- longueur de la base séparant les deux antennes visant une même cellule de résolution qui introduit une décorrélation spatiale due à la légère différence de l'angle d'incidence.

B. Perturbations atmosphériques :

Sur les données interférométriques réelles, les perturbations atmosphériques s'observent sous la forme de franges parasites. Ces perturbations, qui sont dues essentiellement aux modifications des conditions atmosphériques entre les différentes acquisitions, entraînent un changement d'indice de réfraction de l'atmosphère et donc une variation du trajet de l'onde radar. Ainsi une rotation de la phase sur l'interférogramme s'en suit. Ce déphasage

introduit dans les observations de phase est donc une source d'erreur si l'estimation de l'altitude est réalisée directement à partir de ces données.

Dans le travail de thèse (Chaabane, 2004), en s'attaquant à la problématique de l'interférométrie différentielle pour estimer le changement et le déplacement du terrain entre différentes acquisitions, les perturbations atmosphériques sont dans un premier temps localisées et ensuite corrigées. Ces perturbations ont été réparties en deux classes : globales et locales. Les perturbations globales, touchant l'intégralité de la scène imagée, supposent l'homogénéité de la troposphère et introduisent un déphasage linéaire par rapport à la topographie de la scène. Cette hypothèse n'est pas toujours vérifiée dès que des phénomènes météorologiques particuliers tels que les nuages, le changement du taux d'humidité entre les zones de la scène, etc. interviennent. Par conséquent, l'hétérogénéité dans la couche de la troposphère apporte d'autres franges parasites locales indiquant la présence de telles perturbations. L'auteur propose des outils algorithmiques pour pouvoir estimer ces perturbations et les corriger des mesures interférométriques, en vue de fournir des interférogrammes différentiels pour l'estimation des changements ou mouvements du terrain. Nous reviendrons sur ces perturbations atmosphériques dans le chapitre suivant.

Maintenant que nous connaissons la nature des principales perturbations et artefacts dans les mesures interférométriques, les corriger de l'interférogramme brut permet d'obtenir une image de phase en relation directe avec le relief de la scène par l'opération modulo 2π . Nous nous retrouvons alors face à un problème de développement des franges dans cette image. Sa résolution consiste à distinguer entre une mesure de phase de valeur ψ rad et une autre mesure de phase de valeur égale à $(\psi + 2n\pi)$ rad, avec n un entier. Ce problème est connu par le nom de *déroulement de phase*.

3.5 Déroulement de phase InRSO et estimation de MNT

Le problème de déroulement de phase pourrait être simplifié en un premier temps, si nous supposons la contrainte suivante : la phase absolue ϕ ne varie jamais de plus de π entre deux points consécutifs. Ainsi quand ψ varie de plus π la phase absolue a donc changé de cycle, et on augmente ainsi ou diminue la valeur de la phase mesurée par 2π .

Cette contrainte dite de régularité ou condition d'Itoh (Itoh, 1982) entraîne ainsi l'existence d'une solution unique pour ϕ à une translation d'un multiple de 2π près.

En pratique, cette contrainte de régularité est généralement non vérifiée à cause des différentes situations qui entraînent des faux sauts de phase dépassant π , telles que :

- le bruit interférométrique qui pourrait toujours persister même suite à la correction de l'interférogramme des franges parasites,
- les fortes discontinuités sur la zone imagée et la spécificité de la géométrie d'acquisition radar entraînant des phénomènes de repliement de phase et de l'aliasing, surtout observables dans les endroits à fortes pentes par rapport à l'angle d'incidence.

La condition de régularité d'Itoh est alors non respectée. En l'absence de cette condition, le problème devient complexe. Il s'agit bien d'un problème inverse mal-posé. Le développement de franges de proche en proche conduit à des résultats différents suivant le chemin emprunté (Ghiglia *et al.*, 1987). On assiste à un phénomène d'ambiguïté de phase.

Dans la suite, certaines des méthodes développées dans la littérature pour résoudre un

tel problème inverse sont décrites. Nous renvoyons le lecteur à l'ouvrage (Ghiglia et Pritt, 1998) pour une étude détaillée de certaines des approches présentées.

Nous proposons une classification de ces approches similaire à celle proposée dans (Dupont et Berthod, 1994) en méthodes locales et méthodes globales.

Les méthodes locales permettent de reconstruire l'image de phases absolues par un processus de développement de proche en proche. En parcourant l'interférogramme, la phase est déroulée pixel par pixel localement de façon à minimiser la différence de phase entre deux pixels voisins. Cependant, dans le cas pratique avec la présence du bruit, différentes solutions sont obtenues en empreintant des chemins différents. La correction des données des incohérences qui perturbent le déroulement est donc nécessaire.

Initialement, l'approche dite *branch-cut* proposée dans (Goldstein *et al.*, 1988) met en évidence la présence de résidus qui sont les points où la phase interférométrique ne peut pas être déroulée linéairement par voisinage. En les localisant, il est possible de délimiter les zones où le développement de franges par propagation est possible.

Certaines approches reposant sur ce formalisme, telles que celles exposées dans (Huntley, 1989), (Towers *et al.*, 1989), (Judge *et al.*, 1992), (Monti-Guarnieri *et al.*, 1993), (Judge, 1994), (Cusack *et al.*, 1995), (Buckland *et al.*, 1995), (Just *et al.*, 1995), résolvent le problème en associant des branches à ces résidus. D'autres sont guidées par des informations données par des cartes de qualité sur l'interférogramme, soit en exploitant uniquement cette information pour dérouler la phase (Bone, 1991), (Quiroga *et al.*, 1995), (Lim *et al.*, 1995), (Xu et Cumming, 1996), soit en utilisant les cartes pour définir rigoureusement les masques (Prati *et al.*, 1990), (Derauw, 1995). Ces cartes, appelés *mask-cuts* dans (Flynn, 1990), indiquent la qualité de la mesure de phase pour pouvoir développer localement les franges. Elles sont construites par exemple à partir des mesures de phase par pseudo-corrélation (Ghiglia et Pritt, 1998) ou suite à des traitements de type filtrage de l'interférogramme et prise en compte des données d'amplitude (Trouvé, 1996).

Ces méthodes locales restent néanmoins sensibles au bruit de phase caractéristique des données InRSO et à l'information donnée par les cartes de qualité. En cas d'une erreur survenue au cours du processus de développement de franges, cette dernière sera propagée !

Les méthodes globales (ou régionales) reposent principalement sur les trois approches suivantes :

1. La première est la segmentation de l'interférogramme pour localiser les franges. Elle est réalisée en détectant des contours qui correspondent aux lignes de fort contraste (c.-à-d. les limites des franges de phase) (Lin *et al.*, 1992), (Lin *et al.*, 1994). La phase est alors déroulée en ajoutant ou retranchant un multiple de 2π à chaque fois qu'une de ces lignes est franchie. Cette méthode est limitée aux images où la détection des limites de franges est fiable.
2. La deuxième est la minimisation de norme L^p (Ghiglia, 1996) qui consiste en fait à minimiser la différence entre les gradients de la phase mesurée et la solution qu'on cherche.

Le cas de la norme L^2 , qui correspond à l'approche de déroulement par moindres carrées, a été étudié dans (Fried, 1977), (Frost *et al.*, 1979), (Ghiglia et Romero, 1994),

(Fornaro *et al.*, 1996), (Pritt, 1996), (Trouvé, 1996). Un algorithme de minimisation exacte de ce critère a été proposé dans (Bioucas-Dias et Leitao, 2002) reposant sur les algorithmes de flot dans les réseaux. Ensuite, une version plus rapide a été proposée par le même auteur dans (Bioucas-Dias et Valadao, 2007) reposant sur la technique de coupe-minimum sur graphes binaires. En revanche, le point faible de ce critère est sa tendance à lisser la reconstruction. Les discontinuités de la scène peuvent être détruites si des poids qui leur correspondent ne sont pas considérés dans ce critère. L’approche par moindres carrées pondérées est alors une solution possible à ce problème (Ghiglia et Romero, 1994) (Trouvé, 1996).

Le cas de la norme L^1 , qui préserve mieux les discontinuités, a été étudié indépendamment dans (Flynn, 1997) et (Costantini, 1998) tout en proposant des algorithmes d’optimisation exacte reposant aussi sur la théorie des flots dans les réseaux.

Le cas des normes où $0 \leq p < 1$ est encore meilleur pour préserver les discontinuités, sauf que le problème de minimisation devient NP-difficile (Chen, 2001). Des algorithmes de minimisation approximative sont alors nécessaires. Un compromis entre qualité de l’optimum et complexité de l’optimisation devrait être pris en compte.

3. La troisième est l’approche bayésienne reposant sur le formalisme markovien pour régulariser la solution tout en préservant au mieux les discontinuités du relief (Dupont et Berthod, 1994), (Leitao et Figueiredo, 1995), (Labrousse, 1996), (Nico *et al.*, 2000), (Bioucas-Dias et Leitao, 2001), (Ferrauiuolo *et al.*, 2004), (Bioucas-Dias et Valadao, 2007), (Bioucas-Dias *et al.*, 2008), (Valadao et Bioucas-Dias, 2009). L’énergie est formulée selon un modèle d’attache aux données et un modèle *a priori* intégrant des connaissances contextuelles. La reconstruction de phase est obtenue par minimisation de cette fonction d’énergie. Les performances de ce type d’approches sont fortement liées à l’adéquation du modèle établi au problème et aux données interférométriques (c.-à-d. tenir compte du bruit InRSO, des discontinuités, etc.) et à la performance de l’algorithme de minimisation utilisé. Néanmoins, les résultats obtenus par ce type d’approches, comparés à ceux que donnent les approches locales, sont nettement meilleurs en terme de précision, de robustesse et même en complexité algorithmique.

Dans la suite, nous nous focalisons sur les approches bayésiennes de déroulement de phase. Nous décrivons quelques unes des approches déjà proposées dans la littérature, tout en mettant l’accent sur leur apport et limites dans ce domaine.

4 Approches bayésiennes de reconstruction de phase InRSO

Les approches de reconstruction de phase InRSO par estimation MAP nécessitent la définition de modèles de vraisemblance tirés des statistiques sur les données InRSO et des modèles *a priori* décrivant l’information contextuelle sur ces images ainsi que le développement d’algorithmes d’estimation efficace du critère MAP. On propose dans cette section de couvrir en détail ces différentes phases à travers les travaux qui ont adressé ce problème dans la littérature.

4.1 Statistiques des données interférométriques

Les statistiques des estimateurs des données InRSO, au sens du Maximum de vraisemblance, dans le cas de chatoiement pleinement développé ont été étudiées dans différents

travaux (Goodman, 1963), (Lee *et al.*, 1994), (Seymour et Cumming, 1994). Ces statistiques sont tirées de l'analyse multivariée complexe. Nous renvoyons le lecteur à l'ouvrage (Goodman, 2006) pour une étude approfondie. Elles peuvent être développées selon la nature des données RSO que nous disposons : monovues, multivues et polarimétriques.

Remarque : Notons que vu leur complexité, certains auteurs ont développé des modèles approximatifs plus simples à mettre en place (Rodriguez et Martin, 1992), mais qui ne sont valables que sous réserve d'avoir un nombre suffisant de vues et une bonne cohérence (Chaabane, 2004).

Par la suite, on considère deux images RSO réalisation d'un couple interférométrique $Z = [Z_1 Z_2]$ à chatoiement pleinement développé, c.-à-d. respectant chacune le modèle gaussien circulaire symétrique.

Les calculs seront détaillés sur un élément $z = [z_1 z_2]$ du vecteur de réalisations Z . Nous conservons les notations suivantes :

$$z_1 = a_1 e^{j\psi_1}, \quad (4.20)$$

$$z_2 = a_2 e^{j\psi_2}, \quad (4.21)$$

où a_1 et a_2 correspondent respectivement aux amplitudes des images complexes z_1 et z_2 et ψ_1 et ψ_2 leurs phases.

En considérant les lois gaussiennes circulaires complexes des distributions de z_1 et z_2 , les distributions conjointes et marginales des éléments de ces vecteurs sont détaillées dans l'annexe E. Nous ne gardons pour la suite que la distribution sur la phase interférométrique empirique ψ donnée en fonction de la phase interférométrique réelle ϕ et du degré de cohérence γ par l'expression :

$$P(\psi|\phi, \gamma) = \frac{1 - \gamma^2}{2\pi} \frac{1}{1 - \gamma^2 \cos^2(\psi - \phi)} \left(1 + \frac{\gamma \cos(\psi - \phi) \arccos(-\gamma \cos(\psi - \phi))}{\sqrt{1 - \gamma^2 \cos^2(\psi - \phi)}} \right). \quad (4.22)$$

A une valeur de cohérence fixe, cette fonction de vraisemblance est 2π -périodique et atteint ses maxima à $\psi = \phi$ et $\psi = \phi + 2\pi$.

En considérant l'intervalle de taille 2π centré sur la phase absolue ϕ ($[\phi - \pi, \phi + \pi]$), l'espérance et la variance de cette distribution valent respectivement

$$\begin{cases} E[\psi] &= \phi, \\ \sigma_\psi^2 &= E[(\psi - \phi)^2] = \frac{\pi^2}{3} - \pi \arcsin(\gamma) + \arcsin^2(\gamma) - \frac{Li_2(\gamma)}{2}, \end{cases} \quad (4.23)$$

où $Li_2(\cdot)$ est le dilogarithme d'Euler.

La figure 4.7(a) montre la courbe de la densité de probabilité de la phase interférométrique en fixant la phase interférométrique réelle ϕ à 0 et en variant la cohérence γ de 0 à ~ 1 . On remarque que pour une cohérence nulle, c.-à-d., z_1 et z_2 sont deux signaux complètement décorrélés, la phase est distribuée uniformément. Dans le cas inverse, c.-à-d. une cohérence maximale ($\gamma \rightarrow 1$), la distribution se rapproche de celle de Dirac.

Pour ces mêmes valeurs de γ , la figure 4.7(b) trace les valeurs des écarts type σ_ψ associés. On remarque alors la dépendance entre cohérence et bruit interférométrique : plus la cohérence est forte plus l'écart type du bruit interférométrique sur la phase empirique ψ est faible.

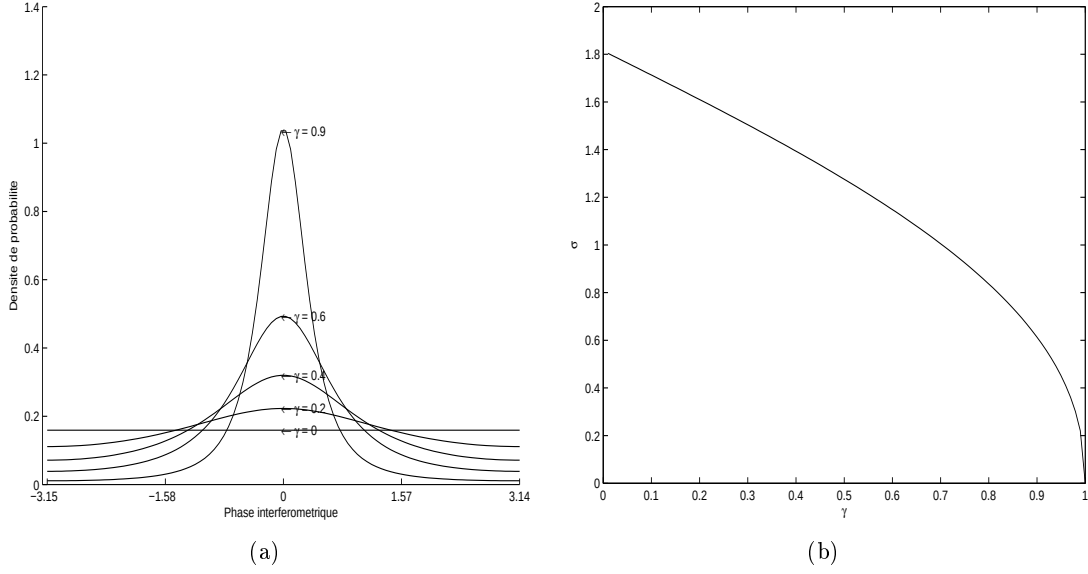


FIGURE 4.7 – Allure des courbes (a) de la loi de densité de probabilité de la phase interférométrique empirique pour une vraie valeur de phase nulle et (b) de l'écart type σ_ψ de la phase interférométrique empirique, en fonction de la cohérence γ .

4.2 Ambiguïté de la phase InRSO

Revenons au problème de déroulement de la phase qui consiste à partir des mesures de l'interférogramme dans l'intervalle $[-\pi, \pi[$ pour reconstruire les altitudes des cibles sur l'image dans leur vrai intervalle $[h_{min}, h_{max}]$, où h_{min} et h_{max} correspondent respectivement aux hauteurs minimale et maximale que peut prendre l'altitude d'une cible. On supposera que ces deux quantités sont *a priori* connues et que les valeurs des cohérences $\{\gamma_p, p \in \mathcal{P}\}$ le sont aussi (estimées).

Une première approche exploitant les statistiques sur la phase InRSO serait l'estimation d'une image de phases dans le nouvel intervalle $[\phi_{min}, \phi_{max}]$ qui maximise la vraisemblance avec les données mesurées. Rappelons que :

$$\phi_{min} = 2\pi \frac{h_{min}}{h_{amb}}, \quad (4.24)$$

$$\phi_{max} = 2\pi \frac{h_{max}}{h_{amb}}. \quad (4.25)$$

Notons par $\tilde{\phi}$ cette image estimée par Maximum de Vraisemblance (MV). Elle est obtenue par :

$$\tilde{\phi} = \arg \max_{\phi} P(\psi|\phi) \quad (4.26)$$

En considérant l'hypothèse d'indépendance des observations individuelles sur les pixels de l'interférogramme, on a :

$$\tilde{\phi}_p = \arg \max_{\phi_p} P(\psi_p|\phi_p); \forall p \in \mathcal{P} \quad (4.27)$$

Dans la figure 4.8, la courbe de la fonction de vraisemblance ($f^{MV} : \phi_p \mapsto P(\psi_p|\phi_p)$) est tracée pour une cible p de vraie hauteur 50m et pour h_p estimée variant sur l'intervalle

$[h_{min}, h_{max}] = [0\text{ m}, 200\text{ m}]$. Les paramètres d'acquisition considérés sont ceux d'un système InRSO avec un capteur ERS-2 : $\{R_1 = 825669\text{ m}; \lambda = 5.66\text{ cm}; \theta = 23^\circ; B^\perp = 150\text{ m}\}$. La cohérence γ est fixée forte à 0.9 pour déduire l'effet du bruit de décorrélation.

Maximiser ainsi f^{MV} serait possible si les cibles de la zone imagée sont d'altitudes inférieures à la hauteur d'ambiguïté qui vaut ici 61m, étant donnée que sur une période de 61m, un seul maximum de cette fonction existe. Si cette contrainte est non vérifiée, l'estimation est ambiguë compte tenu de la périodicité de la fonction et donc de la présence de plusieurs maxima.

L'ambiguïté est donc principalement due au paramètre h^{amb} . Plus il est faible par rapport aux hauteurs des cibles, plus l'ambiguïté de reconstruction des altitudes est élevée. En pratique, ce paramètre dépend essentiellement de la taille de la base (et donc de $B^\perp$) et de la longueur d'onde (λ). En contrepartie, plus il est faible et plus la mesure quantifiée est précise puisque les différents niveaux de phase correspondent à un plus petit intervalle de hauteur.

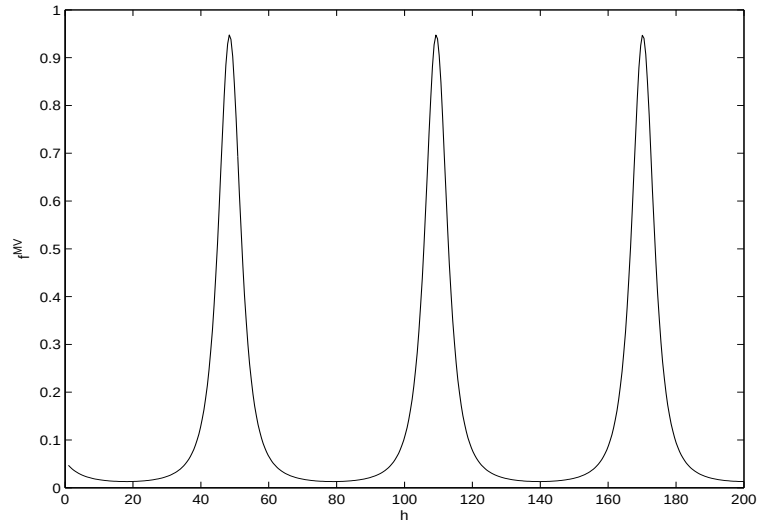


FIGURE 4.8 – Allure de la fonction de vraisemblance f^{MV} sur $[h_{min}, h_{max}]$.

Une approche possible qui permet de réduire et même enlever cette ambiguïté est le recours à plusieurs observations indépendantes de la même cible. Ces observations sont obtenues par des acquisitions InRSO à différentes bases ou fréquences et donc à différents paramètres h^{amb} . Cette approche est connue par le nom de *l'approche multicanal*, initialement introduite dans (Madsen et Zebker, 1992), (Xu *et al.*, 1994), (Ghiglia et Wahl, 1994). L'information de phase apportée par plusieurs canaux (ou observations indépendantes) est exploitée. La phase absolue est ainsi reconstruite sans ambiguïté par maximum de vraisemblance, si un nombre suffisant de canaux est disponible (Lombardini, 1996), (Corsini *et al.*, 1997), (Pascasio et Schirinzi, 2002), (Fornaro *et al.*, 2006).

4.3 Reconstruction de la phase InRSO par l'approche multicanal

Pour illustrer cette approche, nous considérons dans la suite au lieu d'un seul couple interférométrique $[z_1 z_2]$ un ensemble de couples interférométriques $\{[z_1^1 z_2^1], [z_1^2 z_2^2], \dots, [z_1^M z_2^M]\}$, donnant différents interférogrammes. Ces données pourront être générées par deux procédés différents :

- Le premier, fournissant des données interférométriques *multifréquences* (cf. FIG. 4.9), consiste à utiliser plusieurs capteurs opérant à des bandes de fréquences différentes. Dans (Pascazio et Schirinzi, 2002), il a été montré qu’il est possible de disposer de telles données à partir d’un seul capteur, en partitionnant le spectre de Fourier obtenu en sous-bandes sans chevauchement et selon certaines conditions pour réduire l’ambiguïté sur les mesures de phases (Madsen et Zebker, 1992), (Schmitt et Wiesbeck, 1997). En fait, la décomposition du spectre en sous-bandes sur l’axe de visée nécessite des conditions sur les fréquences centrales des sous-bandes pour obtenir de telles mesures de phase. Dans (Xu *et al.*, 1994), reposant sur le théorème des restes chinois, il a été démontré que pour des paires de fréquences choisies (f_i, f_j) vérifiant que $f_i = \frac{p}{q}$ et $f_j = \frac{r}{s}$, avec $\{p, q, r, s\}$ des entiers premiers entre eux, l’ambiguïté est désormais réduite par un certain facteur en fonction de ces paramètres et des ambiguïtés d’origine. Cependant, cela n’implique pas la résolution du problème de déroulement qui nécessite d’enlever complètement l’ambiguïté sur les mesures de phase. Dans (Pascazio et Schirinzi, 2002), le spectre en azimuth est aussi découpé en sous-bandes sans chevauchement pour réduire la variance du bruit sur la phase estimée après déroulement. Cette décomposition génère en réalité différentes vues sur les données, pour chaque fréquence utilisée. En revanche, elle entraîne une perte en résolution spatiale qui est en fait proportionnelle à la bande du spectre considéré (c.f. équation 4.4) (Ulaby *et al.*, 1986). Ainsi un compromis entre le nombre de sous-bandes et la résolution spatiale des images doit être pris en compte lors de cette décomposition.

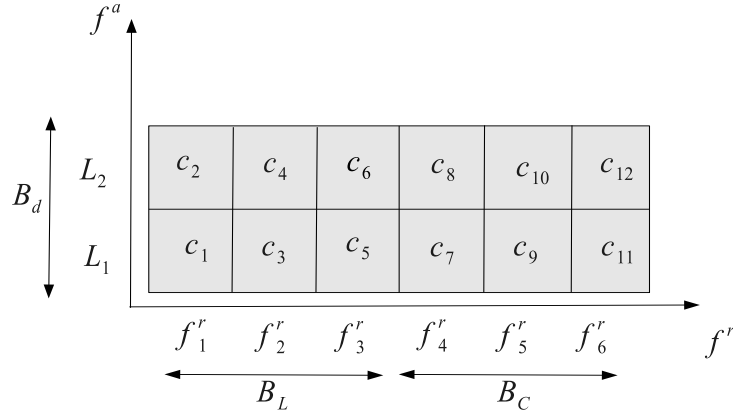


FIGURE 4.9 – Configuration multifréquence avec partitionnement du spectre de Fourier d’une image RSO en 12 sous-bandes $\{c_1, c_2, \dots, c_{12}\}$ en utilisant un partitionnement de deux bandes de fréquences B_L (Bande L) et B_C (Bande C) en trois fréquences différentes chacune $\{f_1^r, f_2^r, f_3^r\}$ et $\{f_4^r, f_5^r, f_6^r\}$ sur l’axe des fréquences en visée f^r et en générant deux vues L_1 et L_2 pour chacune des fréquences considérées dans la bande de Doppler B_d en azimuth.

- Le deuxième, fournissant des données interférométriques *multibases* (cf. FIG. 4.10), procède en considérant des images RSO prises avec des angles de vues légèrement

différents (Lombardini, 1996). Cette technique est le plus souvent considérée dans le cas réel d'une approche multicanal.

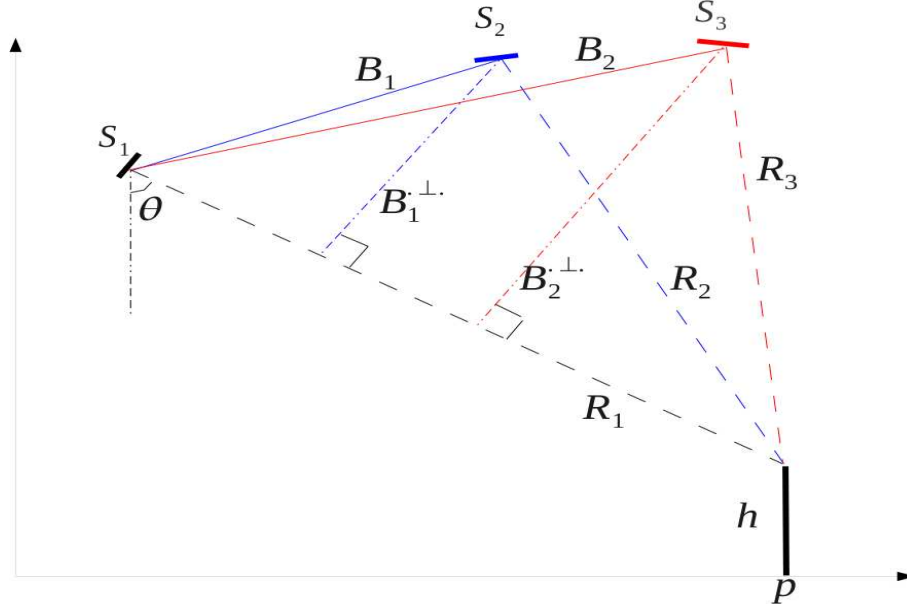


FIGURE 4.10 – Configuration multibase par acquisitions interférométriques en deux bases différentes B_1 et B_2 .

Selon le choix entre les deux configurations, la phase mesurée en un pixel p par un canal c vérifie les équations (4.12) et (4.14). Elle est en relation avec la vraie hauteur du relief par l'expression suivante :

$$\psi_c = \langle 2\pi \frac{h}{h_c^{amb}} \rangle_{2\pi}, \quad (4.28)$$

où h_c^{amb} est la hauteur d'ambiguïté qui varie d'un canal à un autre, et dont l'expression est définie selon le procédé multicanal choisi

$$h_c^{amb} = \begin{cases} \frac{\lambda}{2} \frac{R_1 \sin(\theta)}{B_1^\perp} & \text{dans le cas multibase,} \\ \frac{\lambda_c}{2} \frac{R_1 \sin(\theta)}{B^\perp} & \text{dans le cas multifréquence.} \end{cases} \quad (4.29)$$

Pour chaque canal c , la distribution de phase peut être alors exprimée par la distribution (4.22) en tenant compte de l'indice du canal c :

$$P(\psi_c | \phi_c, \gamma_c) = \frac{1}{2\pi} \frac{1 - \gamma_c^2}{1 - \gamma_c^2 \cos^2(\psi_c - \phi_c)} \left(1 + \frac{\gamma_c \cos(\psi_c - \phi_c) \arccos(-\gamma_c \cos(\psi_c - \phi_c))}{\sqrt{1 - \gamma_c^2 \cos^2(\psi_c - \phi_c)}} \right), \quad (4.30)$$

où, γ_c est la cohérence de l'interférogramme du canal c et $\phi_c = 2\pi \frac{h}{h_c^{amb}}$ est la phase absolue à estimer relativement au canal c .

Comme les canaux sont supposés indépendants, la fonction de vraisemblance multicanal modélisant la distribution de la phase en tenant compte des différentes observations indépendantes est obtenue alors par produit des fonctions de vraisemblances sur les différents canaux, c.-à-d.

$$P(\psi_p|\phi_p, \gamma_p) = \prod_{c=1}^M P(\psi_{c,p}|\phi_{c,p}, \gamma_{c,p}), \quad (4.31)$$

avec M le nombre total de canaux disponibles et $\psi_p = [\psi_{1,p}, \psi_{2,p}, \dots, \psi_{M,p}]$.

Il est à préciser que ce modèle n'est vrai que sous l'hypothèse d'indépendance entre les canaux. En pratique, le recours à l'approche multifréquence en disposant de capteurs opérant à différentes bandes permet d'avoir des mesures considérées statistiquement indépendantes. Cependant, en disposant d'un seul capteur et en décomposant le spectre en sous-bandes ou en utilisant une approche multibase avec ce même capteur, l'indépendance entre les canaux n'est plus toujours vérifiée. Cela nécessite, dans la mesure du possible, un choix judicieux des fréquences ou des bases. Dans le cas multibase, les auteurs de (Meglio *et al.*, 2006) développent le modèle conjoint pour 3 bases (B_{12} , B_{13} et $B_{2,3}$) disposant d'une image maîtresse (S_1) et deux autres esclaves (S_2 et S_3). On montre que pour des petites bases l'hypothèse d'indépendance n'est plus valable. En revanche, pour des grandes bases les deux modèles conjoint et avec indépendance entre les canaux ont un comportement très proche. Dans ce dernier cas de figure, il est alors plus pratique d'exploiter le modèle avec une distribution indépendante entre les canaux, qui s'avère beaucoup moins complexe à déterminer.

La figure 4.11 montre l'allure des fonctions de vraisemblance monocanal pour chacune des bases utilisées pour les 3 acquisitions ainsi que les courbes de la fonction de vraisemblance multicanal avec deux ensuite trois canaux pour une cible donnée p de vraie hauteur 170m. On dispose dans cet exemple de trois bases $\{B_1^\perp = 60m, B_2^\perp = 120m, B_3^\perp = 150m\}$, pour un système InRSO opérant avec les mêmes paramètres capteurs qu'un ERS-2 ($R_1 = 825669m$, $\theta = 23^\circ$ et $\lambda = 5.66cm$). La vraie hauteur de la cible est donc ambiguë pour les 3 canaux puisque les hauteurs d'ambiguïté respectives sont 152m, 76m et 60m, ce que montrent aussi les figures 4.11(a), 4.11(b) et 4.11(c). Elle l'est aussi par maximum de vraisemblance entre les deux premiers canaux. En effet, en reprenant le travail de (Xu *et al.*, 1994), et en se plaçant dans l'hypothèse de l'absence ou de la présence faible de bruit interférométrique, comme les valeurs des bases $B_1^\perp = 60m$ et $B_2^\perp = 120m$ ne sont pas premières entre elles, l'ambiguïté est uniquement réduite au *ppcm* entre 152m et 76m, qui vaut ici 152m. On retrouve alors l'ambiguïté du premier canal. Par contre, en utilisant le premier et le dernier canal, l'ambiguïté est réduite beaucoup plus pour atteindre une hauteur de *ppcm*(152, 60) = 2280m. Ce qui explique la localisation d'un seul maximum de la courbe de vraisemblance multicanal dans l'intervalle $[0m, 200m]$. Exploiter les trois canaux ensemble est encore une solution plus sûre pour réduire d'avantage l'ambiguïté et améliorer en précision le pic correspondant au maximum de vraisemblance multicanal, surtout en présence d'un bruit interférométrique.

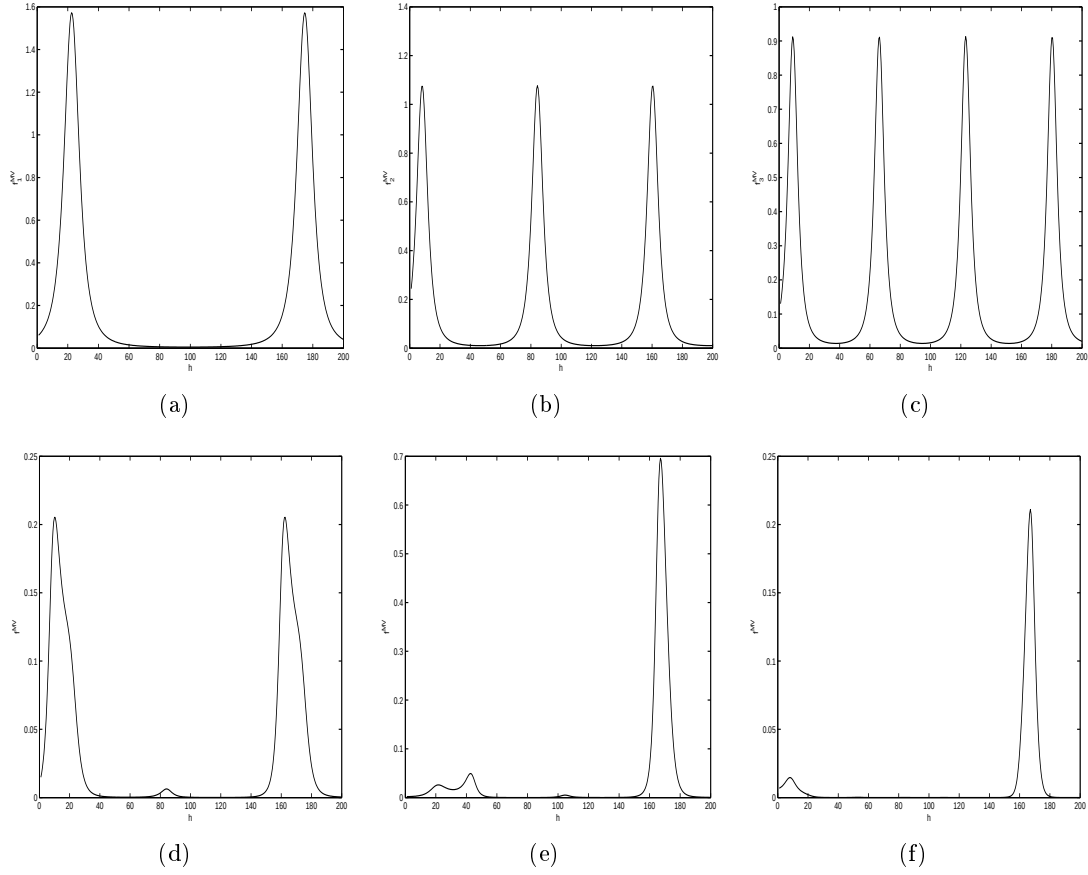


FIGURE 4.11 – Allure des courbes de vraisemblance monocanal pour les trois canaux disponibles et multicanal pour deux et trois canaux ensembles. (a) courbe de vraisemblance associée à la base $B_1^\perp$, (b) courbe de vraisemblance associée à la base $B_2^\perp$, (c) courbe de vraisemblance associée à la base $B_3^\perp$, (d) courbe de vraisemblance multicanal en utilisant les deux premiers canaux ensemble, (e) courbe de vraisemblance multicanal en utilisant le premier et le dernier canal ensemble, (f) courbe de vraisemblance multicanal en utilisant les trois canaux disponibles.

Ainsi en disposant d'un nombre suffisant de canaux et quelque soit l'altitude vraie de la cible dans l'intervalle $[h_{min}, h_{max}]$ où l'ambiguïté est enlevée par l'ensemble de ces canaux, estimer la vraie hauteur d'une cible à partir d'observations indépendantes (mesures de phases) devient un problème moins complexe à résoudre. D'où l'intérêt de cette approche pour la reconstruction du relief.

Récemment les méthodes bayésiennes de reconstruction de la phase InRSO ont été étudiées dans le cas de l'approche multicanal, afin d'aboutir à une reconstruction de la phase absolue déroulée et aussi régularisée. Nous présentons dans la suite deux méthodes récentes proposées dans la littérature pour le déroulement et la régularisation de la phase InRSO par une approche bayésienne. La première se place dans le cas monocanal, où l'ambiguïté sur la phase nécessite un algorithme performant pour la dérouler et la débruiter. La deuxième se place dans le cas multicanal, où l'ambiguïté et le bruit sur la phase sont alors réduits pour obtenir une reconstruction de relief plus précise. Ces méthodes vont nous servir dans les prochains chapitres pour comparer et évaluer les résultats obtenus par nos

approches.

4.4 Algorithmes de reconstruction de phase InRSO

4.4.1 Algorithme de reconstruction de phase par une approche monocanal

Cet algorithme, proposé récemment dans (Valadao et Bioucas-Dias, 2009), repose sur un modèle *a posteriori* pour dérouler et débruiter la phase InRSO simultanément tout en exploitant les performances de la technique d'estimation MAP par coupe-minimum sur graphe.

Le modèle de vraisemblance utilisé est le modèle de vraisemblance monocanal conjoint associé au vecteur $z = [z_1 z_2]$ décrit dans le paragraphe 4.1.

Les potentiels d'interaction proposés dans le modèle *a priori* sont considérés comme efficaces pour préserver les discontinuités ainsi que les zones lisses de la phase. Il s'agit de potentiels de la forme quadratique-tronquée :

$$E_{p,q}(\phi_p - \phi_q) = \begin{cases} (\phi_p - \phi_q)^2 & \text{si } |\phi_p - \phi_q| \leq \pi \\ \pi^2 - \pi^p + |\phi_p - \phi_q|^p & \text{sinon} \end{cases}, \quad (4.32)$$

avec $0 < p < 1$.

Vu la non convexité de ces potentiels, uniquement des optima locaux approximatifs sont obtenus avec les optimiseurs par coupe-minimum sur graphe. L'algorithme d'estimation mis en place s'effectue d'une manière itérative avec un schéma multirésolution adapté, permettant d'améliorer la précision de la solution d'une itération à une autre.

Cette approche évaluée en terme de complexité de calcul est bien efficace. Cependant, le recours à l'approche monocanal présente toujours des limites en présence de données InRSO faiblement corrélées prises sur des zones à fortes discontinuités de profil ; d'autant plus que la minimisation approximative de l'énergie ne permet pas d'exploiter pleinement la robustesse du modèle de régularisation pour bien dérouler la phase dans les endroits ambigus.

4.4.2 Algorithme de reconstruction de phase par une approche multicanal

Cette approche, proposée dans (Ferrauiuolo *et al.*, 2004), permet de reconstruire la phase InRSO en exploitant la technique multicanal.

Le modèle d'attache aux données utilisé repose sur la fonction de vraisemblance multicanal décrite dans le paragraphe 4.3 pour enlever l'ambiguïté sur la phase à dérouler.

Le modèle *a priori* utilisé est le modèle markovien-gaussien local, que l'on notera par *LGMRF* (Saqib *et al.*, 1998), de densité de probabilité :

$$P(\phi|\sigma) \propto \exp \left(-\beta \sum_{(p,q)} \frac{(\Phi(p) - \Phi(q))^2}{2\sigma_{p,q}^2} \right), \quad (4.33)$$

où $\sigma_{p,q}$ est un hyperparamètre local mesurant la corrélation spatiale entre les deux sites voisins p et q . Plus l'hyperparamètre $\sigma_{p,q}$ est fort, plus la probabilité pour que les deux sites p et q aient des valeurs différentes est élevée. Ces hyperparamètres locaux permettent donc d'introduire de l'information concernant la discontinuité sur l'image des phases à reconstruire.

L'estimation de la phase ϕ par MAP est effectuée grâce à un algorithme itératif où chaque itération est composée de deux étapes. La première étape consiste à estimer les

hyperparamètres locaux σ et la deuxième consiste à estimer la phase ϕ connaissant les hyperparamètres locaux du modèle *a priori*.

Les auteurs exploitent l'approche d'estimation par maximum de vraisemblance pour trouver ces hyperparamètres. Cependant l'estimation de σ nécessite la connaissance *a priori* de ϕ . Ainsi l'idée proposée est d'estimer σ à l'itération k en utilisant l'estimation de ϕ à l'itération $k - 1$ par EM (Expectation Minimization). Une fois l'estimation de ces paramètres est faite, l'énergie est ensuite minimisée par l'algorithme ICM. L'algorithme complet converge après un nombre fini d'itérations, quand la phase estimée change très faiblement.

Cette approche, comparée à la précédente, permet d'apporter une meilleure robustesse dans la modélisation mais au prix d'un algorithme d'estimation très coûteux en temps de calcul. En effet, l'ambiguïté sur la phase est cette fois-ci réduite ou enlevée (selon le nombre de canaux utilisés et le degré d'ambiguïté) grâce au recours à l'approche multicanal. Par ailleurs, le modèle *a priori* exploité est considéré comme un modèle de régularisation robuste dans le sens où il tient compte à la fois des discontinuités ainsi que des zones lisses de la phase à reconstruire. Cependant, il nécessite une étape lourde d'estimation des hyperparamètres locaux.

Ces deux approches que nous venons de décrire sont à notre sens parmi les techniques de reconstruction de la phase InRSO qui ont montré de bonnes performances dans la littérature. Elles ont été aussi testées et validées durant nos travaux de thèse sur un ensemble de données et ont montré des résultats intéressants.

Nous notons par ailleurs que des performances similaires pourraient aussi être obtenues par la technique des moindres carrées pondérées. Cependant, pour une telle approche, des poids devront être bien estimés auparavant pour avoir de bons résultats. Ainsi deux étapes sont nécessaires pour une telle approche, comme c'est le cas dans (Trouvé, 1996), où une étape de filtrage de l'interférogramme est au départ réalisée pour à la fois enlever le bruit sur la phase et pouvoir estimer (également avec l'amplitude) les poids qui seront injectés par la suite dans l'étape de déroulement de la phase par moindres carrées.

5 Conclusion

Un aperçu général des méthodes de reconstruction de relief par interférométrie RSO a été proposé. Certains des algorithmes de reconstruction de phase InRSO par des approches markoviennes, proposés récemment dans la littérature, sont présentés et analysés. Ces algorithmes, comparés aux approches classiques locales, sont considérés plus efficaces et robustes du fait du recours à une approche d'estimation globale faisant intervenir des modèles de régularisation locaux. Par ailleurs, nous avons soulevé certains problèmes qui s'avèrent en pratique importants. Ces problèmes touchent essentiellement aux deux critères d'évaluation : la complexité algorithmique et la robustesse de la modélisation. Ces deux critères agissent respectivement sur l'efficacité de l'approche et la qualité de la solution obtenue.

Dans ces travaux de thèse, nous proposons des approches de reconstruction de phase InRSO présentant un compromis entre ces deux critères. La technique multicanal est exploitée et combinée avec des modèles *a priori* locaux et non-locaux aboutissant à des modèles de reconstruction du relief qui sont robustes aux artéfacts des mesures de phase interférométrique. Pour réduire la complexité, nous avons recours à des algorithmes de

minimisation d'énergies markoviennes par coupe-minimum sur graphes. Ces algorithmes qui sont développés dans la première partie de la thèse nous ont permis d'obtenir des optima de bonne qualité tout en gardant une faible complexité algorithmique. Les différents modèles et les algorithmes d'estimation dédiés seront détaillés dans les chapitres suivants.

Chapitre 5

Reconstruction de MNT par une approche markovienne multicanal

Une nouvelle approche de reconstruction de MNT est introduite dans ce chapitre. Elle est fondée d'une part sur un nouveau modèle markovien de déroulement et régularisation de phase simultanément et d'autre part, sur les algorithmes efficaces de minimisation markovienne de premier ordre par coupe-minimum sur graphe.

En premier lieu, le modèle markovien proposé est décrit. Ensuite, une étude du modèle multicanal proposé pour valider son comportement en présence de différents canaux à cohérences variées est réalisée. Cette étude n'étant pas possible théoriquement afin de tenir compte à la fois de la fonction de vraisemblance et de la fonction de régularisation et aussi de la diversité des milieux dans le cas réel, nous proposons une étude empirique. L'impact de la présence de canaux à fort bruit interférométrique sur la précision et la qualité de l'estimation d'un MNT synthétique sur une zone de reliefs est évalué. Ainsi nous jugerons de l'utilité ou non de fixer un critère de sélection des canaux pour bien exploiter le modèle proposé.

Ensuite, pour estimer les hyperparamètres locaux de la fonction de régularisation, l'algorithme de minimisation que nous décrivons et qui sera utilisé pour la reconstruction de l'altitude de la scène apportera une solution simple et efficace à ce problème. Les résultats seront par la suite comparés à ce que donne un algorithme d'estimation robuste des hyperparamètres locaux pour une fonction de régularisation de type LGMRF.

Confrontés à d'autres problèmes qui caractérisent les données réelles, tels que les artéfacts de type perturbations atmosphériques globales ou des décalages de phases différents d'un canal à un autre, qui rendent les données incohérentes pour être exploitées ensemble, on se propose alors d'ajuster le modèle multicanal pour tenir compte de telles perturbations. Un algorithme rapide de correction automatique de ces données tout en estimant conjointement la vraie altitude de la scène sera proposé.

Enfin, des résultats de la reconstruction de MNT sur des données InRSO réelles seront présentés et analysés. La qualité de la reconstruction sera comparée à celle d'un MNT SRTM (Shuttle Radar Topography Mission). Une discussion autour des problèmes d'ordre pratique rencontrés nous permettra de proposer des stratégies pour mieux exploiter les outils que nous proposons dans ces travaux de thèse sur des données interférométriques réelles quelconques.

Une partie de ce travail a été publiée dans la revue *IEEE Geoscience and Remote Sensing Letters*. Un premier article montre la robustesse du modèle élaboré comparé à d'autres modèles markoviens et l'efficacité de l'algorithme de reconstruction en utilisant des optimiseurs par coupe-minimum sur graphe de l'état de l'art (Ferraioli *et al.*, 2009).

Un deuxième article introduit l'algorithme de minimisation efficace mixte pour améliorer les performances de l'approche ainsi que la précision sur la phase reconstruite est en cours de révision (Shabou *et al.*, 2010a).

1 Modèle MCPU+TV

Afin de résoudre le problème de la reconstruction de la phase InRSO tout en tenant compte de certaines difficultés qui peuvent se présenter, comme le bruit interférométrique, les fortes discontinuités du milieu à reconstruire et la présence de zones de faibles cohérences, nous proposons un modèle d'énergie qui permet d'enlever l'ambiguïté sur la phase InRSO et de la régulariser. Ce modèle repose sur la vraisemblance multicanal de la phase InRSO et le modèle de régularisation par variation totale. Cela présente les avantages suivants :

- Disposant d'un ensemble d'interférogrammes sur la même zone (multibases ou multifréquences), le modèle de vraisemblance permet de reconstruire approximativement les profils de la scène imagée en retrouvant les multiples de 2π pour dérouler la phase. Le recours à différentes observations provenant des différents canaux permet d'apporter d'une part plus d'information au niveau des zones à faible cohérence, très bruitées et difficiles à reconstruire et d'autre part de réduire l'ambiguïté de la reconstruction de l'altitude des zones élevées ou présentant des fortes discontinuités.
- A ce modèle de vraisemblance, nous ajoutons une contrainte de régularité sur la phase à reconstruire en utilisant la TV comme *a priori*. Ce dernier modèle permet de conserver les discontinuités du milieu et d'exploiter les algorithmes efficaces de minimisation par coupe-minimum sur graphe.

Le modèle global proposé pour estimer l'image d'altitude h de la scène se présente alors comme suit :

$$E(h|\psi) = - \sum_p \sum_c \log(P(\psi_{c,p}|h_p)) + \sum_{(p,q)} w_{p,q} |h_p - h_q|, \quad (5.1)$$

où

$$P(\psi_{c,p}|h_p) = \frac{1}{2\pi} \frac{1 - \gamma_{c,p}^2}{1 - f_{c,p}^2(h_p)} \left(1 + \frac{f_{c,p}(h_p) \arccos(-f_{c,p}(h_p))}{\sqrt{1 - f_{c,p}^2(h_p)}} \right), \quad (5.2)$$

avec

$$f_{c,p}(h_p) = \gamma_{c,p} \cos \left(\psi_{c,p} - \frac{2\pi}{h_c^{amb}} h_p \right), \quad (5.3)$$

$\gamma_{c,p}$ est la cohérence empirique de l'interférogramme du canal c au site p , calculée à partir des couples RSO formant l'interférogramme ψ_c , h_c^{amb} est la hauteur d'ambiguïté relative au canal c et $w_{p,q}$ est un hyperparamètre local ajustant la régularisation localement entre deux sites voisins (p, q) , si une information de discontinuité est fournie. Généralement, en l'absence d'une telle information, cet hyperparamètre est fixé globalement sur toute l'image, pondérant ainsi l'effet de la régularisation par rapport à la vraisemblance aux données. Les autres notations présentes dans cette forme sont les mêmes que celles du chapitre précédent.

Le modèle établi repose sur deux types de paramètres :

- le nombre des canaux nécessaires pour dérouler correctement la phase InRSO,
- les hyperparamètres locaux utiles afin de mieux préserver les discontinuités du milieu.

L'estimation de ces paramètres permet d'obtenir une approche non-supervisée de reconstruction de la phase InRSO. Cependant, il s'avère qu'en pratique, avec la présence d'énormes variations entre différentes scènes et types de données InRSO, une telle estimation reste toujours une tâche complexe.

Pour ce faire, on se propose au début d'étudier l'influence des canaux à différentes cohérences sur la qualité de l'estimation de l'altitude pour pouvoir conclure sur le nombre et le choix des canaux. Ensuite, l'algorithme de minimisation de la fonction d'énergie établie dans (5.1) sera présenté. Cet algorithme apportera une solution efficace pour estimer les hyperparamètres locaux.

2 Étude de l'influence de la cohérence sur l'estimation multicanal

Dans cette étude, partant d'un modèle numérique de terrain, un ensemble d'interférogrammes multibases et multipasses selon les paramètres du satellite ERS-2 est généré. La zone présente trois milieux différents à cohérences variables. Le facteur de décorrélation est amplifié d'un canal à l'autre différemment entre les milieux.

Nous allons considérer indépendamment deux types de dégradations. La première série de dégradations est due uniquement à l'augmentation de la longueur de la base, tout en considérant une bonne cohérence temporelle sur les différents canaux. La deuxième série est plutôt due à une décorrélation temporelle entraînant un bruit de phase. Cette dernière série est générée uniquement à partir de deux bases.

Pour chacune des deux séries, nous reconstruisons itérativement l'altitude de la scène en ajoutant à chaque itération un nouvel interférogramme plus bruité que les précédents et en enregistrant à chaque fois l'erreur moyenne de reconstruction. Les résultats seront moyennés sur une dizaine de réalisations de la même expérience. Enfin, l'évolution de l'erreur sera étudiée pour conclure sur l'effet de la dégradation de la cohérence du canal sur l'estimation multicanal globale du relief.

Dans la figure 5.1, le MNT synthétique utilisé ainsi que le premier interférogramme (de plus forte cohérence) et le dernier interférogramme (de plus faible cohérence) pour chacune des deux séries sont présentés. La carte de cohérences associée à la scène montre bien la présence de trois milieux différents. Cela pourrait refléter un cas réel en présence de divers milieux dans une même scène : végétations, lacs, zones urbaines, etc.

A chaque canal, du premier au dernier (30 canaux sont utilisés), il s'agit des mêmes cartes de répartition des cohérences utilisées pour générer les deux séries d'interférogrammes. Cela nous permettra donc de comparer aussi l'impact d'une décorrélation temporelle à celle spatiale sur la qualité de la reconstruction par l'approche multicanal.

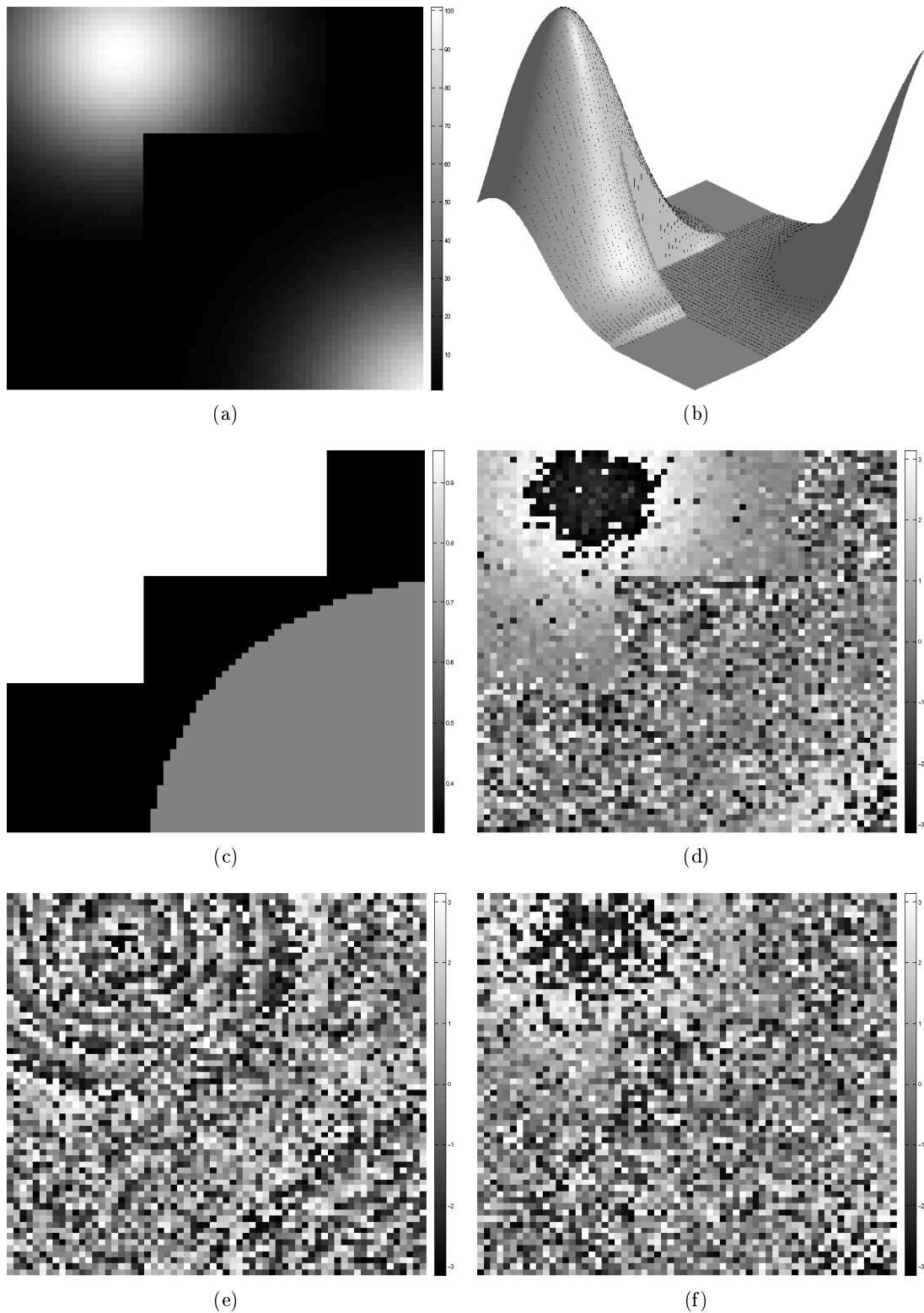


FIGURE 5.1 – Données interférométriques à partir d'un MNT synthétique. (a) Le MNT d'origine, (b) sa visualisation en 3D, (c) la carte de cohérences construite par milieu, (d) le premier interférogramme (le moins bruité) pour les deux séries de dégradations, les derniers interférogrammes (les plus bruités) pour (e) la première série (mutibase), et (f) la deuxième série (multipasse).

La figure 5.2 illustre l'évolution des erreurs moyennes de reconstruction pour les deux différentes séries lorsqu'on augmente le nombre de canaux pris en compte.

Nous remarquons que pour les deux types de dégradations, l'impact des canaux à faibles cohérences est plutôt positif sur la qualité de la reconstruction. En ajoutant autant d'interférogrammes à cohérences réduites, l'erreur de reconstruction diminue. Cela s'explique par le fait que le recours aux interférogrammes apporte une information de précision sur la reconstruction plus importante que le bruit qui les caractérise. De plus, le recours à la régularisation permet de réduire l'information de bruit non nécessaire au déroulement de la phase. Cependant, cet effet reste observable jusqu'à un facteur de décorrélation limite, au delà duquel un interférogramme très bruité n'apporte plus de l'information utile pour la reconstruction et peut même entraîner une légère dégradation des performances.

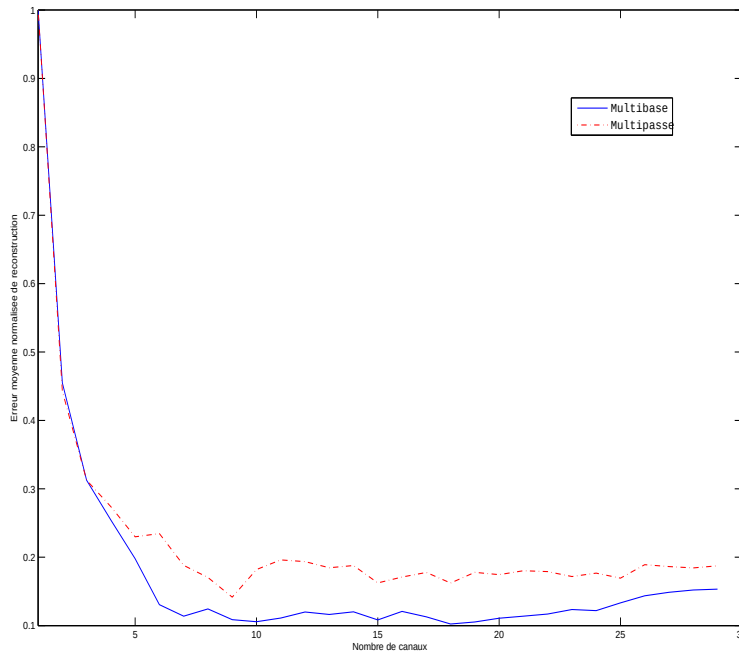


FIGURE 5.2 – Effet de la cohérence sur la qualité de l'estimation du MNT par une approche multicanal. En trait continu, uniquement la décorrélation spatiale est amplifiée en ajoutant des canaux à bases plus longues que les précédentes. En trait discontinu, en utilisant uniquement deux bases, le facteur de décorrélation est amplifié à chaque simulation pour faire correspondre les cohérences de cette série de reconstructions à celles de la première série.

Cette simple étude empirique nous laisse conclure que l'influence du choix ou de la sélection d'interférogrammes à exploiter par le modèle multicanal, parmi l'ensemble d'interférogrammes disponibles, est de faible impact sur la qualité finale de la reconstruction. Néanmoins, dans le cas de données réelles, il serait préférable de ne considérer que les canaux à faible décorrélation temporelle pour réduire le plus possible un impact négatif sur la reconstruction et aussi réduire l'utilisation de la mémoire physique nécessaire pour le stockage des mesures de phase à prendre en compte dans la fonction d'énergie.

3 Minimisation markovienne et estimation des hyperparamètres locaux

Ayant défini les observations à considérer dans le modèle de l'énergie, le problème de reconstruction se ramène à minimiser la fonction d'énergie établie. Ce problème de minimisation se caractérise par :

- des données volumineuses en pratique,
- une fonction d'énergie fortement non-convexe présentant plusieurs optima locaux,
- une information physique à reconstruire nécessitant une bonne précision.

Ainsi sa bonne résolution nécessitera un algorithme d'optimisation suffisamment efficace pour résoudre ces différentes problématiques.

En reprenant l'étude faite dans la première partie du manuscrit, l'optimisation mixte FastDC semble bien adaptée à ce problème d'optimisation. Dans la suite, les étapes de l'algorithme de reconstruction de la phase InRSO sont détaillées.

3.1 Étape 1 : optimisation discrète

Au cours de cette première phase d'optimisation, les mouvements de partitions larges et multi-étiquettes (MPLM) sont exploités. Une précision TF sur l'image de relief à reconstruire est au préalable fixée. Pour les mouvements de type extension ou échange, cette précision sert à quantifier l'intervalle d'étiquettes $[h_{min}, h_{max}]$. Pour les mouvements de type saut, elle sera utile pour définir les étiquettes itérativement. Le modèle *a priori* utilisé est la TV, où les hyperparamètres locaux sont fixés à une constante de régularisation :

$$w_{p,q} = \beta, \forall (p, q). \quad (5.4)$$

Dans la première partie du manuscrit de thèse, on a pu constater que les stratégies de mouvement de partitions par extension ou saut multi-étiquettes sont les plus performantes.

L'extension multi-étiquettes, nécessitant une quantification au préalable de l'ensemble des étiquettes, serait plus adaptée aux reliefs riches en structures constantes par morceaux, tel est le cas des milieux urbains. En présence de telles zones, la quantification des niveaux d'altitudes est d'un impact faible sur la qualité de l'estimation.

En revanche, face à de scènes naturelles telles que des zones montagneuses, on préfère un mouvement multi-étiquette de type saut, puisque ce dernier offre plus de liberté de choix des étiquettes dans $[h_{min}, h_{max}]$. D'autant plus qu'il est possible d'accélérer la convergence de l'algorithme d'optimisation reposant sur ce type de mouvements en proposant une stratégie de parcours à schéma multirésolution. Ainsi on explore mieux l'ensemble des configurations possibles grâce à ce mouvement par saut tout en accélérant la convergence vers de bons optima grâce à la technique multirésolution. Un exemple qui illustre cette dernière stratégie de parcours est présenté dans la figure 5.3. L'algorithme complet est présenté dans la figure 5.4.

cycle 1, ($\delta_{max} = 20, T = 5$)

$$\begin{array}{l} \mathcal{L}_4^{(1)} = \\ \mathcal{L}_4^{(2)} = \end{array} \begin{array}{|c|c|c|c|c|} \hline -10 & -5 & +0 & +5 & +10 \\ \hline -20 & -15 & +0 & +15 & +20 \\ \hline \end{array}$$

cycle 2, ($\delta_{max} = 10, T = 2.5$)

$$\begin{array}{l} \mathcal{L}_4^{(1)} = \\ \mathcal{L}_4^{(2)} = \end{array} \begin{array}{|c|c|c|c|c|} \hline -5 & -2.5 & +0 & +2.5 & +5 \\ \hline 10 & -7.5 & +0 & +7.5 & +10 \\ \hline \end{array}$$

cycle 3, ($\delta_{max} = 6.67, T = 1.67$)

$$\begin{array}{l} \mathcal{L}_4^{(1)} = \\ \mathcal{L}_4^{(2)} = \end{array} \begin{array}{|c|c|c|c|c|} \hline -3.33 & -1.67 & +0 & +1.67 & +3.33 \\ \hline -6.67 & -5 & +0 & +5 & +6.67 \\ \hline \end{array}$$

cycle 4, ($\delta_{max} = 5, T = 1.25$)

$$\begin{array}{l} \mathcal{L}_4^{(1)} = \\ \mathcal{L}_4^{(2)} = \end{array} \begin{array}{|c|c|c|c|c|} \hline -2.5 & -1.25 & +0 & +1.25 & +2.5 \\ \hline -5 & -3.75 & +0 & +3.75 & +5 \\ \hline \end{array}$$

cycle 5, ($\delta_{max} = 4, T = (1 = TF)$)

$$\begin{array}{l} \mathcal{L}_4^{(1)} = \\ \mathcal{L}_4^{(2)} = \end{array} \begin{array}{|c|c|c|c|c|} \hline -2 & -1 & +0 & +1 & +2 \\ \hline -4 & -3 & +0 & +3 & +4 \\ \hline \end{array}$$

FIGURE 5.3 – Exemple de parcours d'étiquettes par un mouvement type δ_4 -saut avec un schéma multirésolution, partant de la précision $T = 5$ à la précision finale $TF = 1$. Les ensembles d'étiquettes sont présentés pour chaque cycle.

En ce qui concerne la stratégie de parcours par des mouvements d'extension multi-étiquette, nous gardons le même schéma d'optimisation que celui présenté dans le second chapitre de ce manuscrit.

Remarque : il faut mentionner que l'optimisation par sauts à schéma multirésolution est en général moins optimale que l'optimisation par le mouvement de saut tel qu'il a été présenté au premier chapitre, où à chaque itération la nouvelle étiquette proposée à un site donné est obtenue en augmentant ou diminuant d'une unité l'étiquette courante et ainsi de suite jusqu'à atteindre une étiquette limite. Cette dernière stratégie permet en réalité d'explorer beaucoup plus de configurations et donc d'avoir des propriétés de convergence plus intéressantes. Cependant, en pratique, il s'est avéré que le schéma multirésolution aboutit souvent au même optimum que ce que donne un schéma normal de saut, mais avec un temps de calcul beaucoup plus réduit pour une même précision donnée. La même constatation a été aussi faite dans (Valadao et Biucas-Dias, 2009), où un autre schéma multirésolution a été adopté pour optimiser la fonction d'énergie proposée.

Le résultat final de l'optimisation est une image d'altitude reconstruite à partir de données de phases interférométriques, à la précision TF , que l'on notera par $\tilde{h}^{(1)}$.

```

Initialisation de  $\tilde{h}$ ,  $T$  et  $\delta_{max}$  ;
Pour chaque cycle  $c$  faire
     $E_c := E(\tilde{h})$  ;
    [Création de l'ensemble d'étiquettes de taille  $m$  par multirésolution]
     $\mathcal{L}_m = \text{MultiResolution}(m, \delta_{max}, T)$  ;
    fin := faux ;
    Répéter ;
        [Recherche du mouvement de partition optimal]
        
$$h_{opt} = \arg \min_h E(h) ,$$

        Si ( $E(h_{opt}) < E(\tilde{h})$ ) Alors
            |  $\tilde{h} := h_{opt}$  ;
        Sinon
            | fin := vraie ;
        Fin Si
    jusqu'à ce que (fin == vraie)
    [Vérifier la convergence sinon réduction des sauts et de la précision par multirésolution]
    Si ( $E(\tilde{h}) < E_c$  ET  $T > TF$ ) Alors
        |  $\delta_{max} := \frac{\delta_{max}}{c}$  ;  $T := \frac{T}{c}$  ;
    Sinon
        | retourner  $\tilde{h}$  ;
    Fin Si
Fin Pour
    
```

FIGURE 5.4 – Algorithme d'optimisation multi-étiquette par sauts multirésolutions.

3.2 Étape 2 : estimation des hyperparamètres locaux

L'estimation des hyperparamètres locaux qui caractérisent les discontinuités de la scène à reconstruire est une problématique d'une grande importance en imagerie. En effet, le recours à des modèles *a priori* non-locaux implique une régularisation globale de l'image pouvant détruire certaines structures fines. Par exemple, un *a priori* de type quadratique, malgré son adéquation avec les zones lisses, introduit du flou au niveau des discontinuités et bords des objets. Un *a priori* de type variation totale, malgré son adéquation avec les zones constantes par morceaux, introduit une perte de contraste entre deux zones adjacentes.

Des études ont été menées dans ce cadre proposant des fonctions de régularisation exprimant implicitement les discontinuités que pourrait avoir une image naturelle (Besag, 1986; Charbonnier, 1994). Sauf que de tels modèles *a priori* ne sont pas adaptés à la plupart des optimiseurs par coupe-minimum sur graphes, vu leur non sous-modularité, et en particulier à notre optimiseur MPLM qui nécessite la propriété de convexité de la fonction multi-étiquette *a priori*.

D'autres méthodes existent et qui reposent cette fois-ci sur les modèles de régularisation

globale, mais pondérés par un champ exprimant les discontinuités de l'image à reconstruire. Ce champs est communément appelé *processus de lignes* et introduit initialement dans (Geman et Geman, 1984). Son estimation conjointement avec l'image à reconstruire est donc nécessaire durant l'algorithme de minimisation. Les approches d'estimation, tel que l'algorithme *EM*, sont coûteuses en temps de calcul. D'ailleurs, un exemple de travaux utilisant cette technique d'estimation pour la reconstruction de la phase InRSO, montrant le coût de l'estimation de ces hyperparamètres, a été présenté dans le chapitre précédent (Ferraiuolo *et al.*, 2004).

Dans notre cas, nous proposons de les estimer avec un coût nettement réduit. En effet, l'image d'altitudes déjà reconstruite ($\tilde{h}^{(1)}$) est une image de structures élevées de la zone reconstruite à une précision *TF*. Ainsi toutes les structures ayant une élévation supérieure à la précision fixée sont reconstruites fidèlement par rapport au modèle utilisé, sous réserve que l'optimum atteint est de bonne qualité (ce qui est généralement le cas avec les optimiseurs MPLM). De plus, en prenant le scalaire β (5.4) du modèle markovien utilisé comme étant une constante faible, l'effet de perte de contraste par TV sera bien réduit au prix d'un taux de bruit qui caractérisera l'image finale $\tilde{h}^{(1)}$. Cela ne posera pas de problèmes, tant qu'une deuxième étape de régularisation sera effectuée partant de cette configuration. Ainsi il est possible à partir de $\tilde{h}^{(1)}$ d'estimer les discontinuités à préserver sur la solution finale en effectuant un seuillage. Le seuil défini décidera parmi les discontinuités qui existent dans $\tilde{h}^{(1)}$ celles qui correspondent à des structures à préserver et celles qui correspondent au bruit non encore éliminé. On parle alors de fortes discontinuités dans le premier cas et de faibles discontinuités dans le deuxième. Notant par μ ce seuil, les hyperparamètres locaux sont alors être estimés de façon binaire comme suit :

$$\begin{cases} w_{p,q} = 0 & \text{si } |\tilde{h}_p^{(1)} - \tilde{h}_q^{(1)}| > \mu, \\ w_{p,q} = 1 & \text{sinon.} \end{cases}$$

Ainsi un simple parcours de l'image reconstruite $\tilde{h}^{(1)}$ permet de localiser les fortes discontinuités à préserver dans la reconstruction finale. Le paramètre μ est fixé manuellement ou automatiquement en étudiant l'histogramme des discontinuités de $\tilde{h}^{(1)}$ (image des gradients de $\tilde{h}^{(1)}$). Cette méthode sera discutée dans les expériences et comparée à un estimateur robuste des hyperparamètres locaux.

3.3 Étape 3 : optimisation continue

Ayant ajusté le modèle de régularisation par les hyperparamètres locaux estimés, l'optimisation continue par descente de gradient est effectuée. Cette étape nécessite la connaissance du gradient de la fonction d'énergie à minimiser.

La fonction d'attache aux données est différentiable, son gradient est donné par :

$$\begin{aligned} \nabla E_p(h_p) &= \sum_c -\gamma_{p,c} \frac{2\pi}{h_c^{amb}} \sin(\psi_{p,c} - \frac{2\pi}{h_c^{amb}} h_p) \left(\frac{3f_{p,c}(h_p)}{1 - f_{p,c}^2(h_p)} \right. \\ &\quad \left. + \frac{\text{acos}(-f_{p,c}(h_p))}{f_{p,c}(h_p)\text{acos}(-f_{p,c}(h_p)) + \sqrt{1 - f_{p,c}^2(h_p)}} \right), \end{aligned} \quad (5.5)$$

avec

$$f_{p,c}(h_p) = \gamma_{p,c} \cos(\psi_{p,c} - \frac{2\pi}{h_c^{amb}} h_p). \quad (5.6)$$

La fonction de régularisation par TV, étant non-différentiable, est approximée selon la même technique utilisée que celle dans le chapitre 3.

A la convergence de l'algorithme, la solution finale $\tilde{h}$ est obtenue, proposant ainsi une reconstruction efficace de la phase absolue avec une meilleure précision que celle de $\tilde{h}^{(1)}$.

L'algorithme dans la figure 5.5 illustre ces trois étapes pour la reconstruction d'un MNT.

1. Optimisation MPLM : $\tilde{h}^{(1)} = \underset{h}{\operatorname{argmin}} E(h)$; (Hyperparamètres locaux constants (5.4))
2. Estimation des hyperparamètres locaux $\{w_{p,q}\}_{(p,q)}$ à partir de $\tilde{h}^{(1)}$ en se fixant un seuil μ ;
3. Optimisation par descente de gradient partant de $\tilde{h}^{(1)}$: $\tilde{h} = \underset{h}{\operatorname{argmin}} E(h)$;

FIGURE 5.5 – Algorithme de reconstruction de MNT.

Pour la suite, cette approche de résolution sera appelée *FastDC(MCPU+TV)*.

4 Expériences sur des données de synthèse

Nous étudions la validité de l'approche proposée sur des données InRSO selon les critères suivants :

- efficacité en temps de calcul et en consommation de mémoire,
- robustesse du modèle face aux problèmes de bruit et discontinuités de phase,
- précision sur l'image de reliefs reconstruite,
- simplicité dans l'estimation des hyperparamètres locaux pour la régularisation.

Le jeux de données sur lequel est testée l'approche est une variété de données interférométriques simulées, multibases et multipasses, présentant différentes structures, cohérences et ambiguïtés de reconstruction. Ces données sont générées avec les mêmes paramètres que ceux du capteur RSO embarqué sur le satellite ERS-2.

La cohérence sera considérée constante sur toute l'image pour chaque canal et dépendra de la décorrélation spatiale due à la longueur de la base et d'un facteur de décorrélation dû à certaines perturbations décrites au paragraphe 3.4 du chapitre 4 : temporelle, erreurs de recalage et bruit thermique. Ainsi la cohérence en chaque canal sera exprimée sous la forme suivante :

$$\gamma_c = \eta_c \left(1 - \frac{B_c^\perp}{B_{critique}^\perp} \right), \quad (5.7)$$

avec η_c un facteur de décorrélation qui dépend du canal.

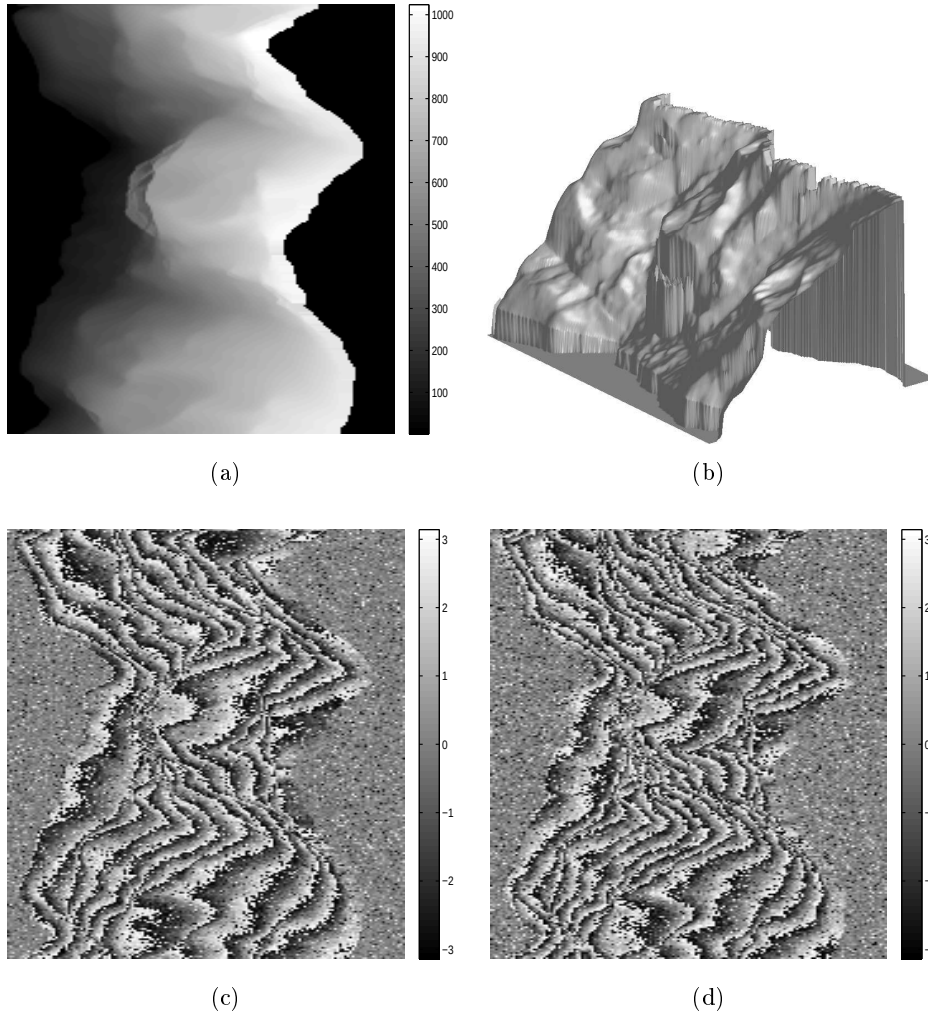
Dans les différentes séries de données InRSO, de fortes discontinuités caractérisent la scène, violant ainsi la condition de régularité d'Itoh. De plus, différents types de structures (constantes ou lisses par morceaux) sont présentes pour s'approcher du cas réel de scènes naturelles. Les longueurs de bases sont choisies de façon à avoir différentes ambiguïtés de déroulement de franges.

Les résultats obtenus par l'approche proposée seront analysés dans un premier temps, ensuite comparés à ceux obtenus par d'autres approches dans la littérature, particulièrement celles présentées dans le chapitre de l'état de l'art précédent.

4.1 Validation du modèle *a priori*

Cette première expérience est réalisée afin de comparer le modèle proposé MCPU+TV au modèle MCPU+LGMRF en vu de montrer la précision de l'*a priori* considéré malgré sa simplicité, comparé au modèle LGMRF, et aussi son apport en efficacité pour l'algorithme d'optimisation et d'estimation des hyperparamètres.

Dans la figure 5.6, les données interférométriques simulées sont présentées. La zone correspond à une zone montagneuse présentant différentes structures lisses et discontinues. Cela nous permettra de valider le recours au modèle TV avec des hyperparamètres locaux (discontinuités) dans ce cas de scènes naturelles.



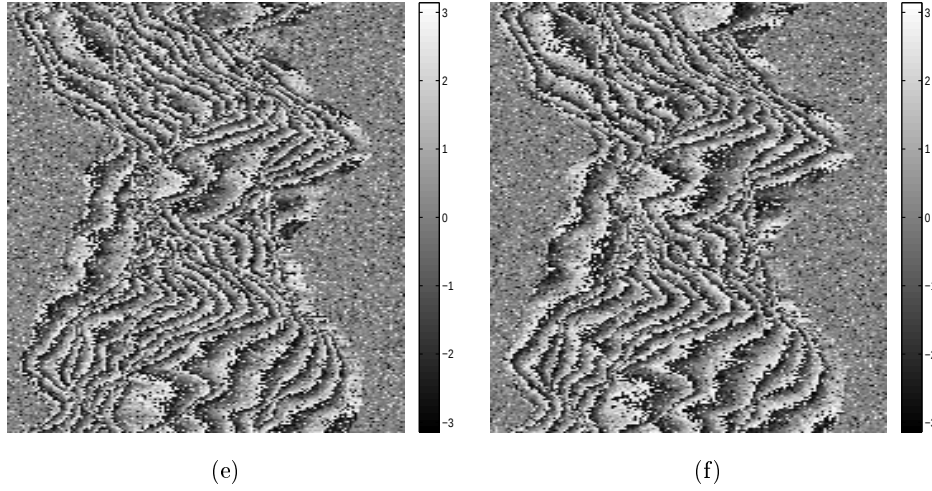


FIGURE 5.6 – Données InRSO simulées à partir d'un MNT synthétique d'une zone montagneuse. (a) Le MNT d'origine, taille de l'image = 250×160 , (b) sa visualisation en 3D, les interférogrammes générés selon les paramètres (base et facteur de décorrélation) (c) (d) (e) et (f) .

Les différents canaux présentent des fortes ambiguïtés de phase à cause des discontinuités du profil et du fort bruit interférométrique, rendant une approche de déroulement monocanal difficile. D'ailleurs, dans la figure 5.7, un exemple de déroulement de l'interférogramme le moins ambigu par l'approche de minimisation exacte de la norme L^1 proposée dans (Bioucas-Dias et Valadao, 2007) est présenté.

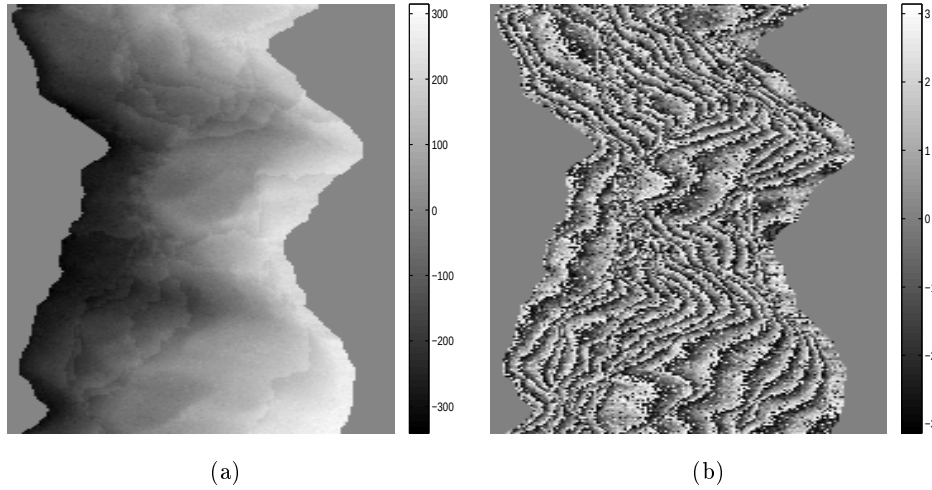


FIGURE 5.7 – Reconstruction du profil par l'approche de déroulement de phase PUMA appliquée sur le premier canal de la figure 5.6(c). (a) Profil reconstruit, (b) son réenroulement dans $[-\pi, \pi[$.

Cette dernière approche de déroulement de la phase InRSO, connue par le nom *PUMA*, permet d'estimer la phase absolue en minimisant exactement et de façon itérative la norme L^1 de la différence entre les gradients de la phase et ceux de l'image estimée. Grâce à la technique de coupe-minimum sur graphe, sa résolution est très rapide et non coûteuse en mémoire. A titre illustratif, le résultat obtenu dans la figure 5.7 est obtenu en 2s. Cepen-

dant, des problèmes de déroulement persistent dans la reconstruction finale. Une telle approche n'est exploitable efficacement que sur des données InRSO non fortement bruitées avec des franges espacées et de faible ambiguïté. En revanche, même si un déroulement est possible, la reconstruction obtenue à partir de telles données est moins précise comparée à ce que donnera une approche multicanal, où des interférogrammes formés par des grandes bases offrent une information de précision plus riche pour la reconstruction de scènes naturelles.

Les résultats de la reconstruction du MNT par les deux modèles MCPU+TV et MCPU+LGMRF sont présentés dans la figure 5.8. Les images des hyperparamètres locaux estimées pour chacun d'eux sont présentées dans la figure 5.9. L'image de discontinuités donnée par notre approche est une image binaire construite automatiquement avec seuillage de l'image des gradients de la solution $h^{(1)}$. Le seuil est choisi de façon à pouvoir conserver les discontinuités au delà d'une hauteur donnée. En dessous de cette hauteur, la régularisation est alors possible entraînant un lissage de la zone. L'image de discontinuités obtenue par l'approche MCPU-LGMRF est quant à elle une image en niveaux de gris obtenue à partir des statistiques locales de l'image des profils reconstruite itérativement par un algorithme de type EM.

Dans le tableau 5.1, les résultats numériques et quantitatives de la reconstruction sont données : le temps de calcul exprimé en minutes : secondes et l'erreur quadratique moyenne normalisée par rapport à la vérité du terrain : $EQMN = \frac{\|h - h_{vraie}\|^2}{\|h_{vraie}\|^2}$.

TABLE 5.1 – Performances des approches de reconstruction FastDC(MCPU+TV) et MCPU+LGMRF en terme de temps de calcul et d'erreur quadratique moyenne normalisée de reconstruction.

Algorithme	Temps de calcul (mn : s)	EQMN
FastDC(MCPU+TV)	02 : 35	0.31
MCPU+LGMRF	31 : 02	0.30

Remarque : Nous précisons que l'approche (MCPU-LGMRF) est par ailleurs sensible à l'étape de l'estimation des hyperparamètres locaux à partir d'une connaissance incomplète sur les données d'altitude. Nous avons présenté ici uniquement les résultats sur la série d'interférogrammes où cette approche a pu donner des résultats satisfaisants. Dans le cas général, des problèmes de déroulement sont possibles, s'il n'y a pas assez d'interférogrammes. En pratiques, des données multivues sont générées pour remédier à ces problèmes.

On remarque que les deux approches avec des modèles *a priori* différents donnent pratiquement les mêmes résultats. Une légère précision est apportée par le modèle LGMRF qui est naturellement un modèle plus proche de la réalité des scènes réelles. Cependant, l'estimation de ses paramètres est une étape coûteuse en temps de calcul et fortement dépendante de la qualité du MNT reconstruit à la première itération. Le recours au modèle TV simplifie notablement cette étape d'estimation puisque ce modèle permet à la fois de préserver les fortes discontinuités de la scène tout en régularisant les données. De plus, la pondération de ce modèle contextuel global par des hyperparamètres binaires locaux précisant les discontinuités à préserver permet d'améliorer la qualité de la régularisation en réduisant l'effet de perte de contraste caractéristique de la minimisation de la variation totale.

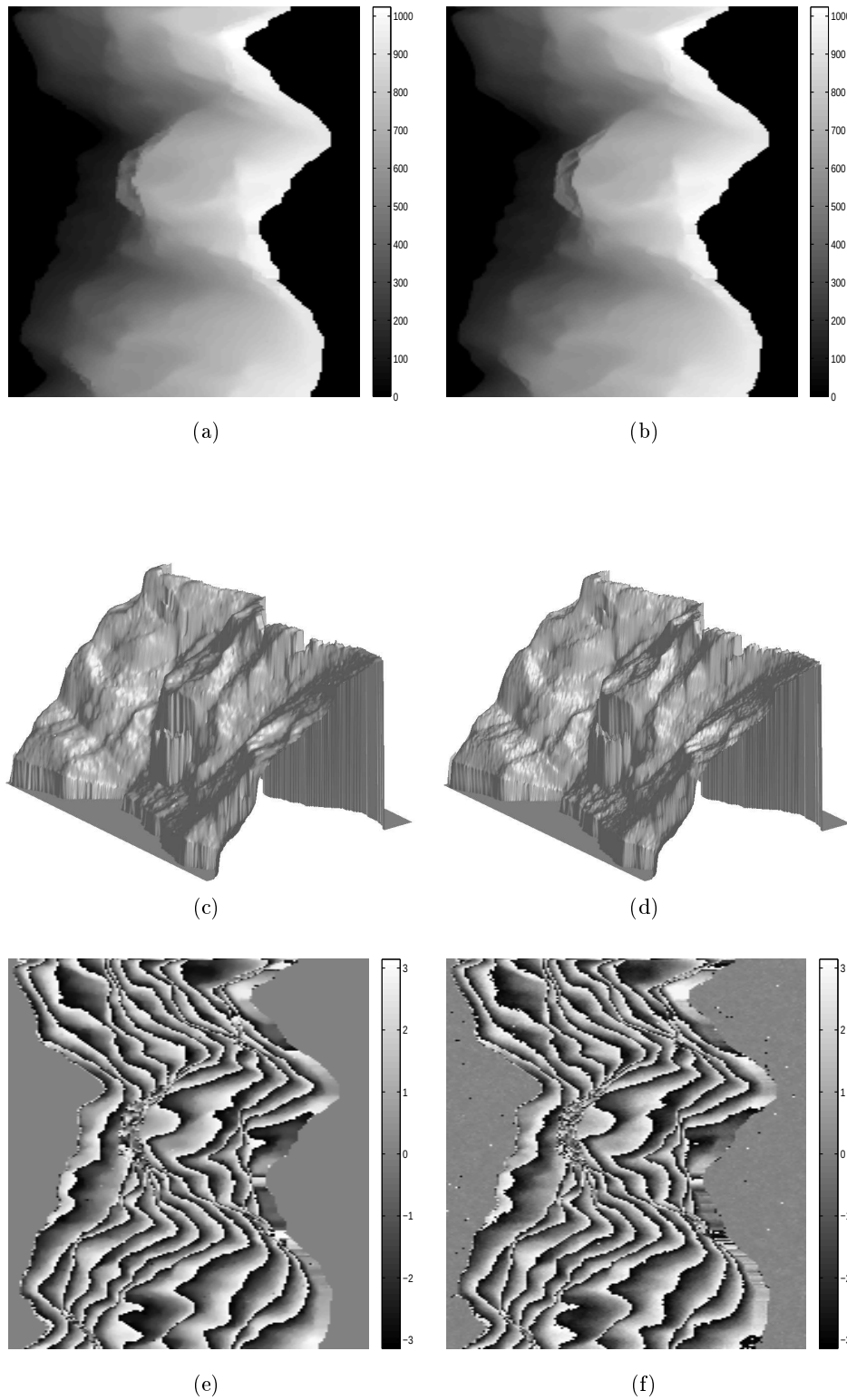


FIGURE 5.8 – Reconstruction du MNT par les deux modèles MCPU+TV et MCPU+LGMRF. Le MNT reconstruit par (a) MCPU+TV, et (b) MCPU+LGMRF, leurs visualisations respectives en 3D dans (c) et (d) et leurs réenroulements dans $[-\pi, \pi[$ (e) et (f) .

Il est important de noter qu'un lissage plus fort pourrait être aussi obtenue par notre approche pour s'approcher plus du résultat donné par le modèle LGMRF et cela en réglant le paramètre de la forme de TV différentiable dans l'étape de l'optimisation continue, comme il a été expliqué dans le chapitre 3. Cependant, nous avons préféré illustrer le résultat obtenu en conservant au mieux la forme de TV utilisée dans l'étape d'optimisation discrète.

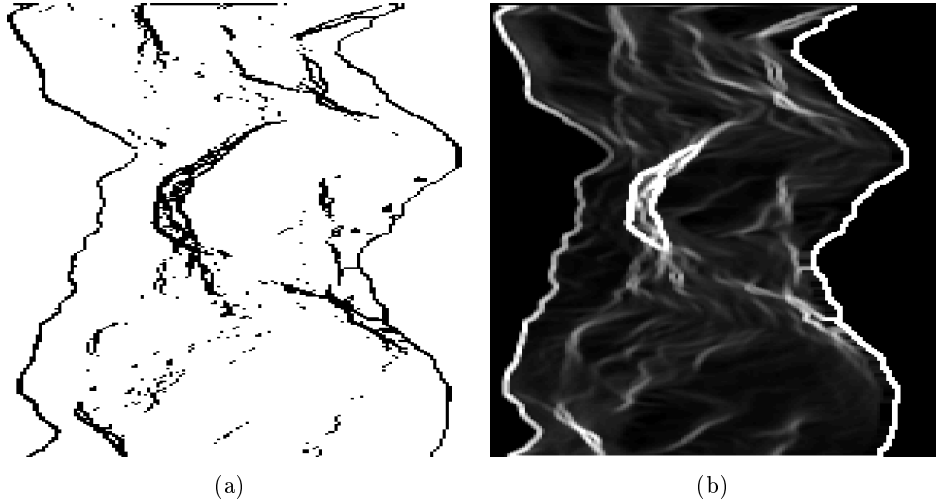


FIGURE 5.9 – Hyperparamètres locaux estimés par les deux approches (a) MCPU+TV et (b) MCPU+LGMRF.

4.2 Étude de l'algorithme d'optimisation

Cette seconde expérience montre l'apport des optimiseurs MPLM pour minimiser efficacement la fonction d'énergie markovienne proposée tout en convergeant vers des minima de meilleures qualités comparé à ce que donnent les optimiseurs binaires par coupure minimale sur graphe, tels que celui utilisé dans l'approche CAPE (Valadao et Bioucas-Dias, 2009). Dans cette dernière approche, l'algorithme d'optimisation utilisé repose sur les mouvements de partitions binaires de type saut, avec un schéma multirésolution. A chaque cycle de l'algorithme de minimisation, le pas du saut est divisé par deux jusqu'à une précision fixée sur les données. L'algorithme est itératif et très rapide, où à chaque itération un graphe binaire est construit et le flot-maximum y est calculé. Cependant, aucune garantie de convergence vers un optimum local de bonne qualité n'est assurée.

Remarque : pour pouvoir comparer les minimiseurs, nous avons considéré le même modèle multicanal (MCPU+TV) dans cette expérience.

Les données interférométriques simulées sont présentées dans la figure 5.10. Comme l'expérience précédente, nous avons généré des données interférométriques ambiguës pour une reconstruction monocanal. Le résultat d'un déroulement monocanal de l'interférogramme le moins ambigu par PUMA est présenté dans la figure 5.11.

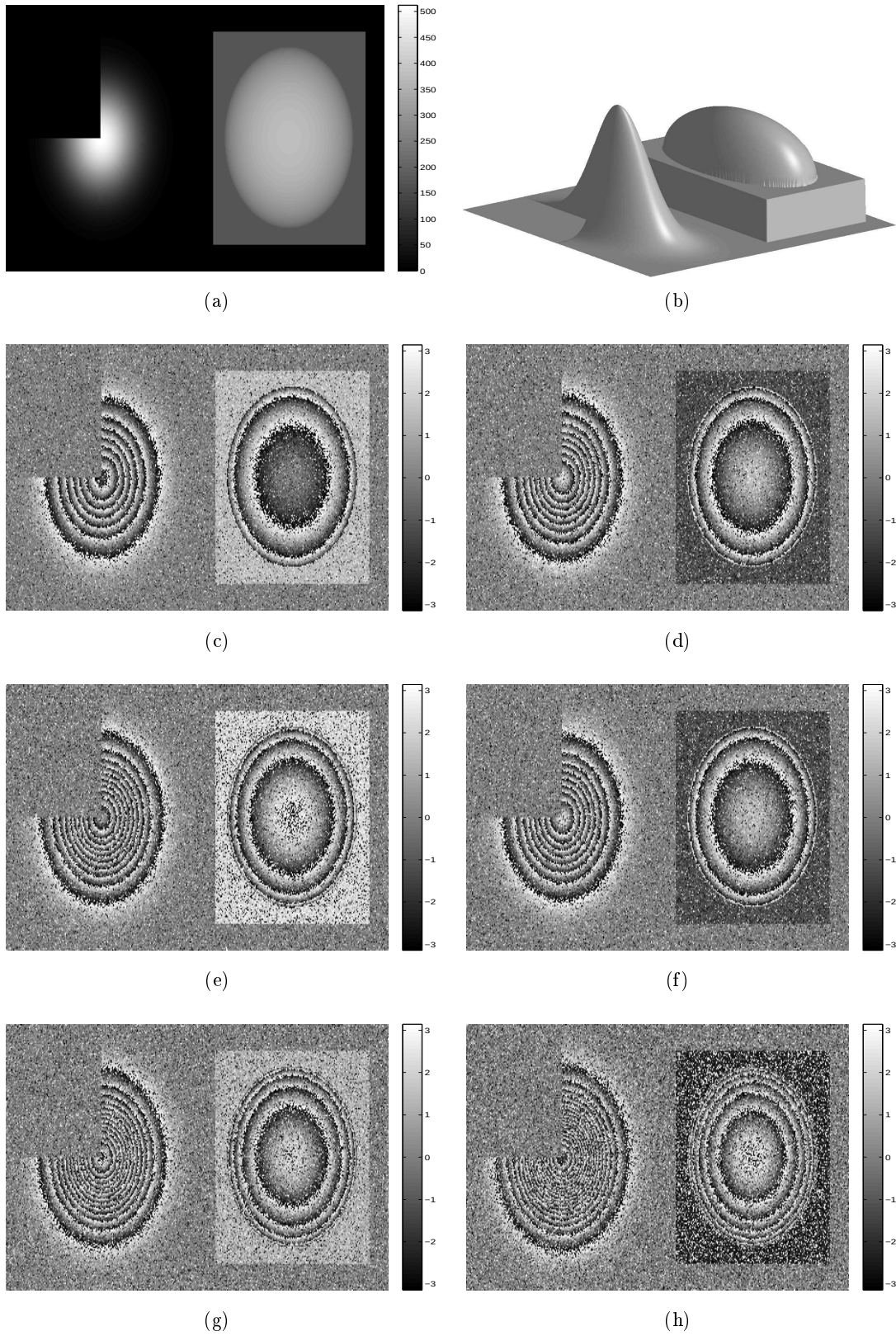


FIGURE 5.10 – Données InRSO simulées à partir d'un MNT synthétique. (a) Le MNT d'origine, taille de l'image = 200×400 , (b) sa visualisation 3D, les interférogrammes générés selon les paramètres (base et facteur de décorrélation) (c)...(h) .

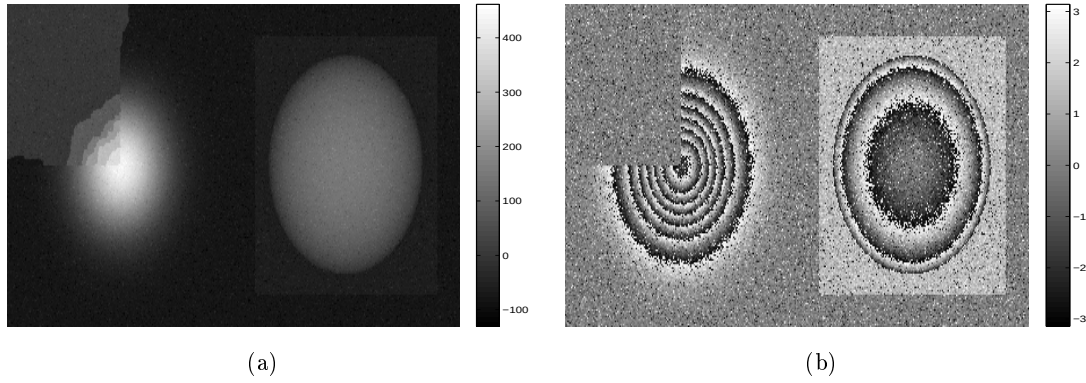


FIGURE 5.11 – Reconstruction du profil par l’approche de déroulement de phase PUMA appliquée sur le premier canal de la figure 5.10(c). (a) Profil reconstruit, (b) son réenroulement dans $[-\pi, \pi]$.

Les résultats de reconstruction du MNT par l’optimisation mixte FastDC et l’algorithme d’optimisation par des sauts binaires sont présentés dans la figure 5.12.

Le tableau 5.2 illustre les résultats numériques de la reconstruction : temps de convergence exprimé en minutes :secondes, mémoire physique en Mega Octets occupée durant une itération de l’algorithme d’optimisation et nécessaire pour la construction du graphe et le minimum de l’énergie atteint à la convergence.

TABLE 5.2 – Performances des optimiseurs MPLM et MPLB par sauts en terme de temps de calcul, occupation en mémoire et minimum atteint de l’énergie.

Algorithme	Temps de calcul (mn :s)	Mémoire (Mo)	E_{min}
Sauts binaires	00 : 47	0.95	$1.3091e + 05$
FastDC	03 : 52	2.44	$1.0867e + 05$

Les résultats montrent que l’optimisation mixte proposée apporte les contributions suivantes :

- un minimum de meilleure qualité en terme d’énergie est obtenu grâce au recours à des mouvements de partitions multi-étiquettes,
- la précision de la solution est aussi bonne comparée à l’image d’origine, ceci grâce à l’estimation des hyperparamètres locaux suivie d’une optimisation continue,
- une complexité algorithmique acceptable relativement à la taille des données comparée à celle de l’algorithme par sauts binaires. Nous précisons que pour ces données, une optimisation exacte n’a pas été possible vu leur taille et celle de l’ensemble des étiquettes. On est confronté aux limites d’une mémoire physique de 4 GO. D’où l’intérêt des optimiseurs multi-étiquettes pour pouvoir converger vers des optima proches d’un optimum global, sans soucis de mémoire physique.

Ainsi un compromis entre qualité de reconstruction, précision et complexité algorithmique est assuré par l’algorithme FastDC(MCPU+TV).

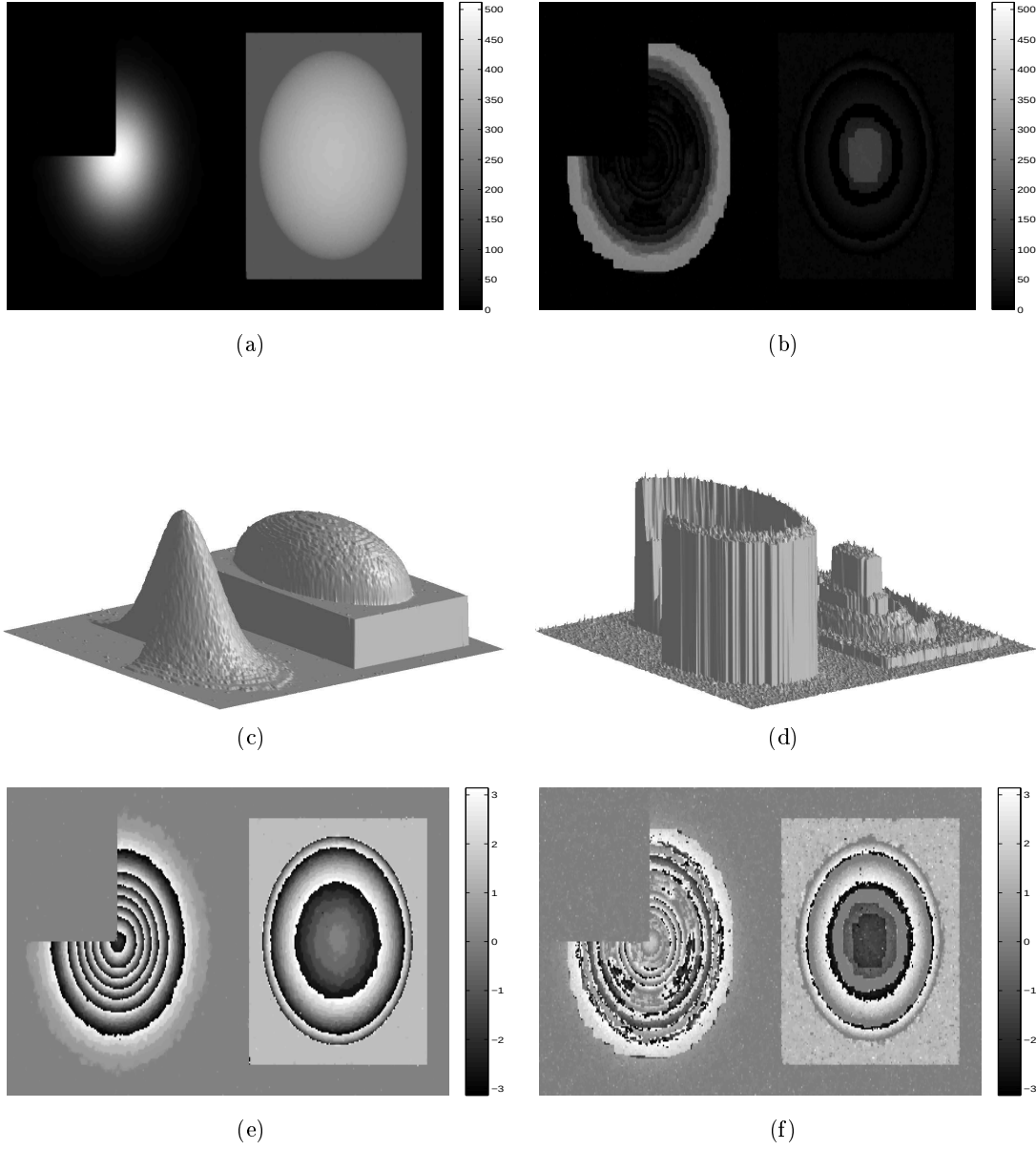


FIGURE 5.12 – Reconstruction du MNT par les deux optimiseurs FastDC et l’algorithme par sauts binaires avec un schéma multirésolution. Minima obtenus par les deux algorithmes respectivement dans (a) et (b), leurs visualisations respectives en 3D dans (c) et (d) et leurs réenroulements dans $[-\pi, \pi[$: (e) et (f) .

5 Prise en compte des perturbations globales

Dans le cadre des approches multicanaux, une décorrélation de l'information d'altitude apportée par les différents interférogrammes empêcherait systématiquement le recours direct au modèle proposé. Une étape d'estimation des perturbations susceptibles de les entâcher, comme les perturbations atmosphériques, est au préalable nécessaire.

Dans cette partie, nous nous intéressons spécialement aux perturbations globales qui pourraient toucher les mesures de phase et qui dépendent linéairement de la topographie de la scène. Cette classe de perturbations peut caractériser les perturbations atmosphériques globales affectant les mesures interférométriques, comme il a été discuté dans le travail de thèse (Chaabane, 2004), ou aussi des problèmes de décalage de phases qui caractérisent aussi les données interférométriques multicanaux réelles, les phases étant enroulées à une constante près (Ferraioli *et al.*, 2008). On se propose ainsi d'ajuster le modèle de vraisemblance multicanal (5.1) pour pouvoir tenir compte de ces problèmes. Ensuite un algorithme rapide d'estimation de l'altitude pour traiter de telles données est introduit. Enfin, l'approche intégrale est testée sur un jeu de données interférométriques multipasses et multi-bases, synthétiques, et où une perturbation de phase globale est introduite sur chaque canal indépendamment.

5.1 Modèle de vraisemblance avec perturbations de phase globales

Pour un interférogramme donné, un modèle simplifié des perturbations globales qui s'ajoutent aux mesures de phase pourrait s'exprimer selon une relation affine avec la topographie du milieu comme suit :

$$\phi_{perturb} = \alpha h + \beta, \quad (5.8)$$

avec $\phi_{perturb}$ la mesure de phase absolue d'une perturbation globale sur toute l'image, $\alpha \in [-1, 1[$ et $\beta \in [-\pi, \pi[$ les paramètres de ce modèle.

En prenant $\alpha = 0$ dans ce modèle, nous retrouvons les perturbations de type décalage de phases étudiées dans (Ferraioli *et al.*, 2008).

Ainsi en tenant compte de cette information additionnelle à la phase topographique, la mesure de phase interférométrique s'écrit comme :

$$\psi = \langle \phi_{top} + \phi_{perturb} \rangle_{2\pi}, \quad (5.9)$$

avec ϕ_{top} la phase topographique absolue.

Il est clair qu'en disposant en plus des interférogrammes d'un modèle numérique de terrain de la même zone avec une même résolution et projeté selon la même géométrie radar, l'estimation des paramètres α et β pour chacun des canaux est désormais possible. Mais alors, l'estimation du relief n'est plus utile, puisqu'il est connu. Nous considérons dans la suite qu'une telle donnée ne nous est pas fournie, ou qu'elle n'est pas assez précise pour une bonne estimation. Nous proposons ainsi un nouveau modèle de vraisemblance multicanal pour estimer conjointement l'altitude et les paramètres de la perturbation sur chaque canal.

En partant du même raisonnement que celui du chapitre précédent pour parvenir au modèle de vraisemblance multicanal, on considère que pour chacun des pixels des canaux disponibles, l'observation de phase a pour expression la forme suivante :

$$\psi_{p,c} = \langle \frac{2\pi}{h_c^{amb}} h_p + \alpha_c h_p + \beta_c \rangle_{2\pi}, \quad (5.10)$$

avec h_c^{amb} l'altitude d'ambiguïté du canal c , (α_c, β_c) les paramètres de la perturbation au canal c et h_p l'altitude vraie de la cible p .

Ainsi pour chaque canal, la distribution de la phase s'écrit comme étant :

$$P(\psi_{p,c}|h_p, \alpha_c, \beta_c) = \frac{1}{2\pi} \frac{1 - \gamma_c^2}{1 - f_{p,c}^2(h_p)} \left(1 + \frac{f_{p,c}(h_p) \arccos(-f_{p,c}(h_p))}{\sqrt{1 - f_{p,c}^2(h_p)}} \right), \quad (5.11)$$

avec

$$f_{p,c}(h_p, \alpha_c, \beta_c) = \gamma_c \cos \left(\psi_{p,c} - \left(\frac{2\pi}{h_c^{amb}} + \alpha_c \right) h_p - \beta_c \right). \quad (5.12)$$

En supposant l'indépendance entre les observations de phases multicanaux, le modèle de vraisemblance multicanal se réécrit alors sous la forme :

$$P(\Psi|h, \alpha, \beta) = \prod_{c=1}^M P(\psi_{p,c}|h_p, \alpha_c, \beta_c), \quad (5.13)$$

avec Ψ le vecteur collectant toutes les observations de phases multicanaux, h l'image d'altitude, α et β les vecteurs collectant les paramètres des perturbations globales sur les différents canaux.

5.2 Algorithme d'estimation conjointe altitude et perturbations globales

Afin d'estimer correctement l'altitude conjointement avec les paramètres des perturbations globales, nous proposons un algorithme itératif estimant à chaque itération successivement l'altitude par maximum *a posteriori* en fonction des paramètres estimés, et ensuite les paramètres par un maximum de vraisemblance en exploitant la nouvelle altitude estimée. Une telle approche d'estimation est non-exacte. Cependant, en présence de plusieurs canaux et une image d'altitude régularisée, l'erreur d'estimation est réduite. En effet, la régularisation de l'image d'altitude permet à la fois de réduire la quantité de bruit sur les données et d'apporter une information contextuelle contraignant l'estimation des paramètres de perturbations globales à s'ajuster au mieux avec le profil de la scène. De plus, le recours à plusieurs canaux améliore la précision de l'estimation de ces paramètres et réduit davantage le bruit interférométrique sur les mesures de phase.

Ainsi l'étape d'estimation de l'altitude connaissant les observations de phases et les paramètres atmosphériques s'effectue exactement avec le même algorithme 5.5. Le résultat de cette étape est injecté dans le modèle de vraisemblance et l'étape d'estimation des paramètres (α_i, β_i) , $i \in \{1, 2, \dots, M\}$ s'effectue par un maximum de vraisemblance multicanal par rapport aux observations de phases connaissant l'altitude. L'estimation par maximum de vraisemblance multicanal est réalisée par un algorithme itératif rapide à deux phases, reposant sur l'optimisation exacte par coupe-minimum sur graphe. Initialement, les paramètres de perturbations sont nuls. Une première phase consiste à estimer exactement les bons paramètres α_i connaissant l'altitude et les paramètres β_i courants. La deuxième phase consiste à estimer les bons paramètres β_i connaissant l'altitude et les paramètres α_i courants. Ces deux phases reprennent jusqu'à la convergence. L'estimation de l'altitude est reprise connaissant les nouveaux paramètres estimés. Ainsi de suite, jusqu'à la convergence de l'algorithme complet.

La figure 5.13 illustre l'algorithme complet d'estimation de l'altitude et des paramètres des perturbations globales. Nous appelons par la suite une itération complète, formée de l'estimation de l'altitude suivie des paramètres, un cycle.

```

Initialisation de  $h^{(0)}$ , des paramètres  $(\alpha^{(0)}, \beta^{(0)})$  de tous les canaux et le nombre maximal
de cycles ;
Pour Chaque cycle  $t$  de 1 au nombre maximal de cycles faire
    [Estimation de l'altitude par FastDC partant de  $h^{(t-1)}$  et connaissant  $(\alpha^{(t-1)}, \beta^{(t-1)})$ ]
     $\tilde{h} = \underset{h}{\operatorname{argmin}} E(h|\Psi, \gamma, \alpha^{(t-1)}, \beta^{(t-1)})$ ;
    [Mise à jour des altitudes]
     $h^{(t)} = \tilde{h}$ ;
    [Estimation des paramètres partant de  $(\alpha^{(t-1)}, \beta^{(t-1)})$  et connaissant  $h^{(t)}$ ]
     $(\tilde{\alpha}, \tilde{\beta}) = \underset{(\alpha, \beta)}{\operatorname{argmax}} P(\alpha, \beta|\Psi, h^{(t)}, \gamma)$ ;
    [Mise à jour des paramètres]
     $(\alpha^{(t)}, \beta^{(t)}) = (\tilde{\alpha}, \tilde{\beta})$ ;
    Si ( $h^{(t)} == h^{(t-1)}$ ) Alors
        | retourner  $\tilde{h}$ ,  $\tilde{\alpha}$  et  $\tilde{\beta}$ ;
    Fin Si
Fin Pour

```

FIGURE 5.13 – Algorithme de correction des données et de reconstruction de MNT.

5.3 Résultats de reconstruction

Dans cette expérience, on se propose d'étudier la robustesse du nouveau modèle proposé tenant compte des perturbations globales de types linéaires avec la topographie.

Dans la figure 5.14, un ensemble d'interférogrammes multibases et multipasses est généré à partir d'un MNT synthétique. Ces mesures interférométriques sont composées à la fois de la topographie de la scène, d'un bruit interférométrique et des perturbations globales selon le modèle linéaire (5.8). Les paramètres des perturbations varient d'un canal à un autre indépendamment.

Dans la figure 5.15, les résultats de la reconstruction du MNT à chaque itération de l'algorithme d'estimation sont présentés. On remarque à la première itération, où les paramètres de perturbations sont initialement nuls, la difficulté que rencontre le modèle multicanal initial pour estimer correctement l'altitude. En avançant dans les itérations, les paramètres sont estimés corrigeant ainsi les artéfacts sur les données d'origine. D'où l'intérêt d'une estimation itérative de ces perturbations qui sont correctement éliminées à l'itération finale de l'algorithme. La reconstruction finale est satisfaisante.

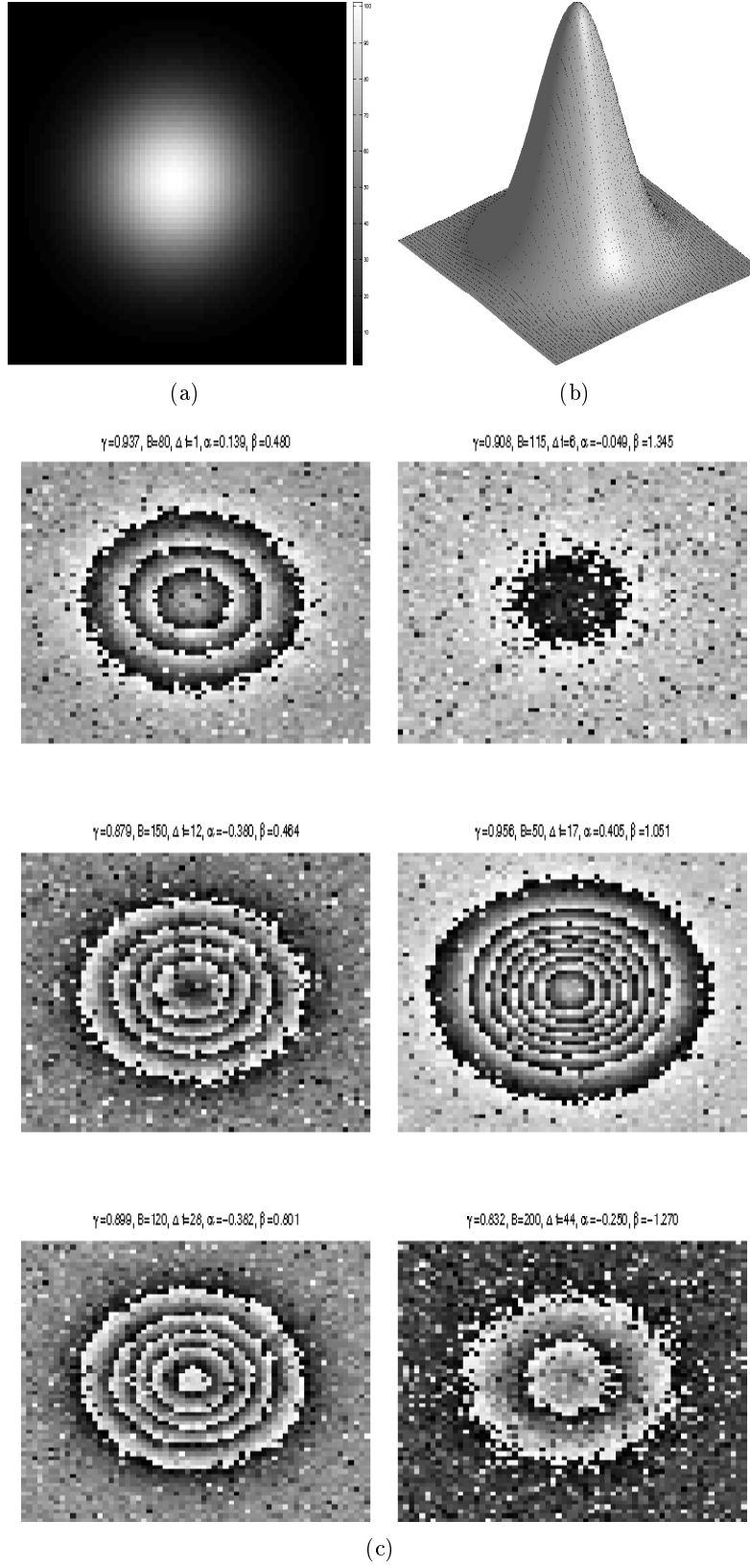


FIGURE 5.14 – Interférogrammes simulés à partir d'un MNT synthétique. (a) Le MNT d'origine, taille de l'image = 64×64 , (b) sa visualisation 3D, (c) ensemble d'interférogrammes générés selon les paramètres (cohérence, base, facteur de décorrélation et paramètres des perturbations globales).

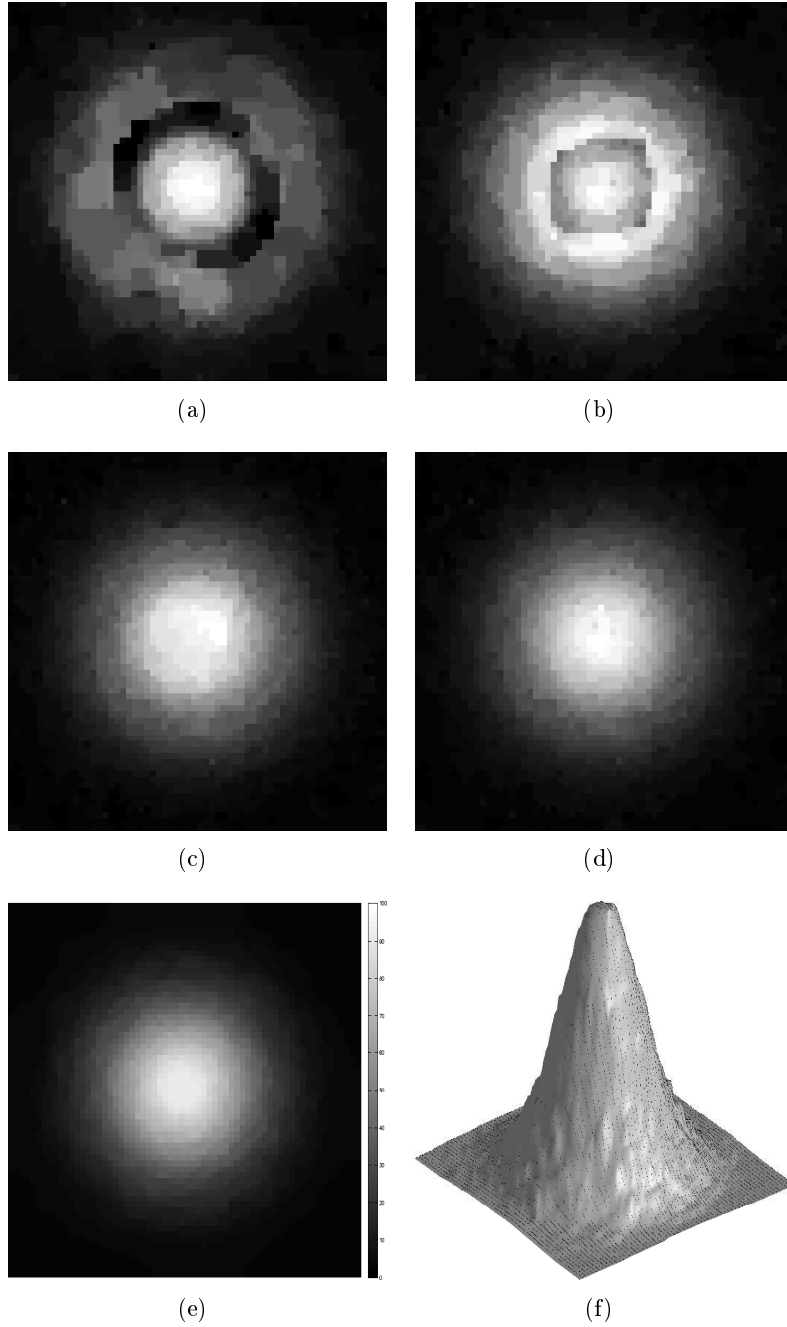


FIGURE 5.15 – Résultats de reconstruction du profil à différentes itérations de l’algorithme d’estimation des paramètres des perturbations globales. (a) itération 1, (b) itération 2, (c) itération 3, (d) itération 4, (e) reconstruction finale et (f) sa visualisation en 3D.

6 Résultats sur des données réelles

Nous passons maintenant à la validation de l’approche proposée sur des données interférométriques réelles. Les données considérées sont en bande C acquises sur la zone de Serre-Ponçon par le satellite ERS-2. Le contenu de la zone est riche en structures de relief permettant de tester la robustesse de l’approche pour dérouler et régulariser la phase InRSO.

Le produit interférométrique fourni est formé des images suivantes :

- une série d’interférogrammes multipasses avec différents intervalles de temps d’acquisition et différentes bases. Ces interférogrammes recalés entre eux sont corrigés des franges orbitales.
- les images de cohérence correspondantes qui indiquent la fiabilité des mesures de phases,
- les images d’amplitudes des couples interférométriques qui sont moyennées afin de réduire le bruit de chatoiement.

Nous disposons aussi du MNT SRTM de la zone étudiée, de résolution 90m, recalé et projeté dans la même géométrie radar que celle des interférogrammes. Dans la suite, ce MNT sera considéré comme étant une vérité terrain. Les résultats obtenus par notre approche seront donc comparés à ce MNT pour juger la précision et la qualité de la reconstruction.

Dans le tableau 5.3, nous présentons les paramètres d’une série d’interférogrammes sur laquelle l’approche est testée. Cette série est caractérisée par des acquisitions à intervalles de temps d’une journée. Ainsi une bonne cohérence temporelle les caractérise. Différentes longueurs de bases sont disponibles.

TABLE 5.3 – Paramètres des interférogrammes utilisés. La case Date indique les dates de la première acquisition (D1) et de la deuxième (D2). La case Base indique la longueur de la base perpendiculaire en mètres. L’altitude d’ambiguïté en mètres est calculée en fonction des paramètres du satellite ERS-2

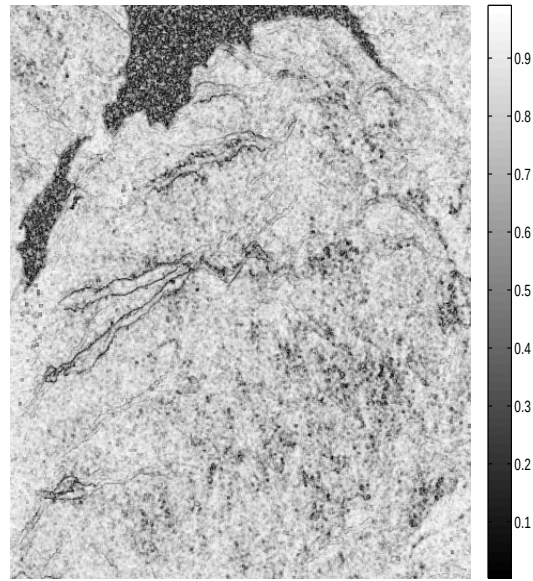
Numéro du canal	Date(D1,D2)	Base	Altitude d’ambiguïté
1	(10-03-1996,11-03-1996).	15.02 m	608 m
2	(09-07-1995,10-07-1995)	28.52 m	320 m
3	(13-08-1995,14-08-1995)	51.54 m	177 m
4	(14-04-1996,15-04-1996)	97.45 m	93.7 m
5	(22-10-1995,23-10-1995)	111.97 m	81.54 m
6	(26-11-1995,27-11-1995)	151.28 m	64.62 m

A partir du MNT SRTM disponible, les hauteurs sur la zone d’étude varient entre 700m et 2300m et la discontinuité maximale enregistrée est de 640m. Ainsi connaissant les altitudes d’ambiguïté considérées, des problèmes de déroulement de phase sont possibles en utilisant uniquement le premier canal, de petite base. Le recours à d’autres canaux est nécessaire pour enlever l’ambiguïté dans ce cas de figure. De plus, une meilleure précision apportée par les grandes bases caractérisera le MNT reconstruit.

Dans la figure 5.16, l’ensemble des interférogrammes et cartes de cohérences correspondant respectivement à ceux du tableau sont présentés.



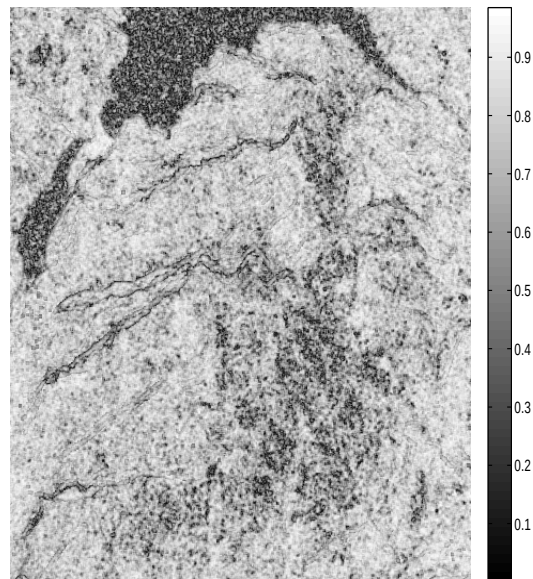
(a)



(b)



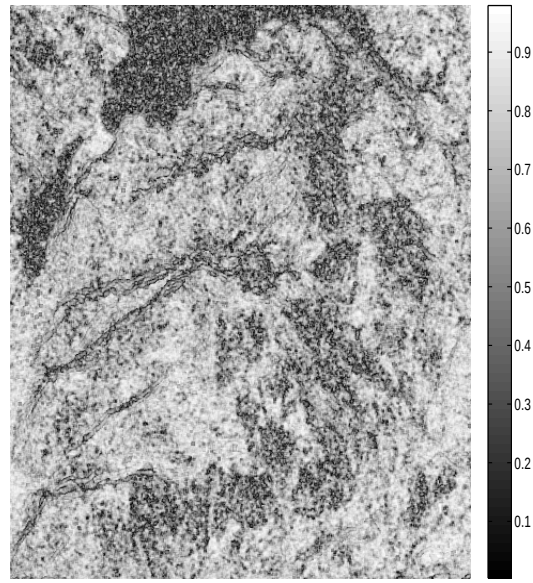
(c)



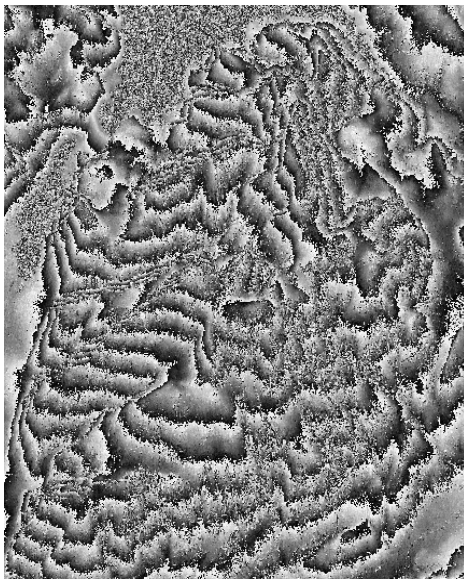
(d)



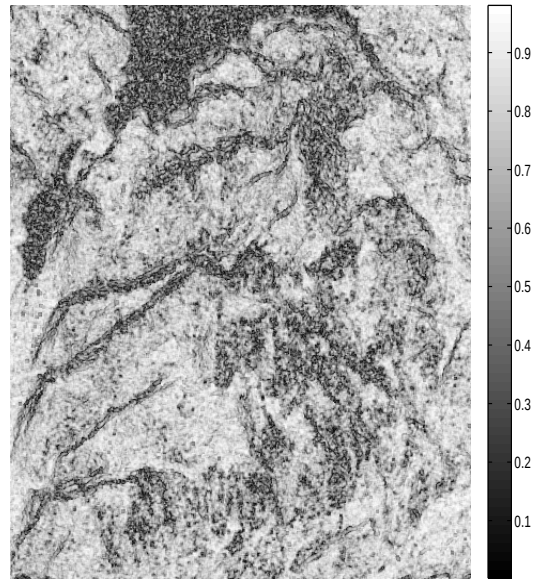
(e)



(f)



(g)



(h)

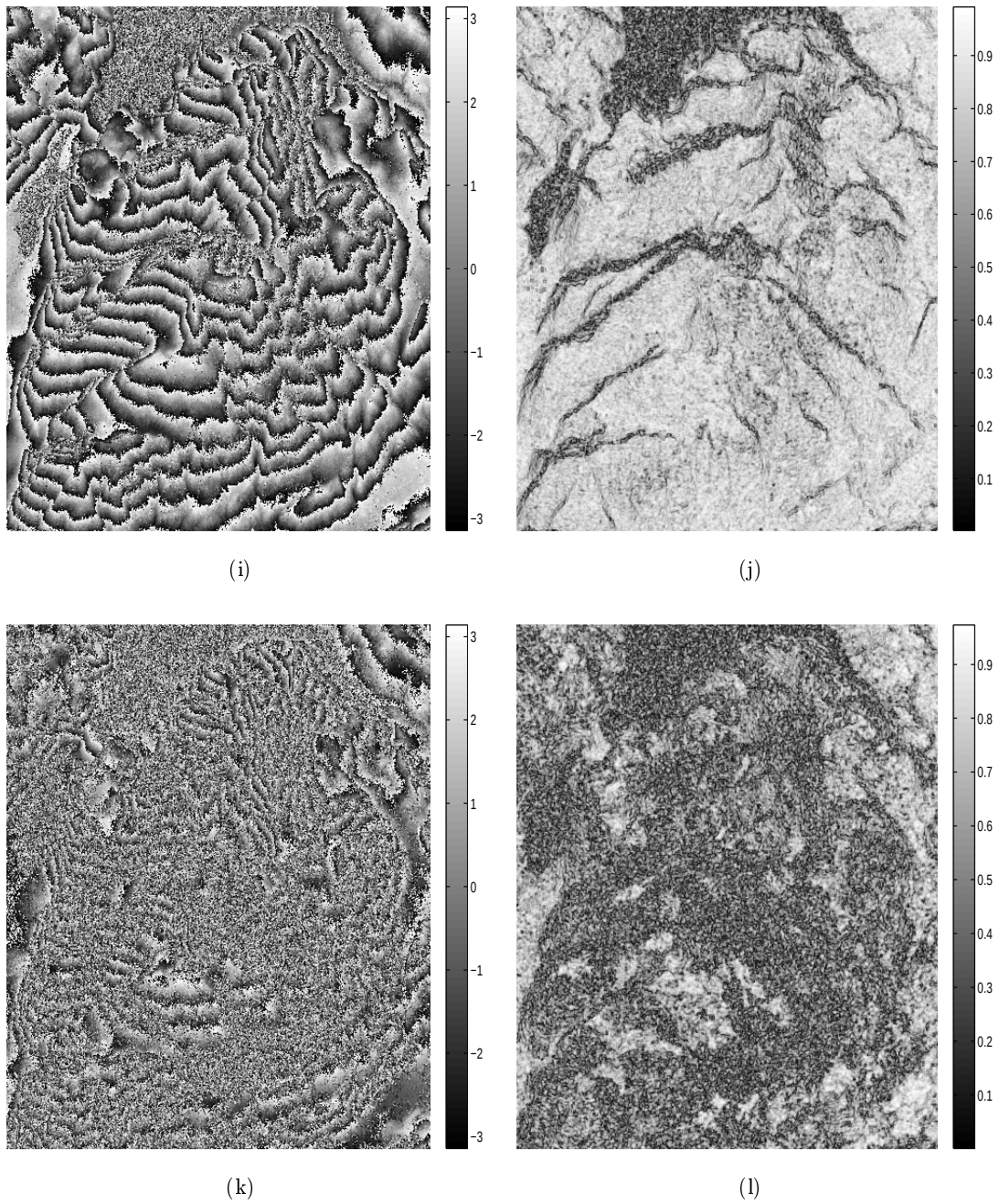


FIGURE 5.16 – Les interférogrammes ainsi que les cartes de cohérences utilisés, taille de l'image = 550×430 . (a) Interférogramme numéro 1 et (b) la carte de cohérence correspondante, (c) interférogramme numéro 2 et (d) la carte de cohérence correspondante, ... (k) interférogramme numéro 6 et (l) la carte de cohérence correspondante.

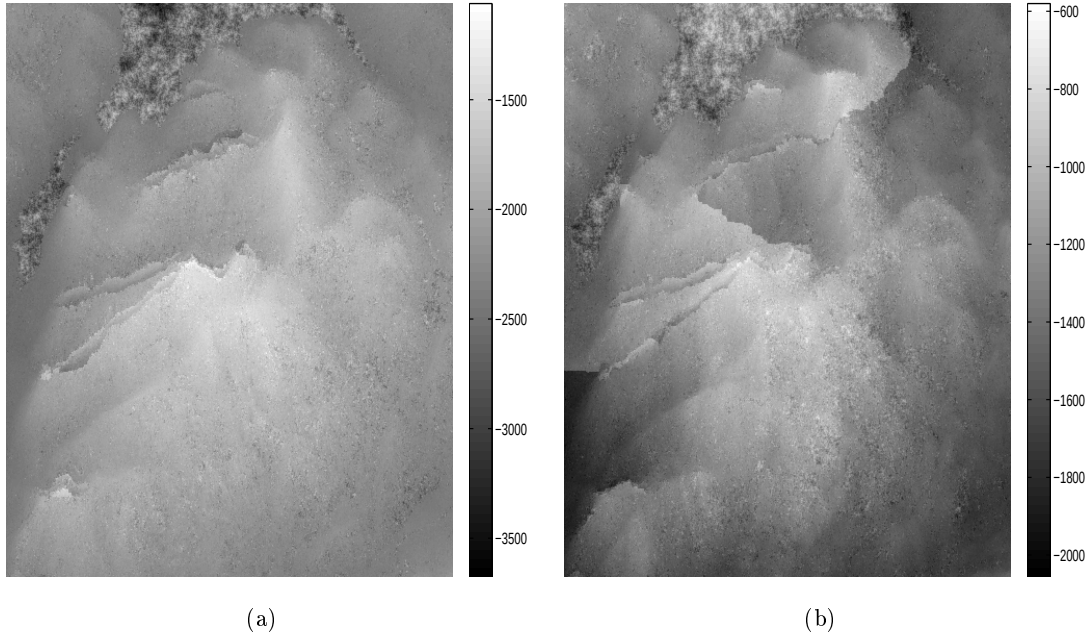


FIGURE 5.17 – Résultat de déroulement monocanal par l’approche PUMA. (a) Phase déroulée du premier canal, (b) phase déroulée du second canal.

La reconstruction obtenue par l’approche monoconal PUMA appliquée sur les interférogrammes les moins ambigus est présentée dans la figure 5.17, montrant bien la faible précision du résultat obtenu ainsi que les quelques problèmes de déroulement qui apparaissent dans les zones à forte discontinuité.

Le déroulement monocanal des autres interférogrammes est encore plus ambigu. Le recours à l’approche multicanal semble alors être une meilleure alternative pour améliorer la précision du MNT reconstruit d’une part et d’autre part enlever l’ambiguïté sur les zones de discontinuité. Ainsi on se propose dans un premier temps de vérifier la présence de perturbations de type atmosphérique sur les données, qui rendent les mesures de phase multicanaux incohérentes pour le modèle multicanal proposé.

Considérons alors les 6 canaux dans l’approche multicanal et une partie de la zone d’étude caractérisée par des fortes discontinuités sur le MNT SRTM ainsi que par la présence de deux milieux différents : le lac et une partie d’une montagne. La figure 5.18 illustre les portions d’interférogrammes qui correspondent à cette zone.

La reconstruction du MNT de cette zone par le modèle multicanal (5.1) partant directement des données d’origine est présentée dans la figure 5.19(a). On remarque clairement l’échec de la reconstruction. Cela pourrait s’expliquer donc par une présence de phénomènes perturbateurs dans les différents canaux qui les décorrèlent. Ainsi on se propose de corriger ces erreurs des interférogrammes d’origine en exploitant l’algorithme d’estimation conjointe altitude/perturbations globales.

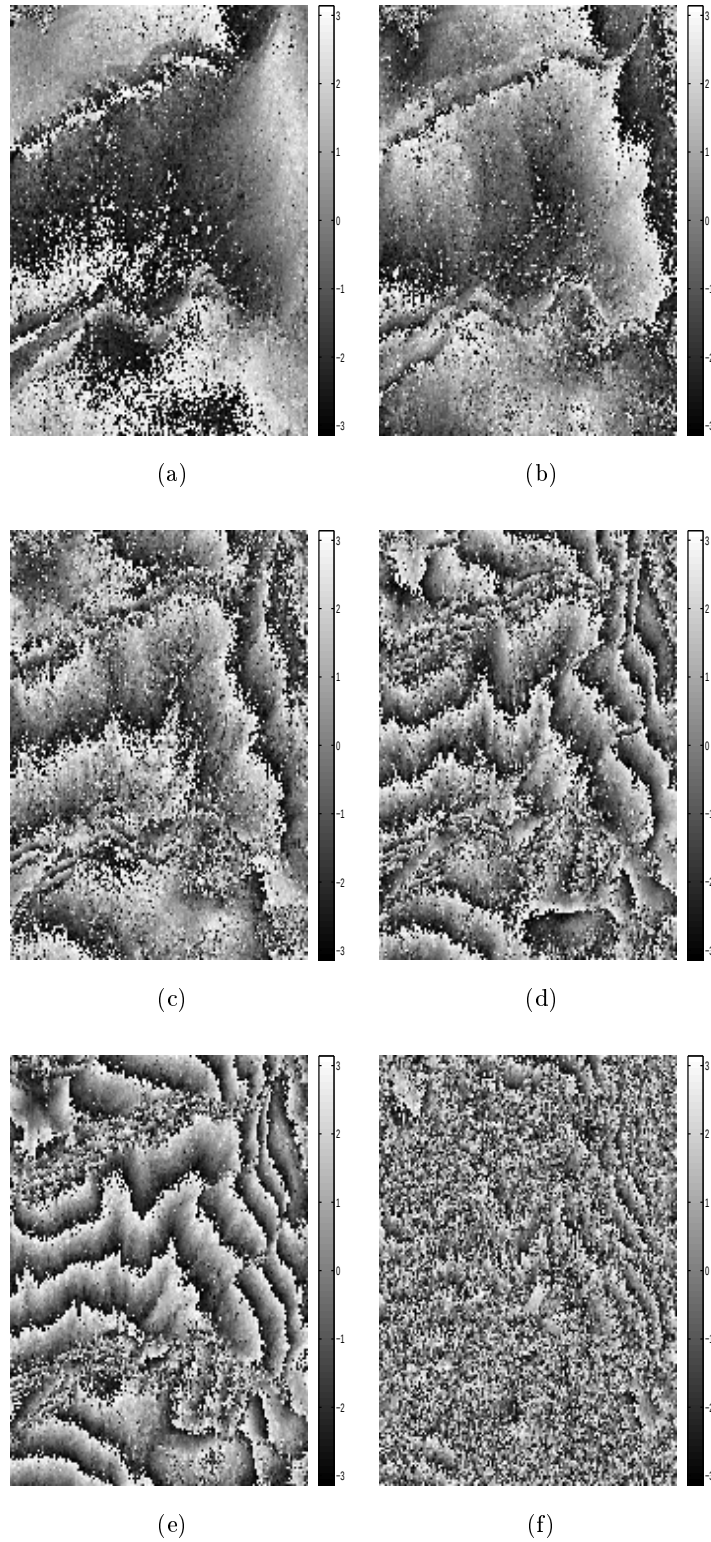


FIGURE 5.18 – Les interférogrammes réels sur une petite zone de Serre-Ponçon, taille de l'image = 200×150 . (a)... (f).

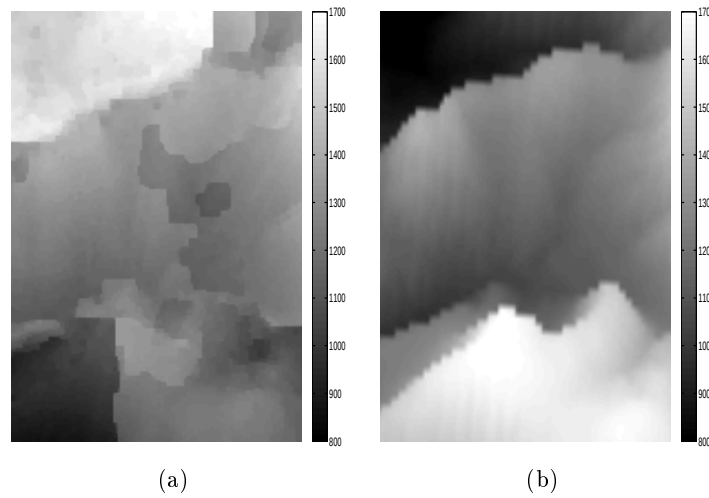


FIGURE 5.19 – Reconstruction du MNT avec le modèle multicanal sans prise en compte des perturbations globales. (a) Le MNT reconstruit et (b) le MNT SRTM.

Nous précisons que l'optimisation de l'énergie sur cette petite zone a été réalisée par le minimiseur exact d'Ishikawa, pour s'assurer que les problèmes de reconstruction ne sont pas dus à la qualité de l'optimum atteint mais plutôt à d'autres sources, telles que les perturbations atmosphériques sur les données.

Deux stratégies peuvent être considérées pour exploiter l'algorithme d'estimation :

1. La première stratégie consiste à appliquer l'algorithme de correction et d'estimation automatique des paramètres des perturbations et de la topographie de la scène en partant d'une image d'altitudes et de vecteurs de paramètres tous nuls. Ainsi on se met dans le cas où on ne dispose d'aucune information sur la zone étudiée.
2. La seconde stratégie, que nous recommandons lorsqu'un MNT d'une résolution plus faible est disponible sur la même scène, consiste à l'exploiter comme une information de départ pour estimer les paramètres des perturbations globales. Ensuite, un MNT de meilleure résolution est reconstruit. Cette situation est la plus réaliste et la plus performante en pratique, puisqu'elle permet ainsi d'éviter une convergence vers un optimum de mauvaise qualité comparée à la première stratégie. De plus, des MNT SRTM de résolution planimétrique de 90m et altimétrique de 16m sont disponibles couvrant toute la surface de la terre ¹.

Les deux stratégies sont testées dans notre étude. Dans la figure 5.20, les résultats de l'algorithme d'estimation itérative des paramètres d'une perturbation globale conjointement avec l'altitude, reposant sur la première stratégie, sont présentés. Ainsi en partant d'un vecteur de paramètres nuls, un MNT est reconstruit donnant le résultat déjà illustré dans la figure 5.19(a). Cette estimation sert ensuite comme un point de départ pour calculer les paramètres de perturbations. L'algorithme reprend jusqu'à la convergence, où aucun changement n'est possible. La reconstruction obtenue est globalement satisfaisante.

1. Lien pour récupérer ces données MNT SRTM : <http://srtm.csi.cgiar.org/>

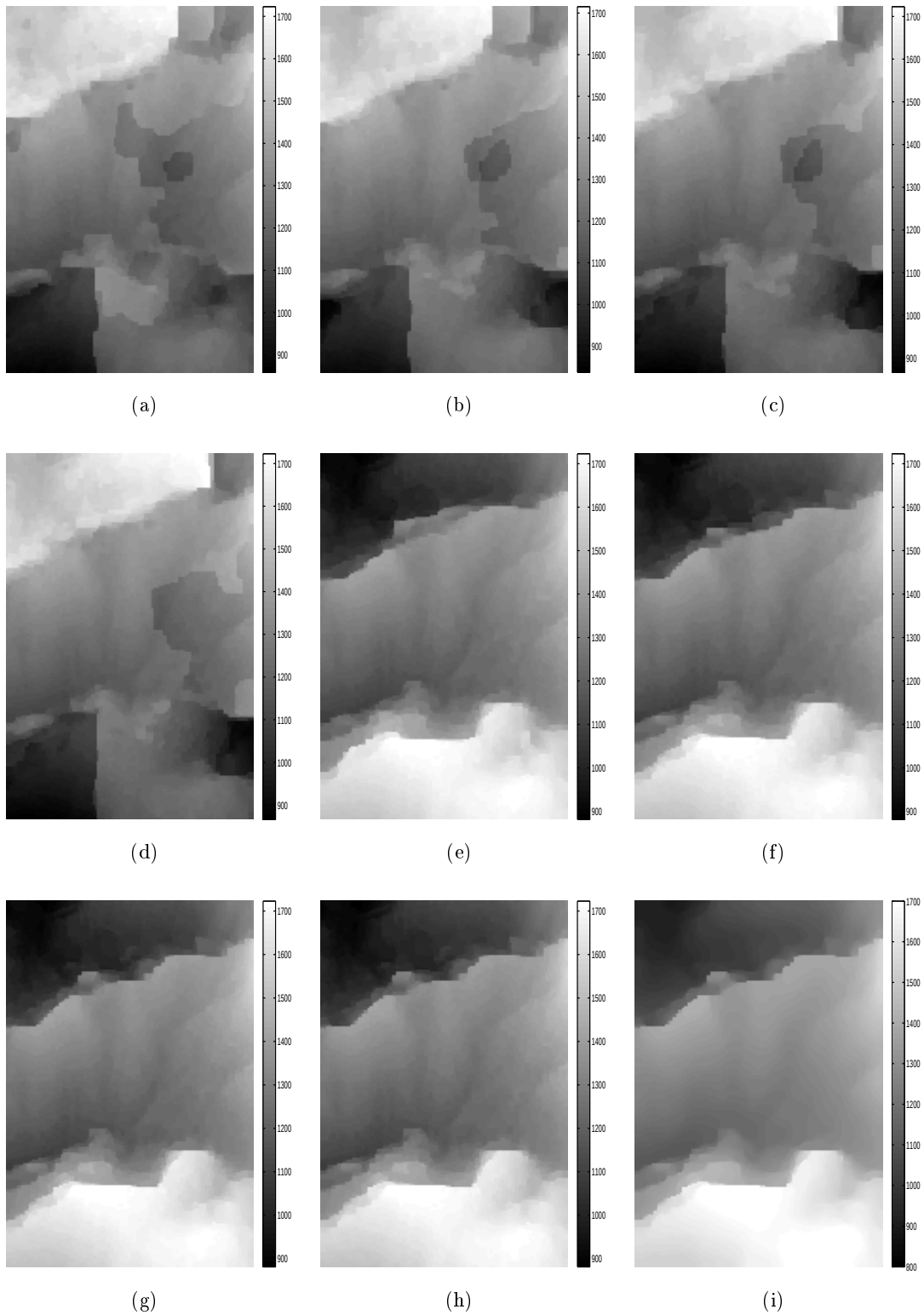


FIGURE 5.20 – Reconstruction du MNT avec correction des perturbations globales par la première stratégie. MNT reconstruit aux cycles (a) 1, (b) 2, ... (h) 11. (i) La reconstruction finale par FastDC.

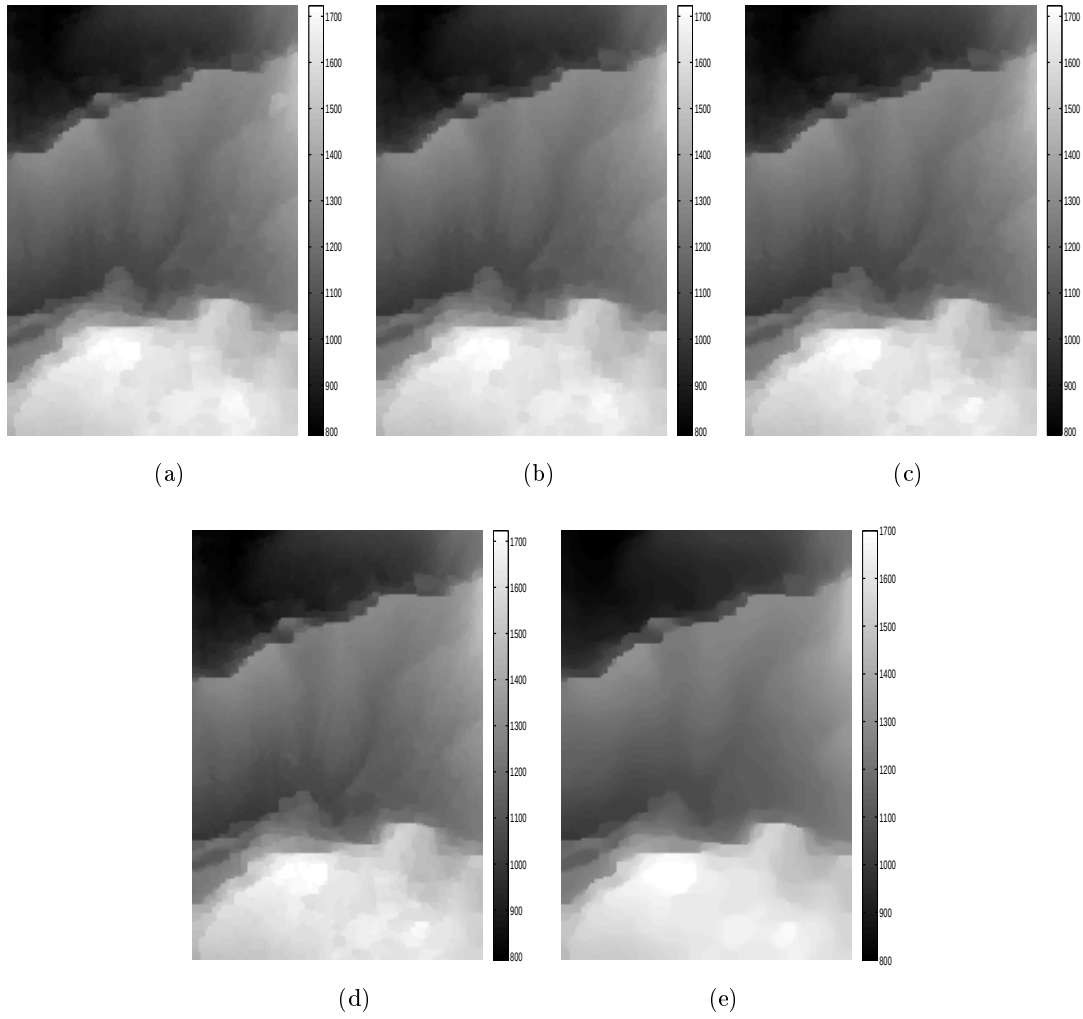


FIGURE 5.21 – Reconstruction du MNT avec correction des perturbations globales par la seconde stratégie. MNT reconstruit aux cycles (a) 1, (b) 2, ..., (d) 4. (a) La reconstruction finale par FastDC.

Dans la figure 5.21, les résultats de l'estimation itérative par la seconde stratégie sont présentés. En partant du MNT SRTM, l'estimation des paramètres à la première itération est donc plus robuste. Ce qui accélère beaucoup plus la convergence vers une reconstruction meilleure que la précédente.

En gardant ce dernier résultat, les bons paramètres des perturbations globales qui ont été estimés sont :

$$\begin{cases} \alpha = \{0.01; 0; 0.01; -0.02; -0.01; 0\} \\ \beta = \{-1.64; 0.91; 0.09; 3.05; 2.14; 2.75\} \end{cases}$$

Ces perturbations sont donc faiblement dépendantes de la pente mais avec un facteur constant qui varie fortement d'un canal à l'autre.

Par ailleurs, nous visualisons clairement dans la figure 5.22 l'apport de la méthode multicanal proposée en terme de précision en comparant les deux MNT : estimé et SRTM. Ce qui laisse conclure d'une part à l'intérêt du recours à une approche multicanal et de l'autre part à la présence de perturbations atmosphériques globales sur les données interférométriques dont nous disposons, correctement corrigées par la méthode proposée.

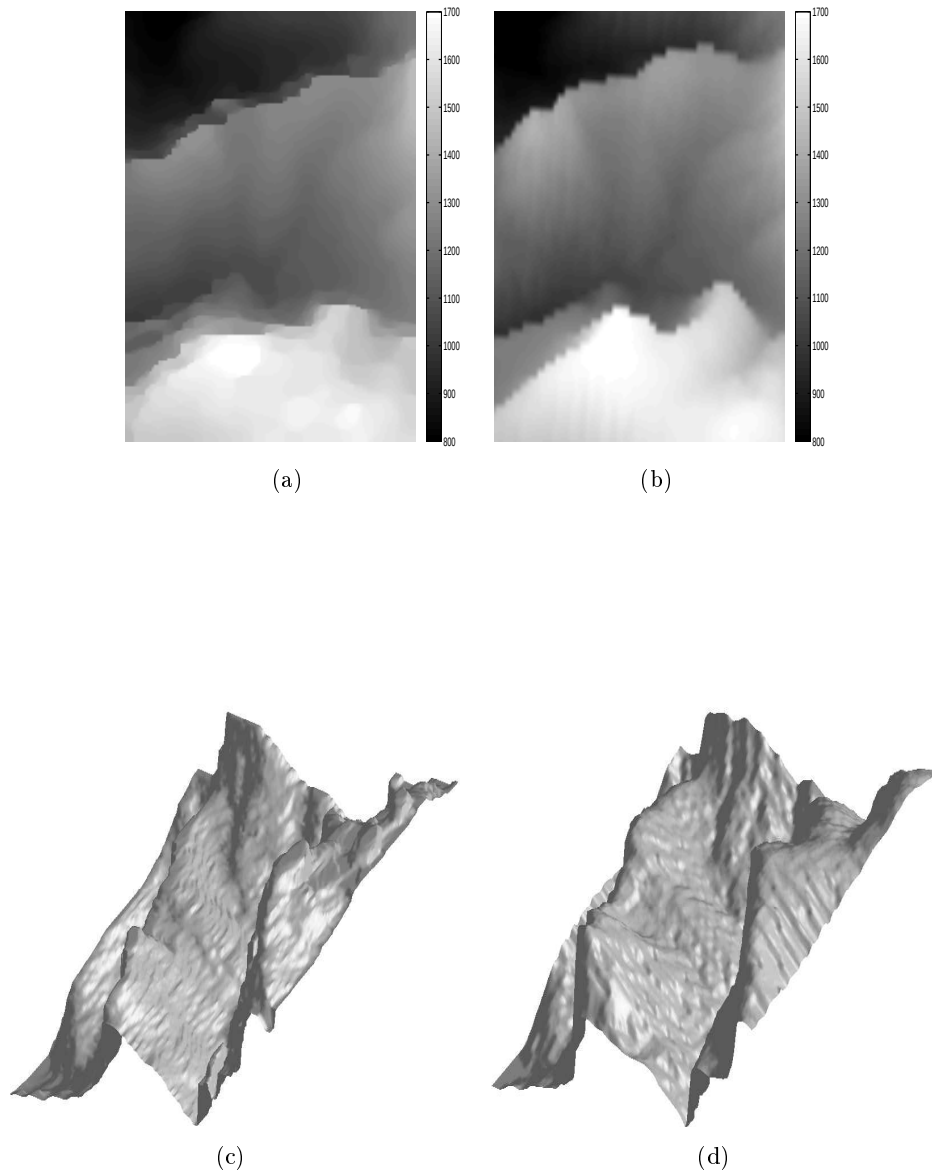


FIGURE 5.22 – Comparaison entre le MNT reconstruit et celui SRTM. (a) MNT reconstruit par notre approche, (b) MNT SRTM de la même zone, et leurs visualisations respectives en 3D dans (c) et (d).

Etant parvenu à valider l'approche sur une zone difficile à dérouler et ayant obtenu une reconstruction d'un MNT de bonne qualité, nous appliquons l'algorithme sur la totalité de la zone d'étude qui correspond à une montagne de la zone de Serre-Ponçon.

La figure 5.23 illustre les reconstructions obtenues par notre approche, avec la première et la seconde stratégie. Les visualisations en 3D de ces MNT sont données dans la figure 5.24. Les résultats sont comparés au MNT SRTM. L'optimisation MPLM est effectuée avec 32 étiquettes. Les reconstructions sont globalement satisfaisantes. Quelques problèmes apparaissent aux bords des images qui sont essentiellement dus à l'absence d'une information contextuelle suffisante pour reconstruire ces zones. De plus, le recours à une optimisation approximative n'assure pas tout le temps une convergence vers un optimum global. Un ensemble d'étiquettes plus large pourrait améliorer davantage la reconstruction.

Le temps de calcul mis par l'algorithme pour corriger les données et reconstruire l'alti-

tude de la scène par la première stratégie est de 66 : 20 (mn :sec)². Par la seconde stratégie, comme la convergence est plus rapide grâce à la bonne initialisation par le MNT SRTM, le temps de calcul mis par l'algorithme de correction et d'estimation d'altitudes est de 42 : 16 (mn :sec).

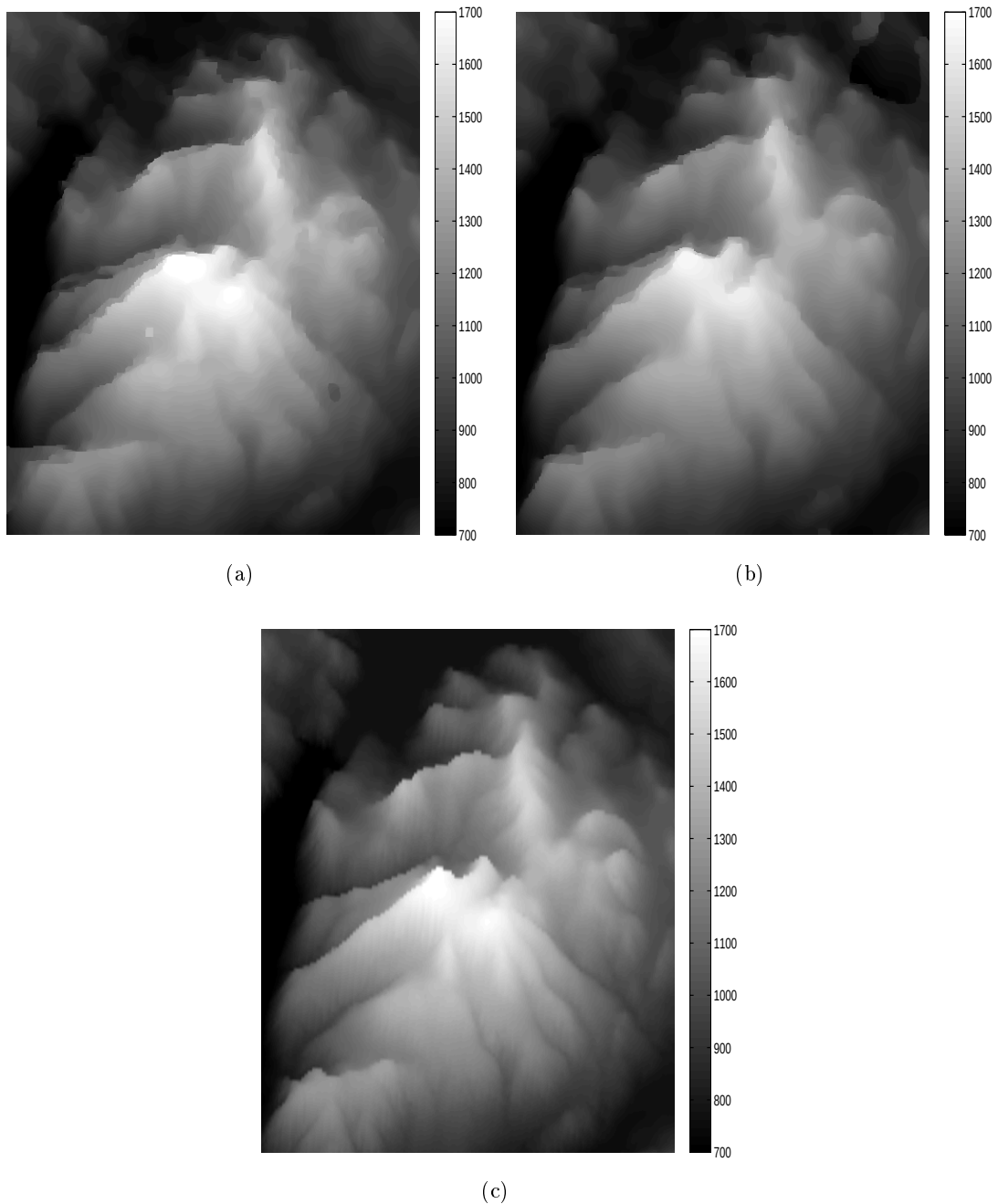


FIGURE 5.23 – Comparaison entre le MNT reconstruit par la première stratégie, la seconde et celui SRTM. (a) MNT reconstruit par la première stratégie, (b) par la seconde stratégie, et (c) le MNT SRTM de la même zone.

2. Processeur de la machine : Intel Core 2 duo, 2.66GHz.

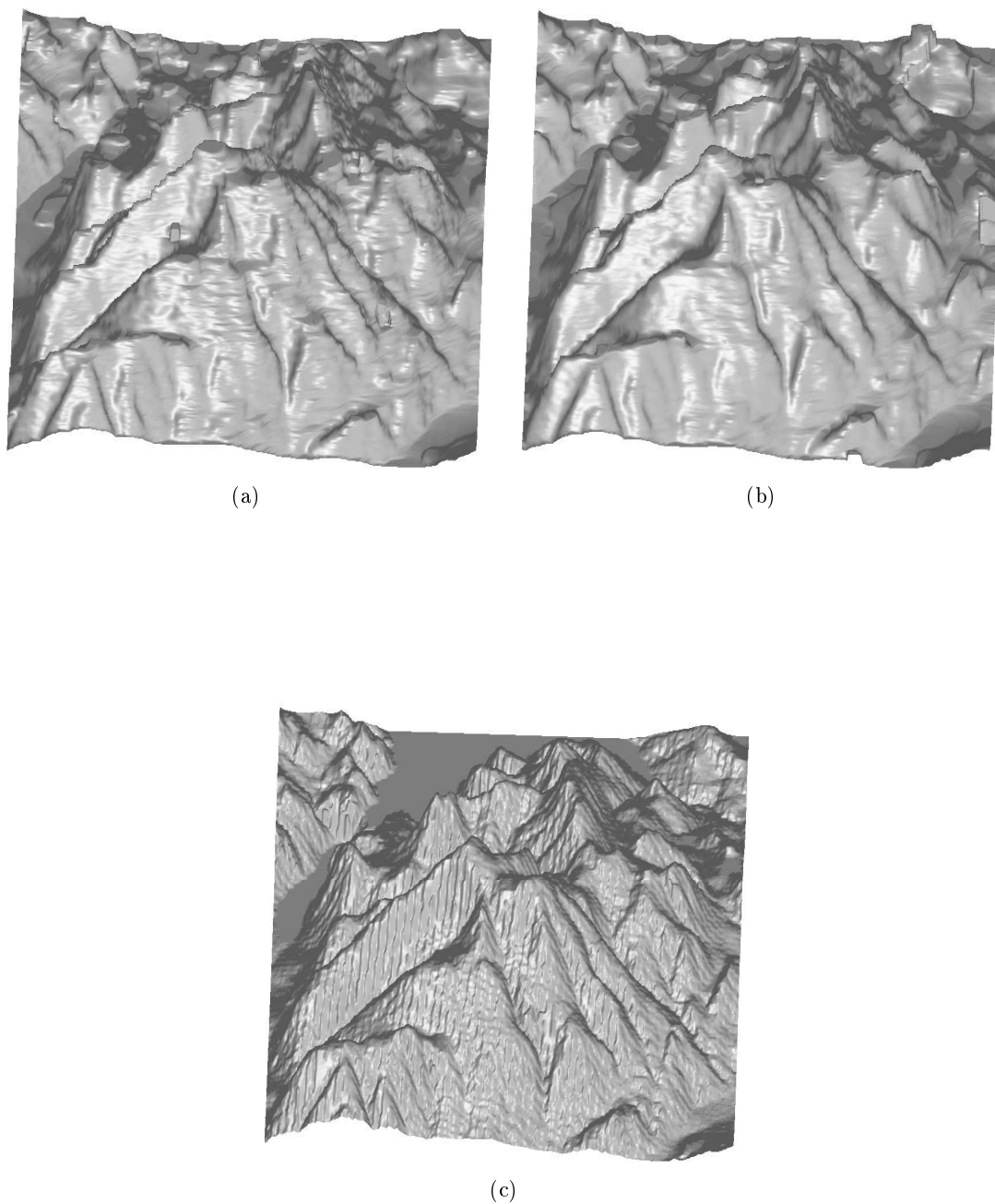


FIGURE 5.24 – Visualisation 3D des MNT obtenus. (a) MNT reconstruit par la première stratégie, (b) par la seconde stratégie, et (c) le MNT SRTM de la même zone.

Par les deux stratégies, la reconstruction est donc satisfaisante. Toutefois, on s'aperçoit de quelques limites de l'approche proposée :

1. En terme de tailles de données réelles à traiter, malgré la réduction du nombre d'étiquettes utiles pour optimiser efficacement l'énergie, en passant de l'ensemble [700, 1700] de taille 1000 à des ensembles d'étiquettes de tailles 32, la construction des graphes multi-étiquettes à 32 couches reste pénalisant pour reconstruire de très grandes zones, typiquement sur des données réelles volumineuses. Actuellement, le recours à des reconstructions successives de zones adjacentes est une alternative possible pour produire un MNT sur une très grande zone. L'effet de blocs qui caractérisera la reconstruction finale par une telle approche pourrait être éliminé par recours à l'information d'un MNT SRTM disponible sur la même zone dans l'étape de l'estimation des paramètres perturbateurs. De plus, cela permettra de paralléliser la méthode réduisant ainsi le temps de calcul nécessaire pour reconstruire la totalité de la zone.
2. Concernant le temps de calcul qui paraît important pour la reconstruction de la zone d'étude, il est à noter qu'il s'agit du temps de calcul mis par l'algorithme complet d'estimation conjointe des perturbations atmosphériques et de l'altitude. Ainsi un nombre d'itérations est nécessaire pour converger vers la bonne solution. Connaissant les paramètres des perturbations, le temps que met l'algorithme de reconstruction du MNT uniquement est d'une dizaine de minutes pour une telle zone.
3. Comparé au MNT SRTM, le MNT reconstruit semble présenter un aspect différent sur certaines frontières des zones élevées du milieu, qui correspondent en fait à des fortes discontinuités de zones de vallées. Ces zones, étant très étroites, et avec l'angle de visée du radar ERS-2, elles sont en partie cachées. Ainsi l'apparition d'une ambiguïté autour de ces lignes dans le MNT reconstruit par notre méthode s'explique par la présence de ces phénomènes dans tous les canaux que nous disposons. Sa correction serait alors quasiment impossible par le modèle proposé, si on se place dans la même géométrie radar de ces données.

Il est à noter par contre que sur le MNT fourni par la mission SRTM ces problèmes n'apparaissent pas, vu que son angle de visée est différent de celui du capteur ERS-2. Ces zones de vallée particulièrement sont complètement non visibles. D'où l'absence des problèmes de discontinuité à ces zones.

L'information apportée alors par le MNT reconstruit est plus précise et plus riche pour une interprétation de la scène imagée.

7 Conclusion

En conclusion, nous avons réussi dans ce chapitre à mettre en place un modèle markovien et un algorithme d'optimisation dédié pour reconstruire le modèle numérique de terrain avec une bonne précision à partir de données interférométriques. Face aux différentes perturbations qui caractérisent les mesures de phase InRSO, telles que les décorrélatons spatiales et temporelles, le bruit thermique et les perturbations atmosphériques globales de types linéaires avec la topographie, l'algorithme proposé a pu surmonter ces différents problèmes tout en gardant un aspect automatique, non supervisé et sans beaucoup pénaliser sa complexité. Sur des données de synthèse et réelles, les résultats obtenus sont suffisamment

satisfaisants pour pouvoir considérer le MNT reconstruit comme étant une bonne estimation des altitudes d'une scène naturelle. Cela est d'une grande importance en imagerie radar, où partant de MNT de bonne qualité, d'autres applications pourraient être réalisées efficacement. Nous citons à titre indicatif : l'interférométrie différentielle pour l'étude des déformations et déplacements des terrains, la reconstruction 3D de modèles numériques d'élévations dans des scènes urbaines, etc.

Dans l'approche proposée, parmi l'ensemble des paramètres qui influent directement sur la qualité de la reconstruction nous avons noté les hyperparamètres locaux. Ce derniers apportent au modèle contextuel une information supplémentaire sur les discontinuités à préserver sur l'image. La précision d'une telle information est capitale en présence de données caractérisées par des objets fins et structurés, comme c'est le cas des images acquises sur des zones urbaines. La méthode utilisée précédemment, qui consiste à estimer ces paramètres binaires à partir de la solution donnée par l'optimiseur MPLM, semble moins robuste dans ce cas de données. L'estimation serait très sensible à la qualité de la reconstruction obtenue par l'optimisation MPLM.

Nous nous proposons alors dans le prochain chapitre d'exploiter une autre information fournie par le produit interférométrique pour améliorer en robustesse l'approche de reconstruction de MNT. L'information contextuelle sera enrichie par l'information qu'apportent les images d'amplitude sur la même zone. Un nouveau modèle de reconstruction de MNT conjointement avec l'amplitude est alors introduit et étudié. Nous proposons par ailleurs de nouveaux optimiseurs efficaces qui lui sont dédiés. La méthode sera validée sur des données centrées sur des scènes urbaines.

Chapitre 6

Reconstruction conjointe de la phase InRSO et de l'amplitude RSO par une approche markovienne multicanal

1 Introduction

Dans ce chapitre, nous proposons une approche de reconstruction de la phase interférométrique conjointement avec l'amplitude radar. Cette méthode, comparée à celle proposée précédemment, permet de mieux préserver les structures des objets de la scène dans le modèle numérique de terrain reconstruit, si les données d'amplitude apportent une information complémentaire. Concrètement, elle est plus adaptée au problème de la reconstruction de la phase InRSO dans un milieu urbain, où les images d'amplitude se caractérisent par une forte localisation des bords des bâtiments et des structures réfléchissantes de la scène.

Puisque la reconstruction de la phase InRSO nécessite un choix judicieux des paramètres de régularisation afin d'éliminer le maximum de bruit sans pour autant détruire les structures présentes dans la scène, le recours à l'information d'amplitude est alors une solution avantageuse. Désormais, l'information de discontinuités n'est plus estimée, comme c'est le cas dans les approches par estimation des hyperparamètres locaux, elle est également apportée directement par l'image d'amplitude. La technique multicanal est aussi exploitée dans ce cadre pour pouvoir à la fois enlever l'ambiguïté sur la phase absolue et apporter plus d'observations sur l'amplitude. Cela permettra de réduire le bruit de chatoiement et assurer ainsi une meilleure localisation des vraies discontinuités de la scène.

Dans un premier temps, le modèle markovien multicanal permettant de reconstruire la phase interférométrique et régulariser l'amplitude conjointement est présenté.

Comme la solution à estimer est formée de deux images, la phase et l'amplitude, un algorithme d'optimisation vectorielle est alors plus approprié. Ainsi nous proposons dans un deuxième temps une étude sur des optimiseurs multi-étiquettes par coupe-minimum sur graphes reposant sur des mouvements de partition binaires et conjoints phase-amplitude. Nous montrons par ailleurs que l'optimisation conjointe par des mouvements d'extension permet d'explorer un espace de configurations plus grand que par des mouvements de saut, aboutissant ainsi à de meilleurs optima locaux.

Enfin, nous testons l'approche sur des données InRSO simulées et des données réelles de milieux urbains. Des comparaisons entre les résultats obtenus et ceux que donne l'approche

MCPU+TV sont aussi réalisées.

Ce travail a été publié dans la conférence internationale International Geoscience and Remote Sensing Symposium (IGARSS 2010) (Shabou *et al.*, 2010b).

2 Modèle markovien conjoint (MCPAR+TV)

Considérons dans un premier temps le cas monocal. Les auteurs de (Denis *et al.*, 2009a) proposent deux modèles markoviens différents permettant de régulariser conjointement la phase interférométrique et les données d'amplitude d'une même zone.

Le premier modèle exploite la technique conjointe au niveau de l'*a priori* uniquement, les informations de phase et d'amplitude étant supposées indépendantes pour les probabilités conditionnelles ; tandis que le deuxième l'exploite pour modéliser l'information *a priori* mais aussi la vraisemblance des données de phase et d'amplitude en utilisant la distribution exacte des données InRSO.

La forme générale de l'énergie totale considérée est la suivante :

$$E(\phi, a) = \sum_p E_p(\psi_p, e_p^{(1)}, e_p^{(2)}, \gamma_p | \phi_p, a_p) + \sum_{(p,q)} E_{p,q}(\phi_p, \phi_q, a_p, a_q), \quad (6.1)$$

où ϕ et a sont respectivement l'image de phase et l'image d'amplitude régularisées et ψ , $e^{(1)}$ et $e^{(2)}$ et γ sont respectivement l'interférogramme, les deux images d'amplitude et l'image de cohérence du couple interférométrique.

La fonction de régularisation est la même pour les deux modèles. Elle est choisie de façon à conserver toutes les discontinuités qui sont présentes en amplitude et en phase. Pour cela, un opérateur *max* entre deux *a priori* de type TV sur les deux données est utilisé. La forme de la fonction de régularisation introduite est donc la suivante :

$$E_{p,q}(\phi_p, \phi_q, a_p, a_q) = \max(|\phi_p - \phi_q|, \eta|a_p - a_q|), \quad (6.2)$$

avec η un facteur de normalisation entre les deux dynamiques des ensembles des étiquettes de la phase et de l'amplitude.

Concernant les modèles de vraisemblance, le premier modèle proposé est non-exact au sens de la distribution de la loi de vraisemblance conjointe phase-amplitude. Il est constitué de deux distributions indépendantes. La première est une distribution Nakagami sur les données d'amplitude et la deuxième est une distribution normale sur les données de phase. Le modèle résultant est le suivant :

$$E_p(\psi_p, e_p^{(1)}, e_p^{(2)} | \phi_p, a_p, \gamma_p) = \frac{(\psi_p - \phi_p)^2}{\sigma_{\phi_p}^2} + M \frac{e_p^2}{a_p^2} + 4 \log a_p, \quad (6.3)$$

avec M le nombre de vues sur les observations et σ_{ϕ_p} l'écart-type au site p défini par $\sigma_{\phi_p}^2 = \frac{1-\gamma_p^2}{2M\gamma_p^2}$. L'image d'amplitude e est considérée à la place des deux images $e^{(1)}$ et $e^{(2)}$ et elle est définie par $e = \sqrt{\frac{1}{2}(e^{(1)})^2 + \frac{1}{2}(e^{(2)})^2}$.

Comme indiqué dans l'article (Denis *et al.*, 2009a), ce premier modèle approximatif n'est valable qu'en présence de données multivues, avec un nombre de vues supérieur à 6. Cependant, il ne permet en aucun cas de dérouler la phase. Il s'agit uniquement d'un modèle de régularisation.

Le deuxième modèle proposé dans le même travail est un modèle exact selon la distribution conjointe phase-amplitude en interférométrie RSO. La fonction de vraisemblance

est tirée des statistiques de la distribution de la matrice de covariance empirique entre les données RSO du couple interférométrique. En reprenant les notations de l'annexe E, la matrice de covariance du vecteur Z formé des deux composantes z_1 et z_2 (les deux images du couple interférométrique) est donnée par :

$$C = Z^t Z^* = \begin{pmatrix} (e^{(1)})^2 & e^{(1)}e^{(2)}e^{j\psi} \\ e^{(1)}e^{(2)}e^{-j\psi} & (e^{(2)})^2 \end{pmatrix}, \quad (6.4)$$

avec $e^{(1)}$ et $e^{(2)}$ étant respectivement les amplitudes des complexes z_1 et z_2 et ψ le déphasage entre eux.

Le calcul analytique de la distribution conjointe des éléments de cette matrice donne :

$$P(\psi, e^{(1)}, e^{(2)} | \phi, a, \gamma) = \frac{2e^{(1)}e^{(2)}}{\pi a^4(1-\gamma^2)} \exp \left(-\frac{1}{a^2(1-\gamma^2)} \left(\frac{(e^{(1)})^2}{a^2} + \frac{(e^{(2)})^2}{a^2} - \frac{2\gamma e^{(1)}e^{(2)} \cos(\psi - \phi)}{a^2} \right) \right). \quad (6.5)$$

Ainsi la fonction de vraisemblance pour tout pixel p est donnée par :

$$E_p(\psi_p, e_p^{(1)}, e_p^{(2)} | \phi_p, a_p, \gamma_p) = 4 \log a_p + \frac{1}{a_p^4(1-\gamma_p^2)} \left((e_p^{(1)})^2 + (e_p^{(2)})^2 - 2e_p^{(1)}e_p^{(2)}\gamma_p \cos(\psi_p - \phi_p) \right). \quad (6.6)$$

Ce modèle de vraisemblance, contrairement au précédent, tient compte à la fois de l'enroulement de la phase et du bruit caractéristique des données interférométriques. De plus, il est aussi valable autant en monovue qu'en multivue ; dans ce dernier cas, il suffirait de multiplier l'expression par un facteur M (le nombre de vues). Cependant, il s'agit d'une fonction de vraisemblance non-convexe, périodique, difficile à minimiser. Des problèmes d'optimisation ont été d'ailleurs illustrés dans les résultats de l'article (Denis *et al.*, 2009a).

Dans notre étude, nous généralisons ce modèle pour une application nécessitant le déroulement de la phase dans un contexte multicanal et sa régularisation conjointement avec l'image d'amplitude. Ainsi l'information apportée par l'amplitude sera d'une importance double. D'une part, elle permettra de mieux localiser les discontinuités à préserver durant l'étape de régularisation de la phase, et vice versa. D'autre part, elle apportera une information complémentaire pour mieux dérouler la phase dans les zones à très faible cohérence, où les structures des objets de la scène deviennent ambiguës et difficiles à reconstruire.

2.1 Modèle de vraisemblance

Le modèle de vraisemblance que nous considérons dans notre approche est un modèle multicanal, défini à partir de la distribution conjointe phase-amplitude donnée dans (6.5).

Nous supposons par ailleurs l'indépendance entre les canaux dans le modèle de vraisemblance multicanal. Sans cette hypothèse, la distribution de la matrice de covariance faisant intervenir tous les canaux disponibles devrait être calculée. La tâche de déterminer analytiquement une telle distribution étant très complexe, l'hypothèse d'indépendance entre les canaux est alors considérée par la suite dans la distribution conjointe multicanal, donnant la forme suivante :

$$P(\Psi, E^{(1)}, E^{(2)} | \phi, a, \Gamma) = \prod_c P(\psi_{p,c}, e_c^{(1)}, e_c^{(2)} | \phi, a, \gamma_c), \quad (6.7)$$

avec $P(\psi_p, e_c^{(1)}, e_c^{(2)}, \gamma_c | \phi, a)$ la distribution conjointe phase-amplitude associée au canal c , donnée dans (6.5), Ψ , $E^{(1)}$, $E^{(2)}$ et Γ respectivement des vecteurs collectant les différentes observations de phases et amplitudes et les cohérences sur tous les canaux disponibles.

Le recours à ce modèle multicanal, avec une distribution conjointe phase-amplitude à chaque canal, permettra d'enlever l'ambiguïté de phase, comme indiqué au chapitre précédent, et d'apporter plus d'observations (supposées indépendantes) sur les images d'amplitude, réduisant ainsi le bruit du chatoiement présent sur ces données. Ainsi la maximisation de la vraisemblance en utilisant ce modèle donne en résultat une image de phase déroulée et une image d'amplitude moins bruitée.

2.2 Modèle *a priori*

Comme il a été indiqué dans le chapitre précédent, les données de phase fortement décorréliées (décorrélation spatiale, et/ou aussi décorrélation temporelle) sont difficilement exploitables pour reconstruire l'information d'altitude. Un nombre important de canaux est nécessaire pour pouvoir dérouler la phase en présence de fortes discontinuités. De plus, l'estimation des hyperparamètres locaux de discontinuités intervient aussi dans la qualité de la reconstruction. Une information sur les structures qui constituent la zone, si elle existe, permettra donc de mieux les localiser et donc de mieux reconstruire l'image d'altitudes. Ainsi exploiter les discontinuités présentes sur les images d'amplitude régularisées est une alternative possible pour améliorer la qualité de la reconstruction de la phase interférométrique. D'où l'intérêt d'un modèle *a priori* conjoint qui décrit le contexte global de l'image de phase reconstruite en reposant sur l'image de phase elle-même et sur l'image d'amplitude.

Le modèle *a priori* proposé dans (Denis *et al.*, 2009a) nous semble bien adapté et nous considérons l'énergie de régularisation suivante :

$$E_{p,q}(\phi_p, \phi_q, a_p, a_q) = \max(|\phi_p - \phi_q|, \eta|a_p - a_q|). \quad (6.8)$$

3 Minimisation markovienne vectorielle

La minimisation de la fonction d'énergie globale définie précédemment est un problème difficile pour les deux raisons suivantes :

- la fonction est non-convexe,
- pour chaque pixel p , deux variables sont à estimer conjointement : l'amplitude et la phase.

Le recours aux approches de minimisation par coupe-minimum semble le plus approprié à ce problème pour une minimisation efficace.

Les optimiseurs multi-étiquettes exacts, tels que celui d'Ishikawa ou de Darbon qui sont présentés dans le premier chapitre de ce manuscrit, ne pourraient pas être utilisés pour ce type de fonctions. En effet, la solution à estimer dans notre problème est une image vectorielle, où chaque pixel est affecté d'un vecteur d'étiquettes dans $\mathcal{L}_\phi \times \mathcal{L}_a$. Il n'est donc plus possible d'ordonner les étiquettes, condition nécessaire pour la construction des graphes des optimiseurs exacts.

De même, les optimiseurs multi-étiquettes proposés dans cette thèse ne sont plus utilisables dans ce cas de problème d'optimisation pour la même raison.

Ainsi on se ramène aux approches d'optimisation par des mouvements de partitions binaires, où l'ordre n'est plus nécessaire pour la construction des graphes. Nous rappelons

que pour cette famille d'optimiseurs, des mouvements binaires optimaux sont effectués itérativement jusqu'à la convergence de l'algorithme d'optimisation. Un mouvement binaire implique que deux états sont possibles pour chaque site de la solution : chaque site a la possibilité soit de changer son état (ou étiquette) courant, soit de le garder. Un mouvement global et optimal de la configuration est obtenu en construisant un graphe binaire, où une coupe-minimum sur ce graphe répartit les noeuds en deux classes d'étiquettes : la classe des noeuds liés à la source et la classe des noeuds liés au puits, où la source indique un changement d'état et le puits le non-changement (ou le contraire). La condition nécessaire et suffisante pour la construction d'un tel graphe est la sous-modularité des termes de l'énergie à minimiser.

L'aspect vectoriel de notre estimation, dans le sens où une image vectorielle formée de la phase et de l'amplitude est à estimer, nécessite une définition claire d'un mouvement de partition binaire dans le cas vectoriel. Ensuite, un algorithme adapté à ce type de mouvement devra être mis en place.

3.1 Mouvements binaires vectoriels

Indiquons pour la suite les ensembles discrets des étiquettes de la phase et de l'amplitude par $\mathcal{L}_\phi$ et $\mathcal{L}_a$ respectivement.

Dans le cas non-vectoriel, un mouvement binaire pour un site donné de l'image peut être exprimé de la façon suivante :

$$x_p = x_p^{(0)}(1 - e_p) + x_p^{(1)}e_p, \quad (6.9)$$

où x_p est la configuration du pixel p après le mouvement binaire, $x_p^{(0)}$ est la configuration de ce pixel si l'état binaire e_p après mouvement vaut 0 et $x_p^{(1)}$ est la configuration de ce même pixel si l'état binaire e_p après mouvement vaut 1.

A partir de cette écriture, nous reconnaissons les différents mouvements binaires présentés dans le premier chapitre de la thèse :

- une α -extension implique $x_p^{(0)}$ est l'étiquette courante du site p et $x_p^{(1)}$ est la nouvelle étiquette α proposée ;
- un $\alpha\beta$ -échange implique $x_p^{(0)}$ est l'étiquette α et $x_p^{(1)}$ est l'étiquette β (ou le contraire), sachant que les sites considérés par ce changement sont ceux dont les étiquettes sont soit α soit β pour rester dans le cadre d'un mouvement binaire ;
- un δ -saut implique $x_p^{(0)}$ est l'étiquette courante du site p et $x_p^{(1)}$ est la nouvelle étiquette proposée en augmentant ou diminuant l'étiquette courante par la quantité δ , sans dépasser les bornes maximale et minimale des étiquettes possibles.

Partant de cette définition d'un mouvement binaire, les auteurs de (Denis *et al.*, 2009a) considèrent le même raisonnement dans le cas vectoriel où l'ensemble d'états possibles est toujours $\{0, 1\}$ donnant ainsi des configurations vectorielles du vecteur $\mathbf{x}_p = [\phi_p, a_p]^T$ selon l'expression suivante :

$$\mathbf{x}_p = \mathbf{x}_p^{(0)}(1 - e_p) + \mathbf{x}_p^{(1)}e_p, \quad (6.10)$$

où $\mathbf{x}_p^{(0)} = [\phi_p^{(0)}, a_p^{(0)}]^T$ est la configuration du vecteur $[\phi, a]^T$ si l'état binaire e_p après mouvement vaut 0 et $\mathbf{x}_p^{(1)} = [\phi_p^{(1)}, a_p^{(1)}]^T$ est la configuration de ce même vecteur si l'état binaire e_p après mouvement vaut 1.

Une autre façon de définir un mouvement binaire dans le cas vectoriel, qui n'a pas été mentionnée dans (Denis *et al.*, 2009a), consiste à considérer qu'un site quelconque de l'image de phase possède un ensemble d'états qui soit binaire $\{0, 1\}$ mais aussi indépendant de celui du site correspondant dans l'image d'amplitude. Formellement, on considérera le système de configurations suivant :

$$\begin{cases} \phi_p = \phi_p^{(0)}(1 - e_p^\phi) + \phi_p^{(1)}e_p^\phi, \\ a_p = a_p^{(0)}(1 - e_p^a) + a_p^{(1)}e_p^a, \end{cases} \quad (6.11)$$

où cette fois-ci un vecteur d'états binaires $[e^\phi, e^a]^T$ est considéré à la place d'un seul état binaire e pour la phase et l'amplitude ensemble, autorisant ainsi plus de configurations possibles par un seul mouvement vectoriel. Plus généralement, dans le cas vectoriel, chaque composante du vecteur a son propre mouvement binaire.

Pour simplifier l'appel à ces deux définitions de mouvement vectoriel binaire dans la suite du chapitre, nous appelons le premier mouvement vectoriel binaire incomplet et le second mouvement vectoriel binaire complet, dans le sens où le second explore plus de configurations d'états possibles en un seul mouvement. Les deux tableaux de la figure 6.1 montrent les différents états possibles par ces deux différentes définitions de mouvement binaire vectoriel.

e_p	$\mathbf{x}_p$	$\mathbf{e}_p$	$\mathbf{x}_p$
0	$[\phi_p^{(0)}, a_p^{(0)}]^T$	$[0, 0]^T$	$[\phi_p^{(0)}, a_p^{(0)}]^T$
1	$[\phi_p^{(1)}, a_p^{(1)}]^T$	$[0, 1]^T$	$[\phi_p^{(0)}, a_p^{(1)}]^T$
		$[1, 0]^T$	$[\phi_p^{(1)}, a_p^{(0)}]^T$
		$[1, 1]^T$	$[\phi_p^{(1)}, a_p^{(1)}]^T$

(a)
(b)

FIGURE 6.1 – Configurations possibles par les deux définitions d'un mouvement vectoriel binaire. États possibles par un mouvement vectoriel binaire (a) tel qu'il a été défini dans (Denis *et al.*, 2009a), (b) selon la deuxième définition.

Ayant défini le mouvement de partitions vectoriel binaire, le but est de pouvoir calculer efficacement par la technique de coupe-minimum sur graphe, à une itération donnée de l'algorithme d'optimisation, la configuration optimale par un tel mouvement. Une construction de graphe appropriée est alors nécessaire.

3.2 Mouvement vectoriel binaire optimal

Étant donné que la fonction d'énergie à minimiser est une fonction multi-étiquette, un algorithme de minimisation reposant sur les mouvements binaires nécessite le choix de la stratégie de mouvement à adopter pour explorer l'espace des configurations possibles.

En reprenant l'étude menée dans la première partie de ce manuscrit de thèse, les deux stratégies de type saut et extension semblent les plus performantes et donnent des résultats satisfaisants en terme de compromis entre temps de calcul et qualité de l'optimum. Le mouvement de type échange ne sera donc pas exploré dans cette partie de l'étude.

Étudions en premier lieu le cas du mouvement incomplet, et ensuite le cas du mouvement complet. Enfin, nous concluerons sur le choix à faire entre les deux.

3.2.1 Mouvement vectoriel binaire incomplet

Saut vectoriel :

Partant d'une configuration initiale $\mathbf{x}^{(0)} = [\phi^{(0)}, a^{(0)}]^T$ de l'image vectorielle phase-amplitude, un mouvement vectoriel binaire incomplet de type saut permet d'atteindre la configuration $\mathbf{x}$ selon la relation suivante :

$$\mathbf{x}_p = \mathbf{x}_p^{(0)}(1 - e_p) + \underbrace{(\mathbf{x}_p^{(0)} + \delta)}_{=\mathbf{x}_p^{(1)}} e_p, \quad \forall p \in \mathcal{P}, \quad (6.12)$$

avec δ le pas de saut vectoriel défini par $\delta = [\delta_\phi, \delta_a]^T$.

C'est ce mouvement vectoriel binaire qui a été exploré dans (Denis *et al.*, 2009a).

Pour réaliser alors le mouvement de saut binaire optimal, la construction du graphe nécessaire est exactement la même que celle décrite pour les mouvements de partition binaires non-vectoriels. En effet, comme l'ensemble d'états est toujours un ensemble binaire $e \in \{0, 1\}$, un graphe binaire suffit alors pour obtenir le mouvement optimal. Chaque noeud du graphe correspond à un site p indiquant à la fois le site p de l'image de phase et le même site de l'image d'amplitude. Les deux noeuds spéciaux, la source et le puits, sont ajoutés au graphe pour construire le réseau où le flot-maximum sera calculé. Les capacités sur les arcs sont définies de la même façon que dans le premier chapitre du manuscrit. La figure 6.2 illustre la structure répétitive de ce graphe sur tous les sites voisins (p, q) .

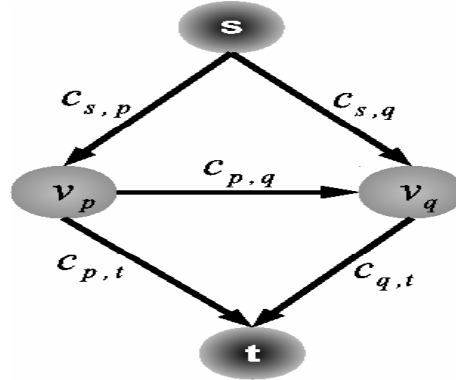


FIGURE 6.2 – Structure répétitive du graphe binaire vectoriel sur toutes les cliques d'ordre 1.

Le calcul en temps polynômial du flot-maximum sur ce graphe nécessite la condition de sous-modularité de *l'a priori* relativement à la nature du mouvement binaire. Comme la fonction de régularisation utilisée, étant une opération maximum entre deux fonctions convexes, est une fonction convexe, l'inégalité de sous-modularité est bien vérifiée :

$$\underbrace{E_{p,q}(\mathbf{x}_p, \mathbf{x}_q) + \underbrace{E_{p,q}(\mathbf{x}_p + \delta, \mathbf{x}_q + \delta)}_{E_{p,q}(\mathbf{x}_p, \mathbf{x}_q)}}_{=2E_{p,q}(\mathbf{x}_p, \mathbf{x}_q)} \leq E_{p,q}(\mathbf{x}_p + \delta, \mathbf{x}_q) + E_{p,q}(\mathbf{x}_p, \mathbf{x}_q + \delta) \quad (6.13)$$

Extension vectorielle :

Partons d'une configuration initiale $\mathbf{x}^{(0)} = [\phi^{(0)}, a^{(0)}]^T$ de l'image vectorielle phase-amplitude. Un mouvement vectoriel binaire incomplet de type extension permet d'atteindre

la configuration $\mathbf{x}$ selon la relation suivante :

$$\mathbf{x}_p = \mathbf{x}_p^{(0)}(1 - e_p) + \underbrace{\alpha}_{=\mathbf{x}_p^{(1)}} e_p, \quad \forall p \in \mathcal{P}, \quad (6.14)$$

avec α l'étiquette vectorielle définie par $\alpha = [\alpha_\phi, \alpha_a]^T$.

Pour réaliser alors le mouvement d'extension binaire optimal, la même construction du graphe que la précédente est réalisée. Les capacités sur les arcs sont définies de la même façon que dans le premier chapitre du manuscrit. Le calcul en temps polynômial du flot-maximum sur ce graphe nécessite la condition de sous-modularité de l'*a priori* relativement à la nature du mouvement binaire. Comme la fonction de régularisation utilisée est une métrique, appelée distance de Tchebychev, l'inégalité de sous-modularité est bien vérifiée :

$$\underbrace{E_{p,q}(\mathbf{x}_p, \mathbf{x}_q) + \underbrace{E_{p,q}(\alpha, \alpha)}_{=0}}_{E_{p,q}(\mathbf{x}_p, \mathbf{x}_q)} \leq E_{p,q}(\mathbf{x}_p, \alpha) + E_{p,q}(\alpha, \mathbf{x}_q) \quad (6.15)$$

3.2.2 Mouvement vectoriel binaire complet

Revenons maintenant au mouvement vectoriel binaire complet.

Saut vectoriel :

Dans le cas du saut et partant d'une configuration initiale $\mathbf{x}^{(0)} = [\phi^{(0)}, a^{(0)}]^T$, le mouvement complet s'obtient par le système de relations suivantes :

$$\begin{cases} \phi_p = \phi_p^{(0)}(1 - e_p^\phi) + \underbrace{(\phi_p^{(0)} + \delta_\phi)}_{=\phi_p^{(1)}} e_p^\phi, \\ a_p = a_p^{(0)}(1 - e_p^a) + \underbrace{(a_p^{(0)} + \delta_a)}_{=a_p^{(1)}} e_p^a. \end{cases} \quad (6.16)$$

Pour réaliser le mouvement de saut optimal dans ce cas de figure, le graphe binaire n'est plus adapté. La construction appropriée devrait inclure d'autres états qui ne sont pas considérés par un mouvement binaire incomplet. Une façon pour obtenir le changement de configuration optimal par ce mouvement consiste à considérer que chaque terme E_p d'attache aux données de l'énergie est une fonction à deux variables (ϕ_p, a_p) et que chaque terme $E_{p,q}$ de régularisation est une fonction à quatre variables $(\phi_p, \phi_q, a_p, a_q)$. Ainsi l'image vectorielle phase-amplitude pourrait être vue comme étant une seule image où mis à part les interactions en rapport avec les cliques d'ordre 0 et 1 du modèle markovien de premier ordre, des interactions supplémentaires existent. La première famille de ces interactions est due à la vraisemblance conjointe et la deuxième famille d'interactions est due à la régularisation conjointe. Le mouvement vectoriel binaire se ramène à un mouvement binaire non-vectoriel faisant intervenir des cliques d'ordre 0, 1 et 3.

Avec cette façon de voir un mouvement vectoriel binaire complet, on se ramène à un cas particulier d'un saut binaire non-vectoriel, qui n'a pas été discuté jusqu'à ce stade du manuscrit, à savoir que le nouveau mouvement de saut non-vectoriel qui découle de la transformation de l'énergie vectorielle en une énergie non-vectorielle diffère d'une partie de l'image à une autre. Pour une partie, le changement implique un saut de l'étiquette

courante d'une valeur δ_ϕ et pour l'autre partie, le changement implique un saut de l'étiquette courante d'une valeur de δ_a .

Pour simplifier l'étude, nous nous mettons dans un cas particulier de ce saut, où il s'agit du même saut pour les deux parties, c.-à-d. $\delta_\phi = \delta_a = \delta$. Le mouvement binaire optimal est alors réalisé en construisant le graphe approprié à une fonction markovienne d'ordre 3. Dans (Zalesky, 2003b), une construction d'un tel graphe est proposée. Pour les cliques d'ordre 3, la structure répétitive dans le graphe sur toutes ces cliques est présentée dans la figure 6.3.

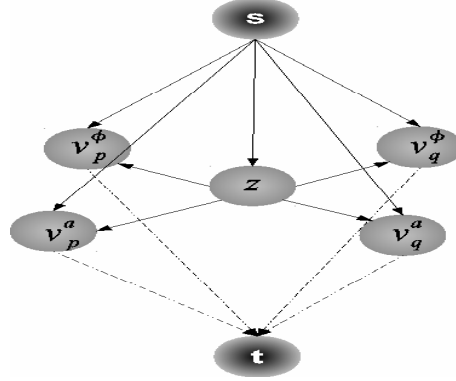


FIGURE 6.3 – Structure répétitive du graphe binaire vectoriel sur toutes les cliques d'ordre 3.

Ainsi un mouvement binaire pour une telle fonction d'énergie d'ordre supérieur accorde une indépendance de changement d'états entre les sites de la phase et ceux de l'amplitude, comme il a été défini pour ce mouvement vectoriel complet. Cependant, on se ramène à un problème d'optimisation d'une fonction d'énergie d'ordre supérieur. La fonction est alors graphe représentable si elle est sous-modulaire.

Nous rappelons la définition de sous-modularité dans le cas général où la fonction markovienne est multi-étiquette et d'ordre quelconque (Flach et Schlesinger, 2004) (Kohli, 2007).

Définition 3.1. Sous-modularité

- Toute fonction f multi-étiquettes définie sur $\mathcal{L}$ à valeurs dans $\mathbb{R}$ à une variable est sous-modulaire.
- Toute fonction f multi-étiquettes à deux variables définie sur $\mathcal{L}^2$ à valeurs dans $\mathbb{R}$, tel qu'un ordre ascendant ou descendant existe entre toute paire (l_1, l_2) d'éléments de $\mathcal{L}$, est sous-modulaire si $f(l_1, l_2) + f(l_1 + 1, l_2 + 1) \leq f(l_1, l_2 + 1) + f(l_1 + 1, l_2)$.
- Toute fonction f multi-étiquettes à plusieurs variables, plus que deux, est sous-modulaire si toutes ses projections sur deux variables sont des fonctions sous-modulaires.

Nous rappelons aussi que les projections d'une fonction f à n variables, avec $n > 2$, sont les fonctions obtenues à partir de f en fixant à chaque fois $n - 2$ variables.

Dans notre cas de fonction d'énergie, on peut montrer que la fonction est non sous-modulaire selon la définition 3.1 puisque la fonction *a priori* ne l'est pas (de même pour la fonction de vraisemblance). En effet, un contre exemple suffit pour le conclure :

Sans perdre de généralités, considérons le même ensemble $\mathcal{L}$ d'étiquettes pour la nouvelle image formée des images de phase absolue et d'amplitude (cela est toujours possible en normalisant les dynamiques des deux images dans un même ensemble). Fixons tout d'abord les variables a_p et ϕ_p aux étiquettes l_1 et l_2 de $\mathcal{L}$ et vérifions la condition de sous-modularité

de la fonction de régularisation projetée, définie uniquement sur les deux variables qui restent, en faisant un pas de saut $\delta = 1$. Le contre exemple suivant suffit pour conclure que cette fonction obtenue par projection est non sous-modulaire :

$$\underbrace{E_{p,q}(l_1, \mathbf{l}_2, l_2, \mathbf{l}_2) + E_{p,q}(l_1, \mathbf{l}_2 + \mathbf{1}, l_2, \mathbf{l}_2 + \mathbf{1})}_{=|\Delta l| + \max(|\Delta l - 1|, 1)} \not\leq \underbrace{E_{p,q}(l_1, \mathbf{l}_2, l_2, \mathbf{l}_2 + \mathbf{1}) + E_{p,q}(l_1, \mathbf{l}_2 + \mathbf{1}, l_2, \mathbf{l}_2)}_{=\max(|\Delta l - 1|, 1) + |\Delta l - 1|}, \quad (6.17)$$

avec $\Delta l = l_1 - l_2$ et les notations en gras correspondent aux configurations des deux variables qui peuvent changer d'états. Cette inégalité n'est pas toujours vraie, il suffit de considérer $\Delta l \geq 1$.

La non sous-modularité de la fonction ramène le problème de la recherche du mouvement vectoriel binaire complet optimal en un problème NP-difficile (Zalesky, 2003b). Sa résolution pourrait alors s'effectuer approximativement grâce aux approches d'optimisation approximative par QPBO, dont on a parlé dans le chapitre 1, sans garantir l'efficacité de l'algorithme d'optimisation ni la bonne qualité de l'optimum vers lequel il convergera.

Extension vectorielle :

Pour l'extension avec un mouvement vectoriel binaire complet, sans reprendre les détails du mouvement comme précédemment, nous montrons directement que la fonction *a priori* n'est pas graphe représentable pour un tel mouvement binaire. Il suffit de donner un exemple où une projection de cette fonction ne vérifie pas l'inégalité de sous-modularité. Fixons les variables a_p et ϕ_p aux étiquettes l_1 et l_2 de $\mathcal{L}$ et vérifions la condition de sous-modularité de la fonction de régularisation projetée, définie uniquement sur les deux variables qui restent, en faisant une extension à l'étiquette l_1 . On obtient :

$$\underbrace{E_{p,q}(l_1, \mathbf{l}_2, l_2, \mathbf{l}_2) + E_{p,q}(l_1, \mathbf{l}_1, l_2, \mathbf{l}_1)}_{=2|\Delta l|} \not\leq \underbrace{E_{p,q}(l_1, \mathbf{l}_2, l_2, \mathbf{l}_1) + E_{p,q}(l_1, \mathbf{l}_1, l_2, \mathbf{l}_2)}_{=|\Delta l|}. \quad (6.18)$$

Cela conclut sa non sous-modularité. Donc le mouvement optimal exact par cette extension vectorielle ne pourrait pas être obtenu par coupe-minimum sur graphe. L'approximation par la technique QPBO est nécessaire.

3.2.3 Choix entre mouvement vectoriel binaire complet ou incomplet

Si on compare au final les deux types de mouvements vectoriels binaires décrits, on a :

- dans le cas de mouvement complet, comme la fonction est non sous-modulaire pour les deux stratégies étudiées, le mouvement binaire optimal ne pourrait pas être obtenu. Un mouvement sous-optimal serait alors obtenu par minimisation approximative ;
- dans le cas de mouvement incomplet, moins de configurations sont possibles par ce mouvement comparé au précédent, sauf que pour ces configurations un mouvement binaire optimal exact est obtenu par coupe-minimum sur un graphe binaire approprié.

Nous avons choisi dans ces travaux de thèse de développer l'étude sur les mouvements vectoriels binaires incomplets pour la raison simple est que l'optimisation itérative de la fonction d'énergie conjointe phase-amplitude se réalisera par des mouvements de partitions optimaux, où le calcul du flot-maximum est réalisé à chaque itération d'une façon très efficace, avec une construction de graphe très réduite (en $O(n)$). Comme l'ensemble d'états possibles par un tel mouvement est moins riche que ce qu'offre un mouvement complet,

nous nous proposons alors d'explorer diverses stratégies possibles de mouvements binaires (par exemple : extension, saut, etc.) dans le but d'en exploiter la meilleure, c'est à dire celle qui explore le mieux possible l'espace des configurations possibles par ces mouvements incomplets.

3.3 Algorithme de minimisation vectorielle

Nous étudions dans le cadre des mouvements vectoriels binaires incomplets, les stratégies de type saut et extension.

3.3.1 Minimisation par des sauts vectoriels :

L'algorithme de minimisation par des sauts vectoriels présenté dans (Denis *et al.*, 2009a) peut être présenté de la façon suivante, en reprenant les notations du paragraphe 3.1 : Partant d'une configuration initiale de la phase et de l'amplitude $([\phi, a])^{(0)}$, à chaque cycle (c) un ensemble d'itérations est réalisé. Chaque itération (i) correspond à un mouvement vectoriel binaire incomplet, où pour chaque site p le mouvement suivant est réalisé :

$$([\phi_p, a_p]^T)^{(i)} = ([\phi_p, a_p]^T)^{(i-1)} (1 - e_p) + \left(([\phi_p, a_p]^T)^{(i-1)} + ([\delta_\phi^{(c)}, \delta_a^{(c)}]^T)^{(i)} \right) e_p, \quad (6.19)$$

avec $\delta_\phi^{(c)}$ et $\delta_a^{(c)}$ le saut de phase et d'amplitude respectivement au cycle (c) . Le mouvement optimal est obtenu par construction du graphe binaire, où la coupe-minimum donne une configuration à énergie inférieure ou égale à sa valeur courante. Huit itérations sont considérées durant un cycle pour explorer les différents sauts possibles par la donnée du vecteur saut $[\delta_\phi^{(c)}, \delta_a^{(c)}]^T$, à savoir : $[\delta_\phi^{(c)}, 0]^T$, $[0, \delta_a^{(c)}]^T$, $[\delta_\phi^{(c)}, \delta_a^{(c)}]^T$, $[-\delta_\phi^{(c)}, 0]^T$, $[0, -\delta_a^{(c)}]^T$, $[-\delta_\phi^{(c)}, -\delta_a^{(c)}]^T$, $[\delta_\phi^{(c)}, -\delta_a^{(c)}]^T$ et $[-\delta_\phi^{(c)}, \delta_a^{(c)}]^T$.

Lorsqu'un cycle s'achève, le pas de saut vectoriel est changé. Pour une stratégie de type dichotomie, le pas de saut est initialement défini par :

$$[\delta_\phi^{(0)}, \delta_a^{(0)}]^T = \left[\frac{L_\phi}{2}, \frac{L_a}{2} \right], \quad (6.20)$$

avec L_ϕ et L_a les étiquettes maximales que peuvent atteindre les configurations des sites de l'image de phase et d'amplitude respectivement.

Ensuite, la quantité de saut est subdivisée à chaque cycle par deux, jusqu'à une précision donnée.

Une telle stratégie, comme déjà mentionné, est très rapide pour minimiser la fonction d'énergie conjointe, mais reste moins appropriée à des fonctions non-convexes ; plus particulièrement dans notre cas, où la fonction de vraisemblance est multimodale et donc fortement non-convexe.

3.3.2 Minimisation par des extensions vectorielles :

L'algorithme de minimisation par des extensions vectorielles binaires se présente de la façon suivante :

Partant d'une configuration initiale des images de phase et d'amplitude, à chaque cycle c un ensemble d'itérations est réalisé. Chaque itération (i) correspond à un mouvement vectoriel binaire incomplet, où pour chaque site p le mouvement suivant est réalisé :

$$([\phi_p, a_p]^T)^{(i)} = ([\phi_p, a_p]^T)^{(i-1)} (1 - e_p) + ([\alpha_\phi, \alpha_a]^T)^{(i)} e_p, \quad (6.21)$$

avec $([\alpha_\phi, \alpha_a]^T)^{(i)}$ le vecteur d'étiquettes phase-amplitude proposé à tous les sites p à l'itération (i) . A chaque cycle, toutes les combinaisons possibles de couples d'étiquettes $(\alpha_\phi, \alpha_a) \in \mathcal{L}_\phi \times \mathcal{L}_a$ sont proposées.

3.3.3 Choix entre les stratégies de saut ou d'extension :

En nous plaçant dans le cadre des mouvements vectoriels binaires incomplets, nous avons indiqué tout au début de cette section que par un seul mouvement optimal moins de configurations sont explorées comparé à ce que réaliserait un mouvement vectoriel binaire complet.

Cette différence s'accroît d'autant plus que le mouvement est large. En effet, si on se place dans le cas d'un seul site en mouvement, un mouvement vectoriel complet phase-amplitude proposerait la meilleure configuration parmi les $4 = 2^2$ possibles du tableau 6.1(a) (si éventuellement une optimisation exacte est possible), tandis que le mouvement incomplet proposerait la meilleure parmi deux. Une configuration optimale (et non la meilleure) parmi les configurations possibles par le mouvement complet pourrait être obtenu par ce mouvement incomplet en effectuant des itérations successives, comme on l'a indiqué dans le paragraphe 3.3.1.

En élargissant le mouvement à deux sites, la différence s'accroît. Un mouvement vectoriel binaire complet propose la meilleure configuration parmi les $16 = 2^4$ configurations possibles; tandis que le mouvement binaire incomplet proposerait la meilleure parmi les $4 = 2^2$ présentées dans le tableau (b) de la figure 6.1. Une configuration optimale (et non l'optimale) parmi les 16 configurations possibles par le mouvement complet peut être obtenue par le mouvement incomplet en effectuant 4 itérations successives.

En continuant ainsi à élargir le mouvement, on s'aperçoit de l'importance de la différence entre la qualité de l'optimum qui serait donné par un mouvement incomplet et ce par un mouvement complet. D'où la question : quelle stratégie de parcours par des mouvements binaires incomplets pourrait explorer plus de configurations, s'approchant ainsi de la qualité de l'optimum que donnerait les mouvements complets ?

Dans la suite, en nous restreignant uniquement aux deux stratégies par saut et extension, nous montrons que par un mouvement vectoriel binaire incomplet selon la stratégie d'extension on explore mieux l'espace des configurations possibles, comparé à la stratégie de saut. La raison est que l'extension par un mouvement incomplet permet de prendre en compte certaines configurations du mouvement complet, non explorées par la stratégie de saut. En d'autres mots, l'espace de configurations atteignables par un saut vectoriel incomplet est inclus dans celui par extension vectorielle. Ainsi le minimum de l'énergie obtenu par une optimisation avec des extensions sera inférieur ou égal au minimum obtenu par une optimisation avec des sauts. Il est alors plus optimal.

Pour une illustration simple, nous allons nous placer dans le cas de deux sites voisins (p, q) . Pour les deux stratégies de mouvements, nous illustrons toutes les configurations binaires possibles obtenues après un mouvement binaire incomplet. On conservera la notation booléenne pour un mouvement binaire dans le cas multi-étiquettes où le 0 indique l'état de non-changement d'étiquette courante et 1 indique l'état d'un changement de cette étiquette.

La figure 6.4 illustre les configurations binaires possibles par les trois itérations de mouvements de saut pour énumérer le maximum possible de configurations réalisables par un mouvement complet. Les configurations en gras sont celles de changement ou non-changement des étiquettes courantes des sites (phase-amplitude).

$\delta = [\delta_\phi, 0]^T$						$\delta = [0, \delta_a]^T$					
e_p	e_q	e_p^ϕ	e_q^ϕ	e_p^a	e_q^a	e_p	e_q	e_p^ϕ	e_q^ϕ	e_p^a	e_q^a
0	0	0	0	0	0	0	0	0	0	0	0
0	1	0	1	0	0	0	1	0	0	0	1
1	0	1	0	0	0	1	0	0	0	1	0
1	1	1	1	0	0	1	1	0	0	1	1

(a) (b)

$\delta = [\delta_\phi, \delta_a]^T$					
e_p	e_q	e_p^ϕ	e_q^ϕ	e_p^a	e_q^a
0	0	0	0	0	0
0	1	0	1	0	1
1	0	1	0	1	0
1	1	1	1	1	1

(c)

FIGURE 6.4 – Configurations binaires possibles par le saut binaire vectoriel incomplet, étant donné un pas de saut vectoriel (δ_ϕ, δ_a) , à l'itération où (a) $\delta = [\delta_\phi, 0]^T$, (b) $\delta = [0, \delta_a]^T$, et (c) $\delta = [\delta_\phi, \delta_a]^T$.

Comparé aux 16 configurations qu'aurait pu donner un mouvement complet, le saut vectoriel ne parvient pas à explorer les 6 configurations de changement suivantes :

e_p^ϕ	e_q^ϕ	e_p^a	e_q^a
0	1	1	0
0	1	1	1
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	0

TABLE 6.1 – Configurations d'états non possibles par un saut vectoriel binaire.

Similairement, nous pouvons illustrer les différentes possibilités d'états dans le cas d'une extension vectorielle. Pour ce type de mouvement, les configurations sont à la fois en relation avec l'étiquetage courant et au vecteur étiquette proposé. Pour le cas de deux sites, il est toujours possible d'énumérer toutes les configurations. En effet, en indiquant par $(l_p^\phi, l_q^\phi, l_p^a, l_q^a)$ des étiquettes quelconques courantes au sites (p, q) de la phase et de l'amplitude et par (α_ϕ, α_a) le couple d'étiquettes proposées, les configurations sont énumérées à chaque fois selon l'étiquetage courant des deux sites par rapport à ce vecteur d'étiquettes. Pour ne pas tout énumérer, nous nous contentons d'illustrer les configurations qui sont possibles grâce à un mouvement d'extension et qui ne le sont pas par un saut vectoriel (il s'agit des configurations listées dans le tableau 6.1), à moins que des itérations successives par ce dernier mouvement ne soient réalisées. Ce qui serait sous-optimal par rapport à ce que donnerait un mouvement en une seule itération.

Par ailleurs, il est aisé de voir que toutes les configurations binaires obtenues par un saut vectoriel, et qui sont présentées dans la figure 6.4, le sont aussi par une extension.

Dans la figure 6.5, les extensions vectorielles binaires permettant de retrouver les configurations listées dans le tableau 6.1 sont présentées. Ces mouvements sont des cas partic-

uliers d'extensions binaires, qui sont récurrents lors d'un processus de minimisation d'une fonction d'énergie par ce type de mouvement.

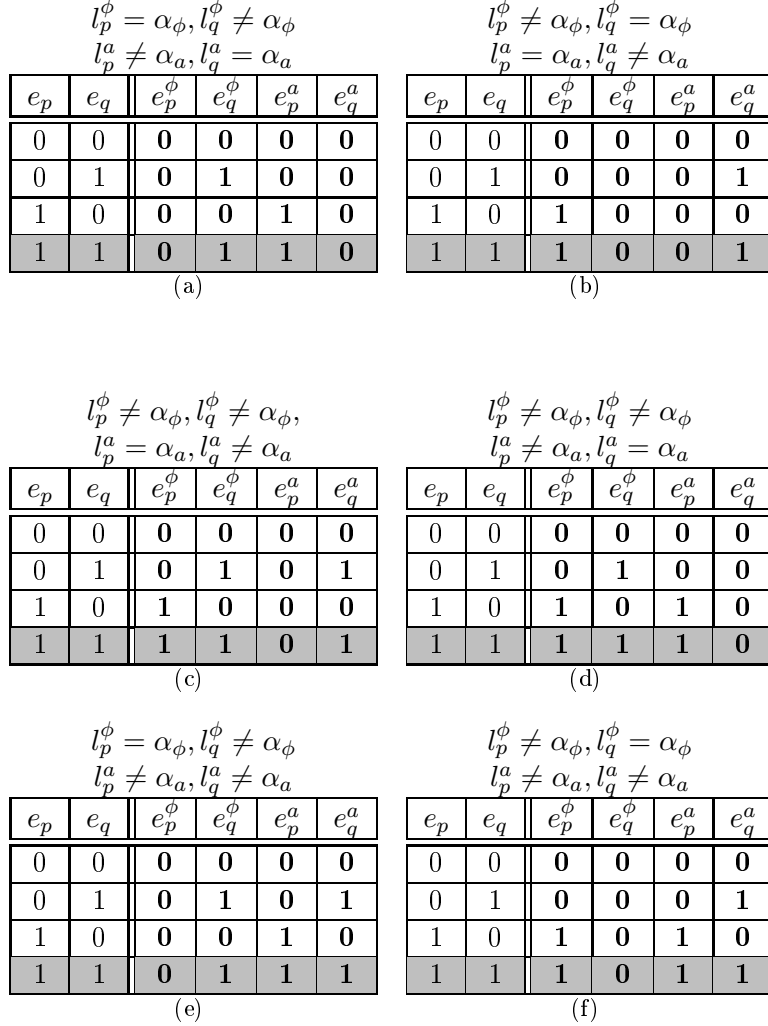


FIGURE 6.5 – Configurations binaires possibles par l'extension binaire vectorielle incomplète en plus de celles obtenues par saut, étant donnée une étiquette vectorielle (α_ϕ, α_a) , à l'itération où (a) $(l_p^\phi \neq \alpha_\phi, l_q^\phi = \alpha_\phi, l_p^a = \alpha_a, l_q^a \neq \alpha_a)$, (b) $(l_p^\phi \neq \alpha_\phi, l_q^\phi = \alpha_\phi, l_p^a = \alpha_a, l_q^a \neq \alpha_a)$, (c) $(l_p^\phi \neq \alpha_\phi, l_q^\phi \neq \alpha_\phi, l_p^a = \alpha_a, l_q^a \neq \alpha_a)$, (d) $(l_p^\phi \neq \alpha_\phi, l_q^\phi \neq \alpha_\phi, l_p^a \neq \alpha_a, l_q^a = \alpha_a)$, (e) $(l_p^\phi = \alpha_\phi, l_q^\phi \neq \alpha_\phi, l_p^a \neq \alpha_a, l_q^a \neq \alpha_a)$, et (f) $(l_p^\phi \neq \alpha_\phi, l_q^\phi = \alpha_\phi, l_p^a \neq \alpha_a, l_q^a \neq \alpha_a)$.

Par conséquent, l'extension binaire incomplète est une meilleure alternative que le saut binaire incomplet pour minimiser la fonction markovienne de notre modèle. En effet, cette stratégie de mouvement permet de mieux explorer l'espace des configurations de changement d'états possibles s'approchant ainsi d'une exploration réalisée par un mouvement complet exact. L'ensemble de configurations possibles par une telle stratégie inclut donc l'ensemble des configurations possible par une stratégie de saut. Cela est d'autant plus claire que le mouvement est plus large, s'étendant à plusieurs sites.

Dans la figure 6.6, le graphe nécessaire pour un mouvement binaire incomplet optimal de type extension est tracé pour tout couple de sites voisins de l'image vectorielle. Les capacités sur les arcs sont aussi précisées.

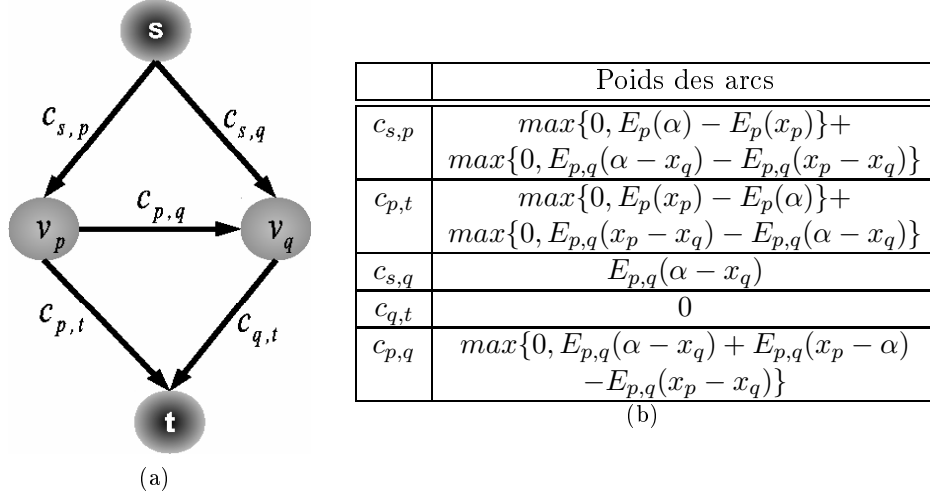


FIGURE 6.6 – Graphe binaire pour un mouvement binaire incomplet de type extension. (a) Structure répétitive du graphe pour tout couple de sites voisins, (b) poids sur les arcs.

```

Initialisation de  $\tilde{x} = (\tilde{\phi}, \tilde{a})$ ;
Pour chaque cycle  $c$  faire
    fin := faux;
    Pour chaque étiquette  $\alpha_\phi \in \mathcal{L}_\phi$  faire
        Pour chaque étiquette  $\alpha_a \in \mathcal{L}_a$  faire
            [Recherche du mouvement de partition optimal]
            
$$x^{(i)} = \arg \min_{x \in \mathcal{X}} E(x),$$

            Si ( $E(x^{(i)}) < E(\tilde{x})$ ) Alors
                 $\tilde{x} := x^{(i)}$ ;
                fin := vrai;
            Fin Si
        Fin Pour
    Fin Pour
    Si (fin == faux) Alors
        retourner  $\tilde{x}$ ;
    Fin Si
Fin Pour

```

FIGURE 6.7 – Algorithme d'optimisation vectorielle par des mouvements d'extension binaires incomplets.

Ainsi nous pouvons conclure sur le choix du mouvement par extension pour optimiser le modèle conjoint phase-amplitude. L'algorithme de minimisation par ce mouvement est présenté dans la figure 6.7. Cet algorithme est donc itératif. A chaque itération, un couple d'étiquettes $(\alpha_\phi, \alpha_a) \in \mathcal{L}_\phi \times \mathcal{L}_a$ est proposé à l'image vectorielle formée de la phase et de

l'amplitude. Un mouvement binaire incomplet de type extension est réalisé. Un cycle fini quand tous les couples d'étiquettes possibles sont proposés. On réitère les cycles jusqu'à la convergence.

4 Évaluations expérimentales sur des données de synthèse

Dans cette partie, des expériences sont menées pour valider la nouvelle approche de reconstruction de modèles numériques de terrain conjointement avec l'information de réflectivité de la scène imagée. L'approche sera validée selon les critères suivants :

- efficacité en temps de calcul et occupation en mémoire,
- robustesse du modèle en présence de zones fortement décorrélées,
- apports par rapport à l'approche proposée au chapitre précédent, c.-à-d. l'avantage d'exploiter l'information d'amplitude à la place d'une étape d'estimation des hyperparamètres locaux.

Le jeu de données simulées correspondra à des zones de structures urbaines. L'image de cohérence est constante par zones. Cela permettra d'évaluer la robustesse du modèle par rapport aux fortes décorrélations qui touchent certaines zones de l'image. Les structures qui figurent dans les images se caractérisent par des fortes discontinuités, rendant l'étape de déroulement de phase ambiguë.

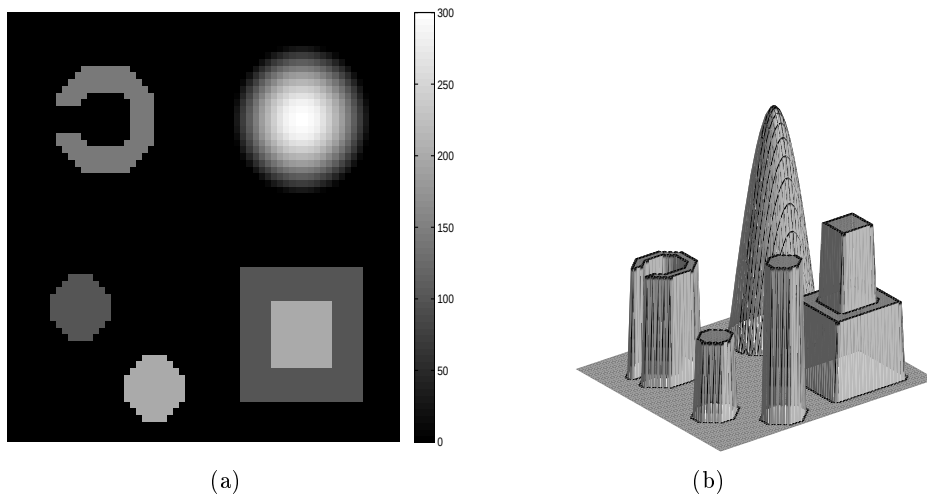
Les résultats obtenus par l'approche conjointe MCPAR+TV seront évalués en premier lieu. Ensuite, ils seront comparés à ceux obtenus par l'approche MCPU+TV.

Nous précisons que pour une meilleure visualisation des images d'amplitude, nous les afficherons dans l'intervalle de niveaux de gris centré sur la moyenne de l'image et de largeur deux fois son écart-type.

4.1 Validation du modèle conjoint

Dans cette expérience, le résultat de la reconstruction de l'altitude d'une scène conjointement avec l'amplitude est présenté et évalué.

Dans la figure 6.8, les données interférométriques simulées sont illustrées. L'image présente différentes structures lisses et constantes par morceaux. Elle est constituée de deux zones à cohérences différentes : une première zone ayant une bonne cohérence globale et une deuxième zone à faible cohérence.



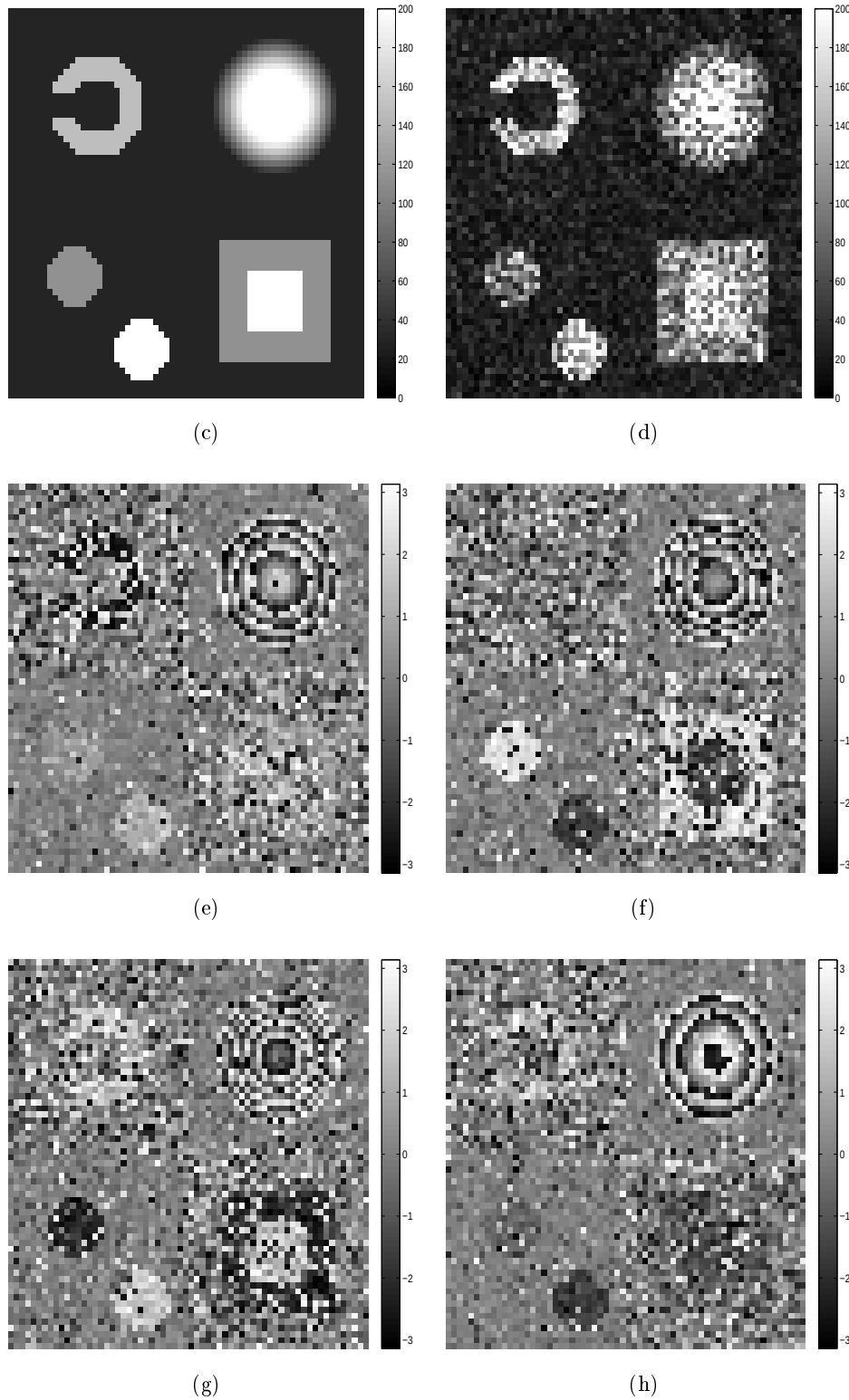


FIGURE 6.8 – Données interférométriques synthétiques d'une zone de structures urbaines. (a) Le MNT de taille 64×64 , (b) sa visualisation 3D, (c) l'image de d'amplitude d'origine, (d) une image de d'amplitude bruitée, et les interférogrammes générés : (e) 1, (f) 2, (g) 3, (h) 4.

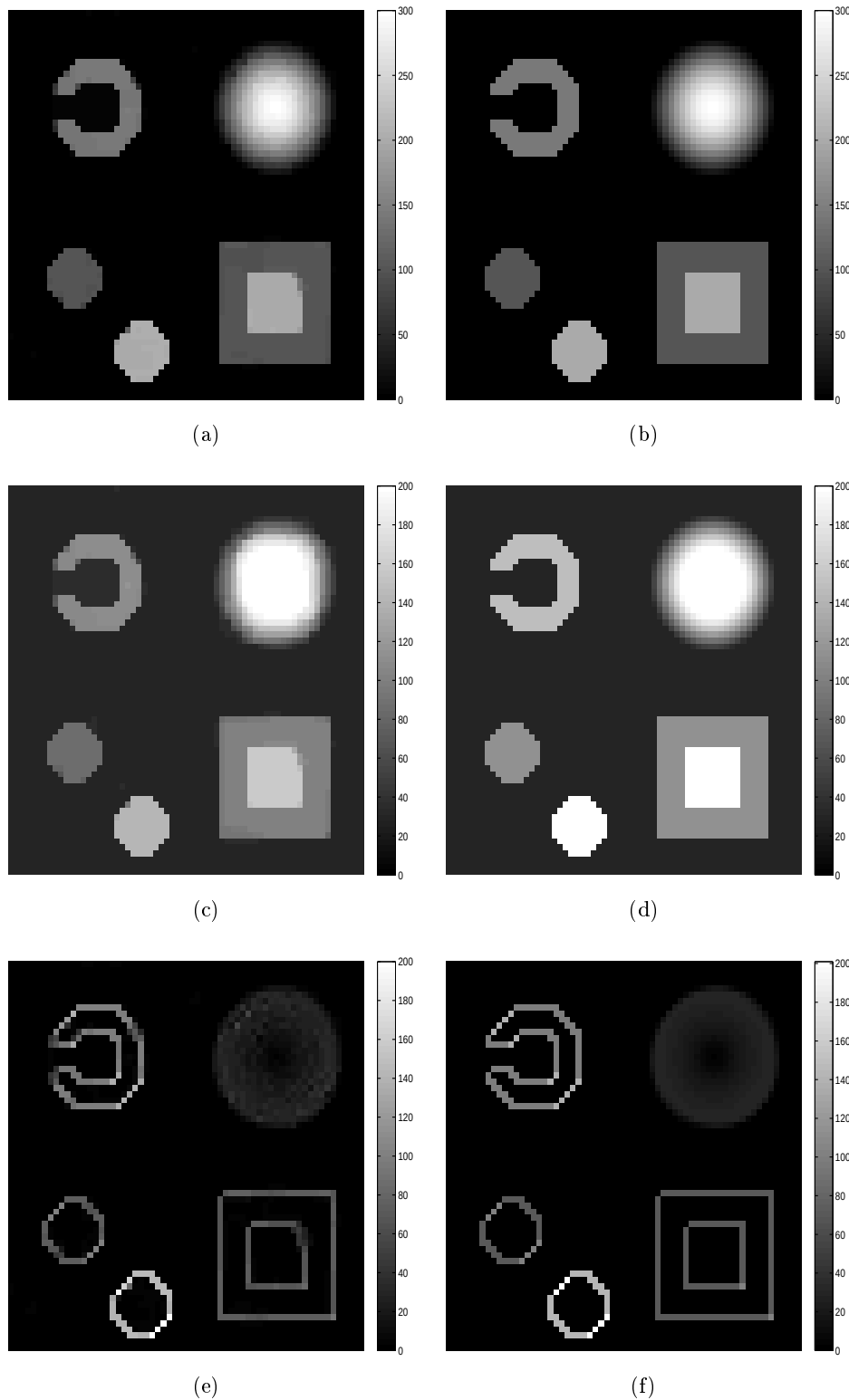


FIGURE 6.9 – Résultat de reconstruction de l'altitude et de l'amplitude. (a) Le MNT reconstruit par notre approche, (b) le MNT d'origine, (c) l'image de réflectivités reconstruite, (d) l'image de réflectivités d'origine, la carte des discontinuités (e) estimée et (f) d'origine.

La reconstruction des images d'altitude et d'amplitude est présentée dans la figure 6.9.

L'algorithme parvient à bien estimer les hauteurs dans la scène tout en préservant le mieux possible les discontinuités des objets qui la constituent. L'image d'amplitude filtrée est aussi de bonne qualité comparée à l'image d'origine (cf. TAB 6.2). La mémoire occupée par le graphe construit à chaque itération est de l'ordre de la taille de l'image de profils, puisque l'optimisation repose sur les mouvements de partitions binaires. Ainsi on conclut que les performances de cette approche sont intéressantes de point de vue complexité et aussi qualité de reconstruction.

TABLE 6.2 – Performances de l'approche de reconstruction par le modèle MCPAR+TV.

Temps de calcul (mn : s)	EQM (altitude)	EQM (amplitude)
2 : 57	4.62	14.13

Dans la zone de très faible cohérence, la qualité de l'estimation est aussi bonne que dans les zones de bonne cohérence. Cela montre l'importance de l'information contextuelle complémentaire apportée par l'image d'amplitude pour mieux discerner les sauts de phase correspondant aux vraies discontinuités de la scène. Ainsi l'ambiguïté du déroulement de la phase est enlevée grâce à une telle information *a priori*.

Dans la même figure, la carte des discontinuités préservées sur l'image d'altitudes reconstruite est illustrée et comparée aux vraies discontinuités du profil d'origine. On remarque clairement que les discontinuités sont fidèlement reconstruites, mis à part quelques problèmes engendrés essentiellement par la présence de profils lisses. En effet, reconstruire une telle information nécessite une fonction *a priori* plus appropriée et une discrétisation très précise de l'ensemble des étiquettes. La variation totale utilisée comme information contextuelle est plutôt adaptée aux objets constants par morceaux. D'où l'apparition de quelques fausses discontinuités sur ces zones lisses. De plus, discrétiser finement les étiquettes pénalise le temps de convergence de l'algorithme. Une solution simple pour remédier à ce problème est le recours à une étape d'optimisation continue, similairement à l'approche FastDC(MCPU+TV). La qualité de l'optimum sera davantage améliorée en précision.

4.2 Comparaison entre les modèles MCPAR+TV et MCPU+TV

L'apport de l'information contextuelle de l'image d'amplitude est mis en valeur par cette expérience où nous comparons les résultats de la reconstruction des reliefs par les approches MCPU+TV et MCPAR+TV.

Avec les mêmes données synthétiques que pour l'expérience précédente, nous illustrons les images de profils reconstruites par MCPU+TV en utilisant dans un premier temps deux canaux, puis en second temps les quatre canaux disponibles ; idem avec l'approche MCPAR+TV. Les résultats figurent dans FIG. 6.10. Nous précisons que l'optimisation du modèle MCPU+TV a été réalisée par l'optimiseur exact d'Ishikawa pour omettre les problèmes qui pourraient être dus à une convergence sous-optimale.

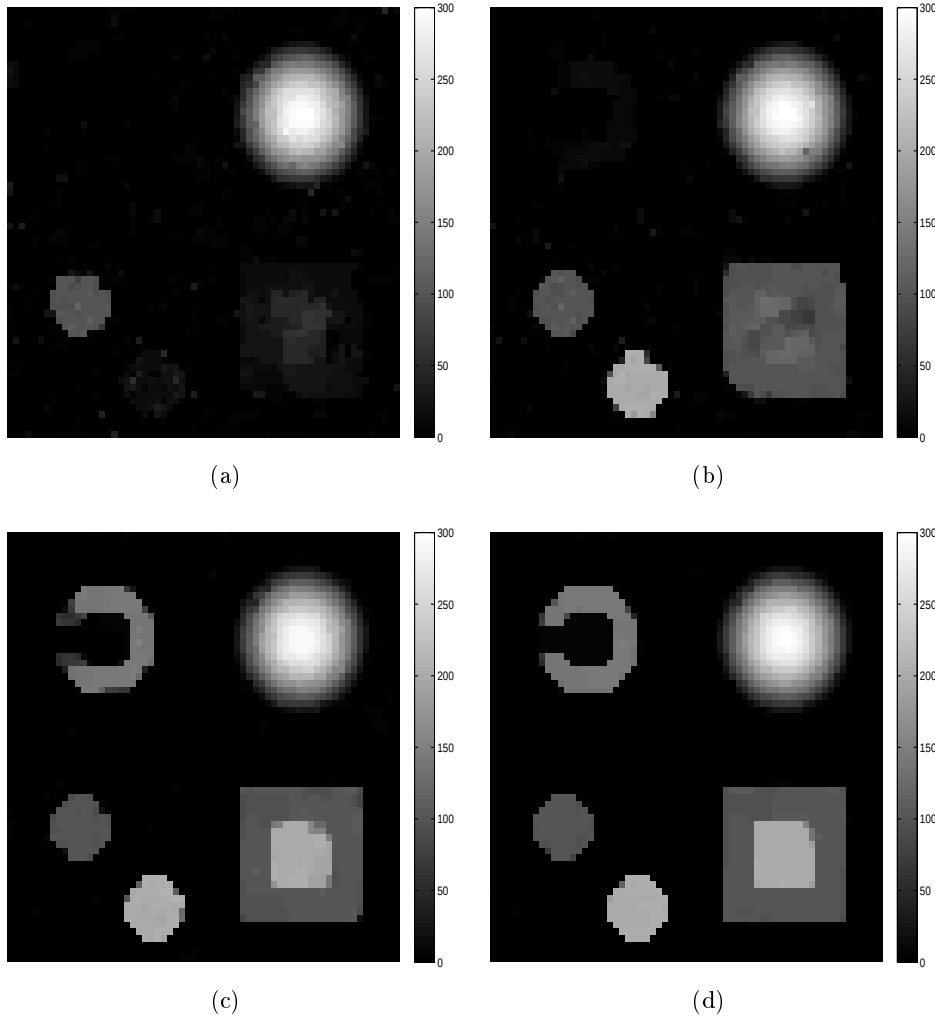


FIGURE 6.10 – Reconstructions de l'image de profil par les deux approches MCPU+TV et MCPAR+TV avec différents nombres de canaux. Reconstruction par MCPU+TV avec (a) deux canaux et (b) quatre canaux. Reconstruction par MCPAR+TV avec (c) deux canaux et (e) quatre canaux.

Avec le modèle MCPU+TV, la reconstruction des structures dans des zones à cohérence faible nécessite un nombre suffisant d'interférogrammes. Autrement, la régularisation a pour effet la destruction de ces structures qui sont considérées comme étant des zones de bruit. Pour cette raison, le recours à deux canaux uniquement n'a pas été suffisant pour reconstruire les objets dans la zone à faible cohérence. Le passage à quatre canaux a permis de retrouver l'information de hauteurs d'une structure de cette zone. Tandis que l'autre structure, qui est plus fine, n'a pas pu être préservée par le modèle.

En visualisant les résultats obtenus par le modèle MCPAR+TV, on s'aperçoit que l'image d'amplitude a pu apporter une information complémentaire aux mesures de phase pour pouvoir localiser ces structures, même en présence de fort bruit interférométrique. Avec deux canaux uniquement, l'image d'altitude reconstruite par MCPAR+TV est satisfaisante. Une meilleure précision a été apportée par les deux autres canaux.

Cela nous permet de conclure que le modèle conjoint est plus approprié aux données interférométriques fortement bruitées, et où les images d'amplitude pourraient intervenir pour préserver les structures de la scène. De plus, l'étape d'estimation des hyperparamètres locaux n'est plus nécessaire dans ce cas, puisque le modèle *a priori* conjoint apporte une solution efficace pour localiser les discontinuités présentes en phase et en amplitude sans procéder à une phase d'estimation.

4.3 Comparaison entre les mouvements de saut et extension vectoriels

Afin de valider le recours aux mouvement vectoriels binaires incomplets par extension, on se propose de comparer dans cette expérience les résultats de reconstruction obtenus par saut et extension.

En ce qui concerne le mouvement par saut, nous illustrons les reconstructions obtenues par la stratégie dichotomique rapide proposée dans (Denis *et al.*, 2009a) et une stratégie de saut normal, c.-à-d. le pas de saut est augmenté d'une unité d'une itération à une autre (c'est le mouvement introduit dans (Bioucas-Dias et Valadao, 2007)).

Les données interférométriques utilisées sont les mêmes que celles des expériences précédentes. Les résultats sont présentés dans la figure 6.11.

Le mouvement par extension offre une meilleure convergence par rapport au mouvement de type saut. Ce dernier, avec une stratégie dichotomique, tombe dans un optimum local de mauvaise qualité pour la simple raison que la stratégie de saut dichotomique, introduite dans (Darbon, 2005), n'est parfaitement appropriée qu'aux fonctions d'énergies convexes. La dichotomie sert alors à accélérer la convergence. Ce qui n'est plus le cas avec la fonction non convexe qu'on minimise dans cette partie.

Avec une stratégie de saut normale, le pas de saut est augmenté à chaque itération, l'optimum local obtenu est de moins bonne qualité que ce que donne l'extension. Cela s'explique par la difficulté que rencontre un mouvement binaire non-complet pour bien minimiser la fonction d'énergie non-convexe. De plus, comme illustré dans les tableaux des configurations binaires par un tel mouvement au paragraphe 3, certains mouvements ne sont pas possibles par un saut vectoriel non-complet. Ces problèmes touchent principalement un mouvement de sites voisins, où l'indépendance de mouvement entre les sites voisins de la phase et de l'amplitude est moins respectée par le saut que par l'extension. L'extension vectorielle permet quant à elle d'explorer plus de configurations et donc d'avoir plus de chances pour converger vers un meilleur optimum local.

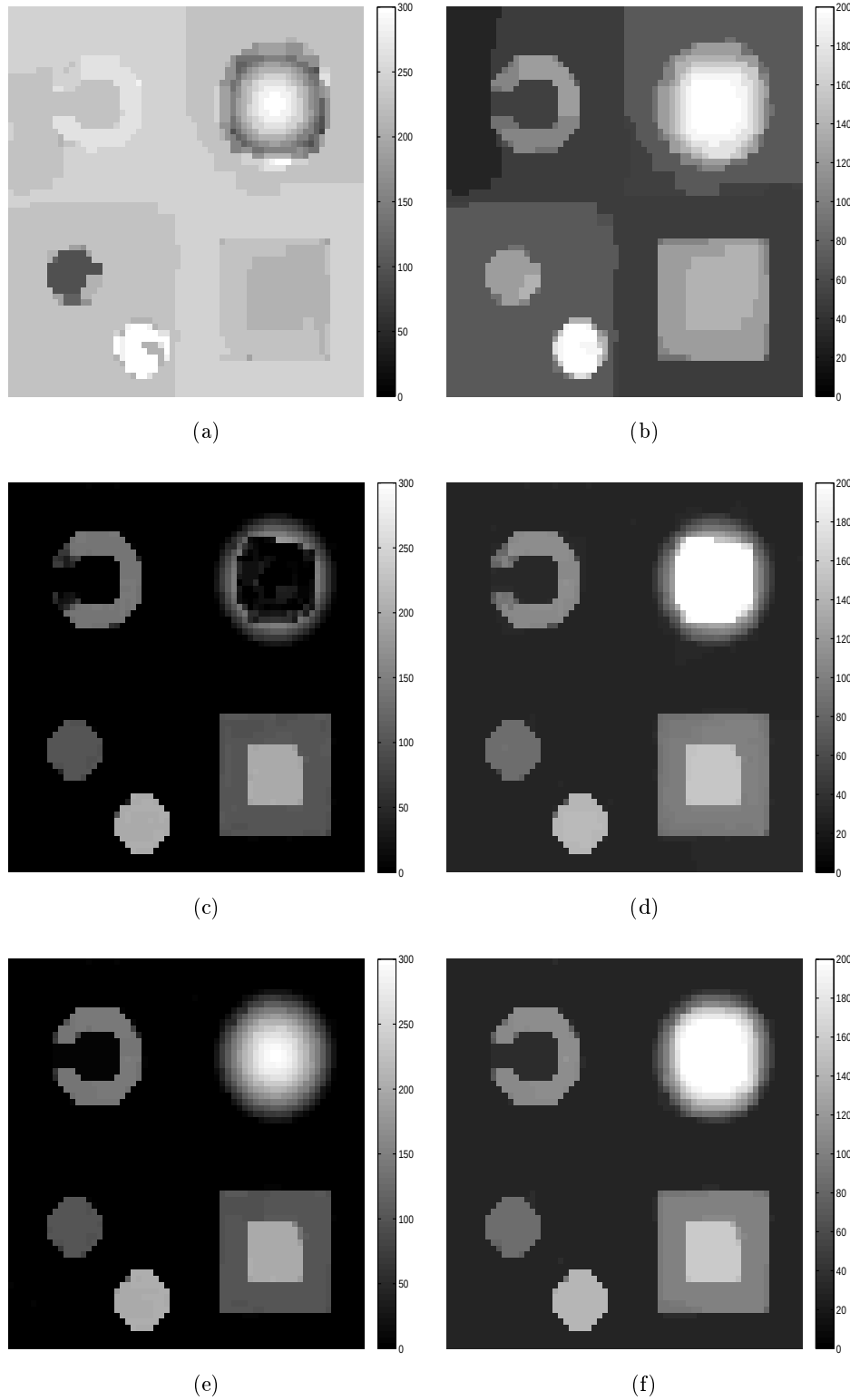


FIGURE 6.11 – Reconstructions de l'image de profil et d'amplitude par l'approche MC-PAR+TV avec des optimiseurs par coupe minimum sur graphe par différents mouvements vectoriels binaires incomplets. Reconstruction par saut dichotomique (a) en altitude et (b) en amplitude. Reconstruction par saut normal (c) en altitude et (d) en amplitude. Reconstruction par extension (e) en altitude et (f) en amplitude.

5 Résultats sur des données réelles

Après avoir réussi à valider l'approche sur des données de synthèse, nous l'expérimentons cette fois-ci sur des données réelles.

Les données InRSO fournies sont des données aériennes, acquises sur la zone de Dresde par un capteur E-SAR en bande L¹. Deux bases sont utilisées de tailles respectives 8.42m et 28.34m. Selon les paramètres d'acquisition de ce capteur, les altitudes d'ambiguïté respectives aux deux bases sont 39m et 11.5m.

Pour chaque base, quatre canaux sont fournis relativement aux quatre polarisations. En tout, 8 interférogrammes sont disponibles sur la zone à reconstruire.

Grâce à cette acquisition aérienne en bande L, les structures urbaines sont clairement présentes dans les images d'amplitude. On s'intéressera par la suite à la reconstruction de l'altitude d'une zone urbaine de Dresde en exploitant l'approche MCPAR+TV. Les résultats seront par ailleurs comparés à ce que donne l'approche de reconstruction de la phase InRSO toute seule par MCPU+TV.

Pour une meilleure qualité de reconstruction, nous effectuons deux phases d'optimisation que ce soit avec le modèle MCPAR+TV ou MCPU+TV. La première phase est une optimisation discrète avec les optimiseurs proposés. La deuxième phase est une optimisation continue permettant d'améliorer en précision la qualité visuelle de la reconstruction. On évitera ainsi les fausses discontinuités (ou bruit) résultant d'une quantification non précise des étiquettes.

Au chapitre précédent, le problème d'estimation des perturbations atmosphériques globales sur les différents canaux des données réelles ERS a été adressé. Un algorithme efficace d'estimation des paramètres de ces perturbations a alors été proposé. Cependant, en ce qui concerne les données aériennes traitées dans ce chapitre, il serait plutôt judicieux de ne considérer que les problèmes de décalage de phases pour de telles perturbations. Les perturbations atmosphériques linéairement dépendantes de l'altitude de la scène n'ont pas de raison d'exister, puisque les trois acquisitions que nous disposons sont d'une part réalisées en une seule journée (01/08/2000, aux heures respectives 11 : 50, 12 : 20, 12 : 49), et donc très proches dans le temps, et sont de l'autre part acquises à une altitude basse de 3390m qui est donc très réduite comparée à celle des capteurs embarqués sur un satellite. Ainsi la relation affine modélisant ces perturbations sera simplifiée et réduite à une constante de phase par canal. On se ramène donc à estimer pour les deux bases utilisées le couple de paramètres (β_1, β_2) de ce nouveau modèle de perturbations globales.

L'algorithme déjà établi au chapitre précédent sera donc appliqué au départ sur les données de phase pour estimer ces paramètres conjointement avec l'altitude. Ces paramètres seront par la suite injectés dans les mesures de phase pour pouvoir appliquer directement l'approche de reconstruction conjointe phase-amplitude sur les données corrigées.

Les figures 6.12 et 6.13 illustrent les données InRSO pour les deux bases respectivement et qui correspondent à la polarisation HH : interférogrammes, images de cohérences et d'amplitudes. Une image optique de la même zone² est aussi montrée dans la figure 6.14. Elle nous sera utile pour vérifier la qualité visuelle de la reconstruction.

1. Nous tenons à remercier le DLR de nous avoir fourni ces données.

2. Image fournie par google map.

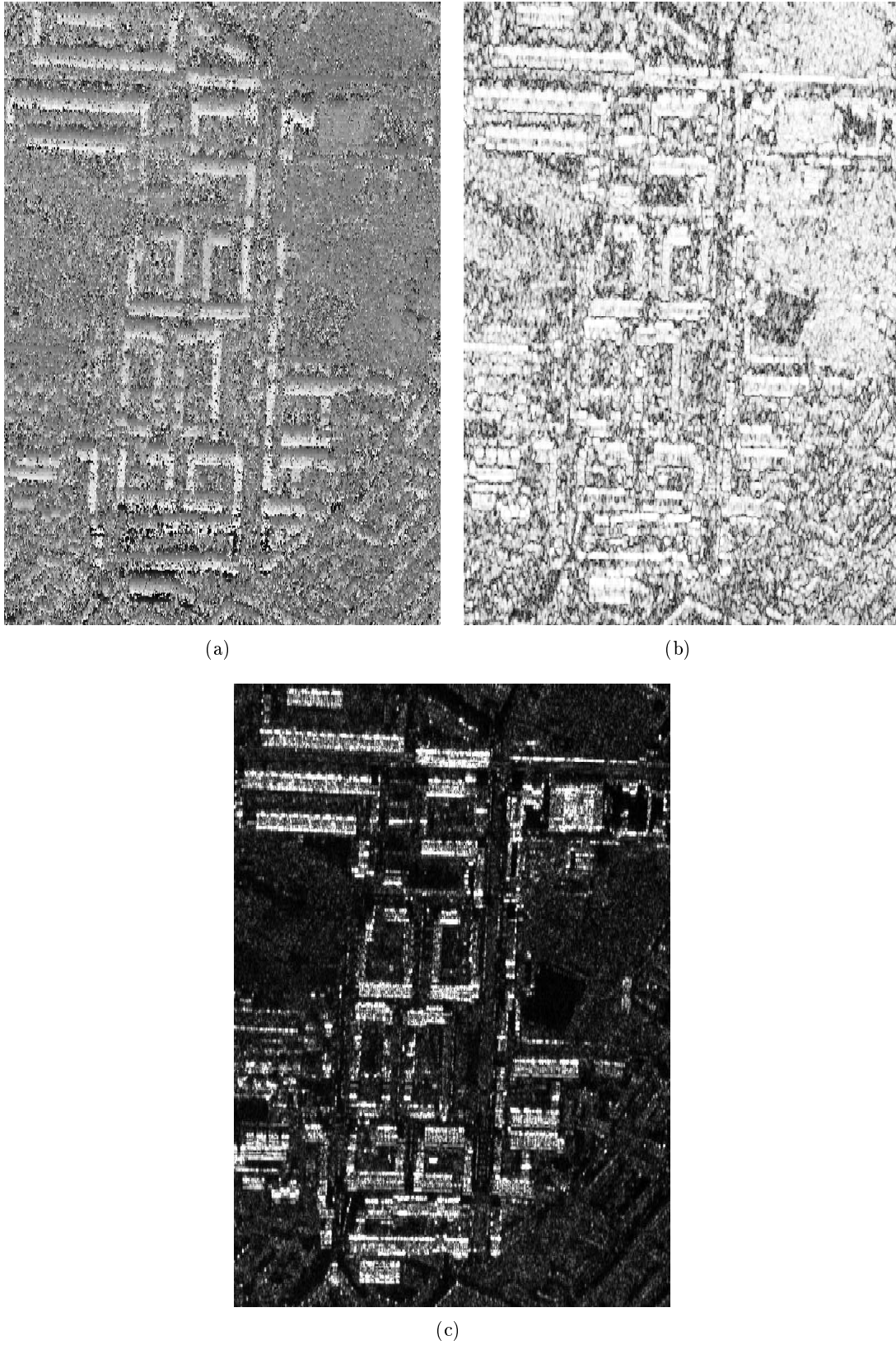


FIGURE 6.12 – Données interférométriques réelles relativement à la première base (8m) pour la polarisation HH ©DLR. La taille de l'image est 700×380 pixels. (a) Interférogramme, (b) image de cohérence (fenêtre de moyennage de taille 3×3) et (c) image d'amplitude.

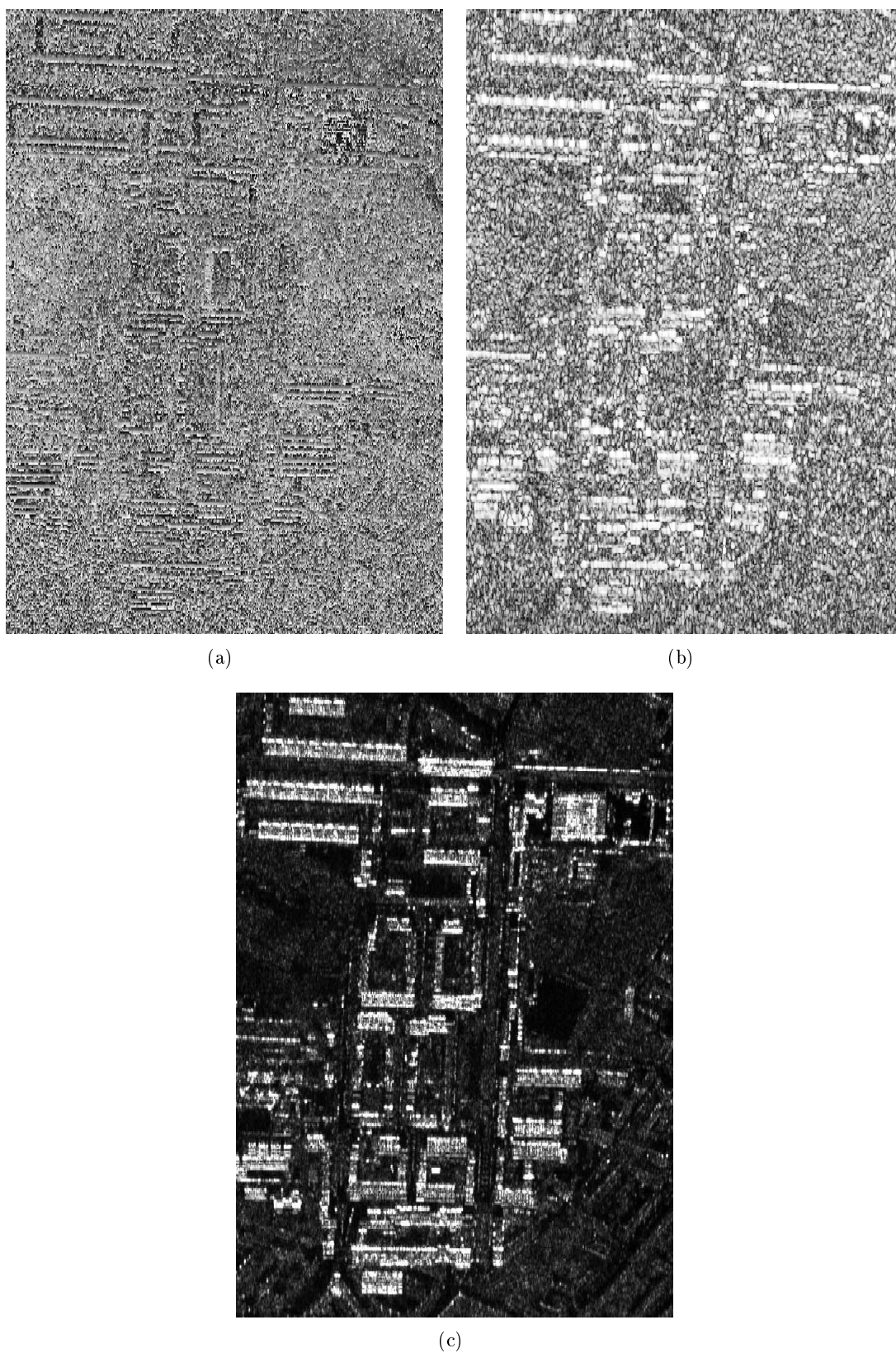


FIGURE 6.13 – Données interférométriques réelles de la même zone relativement à la deuxième base (28m) pour la polarisation HH ©DLR. (a) Interférogramme, (b) image de cohérence et (c) image d'amplitude.



FIGURE 6.14 – Image optique de la même zone ©Google Earth.

Le résultat de l'estimation des paramètres de décalage de phase associés aux deux bases est illustré dans la figure 6.15. On remarque clairement que la reconstruction de la phase à partir des données InRSO d'origine nécessite la correction des décalages de phase. Les paramètres estimés sont $(\beta_1 = 2.66, \beta_2 = 2.01)$. A chaque base, les mêmes paramètres ont été retrouvés pour les quatre polarisations. Ceci confirme encore la précision de l'estimation.

Remarque : Nous précisons que les reconstructions de la phase illustrées ne sont pas re-projetées en géométrie sol, d'où la présence claire des problèmes de repliement qui pourront être éventuellement corrigés par des post-traitements.

La reconstruction finale par l'approche FastDC(MCPU+TV) est présentée dans la figure 6.15(c). Globalement, la reconstruction est satisfaisante, en tenant compte de la géométrie radar, mis à part quelques zones où il a été impossible de bien délimiter les bords des bâtiments. Cela est prévisible pour une telle approche, où préserver les discontinuités des structures fines reste une tâche complexe et délicate. Le résultat de l'optimisation discrète par des mouvements multi-étiquettes n'est pas assez précis pour pouvoir estimer les vraies discontinuités par un simple seuillage et procéder à une optimisation continue afin d'éliminer les fausses discontinuités. On note que le temps mis par l'algorithme pour converger est de 11 : 23 (minutes :secondes)³.

3. Processeur de la machine : Intel Core 2 duo, 2.66GHz.

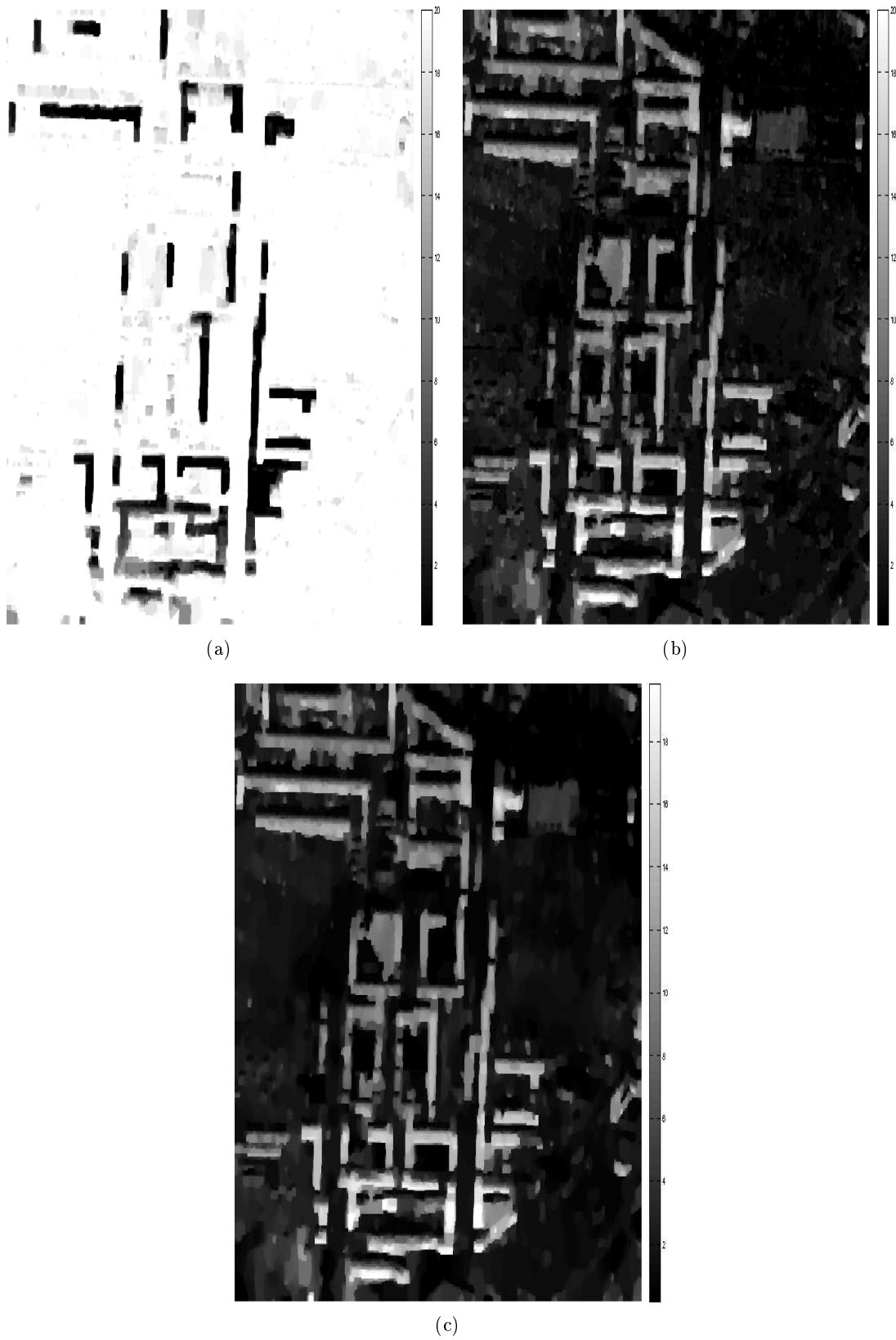


FIGURE 6.15 – Estimation des paramètres de décalage de phase. Reconstruction par MCPU+TV (a) directe à partir des données d'origine, (b) avec estimation conjointe des paramètres de décalage, (c) suivie d'une descente de gradient.

Ainsi cet exemple de zone urbaine montre les limites du modèle MCPU+TV. Un modèle faisant intervenir une information de discontinuités complémentaire est nécessaire pour une meilleure reconstruction.

Le résultat de l'estimation de l'image de phase conjointement avec l'image de réflectivité de la scène par l'approche MCPAR+TV est illustré dans la figure 6.16. Nous précisons par ailleurs qu'un ensemble d'étiquettes de taille réduite est utilisé pour quantifier les étiquettes de l'image d'amplitude. Cela a l'avantage d'accélérer la convergence d'une part, et de mieux localiser les structures des bâtiments d'autre part. Nous accordons ainsi plus d'importance et de précision au modèle d'élévation estimé, qu'à l'image d'amplitude régularisée.

Cette fois-ci, on voit bien l'amélioration de la reconstruction en terme de précision sur les bords des bâtiments. L'information de discontinuités apportée par l'image d'amplitude régularisée a été utile pour mieux préserver les contours des structures urbaines.

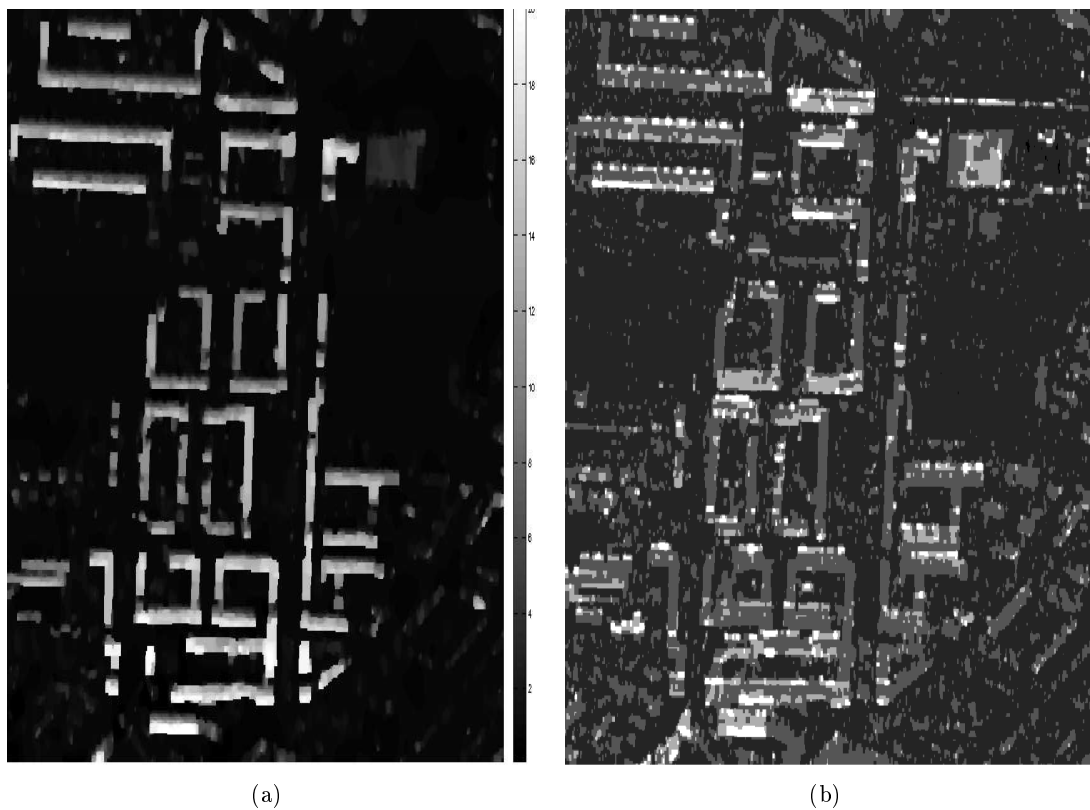


FIGURE 6.16 – Estimation conjointe de la phase et de l'amplitude par MCPAR+TV. (a) Reconstruction du profil, (b) reconstruction de l'amplitude (visualisation dans la palette centrée sur la moyenne de l'image et de largeur trois fois son écart-type).

Dans la figure 6.17, une comparaison visuelle entre les reconstructions obtenues par MCPU+TV et MCPAR+TV sur quelques bâtiments de la zone reconstruite est proposée. Les avantages qui ressortent du recours à un modèle conjoint sont les suivants :

- reconstruction plus fidèle des bords des structures urbaines,
- élimination de certaines fausses structures qui sont dues au bruit interférométrique et qui ne sont pas observables sur les images d'amplitude,
- occupation en mémoire très faible grâce à l'utilisation des mouvements binaires.

Néanmoins, en terme de temps de calcul, l'optimisation vectorielle du critère MCPAR+TV

est plus coûteuse. Cela est dû à un nombre d'itérations plus important pour converger. De plus, deux images sont régularisées simultanément par cette approche. Le temps qui a été mis par l'algorithme pour converger vers la solution de la figure 6.16(a) est de 36 : 22 (minutes :secondes).

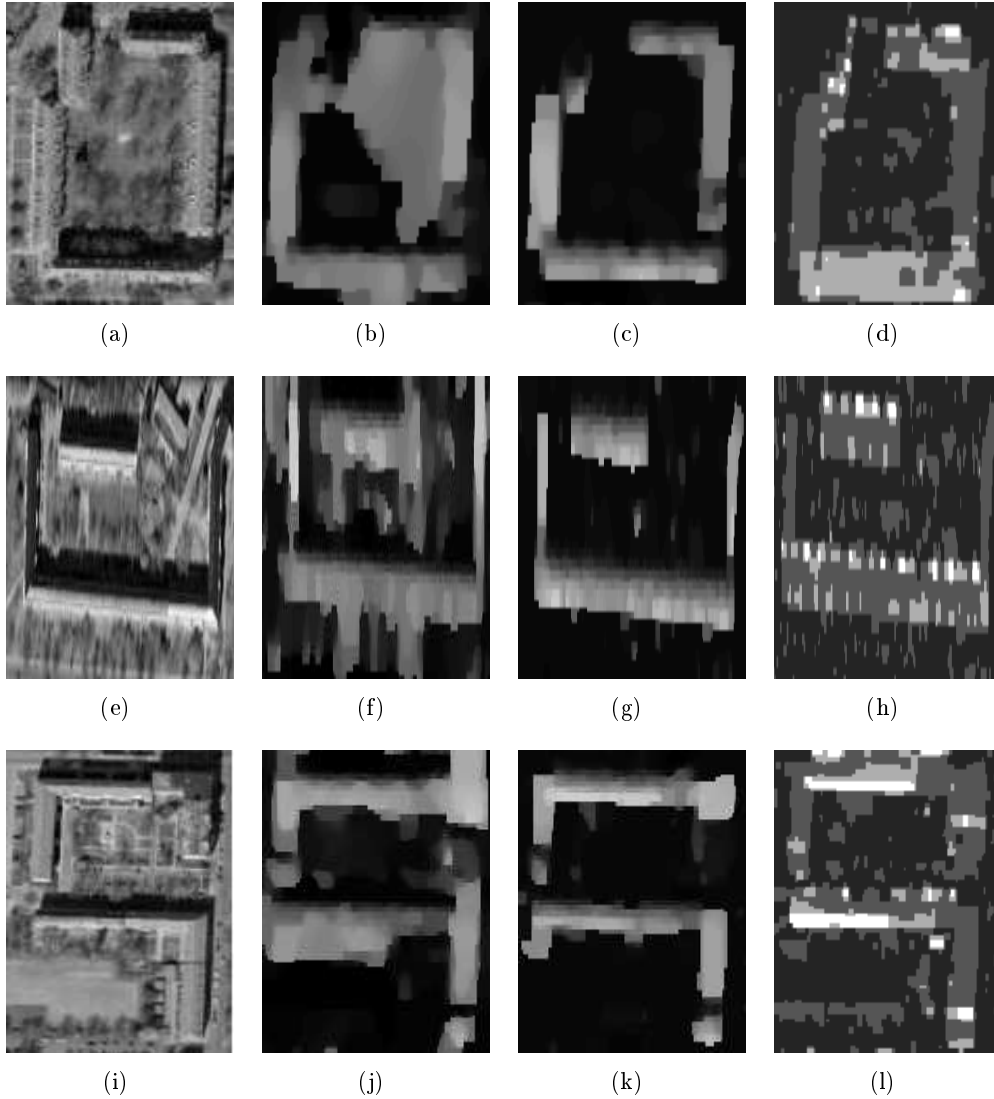


FIGURE 6.17 – Comparaison entre les reconstructions obtenues par MCPU+TV et MCPAR+TV. (a) Bâtiment 1, sa reconstruction par (b) MCPU+TV et (c) MCPAR+TV, (d) l'image d'amplitude régularisée. (e) Bâtiment 2, sa reconstruction par (f) MCPU+TV et (g) MCPAR+TV, (h) l'image d'amplitude régularisée. (i) Bâtiment 3, sa reconstruction par (j) MCPU+TV et (k) MCPAR+TV, (l) l'image d'amplitude régularisée.

6 Conclusion

Le modèle MCPAR+TV introduit et validé dans ce chapitre est un modèle bien adapté à la reconstruction de l'information d'élévation sur des zones urbaines. Les limites du modèle MCPU+TV proposé auparavant, essentiellement liées à l'étape d'estimation des discontinuités de la scène, ont pu alors être surmontées grâce à l'information *a priori*

conjointe apportée par les images de phase et d'amplitude régularisées.

Pour optimiser efficacement la fonction d'énergie établie, une nouvelle famille d'optimiseurs vectoriels par coupe-minimum sur graphe a été explorée. L'optimisation exacte de cette énergie étant impossible par la technique de coupe-minimum sur graphe, un algorithme de minimisation approximative reposant sur un mouvement binaire vectoriel incomplet de type extension a été alors mis en place. Les performances de l'optimisation ainsi que la qualité des optima obtenus dans les évaluations expérimentales confirment l'utilité d'une telle approche conjointe pour reconstruire l'information en hauteur des zones urbaines.

Conclusion et perspectives

Contributions

Dans cette thèse, nous avons abordé dans un premier temps la problématique de la minimisation d'énergies markoviennes de premier ordre avec une attache aux données quelconque et une fonction de régularisation convexe. Cette problématique est d'un intérêt important dans plusieurs applications en imagerie et vision par ordinateur, résolues par optimisation du critère MAP. En raison de l'explosion en complexité des optimiseurs de l'état de l'art face à des fonctions d'énergies en grande dimension, fortement non-convexes et nécessitant une importante précision sur l'estimation, le besoin d'une approche efficace et robuste se ressent. Dans ce cadre, nous avons proposé un algorithme qui permet de converger rapidement vers des optima de bonne qualité tout en ayant le contrôle sur sa complexité. L'approche repose d'une part sur une technique récente d'optimisation discrète en vision par des mouvements de partitions larges et optimaux, où un mouvement optimal est obtenu par coupe-minimum sur un graphe spécifique ; et d'autre part sur l'algorithme d'optimisation continue : descente du gradient.

Par la première technique, mouvement de partition, nous avons proposé de nouveaux mouvements, appelés mouvements de partitions larges et multi-étiquettes (MPLM), qui ont la particularité de bien explorer l'espace des configurations possibles et converger ainsi vers des optima locaux de bonne qualité. Une construction de graphe spécifique à un mouvement large et multi-étiquette a été alors proposée, où le calcul d'une coupe-minimum donne le mouvement optimal. Un algorithme d'optimisation avec différentes stratégies de mouvements a ensuite été mis en place. Pour montrer son apport par rapport aux algorithmes de l'état de l'art, des résultats dans le cadre de la restauration des images corrompues par un bruit impulsif ont été présentés.

Par la deuxième technique, optimisation continue, nous avons voulu accélérer la convergence de l'algorithme d'optimisation discrète précédent tout en améliorant en précision la solution vers laquelle il converge. Ainsi, une stratégie mixte entre MPLM et descente de gradient a été explorée donnant une solution de meilleure qualité. Pour illustrer l'intérêt d'une telle méthode, nous l'avons évaluée dans le cadre de la restauration des images radar à synthèse d'ouverture, où les données sont souvent de grandes dimensions, de forte dynamique radiométrique et nécessitent une bonne précision pour préserver l'information physique ; et où de plus la fonction markovienne nécessaire pour la restauration est non-convexe. Les résultats sont prometteurs ouvrant le champ d'applications de ces algorithmes d'optimisation vers d'autres problématiques en imagerie radar, telle que la reconstruction de modèles numériques de terrain (MNT).

C'est cette problématique qui a été abordée dans un second temps dans cette thèse, comme application illustrant d'une part l'intérêt des approches markoviennes pour reconstruire des MNT de bonne qualité et de l'autre part l'intérêt de l'étude faite dans la partie précédente pour optimiser des énergies markoviennes complexes nécessaires pour une telle

application.

Nous avons tout d'abord proposé un premier modèle markovien, appelé MCPU+TV, exploitant les mesures de phase interférométrique pour estimer le relief de la scène. Le modèle repose à la fois sur des observations multicanaux pour réduire ou enlever l'ambiguïté du déroulement de la phase et sur le modèle de régularisation par variation totale pondéré par un processus ligne pour mieux préserver les discontinuités des objets de la scène. La fonction d'énergie obtenue étant fortement non-convexe, l'optimisation par les MPLM nous a été d'un grand intérêt pour pouvoir la minimiser efficacement. De plus, avec des données de scènes naturelles, une bonne précision sur la reconstruction est nécessaire. D'où l'utilité de la méthode d'optimisation mixte pour fournir un MNT bien résolu et précis. L'approche validée sur des données synthétiques et testée sur des données radar ERS acquises sur une zone de Serre-Ponçon a montré des performances satisfaisantes, malgré la difficulté de la scène et la présence d'artéfacts de différentes natures sur les données. Le MNT reconstruit comparé à celui fourni par la mission SRTM nous a permis d'évaluer la méthode proposée et connaître ses limites.

Une des limites soulevées est la sensibilité de la méthode à l'étape d'estimation des hyperparamètres locaux pour préserver les discontinuités dans l'image. Pour l'améliorer davantage, un deuxième modèle markovien, appelé MCPAR+TV, a alors été proposé dans le dernier chapitre exploitant en plus de l'interférogramme, les images d'amplitude. Ces images peuvent apporter une information complémentaire aux mesures de phase portant sur les discontinuités de la scène. Ainsi, une méthode de reconstruction conjointe phase-amplitude a été mise en place, nécessitant un algorithme d'optimisation approprié aux modèles markoviens conjoints. Une étude d'une famille d'optimiseurs de tels modèles a été ainsi réalisée, aboutissant à une méthode d'optimisation rapide et robuste de l'énergie établie. Cette approche de reconstruction conjointe de MNT et de l'image de réflectivité est avantageuse en présence de données à haute résolution acquises sur des scènes urbaines. Ainsi, elle a été testée et validée sur des données interférométriques acquises par le capteur aérien E-SAR sur une zone de Dresde. Les résultats ont effectivement montré l'avantage d'explorer conjointement l'information de la phase et de l'amplitude, plutôt que l'information de phase seule, comme dans le premier modèle MCPU+TV.

Ces nouvelles approches, introduites tout au long de cette thèse, malgré les encourageantes performances et les résultats atteints, présentent encore quelques limites pour lesquelles, faute de temps, nous n'avons pas été en mesure de leur proposer des solutions efficaces. Toutefois, nous présentons par la suite quelques pistes et réflexions envisageables qu'il serait judicieux d'explorer en profondeur dans de futurs travaux.

Limites et solutions possibles

Optimisation markovienne

L'idée de réaliser des mouvements de partitions larges et multi-étiquettes nous a permis de nous affranchir des limites des optimiseurs par des mouvements larges et binaires en terme de qualité de l'optimum atteint ainsi que des limites des optimiseurs exacts en terme de complexité de calcul. Cependant, un problème qui n'a pas été bien étudié dans ce manuscrit est le choix précis de l'étendue et du parcours de l'ensemble des étiquettes pour s'assurer d'une convergence vers un optimum de bonne qualité. La solution proposée qui consiste à sous-quantifier le vrai ensemble des étiquettes pour pouvoir proposer un nombre m large sans pour autant pénaliser le temps de calcul n'est pas une solution générale. La sous-quantification entraîne bien une perte d'information sur des données réelles qu'il est difficile de récupérer sans disposer d'une information *a priori* suffisante pour l'approximer.

Une solution envisageable à ce problème serait d'exploiter l'approche multi-échelle en imagerie. Cette technique a été explorée récemment dans (Kohli *et al.*, 2010) pour s'affranchir des limites en mémoire des optimiseurs par coupe-minimum sur graphe, spécifiquement en présence de données de grandes dimensions et de modèles d'énergie avec des cliques d'ordre supérieur. Il est alors possible en fusionnant les deux idées de mouvements multi-étiquettes et résolution multi-échelle de minimiser des fonctions non-convexes de très grandes dimensions avec beaucoup moins de contraintes sur le choix du paramètre m . Ce dernier pourrait être choisi d'une manière adaptative, d'une échelle à une autre, de façon à contrôler la taille de la mémoire utilisée tout en s'assurant de la qualité de la convergence. Par exemple, pour les faibles résolutions, où la taille des données serait réduite, la taille de l'ensemble de toutes les étiquettes possibles pourrait être considérée comme valeur de m . Ainsi, on s'assure d'une bonne qualité des optima atteints à ces résolutions. Ensuite, pour les fortes résolutions, comme nous partirons de bonnes solutions, la contrainte sur m est de moins en moins conséquente. Un choix de faibles valeurs de m pourra être alors suffisant pour atteindre un bon optimum. Nous pensons aussi à une stratégie adaptative pour le choix de m en fonction de la résolution courante.

De plus, avec une telle approche, une meilleure exploration de l'ensemble des configurations dans $\mathcal{R}^n$ est réalisée grâce au principe du multi-échelle. Cela rend non nécessaire l'optimisation continue qui pourrait donc servir juste comme une phase finale pour s'assurer de la précision de la solution obtenue.

Enfin, avec des données de très grandes dimensions, spécialement en imagerie radar, la fusion multi-étiquette/multi-échelle apportera une solution élégante pour résoudre les problèmes d'imagerie sans décomposer les données. Tel est le cas pour l'estimation de MNT présentée dans le chapitre 5, où nous avons été amené à ne considérer qu'une zone de la ville de Serre-Ponçon pour pouvoir reconstruire le profil sans soucis de mémoire physique ou de qualité du minimum atteint.

Estimation de MNT précis

Les deux modèles markoviens proposés nous ont permis d’obtenir des modèles numériques de terrain assez précis et adaptés à la nature de la scène reconstruite : urbaine ou non-urbaine. L’optimisation du premier modèle a été réalisée grâce aux algorithmes d’optimisation par MPLM et FastDC et pourrait être encore améliorée en nous reposant sur la technique multi-étiquette/multi-échelle que nous nous proposons d’étudier dans le futur. L’optimisation du deuxième a été réalisée grâce à des algorithmes d’optimisation conjointe qui ont montré de bonnes performances.

Si des problèmes de reconstruction persistent encore, c’est qu’il est plus important d’approfondir l’étude des modèles proposés. Ainsi, nous suggérons dans ce cadre quelques pistes à creuser afin de mieux modéliser le problème de la reconstruction de MNT en imagerie radar à synthèse d’ouverture par une approche markovienne.

Perturbations atmosphériques globales et locales : Nous avons constaté lors de l’étude de la reconstruction du relief à partir de données InRSO réelles que les perturbations globales, apportées par exemple par la contribution atmosphérique dans la mesure de phase, rendent le modèle markovien multicanal proposé non-exploitable en raison de l’incohérence de l’information apportée par les différents canaux. Une estimation puis une correction de ces perturbations sont alors nécessaires. L’algorithme d’estimation conjointe des décalages de phase et de l’altitude qui a été alors proposé est une solution, mais qui est non robuste et sensible à l’initialisation en raison de son aspect itératif. Par ailleurs, il ne permet pas d’estimer des perturbations locales, qu’il est possible de rencontrer en reconstruisant de grandes zones.

Les travaux de thèse (Chaabane, 2004) sont alors une meilleure alternative pour corriger les données dans un premier temps, et pouvoir faire appel dans un second temps au modèle markovien MCPU+TV.

Une autre approche en deux phases, que nous envisageons d’explorer dans le futur, consiste à partir d’un MNT SRTM ou tout MNT approximatif de la zone en question (tel que par exemple la reconstruction du canal le moins ambigu par l’approche rapide PUMA) et d’estimer une carte de perturbations locales en nous reposant sur un modèle bien approprié. La carte générée, injectée ensuite directement dans le modèle MCPU+TV, permet donc de corriger les artéfacts dûs à l’atmosphère, aussi bien globaux que locaux, et de reconstruire un MNT plus précis.

Cliques d’ordre supérieur : En ce qui concerne maintenant le modèle *a priori*, le recours à la variation totale nous a été bénéfique pour exploiter la technique de coupe-minimum sur graphe, où la propriété de sous-modularité est nécessaire, et aussi pour éviter une étape lourde d’estimation de paramètres dans le cas des modèles de régularisation locaux. Cependant, une telle information *a priori* se trouve appropriée à des images constantes par morceaux, et donc moins à des images de scènes naturelles. Dans le cas réel, une information *a priori* plus riche apporte plus de robustesse à la modélisation. Une façon de procéder est de passer à des fonctions d’énergies de régularisation d’ordre supérieur. Nous sommes alors confrontés à des problèmes d’optimisation de ce type de fonctions. Les récents travaux dans (Ishikawa, 2009a) peuvent être d’une grande utilité pour optimiser ce type de fonctions, même en présence d’*a priori* non sous-modulaire. Ainsi, avec moins de contraintes sur les propriétés mathématiques de l’*a priori*, il est possible de mettre en place un modèle plus performant, dont l’optimisation par les approches récentes de coupe-minimum sur graphe pourrait être réalisable efficacement aboutissant ainsi à des solutions

meilleures et plus convaincantes pour un expert ou thématicien en imagerie radar.

Information conjointe radar-optique : En milieu urbain, le deuxième modèle markovien MCPAR+TV qui a été proposé a permis de fusionner l'information d'amplitude avec les mesures de phase du produit interférométrique pour mieux reconstruire les objets de la scène. Cependant, l'information apportée par l'amplitude est limitée par la qualité de cette image et par la réponse des objets qui constituent la scène. En d'autres mots, avec des images d'amplitude très bruitées, où les objets sont difficilement localisables et où des fausses réponses peuvent être apportées par la régularisation, le modèle conjoint ne constitue plus une meilleure alternative par rapport au modèle MCPU+TV.

Une solution possible, dans ce cas de figure, est le recours à une autre information de source différente et moins bruitée, telle qu'une image optique de la même scène. Cependant, on se retrouve face à un problème de géométrie d'acquisition qui diffère entre le capteur optique et le capteur radar, nécessitant alors une étape de recalage fin pour pouvoir exploiter ces données multicapteurs conjointement. Par ailleurs, la projection de l'image optique dans la géométrie du radar, ou vice versa, serait avantageuse pour tenir compte des phénomènes spécifiques à l'acquisition radar lors de la reconstruction, tel que le layover. D'autre part, l'image optique étant généralement beaucoup moins bruitée qu'une image radar permettra de mieux localiser les discontinuités si un *a priori* approprié est proposé dans le modèle conjoint. Une telle approche a été récemment explorée dans (Denis *et al.*, 2009a) donnant des résultats prometteurs pour une étude plus approfondie. La fusion radar-optique, qui constitue un axe de recherche très sollicité par la communauté de la télédétection, surtout avec le lancement des capteurs fournissant des images radar et optique (comme le Pleiades Cosmo-Skymed), apportera ainsi une information riche pour mieux modéliser les problèmes d'imagerie satellitaire, spécifiquement avec des approches markoviennes.

Annexe A

Construction de graphes binaires des algorithmes MPLB

La minimisation des fonctions sous-modulaires de $\mathbb{F}^2$ (c.-à-d. les fonctions markoviennes multi-étiquettes de premier ordre) par les algorithmes reposant sur la technique de mouvements de partitions larges et binaires (MPLB) est réalisée par des constructions itératives de graphes binaires, où une coupe-minimum est calculée à chaque itération donnant la configuration optimale par le mouvement binaire considéré. En effet, à chaque itération un étiquetage binaire est réalisé. On considère pour chaque site p le couple d'étiquettes possibles $(e_p^0, e_p^1) \in \mathcal{L}^2$ que peut prendre x_p . La forme de l'énergie de $\mathbb{F}^2$

$$E(x) = \sum_{p=1}^n E_p(y_p|x_p) + \sum_{(p,q)} E_{p,q}(x_p, x_q), \quad (\text{A.1})$$

peut se mettre aussi sous la forme suivante :

$$\begin{aligned} E(x) &= \tilde{E}(\theta) \\ &= \sum_{p=1}^n \left(E_p(e_p^1)\theta_p + E_p(e_p^0)(1 - \theta_p) \right) \\ &\quad + \sum_{(p,q)} \left(E_{p,q}(e_p^0, e_p^1) + E_{p,q}(e_p^1, e_p^0) - E_{p,q}(e_p^0, e_p^0) - E_{p,q}(e_p^1, e_p^1) \right) (1 - \theta_p)\theta_q \quad (\text{A.2}) \\ &\quad + \sum_{(p,q)} \left(E_{p,q}(e_p^1, e_p^0) - E_{p,q}(e_p^0, e_p^0) \right) \theta_p + \left(E_{p,q}(e_p^1, e_p^1) - E_{p,q}(e_p^1, e_p^0) \right) \theta_q, \quad (\text{A.3}) \end{aligned}$$

avec $\theta = \{\theta_1, \theta_2, \dots, \theta_n\}$ ensemble de variables binaires correspondant aux sites $\mathcal{P}$, $E_p(e_p^i) = E_p(y_p|x_p = e_p^i)$, $i \in \{0, 1\}$ et $E_p(e_p^i, e_q^j) = E_p(x_p = e_p^i, x_q = e_q^j)$, $(i, j) \in \{0, 1\}^2$.

Ainsi, minimiser $\tilde{E}$ (et donc E) revient à trouver la configuration binaire optimale $\tilde{\theta}$. Cela se ramène au problème de minimisation résolu par coupe-minimum sur les fonctions pseudo-boléennes (cf. paragraphe 2.2 du chapitre 1).

Les graphes binaires associés aux trois mouvements binaires décrits dans le chapitre 1 seront illustrés. Nous considérons pour les différentes constructions de graphes un système de voisinage 4-connexité.

α -extension

En se plaçant à une itération donnée de l'algorithme, considérons une étiquette $\alpha \in \mathcal{L}$. Le mouvement d'extension binaire propose à tous les sites (et donc à tous les nœuds du

graphe $\mathcal{G}$ associé à l'image $\mathcal{P}$) de passer à cette nouvelle étiquette ou de garder l'étiquette courante. L'étiquetage binaire est alors :

$$\forall p \in \mathcal{P}, x_p = \begin{cases} e_p^0 = \bar{\alpha}_p, & \text{ou} \\ e_p^1 = \alpha, \end{cases} \quad (\text{A.4})$$

où $\bar{\alpha}_p$ dénote l'étiquette courante au site p .

La construction du graphe est réalisée à partir des deux structures répétitives suivantes :

- chaque nœud v_p du graphe est relié à la source par un arc de données orienté $\overrightarrow{(s, v_p)}$ de poids $E_p(\bar{\alpha}_p)$ et au puits par un arc de données orienté $\overrightarrow{(v_p, t)}$ de poids $E_p(\alpha)$.

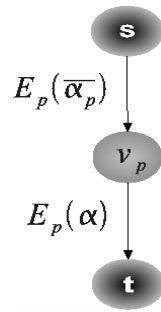


FIGURE A.1 – Structure répétitive du graphe pour tous les sites.

- pour chaque couple de nœuds voisins (v_p, v_q) , nous créons l'arc d'interaction orienté $\overrightarrow{(v_p, v_q)}$ de poids $(E_{p,q}(\bar{\alpha}_p, \alpha) + E_{p,q}(\alpha, \bar{\alpha}_q) - E_{p,q}(\bar{\alpha}_p, \bar{\alpha}_q) - E_{p,q}(\alpha, \alpha))$. Ensuite, un arc de bord orienté $\overrightarrow{(s, v_p)}$ de poids $(E_{p,q}(\alpha, \bar{\alpha}_q) - E_{p,q}(\bar{\alpha}_p, \bar{\alpha}_q))$ si ce poids est positif, sinon $\overrightarrow{(v_p, t)}$ de poids $(-E_{p,q}(\alpha, \bar{\alpha}_q) + E_{p,q}(\bar{\alpha}_p, \bar{\alpha}_q))$, est construit en plus de l'arc de bord orienté $\overrightarrow{(v_q, t)}$ de poids $(E_{p,q}(\alpha, \alpha) - E_{p,q}(\alpha, \bar{\alpha}_q))$ si ce poids est positif, sinon $\overrightarrow{(s, v_q)}$ de poids $(-E_{p,q}(\alpha, \alpha) + E_{p,q}(\alpha, \bar{\alpha}_q))$. Nous remarquons que la condition de sous-modularité est vérifiée par une fonction $E_{p,q}$ métrique, qui intervient dans cette construction pour assurer un arc à poids positif entre les deux nœuds voisins (v_p, v_q) .

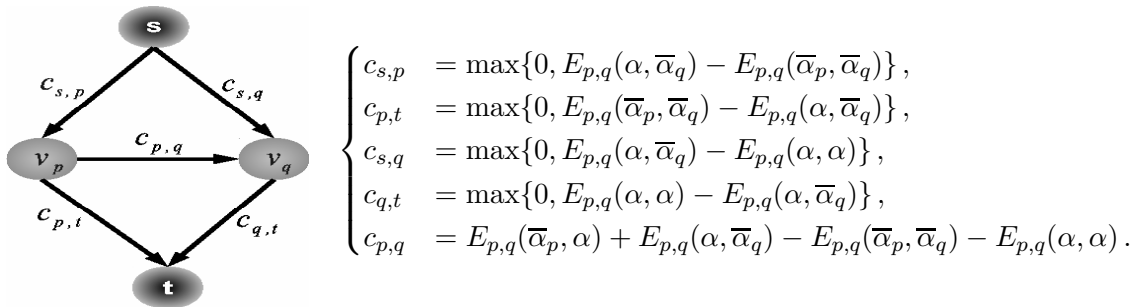


FIGURE A.2 – Structure répétitive du graphe pour chaque couple de sites voisins tel que $p < q$.

La figure A.3 illustre le graphe total correspondant à une image de taille 3×3 .

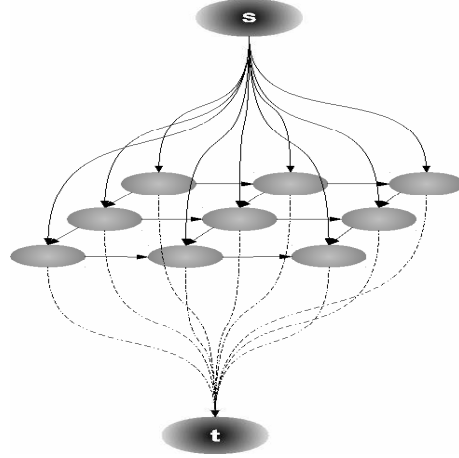


FIGURE A.3 – Exemple de construction de graphe total.

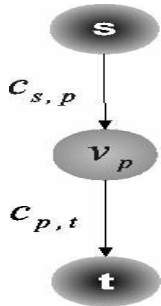
$\alpha\beta$ -échange

En se plaçant à une itération donnée de l'algorithme de minimisation, considérons un couple d'étiquettes $(\alpha, \beta) \in \mathcal{L}^2$. Le mouvement d'échange consiste à proposer aux sites actifs, dont les étiquettes courantes sont dans $\{\alpha, \beta\}$, d'échanger leurs étiquettes. Ces sites sont indiqués par $\mathcal{P}_{\alpha, \beta}$. L'étiquetage binaire est alors :

$$\forall p \in \mathcal{P}_{\alpha, \beta}, x_p = \begin{cases} e_p^0 = \alpha, & \text{ou} \\ e_p^1 = \beta. \end{cases} \quad (\text{A.5})$$

Ainsi, le graphe binaire n'est construit que sur un nombre restreint de nœuds, reposant sur les deux structures répétitives suivantes :

- chaque nœud v_p du graphe, dont l'étiquette courante est dans $\{\alpha, \beta\}$, est relié à la source par un arc de données orienté $\overrightarrow{(s, v_p)}$ de poids $E_p(\alpha)$ et au puits par un arc de données orienté $\overrightarrow{(v_p, t)}$ de poids $E_p(\beta)$. Cependant, d'autres termes de l'énergie globale ne sont fonctions que d'une seule variable et peuvent ainsi être considérés dans cette structure. Il s'agit des termes d'interaction entre une variable de $\mathcal{P}_{\alpha, \beta}$ et une autre n'appartenant pas à cet ensemble. On ajoute alors respectivement aux poids sur les arcs de données $\overrightarrow{(s, v_p)}$ et $\overrightarrow{(v_p, t)}$ les quantités $\sum_{r \notin \mathcal{P}_{\alpha, \beta}}^{(p,r)} E_{p,r}(\alpha, x_r)$ et $\sum_{r \notin \mathcal{P}_{\alpha, \beta}}^{(p,r)} E_{p,r}(\beta, x_r)$.



$$\text{si } p \in \mathcal{P}_{\alpha, \beta} \begin{cases} c_{s,p} &= E(\alpha) + \sum_{r \notin \mathcal{P}_{\alpha, \beta}}^{(p,r)} E_{p,r}(\alpha, x_r) \\ c_{p,t} &= E(\beta) + \sum_{r \notin \mathcal{P}_{\alpha, \beta}}^{(p,r)} E_{p,r}(\beta, x_r). \end{cases}$$

FIGURE A.4 – Structure répétitive du graphe sur tous les sites actifs.

- pour chaque couple de nœuds voisins (v_p, v_q) , tel que les deux nœuds sont actifs, nous créons l'arc d'interaction orienté $\overrightarrow{(v_p, v_q)}$ de poids $(E_{p,q}(\alpha, \beta) + E_{p,q}(\beta, \alpha) - E_{p,q}(\alpha, \alpha) - E_{p,q}(\beta, \beta))$. Ensuite, un arc de bord orienté $\overrightarrow{(s, v_p)}$ de poids $(E_{p,q}(\beta, \alpha) - E_{p,q}(\alpha, \alpha))$ si ce poids est positif, sinon $\overrightarrow{(v_p, t)}$ de poids $(-E_{p,q}(\beta, \alpha) + E_{p,q}(\alpha, \alpha))$, est construit en plus de l'arc de bord orienté $\overrightarrow{(v_q, t)}$ de poids $(E_{p,q}(\beta, \beta) - E_{p,q}(\beta, \alpha))$ si ce poids est positif, sinon $\overrightarrow{(s, v_q)}$ de poids $(-E_{p,q}(\beta, \beta) + E_{p,q}(\beta, \alpha))$. On remarque que la condition de sous-modularité est vérifiée par une fonction $E_{p,q}$ semi-métrique, qui intervient dans cette construction pour assurer un arc à poids positif entre les deux nœuds voisins (v_p, v_q) .

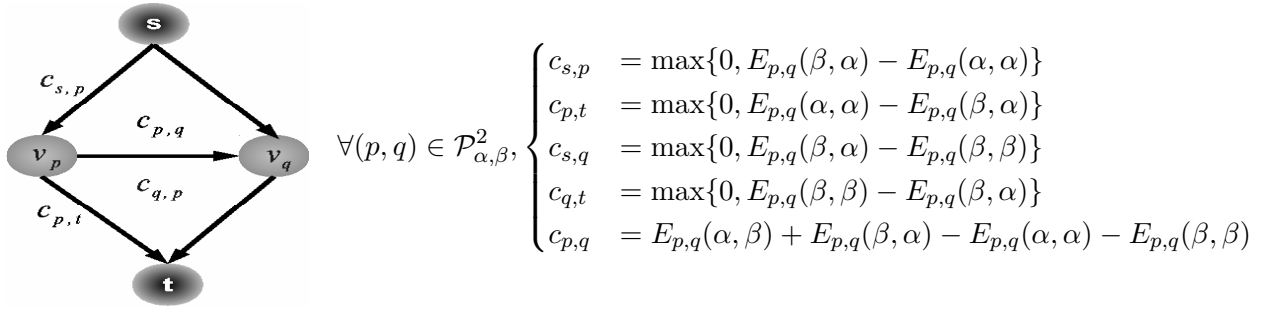


FIGURE A.5 – Structure répétitive du graphe pour chaque couple de sites voisins actifs tel que $p < q$.

La figure A.6 illustre le graphe total correspondant à une image de taille 3×3 .

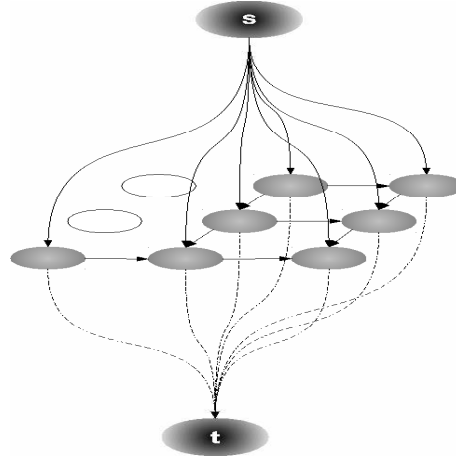


FIGURE A.6 – Construction d'un exemple de graphe total. Les nœuds en blanc correspondent aux sites non actifs.

δ -saut

En se plaçant à une itération donnée de l'algorithme de minimisation, considérons une quantité positive δ à ajouter aux étiquettes courantes de tous les pixels. Le mouvement de saut consiste à proposer aux sites soit de conserver leurs étiquettes courantes soit de les augmenter/diminuer de cette quantité. Si la valeur d'une nouvelle étiquette obtenue dépasse l'ensemble des étiquettes $\mathcal{L}$, le site en question n'est plus actif. Les deux étapes, augmenter

et diminuer, se font successivement. En une étape, prenons par exemple l'augmentation, un étiquetage binaire est réalisé tel que :

$$\forall p \in \mathcal{P}_{+\delta}, x_p = \begin{cases} e_p^0 = \alpha_p, \text{ ou} \\ e_p^1 = \alpha_p + \delta. \end{cases} \quad (\text{A.6})$$

où α_p désigne l'étiquette courante associée au site p et $\mathcal{P}_{+\delta}$ désigne l'ensemble des sites actifs.

Ainsi, le graphe binaire n'est construit que sur un nombre restreint de nœuds, reposant sur les deux structures répétitives suivantes :

- chaque nœud v_p du graphe, dont l'étiquette courante augmentée de δ est dans $\mathcal{L}$, est relié à la source par un arc de données orienté $\overrightarrow{(s, v_p)}$ de poids $E_p(\alpha_p)$ et au puits par un arc de données orienté $\overrightarrow{(v_p, t)}$ de poids $E_p(\alpha_p + \delta)$. Cependant, d'autres termes de l'énergie globale ne sont fonctions que d'une seule variable et peuvent ainsi être considérés dans cette structure. Il s'agit des termes d'interaction entre une variable de $\mathcal{P}_{+\delta}$ et une autre n'appartenant pas à cet ensemble. On ajoute alors respectivement aux poids sur les arcs de données $\overrightarrow{(s, v_p)}$ et $\overrightarrow{(v_p, t)}$ les quantités $\sum_{\substack{(p,r) \\ r \notin \mathcal{P}_{+\delta}}} E_{p,r}(\alpha_p, x_r)$ et $\sum_{\substack{(p,r) \\ r \notin \mathcal{P}_{+\delta}}} E_{p,r}(\alpha_p + \delta, x_r)$.

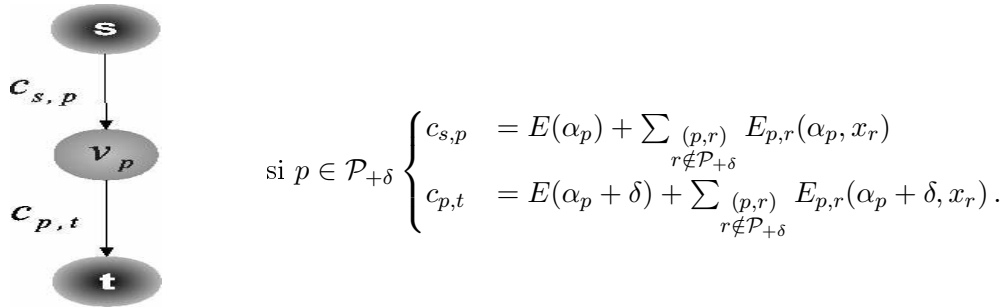


FIGURE A.7 – Structure répétitive du graphe sur tous les sites actifs.

- pour chaque couple de nœuds voisins (v_p, v_q) , tel que les deux nœuds sont actifs, nous créons l'arc d'interaction orienté $\overrightarrow{(v_p, v_q)}$ de poids $(E_{p,q}(\alpha_p, \alpha_q + \delta) + E_{p,q}(\alpha_p + \delta, \alpha_q) - E_{p,q}(\alpha_p, \alpha_q) - E_{p,q}(\alpha_p + \delta, \alpha_q + \delta))$. Ensuite, un arc de bord orienté $\overrightarrow{(s, v_p)}$ de poids $(E_{p,q}(\alpha_p + \delta, \alpha_q) - E_{p,q}(\alpha_p, \alpha_q))$ si ce poids est positif, sinon $\overrightarrow{(v_p, t)}$ de poids $(-E_{p,q}(\alpha_p + \delta, \alpha_q) + E_{p,q}(\alpha_p, \alpha_q))$, est construit en plus de l'arc de bord orienté $\overrightarrow{(v_q, t)}$ de poids $(E_{p,q}(\alpha_p + \delta, \alpha_q + \delta) - E_{p,q}(\alpha_p + \delta, \alpha_q))$ si ce poids est positif, sinon $\overrightarrow{(s, v_q)}$ de poids $(-E_{p,q}(\alpha_p + \delta, \alpha_q + \delta) + E_{p,q}(\alpha_p + \delta, \alpha_q))$. On remarque que la condition de sous-modularité est vérifiée par une fonction $E_{p,q}$ convexe, qui intervient dans cette construction pour assurer un arc à poids positif entre les deux nœuds voisins (v_p, v_q) .

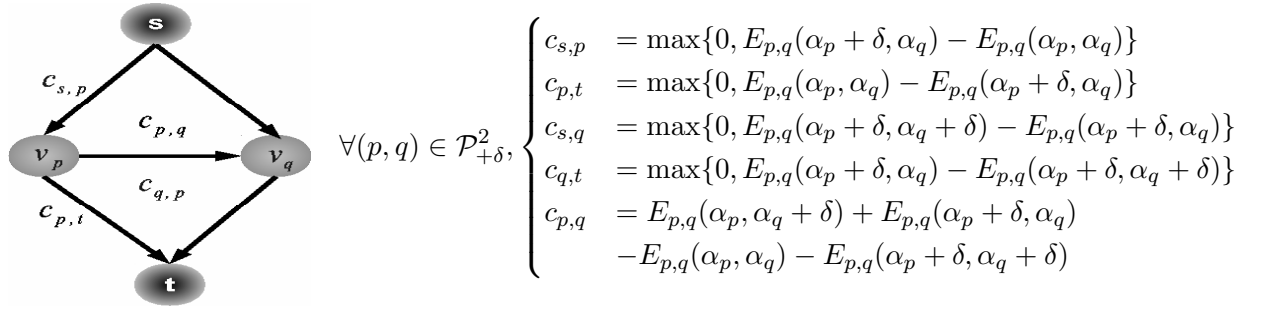


FIGURE A.8 – Structure répétitive du graphe pour chaque couple de sites voisins actifs tel que $p < q$.

La figure A.9 illustre le graphe total correspondant à une image de taille 3×3 .

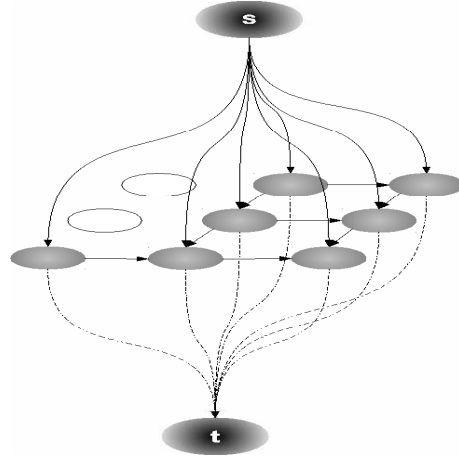


FIGURE A.9 – Exemple de construction du graphe total. Les nœuds en blanc correspondent aux sites non actifs.

Annexe B

Construction du graphe multi-étiquette d'Ishikawa

Dans cette annexe, la construction du graphe d'Ishikawa pour la minimisation exacte des énergies markoviennes sous-modulaires de $\mathbb{F}^2$ est décrite. Les preuves de la construction donnant des minima exacts de ces fonctions sont détaillées dans (Ishikawa, 2003). On considère que l'ensemble discret des étiquettes est un ensemble fini et ordonné d'entiers $\mathcal{L} = \{1, 2, \dots, k\}$.

Soit un graphe orienté $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, où $\mathcal{V}$ est l'ensemble des nœuds du graphe et $\mathcal{E}$ est l'ensemble des arcs, tel que :

- $\mathcal{V} = \left\{ v_{p,i} ; p \in \{1, 2, \dots, n\} ; i \in \{1, 2, \dots, k\} \right\} \cup \{s, t\}$, avec $v_{p,i}$ est un nœud associé au site d'indice p et au étiquette i .
- $\mathcal{E} = \mathcal{E}_D \cup \mathcal{E}_C \cup \mathcal{E}_P$, avec $\mathcal{E}_D$ est l'ensemble des arcs appelés arcs de données, $\mathcal{E}_C$ est l'ensemble des arcs appelés arcs de contraintes et $\mathcal{E}_P$ est l'ensemble des arcs appelés arcs de pénalité.

Le graphe peut être visualisé en 3 dimensions. L'axe vertical z correspond à l'axe des étiquettes et le plan x, y correspond au plan de l'image. Le graphe est donc en couches. La figure B.1 illustre un exemple de construction d'un tel graphe.

Les arcs de données sont les arcs qui correspondent aux termes d'attache aux données dans l'expression de la fonction d'énergie. Ils sont pondérés de la façon suivante :

$$\begin{cases} c(s, v_{p,1}) & = +\infty, \\ c(v_{p,i}, v_{p,i+1}) & = E_p(x_p = i), \forall i = 1..k-1, \\ c(v_{p,k}, t) & = E_p(x_p = k). \end{cases} \quad (\text{B.1})$$

Les arcs allant de la source aux nœuds $\{v_{p,1}\}_{p=1..n}$ sont de poids infini pour éviter de les inclure dans une coupe sur le graphe, puisque aucune étiquette n'est affectée à la source.

Pour contraindre l'algorithme de recherche de coupe de capacité minimale à ne couper qu'une seule fois un arc de données par site, des arcs dits de contraintes sont construits de la façon suivante :

$$c(v_{p,i+1}, v_{p,i}) = +\infty, \forall i = 1..k-1. \quad (\text{B.2})$$

Dans la figure B.2(a), une structure répétitive du graphe 3D dans B.1 qui correspond à une colonne verticale de ce graphe associée à un site p quelconque de $\mathcal{P}$ est présentée.

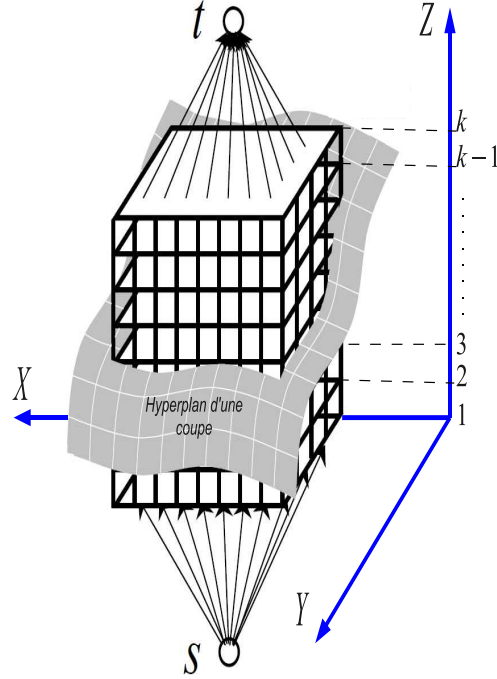


FIGURE B.1 – Graphe en couches d'Ishikawa.

Pour construire les arcs d'interaction correspondants aux termes d'interaction de l'énergie, nous considérons la famille des arcs suivants :

$$\mathcal{E}_P = \left\{ (v_{p,i}, v_{q,j}) ; (p, q) \in \{1, 2, \dots, n\}^2 ; (i, j) \in \{1, 2, \dots, k\}^2 \right\} \quad (\text{B.3})$$

Ainsi, pour deux sites voisins (p, q) , la valeur de la coupe correspondant aux termes d'interaction est égale à la somme des poids de tous les arcs que rencontre la coupe pour passer du nœud $v_{p,i}$ au nœud $v_{q,j}$. En définissant les poids sur ces arcs d'interaction de la façon suivante :

$$c(v_{p,i}, v_{q,j}) = \frac{E_{p,q}(i, j+1) - 2E_{p,q}(i, j) + E_{p,q}(i, j-1)}{2}, \quad (\text{B.4})$$

la somme totale des poids des arcs coupés est alors égale au terme d'interaction entre les deux sites en question p et q à une constante près.

Dans la figure B.2(b), la structure répétitive du graphe 3D dans B.1 des aux arcs d'interaction entre deux sites voisins (p, q) quelconques de $\mathcal{P}$ est présentée.

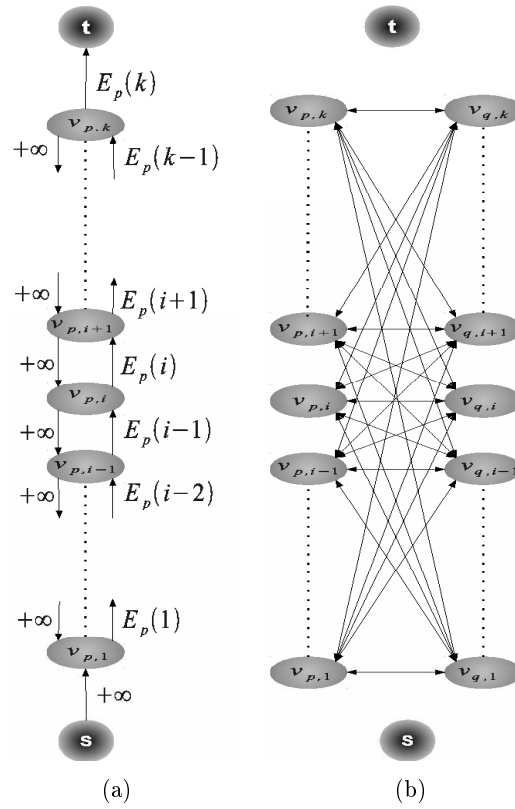


FIGURE B.2 – Illustration des structures répétitives du graphe d'Ishikawa. (a) Arcs de données et de contraintes associés à un site p quelconque, (b) Arcs d'interaction entre deux sites voisins (p, q) quelconques. On précise que pour un *a priori* de type variation totale, uniquement les arcs d'interaction qui sont horizontaux sont présents dans le graphe. Ce qui réduit énormément la taille en mémoire nécessaire pour stocker un tel graphe.

Annexe C

Flot-maximum dans un réseau

Cette annexe est consacré à une brève présentation de quelques algorithmes et références pour le calcul de flot-maximum dans un réseau $(\mathcal{G}, f)$, où $\mathcal{G}$ et f sont respectivement le graphe et la fonction flot définis dans le chapitre 1. Les mêmes notations de ce chapitre seront considérées.

Deux grandes familles d'algorithmes existent dans la littérature : les algorithmes par *saturation de chemins augmentants* et les algorithmes par *réétiquetage*.

Flot-maximum par saturation de chemins augmentants

On définit la capacité résiduelle sur un arc (u, v) , notée par $r(u, v)$, comme étant :

$$r(u, v) = w(u, v) - f(u, v), \quad (\text{C.1})$$

c.-à-d. la quantité de flot que l'on peut faire passer en plus du flot $f(u, v)$ courant dans l'arc (u, v) , de capacité $w(u, v)$.

On appelle le graphe résiduel, que l'on notera par $\mathcal{G}_r$, le graphe formé par les mêmes nœuds que le graphe $\mathcal{G}$ mais en prenant comme arcs ceux dont la capacité résiduelle est strictement positive.

$$\mathcal{G}_r = (\mathcal{V}, \mathcal{E}_r), \quad (\text{C.2})$$

avec

$$\mathcal{E}_r = \{(u, v) \in \mathcal{V} \times \mathcal{V} ; r(u, v) > 0\}. \quad (\text{C.3})$$

Un chemin augmentant est alors défini comme étant un chemin de $\mathcal{G}_r$, de la source s au puits t , constitué d'arcs de capacités positives.

Les algorithmes par saturation de chemins fonctionnent par recherche itérative de chemins augmentant de la source au puits. On ajoute alors autant de flot que possible sur ces chemins jusqu'à saturer leurs arcs à capacités résiduelles minimales. L'algorithme se déroule ainsi :

Initialisation : $f(u, v) = 0 \forall (u, v) \in \mathcal{E}$;

Tant que (Il existe un chemin augmentant H de s à t dans $\mathcal{G}_r$) **faire**

$\delta = \min \{r(u, v) | (u, v) \in H\}$;

Pour $(u, v) \in H$ **faire**

$f(u, v) = f(u, v) + \delta$;

Fin Pour

Fait

Algorithme 1: Algorithme générique de recherche de flot-maximum par saturation de chemins augmentant.

Au bout d'un certain nombre d'itérations (augmentations de flot), il n'est plus possible de trouver un chemin augmentant. On peut montrer alors que le flot obtenu est un flot de valeur maximale.

Cet algorithme a été initialement proposé dans (Ford et Fulkerson, 1956), sa complexité est de $O(|\mathcal{E}||f_{\max}|)$, avec $|f_{\max}|$ la valeur du flot-maximum obtenu et $|\mathcal{E}|$ le nombre d'arcs dans le graphe. Ensuite, des améliorations ont été apportées indépendamment dans (Edmonds et Karp, 1969) et (Dinic, 1969) pour réduire la complexité de l'algorithme, en proposant des stratégies optimales de recherche de chemins augmentants.

La première approche consiste à modifier l'étape de recherche de chemin menant de la source au puits par une recherche de chemin le plus court de la source au puits, dans le graphe résiduel. La complexité de l'algorithme proposé est de $O(|\mathcal{V}||\mathcal{E}|^2)$, où $|\mathcal{V}|$ est le nombre des nœuds, puisque la complexité de recherche de chemin augmentant, par la technique de Breadth First Search (*BFS*), est de $O(|\mathcal{E}|)$ et celle de la recherche de tous les chemins augmentant possibles est de $O(|\mathcal{E}||\mathcal{V}|)$.

La deuxième approche consiste à effectuer l'étape de recherche de chemins augmentant en cherchant ces chemins dans les réseaux de couches définis à partir du réseau $(\mathcal{G}, f)$. Un réseau de couche $(\mathcal{G}_d, f)$ est défini à partir de $(\mathcal{G}, f)$ en prenant les sommets de $\mathcal{G}$ à une distance d de s et les arcs de $\mathcal{G}$ d'extrémités les sommets des réseaux $\{(\mathcal{G}_i, f), i < d\}$. Itérativement, le flot-maximum est obtenu en saturant tous les chemins augmentant des réseaux de couches $\{(\mathcal{G}_i, f), i = 0..|s - t|\}$ ($|s - t|$ dénote la distance qui sépare le puits de la source). La complexité totale de l'algorithme est de $O(|\mathcal{V}|^2|\mathcal{E}|)$. Comparée à celle de l'approche précédente, elle est meilleure pour les graphes denses que pour les graphes creux, c.-à-d. pour les graphes présentant un nombre plus important d'arcs que de nœuds.

D'autres améliorations en complexité, apportées à cette technique de recherche de flot-maximum dans un réseau, ont été proposées dans la littérature. Pour une étude détaillée, nous référons le lecteur à l'ouvrage (Ahuja *et al.*, 1993).

En vision, il a été prouvé expérimentalement que cette technique de recherche de flot-maximum dans un réseau à deux nœuds terminaux s et t (notant deux étiquettes), où chaque nœud du graphe (correspondant à un site de $\mathcal{P}$) est relié à la source et au puits, est bien adaptée. En effet, ces graphes sont peu profonds et la longueur des chemins augmentant est généralement très petite. Dans (Boykov *et al.*, 2001), les auteurs ont utilisé et adapté cette technique à ces graphes pour une recherche optimale de chemins augmentant. En effet, en construisant pour un graphe donné deux arbres en largeur, l'un ayant comme racine la source et l'autre le puits, un chemin augmentant est obtenu quand les deux arbres se touchent. Les capacités résiduelles sur les arcs sont modifiées et la structure des arbres est mise à jour. Cette recherche est réalisée itérativement et l'algorithme converge après un

nombre fini d'itérations. Comme la recherche des chemins augmentant ne repose pas sur la recherche du plus court chemin, la complexité de l'algorithme est fonction ainsi de la valeur du flot-maximum obtenu. Elle est de $O(|\mathcal{V}|^2|\mathcal{E}||f_{max}|)$. Théoriquement, cette complexité est supérieure à celle des algorithmes classiques présentés, mais en pratique, la convergence est plus rapide sur les graphes spécifiques aux problèmes d'étiquetage en vision (Kolmogorov, 2004).

Flot-maximum par réétiquetage

On définit un préflot comme étant un flot ne respectant pas la contrainte de conservation de flots en chaque nœud du réseau. Pour un nœud u fixé, on a ainsi :

$$\sum_{v \in \mathcal{V}} f(v, u) \geq \sum_{v \in \mathcal{V}} f(u, v). \quad (\text{C.4})$$

On notera par $e(u)$ l'excès de flot du sommet u . Cette quantité est toujours positive.

On appellera les nœuds qui ont un excès de flot des nœuds actifs. Ces nœuds poussent le flot en excès vers d'autres nœuds disponibles. Pour connaître la disponibilité des nœuds, une étape d'étiquetage est nécessaire, associant à chaque nœud une étiquette, dite hauteur. Initialement, les hauteurs des nœuds sont fixées à 0, sauf celle de la source qui est fixée à $|\mathcal{V}|$. Ensuite, itérativement, chaque nœud actif pousse le maximum de flot, dans l'arc le reliant à son voisin de hauteur plus petite, sans dépasser la capacité sur cet arc. Si aucun sommet voisin n'est moins élevé, la hauteur du sommet actif est élevée pour pouvoir déverser le flot e . Ces deux opérations élémentaires : *pousser* et *réétiqueter* sont réalisées sur tous les nœuds du graphe selon un ordonnancement quelconque. L'algorithme converge si aucune opération élémentaire n'est possible. L'excès de flot se trouvant à la fin dans le puits correspond au flot-maximum. Le graphe est alors découpé en deux parties : les nœuds dont la hauteur à la fin de l'algorithme est plus grande que la source forment la partition $\mathcal{S}$ et les autres forment la partition $\mathcal{T}$. Ainsi on réalise une coupe-minimum sur le graphe.

L'algorithme comporte ainsi deux phases : phase de poussée de flot (dite *Push*) et phase de réétiquetage des nœuds (dite *relabel*). La phase de poussée de flot se déroule comme suit :

```

[Test si poussée est possible]
Si (u est un nœud en excès et v est un nœud voisin vérifiant  $r(u, v) > 0$  &  $h(u) > h(v)$ )
  Alors
    | Push : envoie  $\min(exces(u), r(u, v))$  de flot de u vers v ;
  Fin Si

```

Algorithme 2: Phase de poussée de flot.

La phase de réétiquetage se déroule comme suit :

```

[Test si réétiquetage est possible]
Si (u est un nœud non terminal en excès et aucun nœud voisin v ne vérifie  $h(u) > h(v)$ )
Alors
    Relabel :  $h(u) = \min\{h(v) + 1 \mid v \in \mathcal{V} \& r(u, v) > 0\}$ ;
Fin Si

```

Algorithme 3: Phase de réétiquetage des nœuds.

Ayant défini ces deux phases, on peut écrire l'algorithme de calcul de flot-maximum par poussée de flot comme suit :

```

Initialiser les préflots et les étiquettes (hauteurs);
Tant que (on peut pousser le flot ou réétiqueter un nœud) faire
    Effectuer le Push/Relabel;
    Mettre à jour les préflots et les étiquettes;
Fait

```

Algorithme 4: Algorithme générique de calcul du flot-maximum par réétiquetage.

Les premiers algorithmes reposants sur cette idée sont introduits indépendamment dans (Malhotra *et al.*, 1977) et (Goldberg et Tarjan, 1988). Le premier est de complexité $O(|\mathcal{E}||\mathcal{V}|^2)$. Le deuxième est plus optimal en agissant sur la priorité entre les actions de poussée ou réétiquetage. Sa complexité est de $O(|\mathcal{V}|^3)$ qui donne de meilleures performances que le premier quand le graphe est plus dense ou complet.

Une deuxième amélioration a été introduite dans (Cherkassky et Goldberg, 1997) qui consiste à ordonner les choix des nœuds en excès pour effectuer l'action de poussée de flot. Deux algorithmes ont été proposés, l'un a une complexité de $O(|\mathcal{V}|^3)$ fondé sur la stratégie *FIFO* (First In First Out), l'autre a une complexité de $O(|\mathcal{V}|^2 \sqrt{|\mathcal{E}|})$ reposant sur la stratégie *HL* (Highest Label). Nous renvoyons encore le lecteur cherchant plus de détails à l'ouvrage (Ahuja *et al.*, 1993).

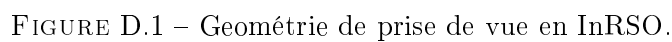
En vision, cette technique de recherche de flot-maximum dans un réseau a été testée sur des graphes pour la minimisation exacte de certaines fonctions d'énergies sous-modulaires de la classe des fonctions $\mathbb{F}^2$ (voir annexe B pour plus de détails sur la structure du graphe) (Roy et Cox, 1998). De plus, cette technique de recherche de flot-maximum se trouve plus simple à implémenter sur une architecture parallèle que la technique de recherche de chemins augmentant (Anderson et Setubal, 1992). Une adaptation possible en GPU appropriée aux problèmes de vision est présentée dans (Vineet et Narayanan, 2008).

Quelle approche choisir en vision ?

Il est difficile de décider laquelle des deux familles de techniques de calcul de flot-maximum est plus appropriée aux graphes utilisés en vision par ordinateur. Toutefois, en pratique, il est observable que pour des graphes non denses, typiquement les graphes binaires de l'annexe A, la technique de recherche de chemins augmentants est appropriée, vu la simplicité de recherche de ces chemins dans ce cas de topologie de graphes. En revanche, pour des graphes denses ou présentant beaucoup d'arcs, typiquement ceux de

l'annexe B, et quand la fonction d'énergie *a priori* est convexe quelconque, la technique par réétiquetage pourrait donner des résultats meilleurs en temps de calcul.

Géométrie de prise de vue en interférométrie Radar



Dans le plan (O,x,z) lié au sol, en négligeant la courbure de la Terre, l'antenne S_1 est à une hauteur H du sol faisant un angle d'incidence θ . L'antenne S_2 située à une distance (ou base) égale à B de S_1 est projetée sur le même repère, donnant les coordonnées suivantes :

Notons par $B^\perp$ la base orthogonale, c.-à-d. la composante orthogonale de la base sur la visée effectuée sous l'angle d'incidence θ , dont la valeur est donnée par :

$$B^\perp = B_z \cos(\theta) + B_x \sin(\theta). \quad (D.2)$$

Dans la figure D.1, deux cibles sont indiquées. La cible P est supposée à une altitude connue qui est nulle. Ce point sert de référence pour les mesures interférométriques qui suivent. La deuxième cible Q est un point dont l'altitude est à déterminer en fonction des paramètres d'acquisition. Cette cible est supposée appartenir au même front d'onde que celui de la cible P , c.-à-d. $S_1 P = S_1 Q$.

Dans le même référentiel lié au sol, ces deux cibles ont pour coordonnées :

$$P(X, 0); Q(X + \delta x, h). \quad (D.3)$$

On notera par la suite les distances antenne-cible comme suit :

$$S_1 P = R_P^1; S_2 P = R_P^2; S_1 Q = R_Q^1; S_2 Q = R_Q^2. \quad (D.4)$$

La distance $\Delta R_1 = R_P^1 - R_P^2$ qui représente la différence de trajet de l'onde radar entre l'antenne S_1 et le point de référence P et l'antenne S_2 et ce même point, n'est liée qu'au trajet de l'onde dans l'atmosphère dû à l'éloignement en distance du capteur (sans compter la topographie de la zone imagée puisque ce point de référence est à une altitude supposée nulle). Cette différence de trajet est donc reliée directement à la différence de phases dite *orbitale*.

La distance $\Delta R_2 = R_Q^1 - R_Q^2$ qui représente la différence de trajet de l'onde Radar entre l'antenne S_1 et le point Q et l'antenne S_2 et ce même point est liée à la fois au trajet de l'onde dans l'atmosphère lié à l'éloignement du point et à la hauteur de la cible et donc à la topographie de la zone imagée. Cette distance est donc reliée à la différence de phase orbitale ainsi qu'à celle *topographique*. Cependant, comme les deux cibles sont sur le même front d'onde et donc à une même distance du capteur, elles ont la même mesure de phase orbitale. En la retirant de la phase interférométrique mesurée sur la cible Q , on retrouve la phase topographique associée à cette cible.

Cherchons donc à calculer ΔR , la différence entre ΔR_1 et ΔR_2 , qui correspond en fait à la différence de marche liée à la topographie.

Sachant que les différences de trajet s'expriment ainsi :

$$\Delta R_1 = \sqrt{X^2 + H^2} - \sqrt{(X - B_x)^2 + (H + B_z)^2}, \quad (D.5)$$

$$\Delta R_2 = \sqrt{(X + \delta x)^2 + (H - h)^2} - \sqrt{(X + \delta x - B_x)^2 + (H - h + B_z)^2}. \quad (D.6)$$

La différence entre ces deux quantités, sous l'hypothèse que H^2 et X^2 sont très grands devant H, X, h, x, B_x et B_z , peut être approximée par un développement limité à l'ordre 2 (Abdelfattah *et al.*, 1998) donnant :

$$\Delta R = \frac{B^\perp}{R_1} (x \cos(\theta) + h \sin(\theta)). \quad (D.7)$$

Sachant que pour les cibles du même front d'onde, l'équation suivante est vérifiée :

$$h \cos(\theta) - x \sin(\theta) = 0, \quad (D.8)$$

la différence de marche reliée à la phase topographique est liée à l'altitude de la cible Q par l'intermédiaire des paramètres d'acquisition selon l'expression suivante :

$$\Delta R = \frac{B^\perp}{R_1 \sin(\theta)} h. \quad (D.9)$$

Par cette relation, on arrive à exprimer l'altitude de la cible en fonction des paramètres d'acquisition InRSO et de la mesure de différence de marche ΔR .

Annexe E

Statistiques sur les données InRSO

Selon le formalisme de Goodman (Goodman, 2006), la réponse d'une cellule de résolution est la somme de N réponses de cibles ponctuelles qui y sont réparties aléatoirement. La réponse d'une cible ponctuelle i est un signal qui s'écrit sous la forme complexe suivante :

$$z_i = a_i \exp(j\psi_i). \quad (\text{E.1})$$

avec $j^2 = -1$.

La réponse totale de la cellule s'écrit alors :

$$z = \sum_{i=1}^N z_i = \sum_{i=1}^N a_i \exp(j\psi_i) = r + jm, \quad (\text{E.2})$$

avec r et m les parties réelles et imaginaires de z .

Sous les hypothèses suivantes :

- a_i et ψ_i sont des variables aléatoires indépendantes,
- les ψ_i sont indépendantes entre elles et suivent la loi uniforme sur $[0, 2\pi]$,
- N est un nombre élevé,

le théorème central limite permet de conclure qu'une réalisation d'un pixel d'une image SLC (cellule de résolution) est considérée comme étant une variable aléatoire complexe $z = r + jm$ qui suit la loi gaussienne complexe circulaire de variance $\sigma^2 = \frac{\mu}{2}$. μ est appelée la réflectivité. Sa densité de probabilité est définie comme étant la densité de probabilité conjointe entre sa partie réelle et sa partie imaginaire. Sous les hypothèses de Goodman, les parties réelle et imaginaire sont indépendantes et distribuées chacune selon une loi normale de même variance μ , c.-à-d.

$$P(r) = \frac{1}{\sqrt{\pi\mu}} \exp\left(-\frac{r^2}{\mu}\right), \quad (\text{E.3})$$

$$P(m) = \frac{1}{\sqrt{\pi\mu}} \exp\left(-\frac{m^2}{\mu}\right). \quad (\text{E.4})$$

L'indépendance entre r et m induit

$$P(z) = P(r, m) = P(r)P(m) = \frac{1}{\pi\mu} \exp\left(-\frac{r^2 + m^2}{\mu}\right). \quad (\text{E.5})$$

En écrivant les variables aléatoires r et m sous les formes suivantes :

$$r = a \cos \psi, \quad (\text{E.6})$$

$$m = a \sin \psi, \quad (\text{E.7})$$

où, a et ψ correspondent respectivement à l'amplitude et la phase du complexe z , la densité de probabilité de z peut être exprimée autrement par un changement de variables. La valeur du Jacobien $|J|$ pour cette transformation vaut a et la densité prend la forme suivante :

$$P(z) = P(a, \psi) = \frac{a}{\pi\mu} \exp\left(-\frac{a^2}{\mu}\right), \quad (\text{E.8})$$

pour $a \geq 0$ et $-\pi \leq \psi < \pi$ (Nous respectons cette hypothèse dans la suite).

La densité marginale associée à la variable amplitude est obtenue en intégrant $P(z)$ sur ψ entre $-\pi$ et π . Les calculs donnent :

$$P(a) = \frac{a}{\mu} \exp\left(-\frac{a^2}{\mu}\right). \quad (\text{E.9})$$

Cela correspond à la loi de Rayeight de moyenne $m_a = \frac{\sqrt{\pi\mu}}{2}$ et d'écart type $\sigma_a = m_a \sqrt{\frac{4}{\pi} - 1}$.

La densité marginale associée à la variable phase est obtenue en intégrant $P(z)$ sur a entre 0 et ∞ . Les calculs donnent

$$P(\psi) = \frac{1}{2\pi}. \quad (\text{E.10})$$

Cela correspond à la loi uniforme sur $[-\pi, \pi]$.

Nous remarquons que les variables a et ψ sont aussi indépendantes puisque

$$P(z) = P(a, \psi) = P(a)P(\psi). \quad (\text{E.11})$$

En interférométrie, nous considérons un couple de réalisations de variables aléatoires complexes gaussiennes circulaires $Z = [z_1 z_2]$. La distribution jointe associée à ce vecteur est donnée par l'expression suivante :

$$P(Z) = \frac{1}{\pi^2 \det C_Z} \exp(-Z^H C_Z^{-1} Z), \quad (\text{E.12})$$

où C_Z est la matrice de covariance complexe définie par

$$C_Z = E\{ZZ^H\} = \begin{pmatrix} E[|z_1|^2] & E[z_1 \bar{z}_2] \\ E[\bar{z}_1 z_2] & E[|z_2|^2] \end{pmatrix}, \quad (\text{E.13})$$

où Z^H est la matrice transconjuguée de Z (conjuguée de la transposée de Z), $\det(C_Z)$ est le déterminant de la matrice C_Z , $E[\cdot]$ est l'espérance mathématique et $\bar{z}_i$ indique la valeur conjuguée du complexe z_i .

Nous rappelons aussi que le coefficient de corrélation complexe entre les deux éléments z_1 et z_2 s'écrit :

$$\rho = \frac{E[z_1 \bar{z}_2]}{\sqrt{E[|z_1|^2]E[|z_2|^2]}} = \gamma e^{j\phi}, \quad (\text{E.14})$$

où γ est le module de ρ et représente la cohérence théorique entre les éléments z_1 et z_2 et ϕ est la phase de ρ et correspond au déphasage entre eux.

En notant par $\mu_1 = E[|z_1|^2]$ et $\mu_2 = E[|z_2|^2]$ les réflectivités respectives de z_1 et z_2 , la matrice de cohérence s'écrit alors :

$$C_Z = \begin{pmatrix} \mu_1 & \sqrt{\mu_1 \mu_2} \gamma e^{j\phi} \\ \sqrt{\mu_1 \mu_2} \gamma e^{-j\phi} & \mu_2 \end{pmatrix}. \quad (\text{E.15})$$

Pour $\gamma \neq 1$ on peut définir l'inverse de C_Z par :

$$C_Z^{-1} = \frac{1}{\mu_1 \mu_2 (1 - \gamma^2)} \begin{pmatrix} \mu_2 & -\sqrt{\mu_1 \mu_2} \gamma e^{j\phi} \\ -\sqrt{\mu_1 \mu_2} \gamma e^{-j\phi} & \mu_1 \end{pmatrix}. \quad (\text{E.16})$$

Calcul fait, la distribution du vecteur Z s'écrit alors :

$$P(z_1, z_2 | \mu_1, \mu_2, \gamma, \phi) = \frac{1}{\pi^2 \mu_1 \mu_2 (1 - \gamma^2)} \exp \left(-\frac{1}{1 - \gamma^2} \left(\frac{z_1 \bar{z}_1}{\mu_1} + \frac{z_2 \bar{z}_2}{\mu_2} - \frac{\gamma(z_1 \bar{z}_2 e^{j\phi} + \bar{z}_1 z_2 e^{-j\phi})}{\sqrt{\mu_1 \mu_2}} \right) \right). \quad (\text{E.17})$$

En reprenant cette fois-ci la décomposition des complexes z_1 et z_2 selon leurs amplitudes et phases, c.-à-d.

$$z_1 = a_1 e^{j\psi_1}, \quad (\text{E.18})$$

$$z_2 = a_2 e^{j\psi_2}, \quad (\text{E.19})$$

avec $a_1, a_2 \geq 0$ et $-\pi \leq \psi_1, \psi_2 < \pi$, et en considérant l'hypothèse que $\mu_1 = \mu_2 = \mu$, la distribution conjointe du vecteur Z peut être recalculée selon ces nouvelles variables aléatoires réelles. La valeur du Jacobien $|J|$ pour cette transformation vaut $a_1 a_2$ et la densité conjointe prend la forme suivante :

$$P(a_1, a_2, \psi_1, \psi_2 | \mu, \gamma, \phi) = \frac{a_1 a_2}{\pi^2 \mu^2 (1 - \gamma^2)} \exp \left(-\frac{1}{\mu(1 - \gamma^2)} (a_1^2 + a_2^2 - 2a_1 a_2 \gamma \cos(\psi - \phi)) \right), \quad (\text{E.20})$$

avec $\psi = \psi_2 - \psi_1$ qui représente le déphasage mesuré entre les deux variables complexes z_1 et z_2 .

Ainsi, nous pouvons déterminer la distribution conjointe entre les variables a_1 et a_2 en intégrant la dernière expression sur les variables ψ_1 et ψ_2 variant de $-\pi$ à π . Les calculs donnent :

$$P(a_1, a_2 | \mu, \gamma) = \frac{4a_1 a_2}{\mu^2 (1 - \gamma^2)} \exp \left(-\frac{a_1^2 + a_2^2}{\mu(1 - \gamma^2)} \right) I_0 \left(\frac{2\gamma a_1 a_2}{\mu(1 - \gamma^2)} \right), \quad (\text{E.21})$$

avec I_0 est la fonction de Bessel modifiée de première espèce, d'ordre 0.

Et la distribution conjointe entre les variables ψ_1 et ψ_2 est obtenue en intégrant la dernière expression sur les variables a_1 et a_2 variant de 0 à ∞ . Les calculs donnent :

$$P(\psi_1, \psi_2 | \gamma, \phi) = \frac{1}{4\pi^2} \frac{1 - \gamma^2}{1 - \gamma^2 \cos^2(\psi - \phi)} \left(1 + \frac{\gamma \cos(\psi - \phi) \arccos(-\gamma \cos(\psi - \phi))}{\sqrt{1 - \gamma^2 \cos^2(\psi - \phi)}} \right). \quad (\text{E.22})$$

En partant de cette distribution, la distribution de la phase interférométrique $\psi = \psi_2 - \psi_1$ est obtenue en effectuant un changement de variable donnant une valeur du Jacobien $|J| = 2\pi$ et la distribution de la phase interférométrique s'écrit comme :

$$p(\psi | \gamma, \phi) = \frac{1}{2\pi} \frac{1 - \gamma^2}{1 - \gamma^2 \cos^2(\psi - \phi)^2} \left(1 + \frac{\gamma \cos(\psi - \phi) \arccos(-\gamma \cos(\psi - \phi))}{\sqrt{1 - \gamma^2 \cos^2(\psi - \phi)^2}} \right). \quad (\text{E.23})$$

Liste des publications

Journaux

- G. Ferraioli, A. Shabou, F. Tupin and V. Pascazio ; **Multichannel Phase Unwrapping with Graph-cuts** ; *IEEE Geoscience and Remote Sensing Letters*, Volume 6, Issue 3, July 2009, Pages : 562-566.
- F. Baselice, G. Ferraioli and A. Shabou ; **Field Map Reconstruction in Magnetic Resonance Imaging Using Bayesian Estimation** ; *Sensors* 10, no. 1, 2010, Pages : 266-279.
- A. Shabou, J. Dabon and F. Tupin ; **A Markovian Approach for InSAR Phase Reconstruction with Mixed Discrete and Continuous Optimization** ; *IEEE Geoscience and Remote Sensing Letters*, in press.
- A. Shabou, J. Darbon and F. Tupin ; **Multilabel Partition Moves for MRF Optimization** ; *Image and Vision Computing*, en révision.

Conférences internationales

- G. Ferraioli, A. Shabou, F. Tupin and V. Pascazio ; **Fast InSAR Multichannel Phase Unwrapping for DEM Generation** ; *International Urban Remote Sensing Conference (URBAN'09)*, Shanghai, China, Mai 2009.
- A. Shabou, J. Darbon and F. Tupin ; **A graph-cut based algorithm for approximate MRF optimization** ; *IEEE International Conference on Image Processing (ICIP'09)*, Cairo, Egypt, November 2009.
- A. Shabou, G. Ferraioli, F. Tupin and V. Pascazio ; **Three Dimensional Reconstruction of Urban Areas Using Jointly Phase and Amplitude Multichannel Images** ; *IEEE International Geoscience and Remote Sensing Symposium (IGARSS'10)*, Hawaii, USA, July 2010.

Bibliographie

- R. ABDELFATTAH, J. M. NICOLAS et M. R. BOUSSEMA : **Topography Optimization for Accurate Interferometric Phase Generation**. In *Proceedings of the Society of Photo-Optical Instrumentation Engineers : Conference Series*, volume 3497, pages 251–256, 1998.
- A. AGARWALA, M. DONTCHEVA, M. AGRAWALA, S. DRUCKER, A. COLBURN, B. CURLESS, D. SALESIN et M. COHEN : **Interactive Digital Photomontage**. *ACM Transactions on Graphics*, 2004.
- R. K. AHUJA, T. L. MAGNANTI et J. B. ORLIN : **Networks Flows : Theory, Algorithms and Applications**. Prentice Hall, 1993.
- A. ALI, A. FARAG et G. GIMELFARB : **Optimizing Binary MRFs with Higher Order Cliques**. pages 98–111, 2008.
- R. J. ANDERSON et J. C. SETUBAL : **On the parallel Implementation of Goldberg’s Maximum Flow Algorithm**. In *Proceedings of the fourth annual ACM symposium on Parallel algorithms and architectures*, pages 168–177, 1992.
- J. BESAG : **On the Statistical Analysis of Dirty Pictures**. *Journal of Royal Statistics Society*, B(48):259–302, 1986.
- J. M. BIOUCAS-DIAS, V. KATKOVNIK, J. ASTOLA et K. EGIAZARIAN : **Absolute Phase Estimation : Adaptive Local Denoising and Global Unwrapping**. *Applied Optics*, 47(29):5358–5369, 2008.
- J. M. BIOUCAS-DIAS et J. M. N. LEITAO : **InSAR Phase Unwrapping : a Bayesian Approach**. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, volume 1, pages 396–400, 2001.
- J. M. BIOUCAS-DIAS et J. M. N. LEITAO : **The ZpM Algorithm for Interferometric Image Reconstruction in SAR/SAS**. *IEEE Transactions on Image processing*, 11(4):408–422, 2002.
- J. M. BIOUCAS-DIAS et G. VALADAO : **Phase Unwrapping via Graph Cuts**. *IEEE Transactions on Image Processing*, 16(3):698–709, 2007.
- A. BLAKE et A. ZISSERMAN : **Visual Reconstruction**. MIT Press, Cambridge, MA., 1987.
- R. BOLTER : **Reconstruction of Man-made Objects from High Resolution SAR Images**. In *Proceedings of IEEE Aerospace Conference*, volume 3, pages 287–292, 2000.

-
- D. J. BONE : **Fourier Fringe Analysis - The Two-dimensional Phase Unwrapping Problem.** *Applied Optics*, 30:3627–3632, 1991.
- J. F. BONNANS, J. C. GILBERT, C. LEMARÉCHAL et C. SAGASTIZABAL : **Optimisation Numérique : Aspects théoriques et pratiques.** Springer-Verlag, 1997.
- E. BOROS et P. L. HAMMER : **Pseudo-boolean Optimization.** *Discrete Applied Mathematics*, 123(1-3):155–225, 2002.
- E. BOROS, P. L. HAMMER et X. SUN : **Network Flows and Minimization of Quadratic Pseudo-Boolean Functions.** Rapport technique, RUTCOR Research Report, 1991.
- E. BOROS, P. L. HAMMER et G. TAVARES : **Preprocessing of Unconstrained Quadratic Binary Optimization.** Rapport technique, RUTCOR Research Report, 2006.
- Y. BOYKOV et V. KOLMOGOROV : **An Experimental Comparison of Min-cut/Max-flow Algorithms for Energy Minimization in Vision.** *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(9):1124–1137, 2004.
- Y. BOYKOV, O. VEKSLER et R. ZABIH : **Fast Approximate Energy Minimization via Graph Cuts.** *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1222–1239, 2001.
- J. R. BUCKLAND, J. M. HUNTLEY et S. R. E. TURNER : **Unwrapping Noisy Phase Maps by Use of a Minimum-cost-matching Algorithm.** *Applied Optics*, 34:5100–5108, 1995.
- P. CARR et R. HARTLEY : **Solving Multilabel Graph Cut Problems with Multilabel Swap.** *Digital Image Computing : Techniques and Applications*, pages 532–539, 2009.
- F. CHAABANE : **Suivi multitemporel des mouvements de terrain par Interférométrie RSO.** Thèse de doctorat, Télécom Paris Tech, 2004.
- A. CHAMBOLLE : **An Algorithm for Total Variation Minimization and Applications.** *Journal of Mathematical Imaging and Vision*, 20(1-2):89–97, 2004.
- P. CHARBONNIER : **Reconstruction d'image : régularisation avec prise en compte des discontinuités.** Thèse de doctorat, Université de Nice Sophia Antipolis, 1994.
- C. W. CHEN : **Statistical-cost Network-flow Approaches to Two-dimensional Phase Unwrapping for Radar Interferometry.** Thèse de doctorat, Stanford University, 2001.
- D. Z. CHEN et X. WU : **Efficient Algorithms for k-Terminal Cuts on Planar Graphs.** *Algorithmica*, 38(2):299–316, 2004.
- B. V. CHERKASSKY et A. V. GOLDBERG : **On Implementing the Push-Relabel Method for the Maximum Flow Problem.** *Algorithmica*, 19(4):390–410, 1997.
- G. CORSINI, M. DIANI, F. LOMBARDINI et G. PINELLI : **Reduction of the Phase-Unwrapping Drawbacks by the Three-antenna Interferometric SAR System.** *In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, volume 4, pages 1536–1541, 1997.
-

-
- M. COSTANTINI : **A Novel Phase Unwrapping Method Based on Network Programming.** *IEEE Transactions on Geoscience and Remote Sensing*, 36(3):813–821, 1998.
- J. C. CURLANDER et R. N. McDONOUGH : ***Synthetic Aperture Radar : Systems and Signal Processing.*** Wiley-Interscience, 1991.
- R. CUSACK, J. M. HUNTLEY et H. T. GOLDBREIN : **Improved Noise-immune Phase-Unwrapping Algorithm.** *Applied Optics*, 34:781–789, 1995.
- R. A. DANA et D. L. KNEPP : **The Impact of Strong Scintillation on Space Based Radar Design. II - Noncoherent Detection.** *IEEE Transactions on Aerospace Electronic Systems*, 22:34–46, 1986.
- J. DARBON : ***Composants Logiciels et Algorithmes de Minimisation Exacte d'Énergies Dédiés au Traitement des Images.*** Thèse de doctorat, École Nationale Supérieure des Télécommunications de Paris, 2005.
- J. DARBON : **Global Optimization for First Order Markov Random Fields with Submodular Priors.** *Combinatorial Image Analysis*, pages 229–237, 2008.
- J. DARBON et M. SIGELLE : **Image Restoration with Discrete Constrained Total Variation Part I : Fast and Exact Optimization.** *Journal of Mathematical Imaging and Vision*, 26(3):261–276, 2006a.
- J. DARBON et M. SIGELLE : **Image Restoration with Discrete Constrained Total Variation Part II : Levelable Functions, Convex Priors and Non-Convex Cases.** *Journal of Mathematical Imaging and Vision*, 26(3):277–291, December 2006b.
- F. DELL'ACQUA, P. GAMBA, U. SOERGEL et U. THOENNESSEN : **Fully 3D and Joint 2D/3D Segmentation of InSAR Data Over Urban Environments.** *In Proceeding of the IEEE/ISPRS Joint Workshop of Remote Sensing and Data Fusion over Urban Areas*, pages 328–331, 2001.
- L. DENIS, F. TUPIN, J. DARBON et M. SIGELLE : **Joint regularization of phase and amplitude of InSAR data : application to 3D reconstruction.** *IEEE Transactions on Geoscience and Remote Sensing*, 47(11), 2009a.
- L. DENIS, F. TUPIN, J. DARBON et M. SIGELLE : **SAR Image Regularization with Fast Approximate Discrete Minimization.** *IEEE Transactions on Image Processing*, 18(7):1588–1600, 2009b.
- D. DERAUW : **Phase Unwrapping Using Coherence Measurements.** *In Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 2584, pages 319–324, 1995.
- E. A. DINIC : **Algorithm for Solution of a Problem of Maximum Flow in a Network with Power Estimation.** *Soviet mathematics - Doklady*, 11:1277–1280, 1969.
- S. DUPONT et M. BERTHOD : **SAR Interferometry and Phase Unwrapping.** Rapport technique, INRIA, 1994.
-

-
- J. EDMONDS et R. M. KARP : **Theoretical Improvements in Algorithmic Efficiency for Network Flow Problems**. *International Conference on Combinatorial Structures and Their Applications*, pages 93–96, 1969.
- P. ELIAS, A. FEINSTEIN et Shannon C. E. : **Note on Maximum Flow Through a Network**. *IRE Transactions on Information Theory*, 2(4):117–119, 1956.
- G. FERRAIOLI, G. FERRAIUOLO et V. PASCAZIO : **Phase-offset Estimation in Multichannel SAR Interferometry**. *IEEE Geoscience and Remote Sensing Letters*, 5(3):458–462, 2008.
- G. FERRAIOLI, A. SHABOU, F. TUPIN et V. PASCAZIO : **Multichannel Phase Unwrapping With Graph-Cuts**. *IEEE Geoscience and Remote Sensing Letters*, 6(3):562–566, 2009.
- G. FERRAIUOLO, V. PASCAZIO et G. SCHIRINZI : **Maximum A Posteriori Estimation of Height Profiles in InSAR Imaging**. *IEEE Geoscience and Remote Sensing Letters*, 1:66–70, 2004.
- B. FLACH et D. SCHLESINGER : **Best Labeling Search for a Class of Higher Order Gibbs Models**. *Pattern Recognition and Image Analysis*, 14(2):249–254, 2004.
- T. J. FLYNN : **Consistent 2D Phase Unwrapping Guided by a Quality Map**. *In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, volume 4, pages 2057–2059, 1990.
- T. J. FLYNN : **Two-Dimensional Phase Unwrapping with Minimum Weighted Discontinuity**. *Journal of the Optical Society of America A*, 14:2692–2701, 1997.
- L.R. FORD et D.R. FULKERSON : **Maximal Flow through a Network**. *Canadian Journal of Mathematics*, 8:309–404, 1956.
- G. FORNARO, G. FRANCESCHETTI et R. LANARI : **Interferometric SAR Phase Unwrapping using Green’s Formulation**. *IEEE Transactions on Geoscience and Remote Sensing*, 34(3):720–727, 1996.
- G. FORNARO, A. M. GUARNIERI, A. PAUCIULLO et F. DE-ZAN : **Maximum Likelihood Multi-baseline SAR Interferometry**. *Proceedings of the IEEE Radar Sonar and Navigation*, 153(3):279–288, 2006.
- D. FREEDMAN et P. DRINEAS : **Energy Minimization via Graph Cuts : Settling What is Possible**. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 939–946, 2005.
- D. L. FRIED : **Least-Square Fitting a Wave-Front Distortion Estimate to an Array of Phase-Difference Measurements**. *Journal of the Optical Society of America*, 67:370–375, 1977.
- R. L. FROST, C. K. RUSHFORTH et B. S. BAXTER : **Fast FFT-based Algorithm for Phase Estimation in Speckle Imaging**. *Applied Optics*, 18:2056–2061, 1979.
- A. K. GABRIEL et R. M. GOLDSTEIN : **Radar Interferogram Made from Two Passes of SIR-B Over the Rocky Mountains in British Columbia, Canada**. *International Journal of Remote Sensing*, 9:835–835, 1988.
-

-
- P. GAMBA, B. HOUSHMAND et M. SACCANI : **Detection and Extraction of Buildings from Interferometric SAR Data**. *IEEE Transactions on Geoscience and Remote Sensing*, 38:611–617, 2000.
- D. GEMAN et S. GEMAN : **Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images**. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.
- D. C. GHIGLIA, G. A. MASTIN et L. A. ROMERO : **Cellular-automata Method for Phase Unwrapping**. *Journal of the Optical Society of America A*, 4:267–280, 1987.
- D. C. GHIGLIA et M. D. PRITT : **Two-Dimensional Phase Unwrapping-Theory, Algorithms, and Software**. New York, Wiley, 1998.
- D. C. GHIGLIA et L. A. ROMERO : **Robust Two-Dimensional Weighted and Unweighted Phase Unwrapping that Uses Fast Transforms and Iterative Methods**. *Journal of the Optical Society of America A*, 11:107–117, 1994.
- D. C. GHIGLIA et D. E. WAHL : **Interferometric Synthetic Aperture Radar Terrain Elevation Mapping from Multiple Observations**. In *Proceedings of the IEEE Digital Signal Processing Workshop*, pages 33–36, 1994.
- L. A. GHIGLIA, D. C. and Romero : **Minimum Lp-norm Two-dimensional Phase Unwrapping**. *Journal of the Optical Society of America A : Optics, Image Science, and Vision*, 13(10):1999–2013, 1996.
- A. V. GOLDBERG et R. E. TARJAN : **A New Approach to the Maximum-Flow Problem**. *Journal of the ACM*, 35(4):921–940, 1988.
- R. GOLDSTEIN, H. A. ZEBKER et C. WERNER : **Satellite Radar Interferometry : Two Dimensional Phase Unwrapping**. *Radio Science*, 23(4):713–720, 1988.
- J. W. GOODMAN : **Statistical Optics**. New York, Wiley, 1985.
- J. W. GOODMAN : **Speckle Phenomena in Optics : Theory and Applications**. Roberts & Company Publishers, 2006.
- N. R. GOODMAN : **Statistical Analysis Based on a Certain Multivariate Complex Gaussian Distribution (an Introduction)**. *The Annals of mathematical statistics*, 34(1):152–177, 1963.
- L. GRAHAM : **Synthetic Interferometer Radar for Topographic Mapping**. In *Proceedings of IEEE*, volume 62, pages 763–768, 1974.
- D. M. GREIG, B. T. PORTEOUS et A. H. SEHEULT : **Exact Maximum A Posteriori Estimation for Binary Images**. *Journal of Roayal Statistical Society*, 51(2):271–279, 1989.
- F. GUICHARD et J. M. MOREL : **Mathematical morphology 'almost everywhere'**. In *Proceedings of the International Symposium on Mathematical Morphology*, pages 293–303, 2002.
- P. L. HAMMER, P. HANSEN et B. SIMEONE : **Roof Duality, Complementation and Persistency in Quadratic 0-1 Optimization**. *Mathematical Programming*, 28:121–155, 1984.
-

-
- T. HEBERT et R. LEAHY : **A Generalized EM Algorithm for 3D Bayesian Reconstruction from Poisson Data Using Gibbs Priors.** *IEEE transactions on medical imaging*, 8(2):194–202, 1989.
- T. C. HU : **Integer Programming and Network Flows.** Addison-Wesley Reading, MA, 1969.
- P. J. HUBER : **Robust Regression : Asymptotics, Conjectures and Monte Carlo.** *The Annals of Statistics*, 1(5):799–821, 1973.
- J. M. HUNTLEY : **Noise-immune Phase Unwrapping Algorithm.** *Applied Optics*, 28(16):3268–3270, 1989.
- H. ISHIKAWA : **Global Optimization Using Embedded Graphs.** Thèse de doctorat, New York University, 2000.
- H. ISHIKAWA : **Exact Optimization for Markov Random Fields with Convex Priors.** *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(10):1333–1336, 2003.
- H. ISHIKAWA : **Higher-Order Clique Reduction in Binary Graph Cut.** pages 2993–3000, 2009a.
- H. ISHIKAWA : **Higher-Order Gradient Descent by Fusion-Move Graph Cut.** *In Proceedings of the IEEE International Conference on Computer Vision*, pages 568–574, 2009b.
- H. ISHIKAWA et D. GEIGER : **Occlusions, Discontinuities, and Epipolar Lines in Stereo.** *In Proceedings of the IEEE European Conference on Computer Vision*, pages 232–236, 1998a.
- H. ISHIKAWA et D. GEIGER : **Segmentation by Grouping Junctions.** *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 125–131, 1998b.
- K. ITOH : **Analysis of the Phase Unwrapping Problem.** *Applied Optics*, 2(14), 1982.
- E. JAKEMAN et P. N. PUSEY : **A Model for Non-Rayleigh Sea Echo.** *IEEE Transactions on Antennas and Propagation*, 24:806–814, 1976.
- T. JUDGE : **A Review of Phase Unwrapping Techniques in Fringe Analysis.** *Optics and Lasers in Engineering*, 21:199–239, 1994.
- T. R. JUDGE, C. QUAN et P. J. BRYANSTON-CROSS : **Holographic Deformation Measurements by Fourier Transform Technique with Automatic Phase Unwrapping.** *Optical Engineering*, 31:533–543, 1992.
- D. JUST, N. ADAM, M. SCHWABISCH et R. BAMLER : **Comparison of Phase Unwrapping Algorithms for SAR Interferograms.** *In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, volume 1, pages 767–769, 1995.
- M. KIRSCHT et K. RINKE : **3D Reconstruction of Buildings and Vegetation from Synthetic Aperture Radar (SAR) Images.** *In Proceedings of IAPR Workshop on Machine Vision Applications*, pages 17–19, 1998.
-

-
- P. KOHLI : *Minimizing Dynamic and Higher Order Energy Functions using Graph Cuts*. Thèse de doctorat, Oxford Brookes University, 2007.
- P. KOHLI, M. P. KUMAR et P. H. S. TORR : **P3 Beyond : Solving Energies with Higher Order Cliques**. pages 1–8, 2007.
- P. KOHLI, M. P. KUMAR et P. H. S. TORR : **P3 & Beyond : Move Making Algorithms for Solving Higher Order Functions**. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(9):1645–1656, 2009.
- P. KOHLI, V. LEMPITSKY et C. ROTHER : **Uncertainty Driven Multi-scale Optimization**. In *DAGM*, volume in press, 2010.
- V. KOLMOGOROV : *Graph Based Algorithms for Scene Reconstruction from Two or More Views*. Thèse de doctorat, Cornell University, 2004.
- V. KOLMOGOROV et C. ROTHER : **Minimizing Nonsubmodular Functions with Graph Cuts - A Review**. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(7):1274–1279, 2007.
- V. KOLMOGOROV et R. ZABIH : **What Energy Functions Can Be Minimized via Graph Cuts ?** *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(2):147–159, 2004.
- V. KWATRA, A. SCHÖDL, I. ESSA, G. TURK et A. BOBICK : **Graphcut Textures : Image and Video Synthesis Using Graph Cuts**. *ACM Transactions on Graphics*, 22(3):277–286, 2003.
- D. LABROUSSE : *Modélisation Markovienne pour le déroulement de phases Interférométriques SAR*. Thèse de doctorat, Université de Nice- Sophia Antipolis, 1996.
- F. W. LEBERL : *Radargrammetric Image Processing*. Artech House, 1990.
- J. S. LEE, K. W. HOPPEL, S. A. MANGO et A. R. MILLER : **Intensity and Phase Statistics of Multilook Polarimetric and Interferometric SAR Imagery**. *IEEE Transactions on Geoscience and Remote Sensing*, 32:1017–1028, 1994.
- J. M. N. LEITAO et M. A. T. FIGUEIREDO : **Interferometric Image Reconstruction as a Nonlinear Bayesian Estimation Problem**. In *Proceedings of the IEEE International Conference on Image Processing*, page 2453, 1995.
- V. LEMPITSKY, C. ROTHER, S. ROTH et A. BLAKE : **Fusion Moves for Markov Random Field Optimization**. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009.
- F. LI et R. GOLDSTEIN : **Studies of Multi-baseline Spaceborne Interferometric Synthetic Aperture Radars**. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, volume 2, pages 1545–1550, 1987.
- S. Z. LI : *Markov Random Field Modeling in Image Analysis*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2001.
- H. LIM, W. XU et X. HUANG : **Two New Practical Methods for Phase Unwrapping**. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, volume 1, pages 196–198, 1995.
-

-
- Q. LIN, J. F. VESECKY et H. A. ZEBKER : **Topography Estimation with Interferometric Synthetic Aperture Radar using Fringe Detection**. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, volume 4, pages 2173–2176, 1991.
- Q. LIN, J. F. VESECKY et H. A. ZEBKER : **New Approaches In Interferometric SAR Data Processing**. *IEEE Transactions on Geoscience and Remote Sensing*, 30(3):560–567, 1992.
- Q. LIN, J. F. VESECKY et H. A. ZEBKER : **Phase Unwrapping Through Fringe-Line Detection in Synthetic Aperture Radar Interferometry**. *Applied Optics*, 33:201–208, 1994.
- F. LOMBARDINI : **Absolute Phase Retrieval in a Three-element Synthetic Aperture Radar Interferometer**. In *Proceedings of the IEEE International Conference of Radar*, pages 309–312, 1996.
- A. LOPES, H. LAUR et E. NEZRY : **Statistical Distribution and Texture in Multilook and Complex SAR Images**. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, pages 2427–2430, 1990.
- S. N. MADSEN et H. A. ZEBKER : **Automated Absolute Phase Retrieval in Across-Track Interferometry**. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, pages 1582–1584, 1992.
- V. M. MALHOTRA, M. P. KUMAR et S. N. MAHESHWARI : **Algorithm of Construction of Maximal Flow in Networks with Complexity of $O(V^2E^{1/2})$ Operations**. *Mathematic Methods of Solution of Economic problems*, 7:117–125, 1977.
- H. MAÎTRE : *Traitement des Images RSO*. Hermes, 2001.
- F. MEGLIO, V. PASCAZIO et G. SCHIRINZI : **Joint Statistical Distribution of Multi-Baseline SAR Interferograms**. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, pages 2549–2552, 2006.
- E. MICHAELSEN, U. SOERGEL et U. THOENNESSEN : **Perceptual Grouping for Automatic Detection of Man-made Structures in High-resolution SAR Data**. *Pattern Recognition Letters*, 27(4):218–225, 2006.
- L. MOISAN : **How to Discretize the Total Variation of an Image**. In *proceedings of the International Congress on Industrial and Applied Mathematics*, volume 7, pages 1041907–1041908, 2007.
- A. MONTI-GUARNIERI, F. PARIZZI, P. PASQUALI, C. PRATI et F. ROCCA : **SAR Interferometry Experiments with ERS-1**. In *Proceedings of the IEE International Geoscience and Remote Sensing Symposium*, volume 3, pages 991–993, 1993.
- G. NICO, G. PALUBINSKAS et M. DATCU : **Bayesian Approaches to Phase Unwrapping : Theoretical Study**. *IEEE Transactions on Signal Processing*, 48(9):2545–2556, 2000.
- J. M. NICOLAS : **Introduction aux statistiques de deuxième espèce : Applications des Logs-moments et des Logs-cumulants à l'analyse des lois d'images radar**. *Traitement du signal*, 19(3):139–167, 2002.
-

-
- R. OTTEN et L. VAN GINNEKEN : *The Annealing Algorithm*. Kluwer Academic Publishers Boston MA., 1989.
- V. PASCAZIO et G. SCHIRINZI : **Multifrequency InSAR Height Reconstruction Through Maximum Likelihood Estimation of Local Planes Parameters**. *IEEE Transactions on Image Processing*, 11(12):1478– 1489, 2002.
- D. PETIT : *Reconstruction du 3D par Interférométrie radar haute résolution*. Thèse de doctorat, Institut de Recherche en Informatique de Toulouse, 2004.
- J. P. PICARD et H. D. RATLIF : **Minimum Cuts and Related Problem**. *Networks*, 5:357–370, 1975.
- C. PRATI, M. GIANI et N. LEURATTI : **SAR Interferometry : A 2D Phase Unwrapping Technique Based On Phase And Absolute Values Informations**. *In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, numéro 20, pages 2043–2046, 1990.
- M. D. PRITT : **Phase Unwrapping by Means of Multigrid Techniques For Interferometric SAR**. *IEEE Transactions on Geoscience and Remote Sensing*, 34(3):728–738, 1996.
- M. QUARTULLI et M. DATCU : **Stochastic Modelling for Structure Reconstruction from High-resolution SAR Data**. *In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, volume 6, pages 4080– 4082, 2003.
- J. A. QUIROGA, E. BERNABEU et A. GONZÁLEZ-CANO : **Phase-Unwrapping Algorithm Based on an Adaptive Criterion**. *Applied Optics*, 34:2560–2563, 1995.
- D. RIVIÈRE : *Apprentissage de la variabilité inter-individuelle de l'anatomie corticale cérébrale pour la reconnaissance automatique des sillons*. Thèse de doctorat, Université d'Evry, 2000.
- E. RODRIGUEZ et J. M. MARTIN : **Theory and Design of Interferometric Synthetic Aperture Radars**. *In Proceedings of IEEE Radar and Signal Processing*, volume 139, pages 147–159, 1992.
- I. G. ROSENBERG : **Reduction of Bivalent Maximization to the Quadratic Case**. *Cahiers du Centre d'Etudes de Recherche Operationnelle*, 17:71–74, 1975.
- C. ROTHER, V. KOLMOGOROV et A. BLAKE : **GrabCut : Interactive Foreground Extraction Using Iterated Graph Cuts**. *ACM Transactions on Graphics*, 23(3):309–314, 2004.
- C. ROTHER, V. KOLMOGOROV, V. LEMPITSKY et M. SZUMMER : **Optimizing Binary MRFs via Extended Roof Duality**. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2007.
- C. ROTHER, S. KUMAR, V. KOLMOGOROV et A. BLAKE : **Digital Tapestry**. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 589–596, 2005.
- S. ROY : **Stereo Without Epipolar Lines : A Maximum-Flow Formulation**. *International Journal of Computer Vision*, 34(2):147–161, 1999.
-

-
- S. ROY et I. J. COX : **A Maximum-Flow Formulation of the N-Camera Stereo Correspondence Problem**. In *Proceedings of the IEEE International Conference on Computer Vision*, page 492, 1998.
- L. I. RUDIN, S. OSHER et E. FATEMI : **Nonlinear Total Variation Based Noise Removal Algorithms**. *Physica D : Nonlinear Phenomena*, 60(1-4):259–268, 1992.
- S. S. SAQUIB, C. A. BOUMAN et K. SAUER : **ML Parameter Estimation for Markov Random Fields with Applications to Bayesian Tomography**. *IEEE Transactions on Image Processing*, 7(7):1029–1044, 1998.
- K. SCHMITT et W. WIESBECK : **An Interferometric SAR Processor Avoiding Phase Ambiguities**. *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, 4, 1997.
- A. SCHUBERT : **Stereo-Assisted Interferometric SAR**. Thèse de doctorat, University of Zürich, Remote Sensing Laboratories, 2004.
- A. SCHUBERT, D. SMALL, F. HOLECZ, E. MEIER et D. NÜESCH : **Extraction of Surface Topography from SAR Stereo Pairs Using an Airborne X-band Sensor : Preliminary Results**. In *Retrieval of Bio- and Geo-Physical Parameters from SAR Data for Land Applications*, volume 475, pages 151–156, 2002.
- M. S. SEYMOUR et I. G. CUMMING : **Maximum Likelihood Estimation for SAR Interferometry**. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, volume 4, pages 2272–2274, 1994.
- A. SHABOU, J. DARBON et F. TUPIN : **A Markovian Approach for InSAR Phase Reconstruction with Mixed Discrete and Continuous Optimization**. *IEEE Geoscience and Remote Sensing Letters*, to appear, 2010a.
- A. SHABOU, G. FERRAIOLI, F. TUPIN et V. PASCAZIO : **Three Dimensional Reconstruction of Urban Areas Using Jointly Phase and Amplitude Multichannel Images**. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, volume to appear, 2010b.
- A. SHABOU, F. TUPIN et J. DARBON : **A Graph-cut Based Algorithm for Approximate MRF Optimization**. In *Proceedings of the IEEE International Conference on Image Processing*, 2009.
- E. SIMONETTO : **Extraction 3D de structures industrielles sur des images RAMSES haute résolution par radargrammétrie**. Thèse de doctorat, Université de Rennes 1, ONERA/ENST-Bretagne, 2002.
- U. SOERGEL, E. MICHAELSEN, A. THIELE et U. THOENNESSEN : **Radargrammetric Extraction of Building Features from High Resolution Multi-aspect SAR Data**. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, pages 3635–3638, 2006.
- U. SOERGEL, K. SCHULZ et U. THOENNESSEN : **Segmentation of Interferometric SAR Data for Building Detection**. In *International Archives of Photogrammetry and Remote Sensing*, volume 33, pages 328 – 335, 2000.
-

-
- U. STILLA, U. SOERGEL et U. THOENNESSEN : **Potential and Limits of InSAR Data for the Reconstruction of Buildings**. In *Proceeding of the IEEE/ISPRS Joint Workshop of Remote Sensing and Data Fusion over Urban Areas*, pages 64–68, 2001.
- R. STOICA : *Processus ponctuels pour l'extraction des réseaux linéiques dans les images satellitaires et aériennes*. Thèse de doctorat, Université de Nice-Sophia Antipolis, 2001.
- W. J. SZAJNOWSKI : **Estimators of Log-Normal Distribution Parameters**. *IEEE Transactions on Aerospace and Electronic Systems*, AES-13(5):533–536, 1977.
- R. SZELISKI, R. ZABIH, D. SCHARSTEIN, O. VEKSLER, V. KOLMOGOROV, A. AGARWALA, M. TAPPEN et C. ROTHER : **A comparative Study of Energy Minimization Methods for Markov Random Fields**. pages 16–29, 2006.
- A. THIELE, E. CADARIO, K. SCHULZ, U. THOENNESSEN et U. SOERGEL : **Building Recognition Fusing Multi-Aspect High-Resolution Interferometric SAR Data**. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, pages 3643–3646, 2006.
- C. TISON : *Interférométrie RSO à haute résolution en milieu urbain : Application au calcul de MNS urbain*. Thèse de doctorat, Télécom Paris Tech, 2004.
- D. P. TOWERS, T. R. JUDGE et P. J. BRYANSTON-CROSS : **A Quasi Heterodyne Holographic Technique and Automatic Algorithms for Phase Unwrapping**. *Society of Photo-optical and Instrumentation Engineers*, 1163:95–119, 1989.
- E. TROUVÉ : *Imagerie interférentielle en radar à ouverture synthétique*. Thèse de doctorat, Télécom Paris Tech, 1996.
- F. ULABY, R. MOORE et A. FUNG : *Microwave Remote Sensing : Active and Passive*. Artech House, 1986.
- G. VALADÃO et J. BIOUCAS-DIAS : **CAPE : Combinatorial Absolute Phase Estimation**. *Journal of the Optical Society of America A*, 26:2093–2106, 2009.
- G. VALADAO et J. M. BIOUCAS-DIAS : **CAPE : Combinatorial Absolute Phase Estimation**. Rapport technique, IT/IST Communications Theory and Pattern Recognition Group, Lisboa, Portugal, 2009.
- O. VEKSLER : *Efficient Graph-Based Energy Minimization Methods in Computer Vision*. Thèse de doctorat, Cornell University, 1999.
- O. VEKSLER : **Graph Cut Based Optimization for MRFs with Truncated Convex Priors**. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2007.
- O. VEKSLER : **Multi-label Moves for MRFs with Truncated Convex Priors**. In *Proceedings of International Conference on Energy Minimization Methods in Computer Vision and Pattern Recognition*, pages 1–13, 2009.
- V. VINEET et P. J. NARAYANAN : **CUDA cuts : Fast Graph Cuts on the GPU**. volume 0, pages 1–8, 2008.
-

-
- C. R. VOGEL et M. E. OMAN : **Iterative Methods for Total Variation Denoising.** *SIAM Journal on Scientific Computing*, 17(1):227–238, 1996.
- Y. WEISS et W. T. FREEMAN : **On the Optimality of Solutions of the Max-Product Belief-Propagation Algorithm in Arbitrary Graphs.** *IEEE Transactions on Information Theory*, 47(2):736–744, 2001.
- G. WINKLER : **Image Analysis, Random Fields and Markov Chain Monte Carlo Methods : A Mathematical Introduction (Stochastic Modelling and Applied Probability).** Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.
- W. XU, E. CHANG, L. KWOH, H. LIM et H WENG : **Phase-unwrapping of SAR Interferogram with Multi-frequency or Multi-baseline.** *In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, volume 2, pages 730–732, 1994.
- W. XU et I. CUMMING : **A Region Growing Algorithm for InSAR Phase Unwrapping.** *In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, volume 4, pages 2044–2046, 1996.
- W. C. YEH : **A Simple Algorithm for the Planar Multiway Cut Problem.** *Journal of Algorithms*, 39(1):68–77, 2001.
- L. YOUNES : **Estimation and Annealing for Gibbsian Fields.** *Annales de l’Institut Henri Poincaré, Section. B, Calcul des Probabilités et Statistique*, 24(2):269–294, 1988.
- B. ZALESKY : **Efficient Determination of Gibbs Estimators with Submodular Energy Functions.** Rapport technique, United Institute of Information Problems, Belarus, 2003a.
- B. ZALESKY : **Efficient Determination of Gibbs Estimators with Submodular Energy Functions.** *ArXiv Mathematics e-prints*, 2003b.
- H. A. ZEBKER et R. GOLDSTEIN : **Topographic Mapping from Interferometric Synthetic Aperture Radar Observations.** *Journal of Geophysical Research*, 91:4993–4999, 1986.
- H. A. ZEBKER, S. N. MADSEN, J. MARTIN, G. ALBERTI, S. VETRELLA et A. CUCCI : **The Topsar Interferometric Radar Topographic Mapping Instrument.** *IEEE Transactions on Geoscience and Remote Sensing*, 30(5):933–939, 1992.