



École Doctorale
d'Informatique,
Télécommunications
et Électronique de Paris

Thèse

présentée pour obtenir le grade de docteur

de Télécom ParisTech

Spécialité : Signal et Images

Hélène SPORTOUCHE

**Extraction et reconstruction
des bâtiments en milieu urbain
à partir d'images satellitaires
optiques et radar à haute résolution.**

Soutenue le 10 décembre 2010 devant le jury composé de :

M. Philippe Réfrégier (Professeur, ECM, Marseille)	Président
M. Jocelyn Chanussot (Professeur, INP, Grenoble)	Rapporteurs
M. Christian Germain (Professeur, ENITA, Bordeaux)	
M. Jean-Marie Nicolas (Professeur, Télécom ParisTech, Paris)	Examineurs
M. Jordi Inglada (Ingénieur, CNES, Toulouse)	
Mme Élisabeth Simonetto (Maître de conférences, ESGT, Le Mans)	
Mme Florence Tupin (Professeur, Télécom ParisTech, Paris)	Directrice de thèse
M. Léonard Denise (Ingénieur, Thales Communications, Massy)	Responsable industriel

Remerciements

Ces travaux de doctorat ont été réalisés dans le cadre d'une convention CIFRE, établie entre le département TSI de Télécom ParisTech (Paris) et le département IMINT de Thales Communications (Massy). Cette thèse m'a donné l'occasion de découvrir à la fois l'univers de la recherche en laboratoire et le monde de l'entreprise et m'a ainsi permis de vivre une double expérience, extrêmement enrichissante sur le plan professionnel comme personnel.

Je tiens ici à remercier toutes les personnes qui m'ont accompagnée, au cours de ces trois années, dans cette grande aventure, scientifique et humaine.

Je remercie tout particulièrement ma directrice de thèse, Mme Florence Tupin, Professeur à Télécom ParisTech, et mon responsable industriel, M. Léonard Denise, Ingénieur chez Thales Communications, qui m'ont tous deux encadrée pendant ces trois années.

J'ai pris un très grand plaisir à travailler avec eux et je les remercie très vivement pour la qualité de leur encadrement, leur investissement permanent, leurs précieux conseils, leur écoute, leur bienveillance et leur très grande disponibilité. Je leur exprime ici ma gratitude pour toute l'aide qu'ils m'ont apportée au cours des différentes étapes de cette thèse. Je les remercie également pour l'enseignement d'exception dont j'ai bénéficié, dans le domaine passionnant de la télédétection et de l'imagerie optique et RADAR, ainsi que pour toutes les conversations fructueuses et motivantes, que nous avons été amenés à avoir autour de tels sujets.

Je tiens aussi à remercier M. Jean-Marie Nicolas, Professeur à Télécom ParisTech, qui a initié ce projet de doctorat et qui a été présent lors de chacune des orientations clefs de ces travaux. Je le remercie vivement pour sa participation, sa gentillesse et nos multiples échanges, toujours très enrichissants.

Je souhaite également remercier l'ensemble des membres de mon jury de soutenance. Je remercie très sincèrement M. Jocelyn Chanussot, Professeur à l'INP Grenoble, ainsi que M. Christian Germain, Professeur à l'ENITA Bordeaux, qui ont accepté de rapporter cette thèse, et dont les remarques constructives et les avis éclairés ont beaucoup participé à la progression de ces travaux et à la finalisation du manuscrit.

Je remercie très vivement M. Philippe Réfrégier, Professeur à l'École Centrale Marseille, qui a présidé mon jury de soutenance, et je lui exprime toute ma reconnaissance pour les conseils avisés qu'il m'a prodigués, lors de mon stage de Master à l'Institut Fresnel, et qui m'ont encouragée à entreprendre une thèse à l'issue de ce stage.

Je remercie également les examinateurs de mon jury, M. Jordi Inglada, Ingénieur au CNES et Mme Élisabeth Simonetto, Maître de conférences à l'ESGT Le Mans, pour toute l'attention et l'intérêt qu'ils ont portés à ces travaux de recherche.

Je remercie par ailleurs Madame Sylvie Le Hégarat-Masclé, Professeur à Polytech Paris Sud, Madame Séverine Dubuisson, Maître de conférences au LIP6 / UPMC, ainsi que Madame Isabelle Bloch, Professeur à Télécom ParisTech, qui ont fait partie de mon jury

d'évaluation à mi-parcours, et dont les remarques judicieuses, lors de ma présentation de fin de deuxième année, m'ont beaucoup aidée.

Je voudrais à présent remercier les deux équipes d'ingénieurs et de chercheurs, qui m'ont chaleureusement accueillie durant ces trois années.

Côté entreprise, un grand merci à l'ensemble du département IMINT, au sein duquel j'ai passé de très bons moments, pleins de rires et de satisfactions et dont je garderai un excellent souvenir. En particulier, je remercie vivement Amandine et Édouard, qui ont suivi pas à pas et avec grand intérêt mes travaux, et dont l'écoute, la gentillesse, les encouragements et la bonne humeur me furent d'une aide précieuse au quotidien. Je remercie également François et Yannick, qui se sont beaucoup intéressés à mes recherches et qui, à de nombreuses occasions, ont mis à contribution, avec patience et générosité, leurs compétences informatiques pour le développement de programmes dans l'environnement "maison". Merci aussi à Élise, Gwen, Noémie et Marie-Stéphanie (qui nous ont réjouis chaque jour autour d'un univers rempli de pâtisseries et gourmandises), merci au dynamique trio Christian-Sébastien-Laurent (pour leurs discours sur le monde, enflammés et toujours très instructifs), merci à Véro (pour l'organisation de toutes mes missions ... jusqu'à Shanghai !) et merci à tous les autres !

Côté laboratoire, un grand merci à tous les membres, étudiants, permanents, et personnels du département TSI, avec lesquels j'ai été ravie de partager chacun de mes vendredis, dans une ambiance extrêmement sympathique et conviviale. En particulier, je remercie vivement Charles, Guillaume, Aymen, Franck, Carolina, Talita, Flora, Benoît, Isabelle, Marine, Patricia et Michel, pour leurs encouragements dans la dernière ligne droite, pour leurs conseils et suggestions lors de ma répétition de soutenance, ainsi que pour toute l'aide qu'ils m'ont apportée, pour l'organisation de mon pot de thèse. Je n'oublierai pas nos fameuses "réunions RADAR", pleines de découvertes quant aux diverses facettes de l'imagerie RADAR et auxquelles j'ai beaucoup apprécié de participer.

J'en profite aussi pour adresser mes remerciements à tous mes amis, qui m'ont encouragée pendant ces trois années, et, en particulier, à Julie et à Céline, pour leur gentillesse et nos discussions enthousiastes sur l'expérience de la thèse.

Enfin, je remercie avec grand plaisir et beaucoup de gratitude mes parents, mon frère et ma sœur, mon mari et toute ma famille, pour leur soutien constant et essentiel durant toutes mes années d'études, pour leur grande patience, leur vif intérêt pour mes travaux, leur confiance absolue et tout leur amour. Un grand merci pour m'avoir suivie avec beaucoup d'optimisme dans cette voie et pour avoir contribué à la réussite de cette thèse.

À mes grands-parents.

Résumé

Ces travaux s'inscrivent dans le cadre de l'extraction et de la reconstruction 3D des bâtiments en milieu urbain et semi-urbain, à partir d'images satellitaires optiques et RADAR à haute résolution.

L'objectif majeur réside dans le développement d'une chaîne complète de traitements semi-automatiques, capable de fournir une reconstruction simple et fiable des bâtiments de la scène, à partir d'une configuration spécifique des données d'entrée, composée d'une image optique et d'une image RADAR, ainsi que d'un modèle numérique de terrain (MNT).

Cette configuration restreinte, particulièrement délicate à traiter mais susceptible d'intervenir en milieu opérationnel, lorsque des couples stéréoscopiques, radargrammétriques ou interférométriques sont indisponibles ou inexploitable, a fait l'objet de peu de travaux jusqu'ici recensés dans la littérature. La proposition d'une démarche dédiée à la gestion d'un tel scénario représente donc un intérêt certain pour plusieurs applications de télédétection (cartographie du paysage urbain, détection de changements, simulation de scènes).

L'enjeu consiste à bénéficier pleinement du contexte de la fusion de données, en exploitant, de façon appropriée, les complémentarités optiques et RADAR, en vue d'une reconstruction de la scène par combinaison d'informations planimétriques et altimétriques.

Notre cadre de travail se limite à celui de la reconstruction de bâtiments simples de type parallélépipédique.

La chaîne proposée se décompose en quatre étapes principales, qui correspondent aux quatre contributions majeures de nos travaux : premièrement, une nouvelle méthode, basée sur l'adaptation d'outils morphologiques, géométriques et radiométriques à la problématique considérée, est proposée pour l'étape de détection des bâtiments potentiels en monoscopie optique panchromatique ; deuxièmement, une méthode originale, dédiée à la superposition fine de primitives homologues caractéristiques du bâtiment, est présentée pour l'étape de projection-recalage des emprises optiques potentielles dans la donnée RADAR ; troisièmement, une approche innovante, combinant deux critères RADAR complémentaires, est développée pour l'étape jointe d'estimation des hauteurs des bâtiments et de validation de leur présence sur l'image RADAR ; et, dernièrement, une méthode, fondée sur l'introduction de scores de qualité, est suggérée pour l'étape de qualification des bâtis reconstruits.

La chaîne complète est mise en œuvre sur des scènes d'étude issues d'un couple de données réelles Quickbird / TerraSAR-X.

Les résultats fournis sont analysés qualitativement et quantitativement (évaluation des performances en terme de détection et de fausse alarme après l'étape de détection des bâtis potentiels et après l'étape de validation, analyse de la précision planimétrique des contours extraits et de la précision altimétrique des hauteurs estimées).

À l'issue de la séquence de procédés proposée, nous obtenons une reconstruction globalement satisfaisante des bâtiments composant la scène.

Mots clés

Extraction du bâtiment, reconstruction du bâtiment, chaîne de traitements semi-automatiques, fusion de données optiques-RADAR, milieu urbain et semi-urbain, haute résolution spatiale, imagerie satellitaire, Quickbird, TerraSAR-X, télédétection.

Abstract

These works take place in the framework of building extraction and 3D building reconstruction in urban and semi-urban areas, from high-resolution optical and SAR satellite images.

The main objective is to develop a complete semi-automatic processing chain, able to provide a simple and reliable reconstruction of buildings on the scene, from a specific configuration of the input data, composed of an optical image and a SAR image, and from a digital terrain model (DTM).

This restricted configuration, particularly delicate to deal with but likely to happen in operational conditions, when stereoscopic, radargrammetric or interferometric couples are not available or unworkable, has been until now studied only in a few works of the literature. Proposing a specific approach to manage such a scenario seems thus very interesting for a lot of applications in remote sensing (cartography of urban areas, change detection, scene simulation).

We aim to fully benefit from the data fusion context, by exploiting, in an appropriate way, the optical-SAR complementarities for the reconstruction of the scene, through the combination of planimetric and altimetric information.

Our research work is limited to the reconstruction of simple parallelepipedic buildings.

The proposed chain is decomposed into four main steps, corresponding to the four principal contributions of our work : first, a new method, based on the adaptation of morphological, geometric and radiometric tools to the current problematic, is proposed for the step of potential building detection in optical panchromatic monoscopy ; secondly, an original method, specifically developed for a fine superposition between homologous characteristic building features, is presented for the step of projection and registration of the potential optical footprints into SAR data ; thirdly, an innovating approach, using the combination of two complementary SAR criteria, is developed for the joint step of building height estimation and validation on the SAR image ; and, lastly, a method, based on the introduction of quality scores, is suggested for the step of qualification of the reconstructed buildings.

The whole chain is applied on studied scenes issued from a couple of Quickbird / TerraSAR-X real data.

The obtained results are qualitatively and quantitatively analyzed (evaluation of the performances in terms of detection and false alarm after the step of potential building detection and after the step of validation, analysis about the planimetric precision of the

extracted outlines and the altimetric precision of the retrieved heights).

At the end of the proposed sequence of processings, a globally satisfying reconstruction of the buildings present on the scene is achieved.

Keywords

Building extraction, building reconstruction, semi-automatic processing chain, optical-SAR data fusion, urban and semi-urban areas, high spatial resolution, satellite imagery, Quickbird, TerraSAR-X, remote sensing.

Table des matières

Résumé	7
Abstract	9
Table des matières	15
Table des figures	18
Liste des tableaux	19
Introduction	21
1 Imagerie satellitaire optique et RADAR à haute résolution	25
1.1 Imagerie optique	25
1.1.1 Généralités	25
1.1.2 Système d'acquisition	25
1.1.3 Caractéristiques radiométriques	26
1.1.4 Caractéristiques spatiales	26
1.1.5 Spécificités géométriques	27
1.1.6 Propriétés statistiques et filtrage des images	27
1.2 Imagerie RADAR à synthèse d'ouverture (RSO)	28
1.2.1 Généralités	28
1.2.2 Formation des images	28
1.2.3 Caractéristiques radiométriques	32
1.2.4 Caractéristiques spatiales	32
1.2.5 Spécificités géométriques	34
1.2.6 Propriétés statistiques et filtrage des images	35
1.3 Nouveaux capteurs satellitaires optiques et RADAR à haute résolution	36
1.4 Comparatif imagerie optique - imagerie RADAR : avantages et inconvénients	37
2 État de l'art en reconstruction 3D du bâtiment à partir d'images optiques et RADAR à haute résolution	41
2.1 Généralités	41
2.2 Reconstruction des bâtiments en imagerie optique	42
2.2.1 Détection des bâtiments et extraction de leurs contours	42
2.2.2 Estimation des hauteurs de bâtiments	45
2.2.3 Génération de MNS ou de MNE et reconstruction 3D des bâtiments	45
2.2.4 Conclusions : Performances et limites	47
2.3 Reconstruction des bâtiments en imagerie RADAR	48

2.3.1	Extraction de primitives caractéristiques et génération d'empreises de bâtiments	48
2.3.2	Estimation des hauteurs de bâtiments	52
2.3.3	Génération de MNT en milieu urbain	57
2.3.4	Génération de MNS ou de MNE et reconstruction 3D des bâtiments	57
2.3.5	Conclusions : Performances et limites	59
2.4	Reconstruction des bâtiments par fusion d'informations optiques-RADAR	60
2.4.1	Recalage de données optiques et RADAR à résolution métrique	60
2.4.2	Reconstruction 3D à partir d'une image optique et d'une image RADAR	62
2.4.2.1	Détection de bâtiments et extraction de leurs contours	63
2.4.2.2	Estimation des hauteurs de bâtiments	64
2.4.2.3	Génération de MNS ou de MNE et reconstruction 3D des bâtiments	65
2.4.3	Reconstruction 3D à partir d'une image optique et d'un couple RADAR	65
2.4.3.1	Détection de bâtiments et extraction de leurs contours	65
2.4.3.2	Estimation des hauteurs de bâtiments	66
2.4.3.3	Génération de MNS ou de MNE et reconstruction 3D des bâtiments	66
2.4.4	Conclusions : Performances et limites	66
2.5	Applications de la reconstruction 3D des bâtiments	67
3	Présentation de la chaîne de traitements proposée	69
3.1	Description de la chaîne complète	69
3.1.1	Objectifs, fonctionnalités et contraintes	69
3.1.2	Choix stratégiques	70
3.1.3	Organisation de la chaîne proposée	70
3.2	Hypothèses de travail	72
3.2.1	Modélisation des bâtiments	72
3.2.2	Conditions d'acquisition	73
3.2.3	Conséquences sur la modélisation des signatures de bâtiments	74
3.3	Présentation des données	77
3.3.1	Jeu de données complet	77
3.3.2	Scènes d'intérêt	82
3.3.3	Vérité terrain	82
4	Extraction des bâtiments en monoscopie optique	87
4.1	Méthode proposée pour la détection de bâtiments	87
4.1.1	Stratégie globale	87
4.1.2	Phase 1 : détection des fenêtres d'intérêt	88
4.1.2.1	Définition des profils morphologiques différentiels	88
4.1.2.2	Présentation de l'algorithme	89
4.1.3	Phase 2 : extraction fine des empreises potentielles	90
4.1.3.1	Présentation de l'algorithme	90
4.1.4	Discussion sur la gestion des paramètres	94
4.2	Application sur les scènes d'études	94
4.2.1	Paramètres utilisés	94

4.2.2	Résultats obtenus	95
4.2.3	Analyse qualitative et quantitative	95
4.3	Étude détaillée de quelques bâtiments d'intérêt	97
4.4	Application sur des scènes complémentaires	99
4.5	Conclusions et améliorations	102
5	Recalage au sol de primitives caractéristiques optiques et RADAR	103
5.1	Notations spécifiques	103
5.2	Rappels : fonctions de localisation directes et inverses	104
5.2.1	Système d'équations optique	104
5.2.2	Système d'équations RADAR	104
5.2.3	Projection depuis une image optique vers une image RADAR	106
5.3	Méthode proposée pour le recalage	106
5.3.1	Stratégie globale	106
5.3.2	Phase 1 : projection des primitives optiques	106
5.3.2.1	Modèles de projection disponibles pour nos données	106
5.3.2.2	Affinage optionnel des paramètres du "modèle joint de projection"	107
5.3.2.3	Mise en œuvre de la projection	107
5.3.3	Phase 2 : affinage du recalage sur l'image RADAR	108
5.3.3.1	Motivations	108
5.3.3.2	Hypothèses et principe	108
5.3.3.3	Optimisation des déplacements	108
5.3.3.4	Régularisation des déplacements	110
5.3.4	Discussion sur la gestion des paramètres	114
5.4	Application sur les scènes d'études	114
5.4.1	Paramètres utilisés	114
5.4.2	Résultats obtenus	115
5.4.3	Analyse qualitative et quantitative	115
5.5	Étude détaillée de quelques bâtiments d'intérêt	119
5.6	Conclusions et améliorations	121
6	Estimation des hauteurs et validation des bâtiments	125
6.1	Méthode proposée pour l'estimation des hauteurs et pour la validation des bâtiments sur l'image RADAR	125
6.1.1	Stratégie globale	125
6.1.2	Génération de partitionnements tests sur une imagerie RADAR	126
6.1.2.1	Rappels : modélisation des signatures RADAR de bâtiments	126
6.1.2.2	Technique utilisée pour la génération de partitionnements tests	126
6.1.3	Phase 1 : utilisation d'un critère RADAR global statistique pour la génération de hauteurs candidates	128
6.1.3.1	Principe	128
6.1.3.2	Modélisation de la distribution d'intensité dans chaque région caractéristique	128
6.1.3.3	Expression analytique de la log-vraisemblance individuelle de chaque région caractéristique	131

6.1.3.4	Estimation des paramètres locaux par maximum de vraisemblance	131
6.1.3.5	Formulation analytique du critère statistique global	132
6.1.3.6	Optimisation du critère statistique pour la génération de hauteurs candidates	133
6.1.3.7	Comparaison avec l'approche développée par Brunner et al.	134
6.1.4	Phase 2 : utilisation d'un critère RADAR de contraste local pour l'estimation des hauteurs finales et pour la validation des bâtiments	136
6.1.4.1	Principe	136
6.1.4.2	Formulation du critère de contraste local	136
6.1.4.3	Utilisation du critère de contraste pour l'estimation des hauteurs finales et pour la validation des bâtiments	138
6.1.5	Justifications sur la complémentarité des deux critères	139
6.1.6	Commentaires sur la propagation des erreurs	139
6.1.7	Discussion sur la gestion des paramètres	139
6.1.8	Validation de la méthode précédente sur des données RADAR simulées	141
6.2	Reconstruction finale des bâtiments sur la scène	142
6.3	Application sur les scènes d'études	144
6.3.1	Paramètres utilisés	144
6.3.2	Résultats obtenus	144
6.3.3	Analyse qualitative et quantitative	144
6.4	Étude détaillée de quelques bâtiments d'intérêt	148
6.5	Conclusions et améliorations	152
7	Qualification de la reconstruction des bâtiments	157
7.1	Méthode proposée pour la qualification des bâtiments reconstruits	157
7.1.1	Introduction de scores de qualité à partir de l'image RADAR	157
7.1.2	Intérêt opérationnel	158
7.2	Application sur les scènes d'études	159
7.2.1	Résultats obtenus	159
7.2.2	Analyse qualitative et quantitative	159
7.3	Conclusions et améliorations	162
8	Extensions	163
8.1	Perspectives de recherche	163
8.1.1	Complexification du modèle de bâtiment	163
8.1.2	Symétrisation de la chaîne de traitements	164
8.1.2.1	Pour l'étape 1 : détection de fenêtres d'intérêt supplémentaires et introduction de valeurs d'orientation précises à partir de l'image RADAR	165
8.1.2.2	Pour l'étape 3 : estimation jointe des hauteurs des bâtiments et validation croisée	168
8.1.2.3	Pour l'étape 4 : qualification jointe des bâtiments reconstruits	170
8.1.2.4	Conclusions et perspectives	170

8.1.3	Proposition d'une variante méthodologique pour la détection des bâtiments dans le cadre de la fusion optique-RADAR	171
8.1.3.1	Présentation	171
8.1.3.2	Stratégie	171
8.1.3.3	Exemple de résultat d'application	173
8.1.3.4	Analyse et conclusions	173
8.1.4	Gestion d'une nouvelle configuration d'entrée : cas d'une image optique et d'un couple RADAR interférométrique	173
8.2	Capitalisation industrielle	175
8.2.1	Projet de développement d'un démonstrateur 3D	175
8.2.1.1	Descriptif	175
8.2.1.2	Utilisation de modules indépendants	176
8.2.1.3	Utilisation de la chaîne de traitements complète	177
8.2.2	Scénarios types d'utilisation de la chaîne de traitements	177
	Conclusion générale	179
	Liste des travaux	183
	Bibliographie	196

Table des figures

1.1	Bandes de fréquence utilisées en imagerie RADAR.	29
1.2	Géométrie du système d'acquisition RADAR.	30
1.3	Modes d'acquisition utilisés en imagerie RADAR satellitaire.	31
1.4	Géométrie image et géométrie terrain.	31
1.5	Déformations géométriques observées en imagerie RADAR.	35
2.1	Estimation de la hauteur des bâtis à partir de la "zone de recouvrement".	53
2.2	Estimation de la hauteur des bâtis à partir de la "zone d'ombre".	54
3.1	Schéma synoptique de la chaîne complète proposée.	71
3.2	Modèle 3D simple de bâtiment adopté pour la reconstruction, représenté dans le référentiel terrestre $(O, \vec{x}, \vec{y}, \vec{z})$	73
3.3	Modélisation de la signature optique d'un bâtiment adoptée pour la reconstruction, représentée dans le référentiel image optique.	74
3.4	Modélisation de la signature RADAR d'un bâtiment adoptée pour la reconstruction, représentée dans le référentiel image RADAR.	75
3.5	Image Quickbird disponible sur Marseille (France)	78
3.6	Image TerraSAR-X disponible sur Marseille (France)	79
3.7	Recouvrement entre l'image optique et l'image RADAR.	80
3.8	Modèle numérique de terrain (MNT) disponible sur Marseille (France).	81
3.9	Scènes d'étude A et B.	83
3.10	Photos vérité terrain de quelques bâtiments composant la scène d'étude B.	84
3.11	Cartes de la vérité terrain pour les scènes d'étude A et B.	85
4.1	Exemple partiel de profil morphologique différentiel associé à un bâtiment.	89
4.2	Schéma synoptique de la phase 1 de l'étape 1.	91
4.3	Schéma synoptique de la phase 2 de l'étape 1.	93
4.4	Carte de localisation des contours optiques extraits pour les scènes d'étude A et B.	96
4.5	Exemples de fenêtres d'intérêt associées à des bonnes détections sur la scène d'étude A.	99
4.6	Exemples de bonnes détections sur les scènes d'étude A et B.	100
4.7	Exemples de fausses alarmes sur les scènes d'étude A et B.	100
4.8	Exemples de non détections sur les scènes d'étude A et B.	100
4.9	Exemples de détections sur la zone de Toulouse (capteur satellitaire).	101
4.10	Exemples de détections sur la zone de Saint-Mandrier (capteur aéroporté).	101
4.11	Exemples de détections sur la zone de Marignane (capteur satellitaire).	101

5.1	Schéma synoptique de la phase 2 de l'étape 2 (sous-phase 1).	111
5.2	Schéma synoptique de la phase 2 de l'étape 2 (sous-phase 2).	113
5.3	Carte de localisation des contours RADAR projetés pour les scènes d'étude A et B (avec modèle initial).	116
5.4	Carte de localisation des contours RADAR projetés pour les scènes d'étude A et B (avec modèle affiné).	117
5.5	Carte de localisation des contours RADAR recalés pour les scènes d'étude A et B.	118
5.6	Exemples de bâtiments recalés après optimisation sans régularisation sur les scènes d'étude A et B.	121
5.7	Exemples de bâtiments recalés après optimisation avec régularisation sur les scènes d'étude A et B.	122
6.1	Génération des zones caractéristiques de la signature du bâtiment. [Partie 1/2]	129
6.2	Génération des zones caractéristiques de la signature du bâtiment. [Partie 2/2]	130
6.3	Schéma synoptique de la phase 1 de l'étape 3.	135
6.4	Positionnement du contraste au niveau de la discontinuité entre la "zone de fond" et la "zone de recouvrement".	137
6.5	Schéma synoptique de la phase 2 de l'étape 3.	140
6.6	Résultats obtenus pour les deux bâtiments testés sur données RADAR simulées.	143
6.7	Cartes des hauteurs estimées des bâtiments validés pour les scènes d'étude A et B.	145
6.8	Résultats de la reconstruction tridimensionnelle des bâtiments sur les scènes d'études A et B.	146
6.9	Résultats obtenus pour deux bâtiments d'intérêt sur données RADAR réelles.	153
6.10	Résultats du partitionnement optimal en régions caractéristiques, obtenu pour le bâtiment 7 sur la scène A.	154
6.11	Résultats du partitionnement optimal en régions caractéristiques, obtenu pour le bâtiment 2 sur la scène B.	155
7.1	Cartes de confiance des bâtiments reconstruits pour les scènes d'étude A et B.	160
8.1	Schéma synoptique de la chaîne symétrique proposée.	166
8.2	Carte de localisation des fenêtres d'intérêt optiques-RADAR détectées pour la scène d'étude B.	167
8.3	Cartes de localisation des contours optiques pour la scène d'étude sur Toulouse (France).	174

Liste des tableaux

1.1	Quelques nouveaux capteurs satellitaires optiques.	38
1.2	Quelques nouveaux capteurs satellitaires RADAR.	39
3.1	Caractéristiques du jeu de données optique-RADAR.	77
3.2	Hauteurs réelles des bâtiments de la vérité terrain pour la scène d'étude A.	86
3.3	Hauteurs réelles des bâtiments de la vérité terrain pour la scène d'étude B.	86
4.1	Paramètres utilisés pour l'étape d'extraction des emprises potentielles.	95
4.2	Résultats individuels de détection des bâtiments sur la scène d'étude A.	98
4.3	Résultats individuels de détection des bâtiments sur la scène d'étude B.	98
5.1	Paramètres utilisés pour l'étape de projection et de recalage des emprises potentielles.	115
5.2	Résultats individuels de recalage des bâtiments sur la scène d'étude A.	120
5.3	Résultats individuels de recalage des bâtiments sur la scène d'étude B.	120
6.1	Paramètres utilisés pour la simulation de bâtiments RADAR.	144
6.2	Paramètres utilisés pour l'étape d'estimation des hauteurs et de validation.	147
6.3	Résultats individuels d'estimation des hauteurs de bâtiments sur la scène d'étude A.	149
6.4	Résultats individuels d'estimation des hauteurs de bâtiments sur la scène d'étude B.	149
6.5	Résultats individuels de validation des bâtiments sur la scène d'étude A.	150
6.6	Résultats individuels de validation des bâtiments sur la scène d'étude B.	150
6.7	Différences altimétriques entre les hauteurs réelles vérité terrain et les hauteurs estimées pour la scène d'étude A.	151
6.8	Différences altimétriques entre les hauteurs réelles vérité terrain et les hauteurs estimées pour la scène d'étude B.	151
7.1	Résultats individuels de qualification des bâtiments reconstruits sur la scène d'étude A.	161
7.2	Résultats individuels de qualification des bâtiments reconstruits sur la scène d'étude B.	161
8.1	Temps de calcul relatif à l'application de la chaîne de reconstruction sur un extrait optique (de taille 256 par 256 pixels).	176

Introduction

Contexte et motivations

La dernière génération de capteurs satellitaires optiques et RADAR à haute résolution (TerraSAR-X, Quickbird, CosmoSkyMed ...) délivre un ensemble de données métriques et submétriques extrêmement riches, à l'origine de nouveaux défis en télédétection.

Avec l'avènement de telles images, le contenu des données à exploiter s'est réellement densifié au cours des dix dernières années. Le passage de la moyenne résolution à la haute résolution s'est accompagné d'une re-définition des enjeux dans les applications civiles et militaires. Au sein d'une image à résolution métrique, la notion d'objet est devenue accessible. Ce changement d'échelle spatiale a nécessité l'adaptation des algorithmes de traitement d'images à de nouvelles problématiques dans le domaine de la reconnaissance et de la reconstruction d'objets en télédétection spatiale et aéroportée (détection de bâtiments [FBB09] ou d'arbres [WZGZ04] en zone urbaine, délimitation de parcelles en zone agricole [DCMG⁺07], ...).

Ceci est particulièrement intéressant en milieu urbain et semi-urbain, où l'analyse d'images optiques et RADAR à haute résolution nous permet aujourd'hui d'identifier individuellement de multiples éléments composant la scène d'intérêt (routes, ponts, maisons) [WS08, STGT07, SD03, Amb05, CBF06] et en particulier, les bâtiments. Le domaine de l'extraction et de la reconstruction tridimensionnelle (3D) des bâtiments occupe naturellement une place essentielle et offre une palette d'applications extrêmement variées (cartographie du paysage urbain et semi-urbain, détection de changements, simulation de scènes).

Plusieurs techniques classiques de restitution 3D ont été proposées dans la littérature pour la reconstruction des bâtiments, en imagerie satellitaire et aéroportée à résolution métrique et submétrique.

En imagerie optique, l'exploitation de couples stéréoscopiques ou multiscopiques a permis d'obtenir des résultats très satisfaisants, montrant une reconstruction 3D précise et détaillée des structures urbaines [CJPDS03, BM99].

En imagerie RADAR, l'analyse des signatures du bâti sur des données à haute résolution a fait l'objet de nombreux travaux [BBFL09, FIRR03, HBC⁺08]. Plusieurs méthodes de reconnaissance des bâtiments et d'estimation des hauteurs ont été développées dans le cadre de la radargrammétrie [SOG05], de l'interférométrie RADAR (InSAR) [GHS00, BFP09, TTM07] ou encore de la polarimétrie RADAR (PolSAR, PolInSAR) [DJHS08, GFFRP05]. En outre, l'utilisation de données RADAR multi-aspects s'est révélée tout à fait pertinente pour l'amélioration des résultats de reconstruction [XJ07, Soe06].

Toutefois, un certain nombre de difficultés peuvent être soulignées. Premièrement,

les couples de données optiques ou RADAR, utilisés dans la plupart des techniques précitées, ne sont pas toujours disponibles ou exploitables en conditions opérationnelles (exemple courant du couvert nuageux en imagerie optique). Deuxièmement, les techniques basées sur l'interférométrie RADAR ou la radargrammétrie, fournissent souvent des résultats encore limités, à l'origine de scènes reconstruites bruitées ou incomplètes. Ceci s'explique notamment par les difficultés propres à l'imagerie RADAR en milieu urbain : fort bruit de speckle, déformations géométriques importantes (zones de recouvrement), présence d'ombres et de fortes discontinuités rendant délicates certaines opérations interférométriques, ou encore présence de structures urbaines en hauteur sans réponse vers le capteur RADAR empêchant une reconstruction par radargrammétrie. Troisièmement, dans le cas des approches par monoscopie, les méthodes proposées nécessitent souvent l'intervention d'un opérateur ou l'introduction de données auxiliaires supplémentaires [BLBG10].

Dans un tel contexte, plusieurs pistes, bénéficiant de la complémentarité des données optiques et RADAR ont été récemment explorées dans le cadre de la fusion de données pour la détection et la reconstruction 3D des bâtiments [TR05, PIS⁺09, WZS10]. Lorsqu'un jeu composé d'un couple RADAR et d'une image optique est disponible, la situation s'est révélée particulièrement bien adaptée à l'extraction de l'information d'élévation à partir des données RADAR et à la localisation planimétrique des bâtiments à partir de l'optique. La configuration d'entrée, plus délicate, composée d'une seule image optique et d'une seule image RADAR, a fait l'objet de peu de travaux jusqu'à ce jour.

Objectifs et contributions

L'objectif de ces travaux est de développer une chaîne complète de traitements semi-automatiques, capable de fournir une reconstruction simple des bâtiments présents sur une scène urbaine ou semi-urbaine, à partir d'un couple de données satellitaires à haute résolution, composé d'une image optique à résolution submétrique et d'une image RADAR à résolution métrique, ainsi que d'un modèle numérique de terrain (MNT).

Cette configuration spécifique des données d'entrée correspond à un scénario susceptible d'intervenir en conditions opérationnelles, pour lequel aucune des techniques classiques de reconstruction 3D ne peut être mise en œuvre (stéréoscopie optique, radargrammétrie, interférométrie RADAR). La proposition d'une démarche dédiée à la gestion d'une telle configuration revêt donc un intérêt majeur pour certaines applications de télédétection.

Il s'agit d'exploiter, dans le contexte de la fusion de données, l'ensemble des informations planimétriques et altimétriques issues des deux types de capteurs complémentaires optiques et RADAR et de les intégrer au sein d'une séquence de procédés semi-automatiques, afin d'obtenir une reconstruction complète, fiable et relativement rapide de la scène d'intérêt.

Notre cadre de travail se limite à celui de la reconstruction de bâtiments simples de type parallélépipédique (emprise rectangulaire, façades verticales et toit plat).

La chaîne de traitements proposée dans ce manuscrit se décompose en quatre étapes successives, respectivement appelées "détection", "projection-recalage", "estimation des hauteurs-validation" et "qualification".

Sur le plan méthodologique, chacune de ces étapes se rapporte à l'une des contributions principales de nos travaux de thèse :

1. La proposition d'une nouvelle méthode de détection des bâtiments potentiels en monoscopie optique panchromatique, lors de la première étape de la chaîne. Le procédé mis en jeu se divise en deux phases : une phase de génération de "fenêtres d'intérêt" est d'abord réalisée à travers une approche "orientée régions", basée sur l'utilisation d'outils morphologiques et géométriques, existants dans la littérature et adaptés à notre problématique spécifique de détection du bâtiment ; une phase d'extraction précise des emprises de bâtiments potentiels est ensuite appliquée à travers une approche "orientée contours", fondée sur l'optimisation d'un critère de discontinuité radiométrique.
2. La présentation d'une méthode originale de projection-recalage des emprises optiques potentielles dans la donnée RADAR, lors de la seconde étape de la chaîne. Le procédé impliqué utilise la connaissance des paramètres capteurs pour la phase de projection et se fonde sur l'optimisation d'un score de recalage suivie d'une régularisation pour la phase d'affinage du recalage (intégration de contraintes spatiales, radiométriques et contextuelles). Il s'agit d'une méthode dédiée à l'appariement de primitives homologues caractéristiques du bâtiment, qui vise une superposition des primitives à l'échelle du mètre (la finesse de recalage souhaitée est supérieure à celle obtenue par les méthodes standard de recalage utilisées en fusion de données).
3. Le développement d'une approche innovante d'estimation des hauteurs et de validation jointe de leur présence, à partir des emprises potentielles recalées et de l'image RADAR, lors de la troisième étape de la chaîne. Le procédé mis en jeu se base sur la combinaison de deux nouveaux critères RADAR complémentaires, dépendants de la hauteur du bâti et définis pour un partitionnement de l'imagette RADAR de bâtiment en régions caractéristiques. Ces critères se réfèrent à une mesure statistique globale de log-vraisemblance et à une mesure locale de contraste.
4. La proposition d'une méthode de qualification de la reconstruction, lors de la quatrième étape de la chaîne. Les scores de qualité introduits témoignent de la fiabilité à accorder à chaque bâtiment reconstruit et sont susceptibles de constituer des indicateurs de confiance en conditions opérationnelles.

Sur le plan applicatif, nos principales contributions sont relatives aux points suivants :

1. La mise en œuvre de la chaîne complète proposée sur un jeu de données réelles satellitaires Quickbird / TerraSAR-X.
2. L'analyse qualitative de la pertinence des résultats fournis sur des scènes d'étude sélectionnées et l'évaluation quantitative des performances de reconstruction.
3. La proposition de démarches explicites à adopter en conditions opérationnelles pour la gestion de problèmes pratiques (exemples : réglage semi-automatique des paramètres des algorithmes, sélection d'un seuil d'intervention ponctuelle d'un opérateur).

Organisation de la thèse

Le présent manuscrit s'articule autour de huit chapitres, qui peuvent être regroupés de la façon suivante :

- les chapitres bibliographiques 1 et 2 dressent un état de l'art relatif à l'imagerie satellitaire optique et RADAR à haute résolution et aux techniques de reconstruction 3D du bâtiment développées dans la littérature ;
 - le chapitre 3 expose globalement la chaîne de traitements proposée (organisation synoptique, fonctionnalités et contraintes, choix stratégiques et hypothèses de travail) et présente le jeu de données tests disponibles ;
 - les chapitres 4 à 7 décrivent successivement les procédés méthodologiques relatifs à chacune des quatre étapes principales de la chaîne ; au sein de chaque chapitre, un schéma similaire est suivi : dans un premier temps, les stratégies employées sont justifiées et les algorithmes utilisés sont détaillés ; dans un second temps, les résultats obtenus sont illustrés sur des scènes d'études et les performances atteintes sont analysées ;
 - le chapitre 8 envisage enfin quelques extensions de nos travaux suivant deux axes : un axe orienté "recherche" visant à généraliser et à renforcer la chaîne de traitements présentée et un axe orienté "industrie" visant à capitaliser les outils développés dans le cadre de notre thèse CIFRE.
-

Chapitre 1

Imagerie satellitaire optique et RADAR à haute résolution

1.1 Imagerie optique

Dans cette section, nous rappelons les bases fondamentales de l'imagerie optique. Les concepts exposés relèvent de plusieurs domaines (optique, traitement du signal, statistiques) et s'appuient sur les ouvrages et références suivants : [KE01, MBM01, RJ98, Réf04, Wal07].

1.1.1 Généralités

Le principe du capteur optique est le suivant : la lumière (photons), réfléchiée par une cible présente à la surface de la Terre et éclairée par le soleil, est captée par des instruments optiques très sensibles puis convertie en signal électrique au travers de détecteurs micro-électroniques photo-sensibles appelés dispositifs à transfert de charge (DTC). Ces détecteurs DTC, disposés suivant une matrice rectangulaire ou le long d'une barette linéaire, convertissent l'intensité lumineuse reçue en un signal électrique, qui est ensuite numérisé afin d'obtenir une image. Le capteur optique est ainsi dit passif.

Les canaux situés sur les satellites optiques recouvrent généralement deux gammes de longueur d'ondes :

- le spectre du visible avec des longueurs d'ondes variant entre 380 et 780 nm ;
- le domaine de l'infra-rouge avec des longueurs d'ondes variant entre 780 nm et 12.7 μm ; on note en particulier l'usage des bandes fréquentielles suivantes : entre 780 et 890 nm pour le proche infra-rouge (NIR, near infra-red), entre 1.58 et 1.75 μm pour l'infra-rouge à ondes courtes (SWIR, short wave infra red) et entre 10.2 et 12.7 μm pour l'infra-rouge thermique (thermal infra red).

1.1.2 Système d'acquisition

La charge utile des systèmes imageurs optiques embarqués sur les satellites est notamment composée d'un dispositif optique et d'un radiomètre constitué de détecteurs DTC mesurant la quantité de lumière issue de la zone imagée.

Deux types de géométrie d'acquisition sont utilisés sur les plateformes satellitaires optiques : les acquisitions à géométrie linéaire (en configuration "line-pushbroom" ou

"line-panoramic", en anglais) et les acquisitions à géométrie surfacique (en configuration "area-frame", en anglais).

Avec une géométrie d'acquisition de type linéaire, les images sont formées ligne par ligne.

L'acquisition de l'image dans une dimension ("along-track") est assurée par le défilement du satellite le long de sa trajectoire.

L'acquisition dans l'autre dimension ("across-track") peut s'effectuer de deux manières :

- en configuration "râteau" (ou "pushbroom") : un système de barettes DTC alignées, permettant d'acquérir simultanément une ligne de pixels de l'image sans mouvement de la part du dispositif optique, est alors adopté (cas des satellites SPOT) ;
- en configuration "panoramique" (ou "panoramic") : un dispositif optique à miroir tournant, permettant d'accéder à une large fauchée, est alors utilisé (cas des satellites LANDSAT).

Avec une géométrie d'acquisition de type surfacique, tous les points d'une large surface rectangulaire au sol sont simultanément imagés (cas du satellite KOSMOS).

Par ailleurs, on distingue deux modes d'acquisition :

- le mode panchromatique, où le capteur mesure la luminance des objets imagés dans une large gamme de fréquences et qui se caractérise par l'obtention de données en noir et blanc finement résolues spatialement ;
- le mode multi-bandes (multi-spectral) où le capteur mesure la luminance des objets imagés dans plusieurs bandes spectrales plus étroites ; par combinaison de celles-ci on obtient alors des images en couleur, généralement moins bien résolues spatialement.

Selon le mode et l'application choisis, le dispositif instrumental à utiliser varie. Par exemple, le satellite optique français Spot 5 lancé en 2002 possède deux nouveaux instruments : le HRVIR (haute résolution visible et infra-rouge) dédié à l'acquisition d'images multi-spectrales et le HRS (haute résolution stéréoscopique) dédié à l'acquisition stéréoscopique d'images panchromatiques.

1.1.3 Caractéristiques radiométriques

Les données optiques satellitaires sont classiquement codées sur 8 ou 16 bits, ce qui correspond respectivement à une gamme de 256 ou 65536 niveaux possibles. La résolution radiométrique d'un capteur se définit par sa capacité à différencier de faibles variations d'intensité reçue ("contrastes détectables") et dépend également de la durée d'éclairément. Une forte résolution radiométrique nécessite un important nombre de niveaux disponibles pour coder l'image (dynamique importante requise). Comme souligné dans [KE01], une importante dynamique de l'image est un *"atout essentiel pour les outils de corrélation automatique"* utilisés dans les applications de stéréoscopie optique, car cela permet de discerner de nombreux détails présentant des différences de radiométrie très faibles.

1.1.4 Caractéristiques spatiales

La résolution spatiale d'un capteur est liée au pouvoir séparateur de l'instrument optique utilisé. Celui-ci est donné par le critère de Rayleigh indiquant le plus petit angle solide que l'instrument est capable de résoudre, en fonction du diamètre de l'optique et de la longueur d'onde considérée. La cellule de résolution spatiale est définie par l'aire

de la surface visible, à une distance donnée, par un élément DTC du système imageur. Cette cellule est déterminée par le champ de vision instantané (CVI) de l'élément DTC : un CVI faible correspond à une résolution spatiale fine.

La fonction de transfert de modulation (FTM) d'un capteur optique, encore appelée fonction de transmission du contraste, mesure sa capacité à transférer un contraste depuis une cible "spécimen" jusqu'au plan image du système d'acquisition (ce contraste se dégradant lorsque la fréquence spatiale des objets imagés augmente). La FTM est définie comme le rapport qui lie l'amplitude d'un signal vu par un capteur à l'amplitude sinusoïdale réelle de la cible imagée, en fonction de la fréquence spatiale de celle-ci. La FTM permet de prédire et de modéliser la qualité des images produites par ce système lorsque celui-ci est soumis à des conditions variables.

1.1.5 Spécificités géométriques

L'acquisition d'une image optique peut être réalisée en visée verticale (on parle alors de prise de vue au nadir) ou en visée oblique.

En visée au nadir, les façades verticales des objets n'apparaissent pas sur l'image. La localisation planimétrique dans l'image des contours au sol d'un objet (par exemple, l'emprise d'un bâtiment) coïncide ainsi avec la localisation planimétrique des contours élevés de cet objet (par exemple, le toit du bâtiment). Le phénomène de parallaxe étant absent, il devient alors plus délicat d'extraire des informations de hauteurs par des mesures manuelles.

Une image acquise avec une visée oblique se caractérise par la présence de plusieurs parties "cachées" c'est-à-dire par la présence de plusieurs objets ou zones de la scène qui ne peuvent être vus par le capteur, étant donné l'angle de visée du système imageur et leur environnement (présence d'objets occultants). Notons qu'en visée oblique, la résolution spatiale n'est pas constante le long de la fauchée.

Le procédé d'ortho-rectification permet de transformer une image présentant une visée oblique en une image de la même scène acquise avec une visée au nadir et nécessite de connaître l'élévation des objets.

1.1.6 Propriétés statistiques et filtrage des images

Tout système de détection optique se caractérise par la présence de deux sources de bruit essentiellement :

- le bruit inhérent au capteur numérique utilisé, souvent modélisé par un bruit additif Gaussien ;
- le bruit de photons, dû à la nature corpusculaire de la lumière, suivant une loi de Poisson.

Plusieurs types de filtres ont été développés dans la littérature pour exploiter les images optiques parmi lesquels figurent les filtres moyenneurs, les filtres médians, les filtres passe-bas, les filtres gaussiens, les filtres à rehaussement de contours (Roberts, Sobell, Prewitt, Laplacien ...) ou encore les filtres morphologiques.

Notons cependant, que contrairement aux données RADAR (voir section 1.2), les images optiques présentent la plupart du temps un rapport initial signal à bruit suffisamment bon pour permettre une interprétation visuelle directe et intuitive de l'image par un photo-interprète.

1.2 Imagerie RADAR à synthèse d'ouverture (RSO)

Dans cette section, nous présentons les bases fondamentales de l'imagerie RADAR à synthèse d'ouverture (RSO). Les notions exposées relèvent de plusieurs domaines (électro-magnétisme, traitement du signal, statistiques) et s'inspirent des ouvrages suivants : [Tis04, Tup97, Cel06, Nic06, Mai01, MS08, Den09, EFM⁺09].

1.2.1 Généralités

Le principe du RADAR imageur est le suivant : une antenne émettrice envoie un train d'ondes électro-magnétiques, dont la rétrodiffusion par la scène est mesurée par une antenne réceptrice. Les ondes utilisées se caractérisent par leur longueur d'onde λ (appartenant aux domaines des micro-ondes) et leur polarisation. Leur propagation se caractérise par la direction de visée et par la largeur du faisceau lumineux (dépendant de l'ouverture d'antenne).

La plupart du temps, une même antenne est utilisée à l'émission et à la réception ; on parle alors de RADAR monostatique.

Le capteur RADAR correspond ainsi à un capteur actif, c'est-à-dire qu'il est lui-même sa propre source d'énergie. Il permet ainsi d'acquérir des images de jour comme de nuit. Les systèmes RADAR présentent l'avantage d'être faiblement touchés par les conditions météorologiques (nuages, pluie...). Les phénomènes perturbateurs proviennent essentiellement de la ionosphère aux basses fréquences et de la troposphère aux hautes fréquences.

Le signal RADAR présente la particularité d'être un signal complexe, dû à la nature cohérente des ondes électro-magnétiques utilisées ; il contient donc deux types d'informations : l'amplitude et la phase.

L'acronyme RADAR, "ratio detection and ranging", se réfère à la capacité du système à déterminer la distance séparant le capteur de la cible imagée à partir de la mesure du temps entre émission et réception.

Dans [Den09], il est fait référence aux trois quantités mesurées directement par un RADAR :

- une énergie (celle rétrodiffusée par une cible) ;
- une distance (celle liée au temps d'aller-retour capteur-cible) ;
- une fréquence (celle liée à l'effet Doppler, induit par le mouvement du porteur lors de l'acquisition).

Les longueurs d'ondes utilisées en RADAR couvrent diverses bandes fréquentielles, présentées en figure 1.1 et s'étalant entre 0,225 GHz et 94 GHz (ce qui correspond à des longueurs d'onde variant entre 133 et 0.3 cm). Les bandes intermédiaires (X, C, S et L) figurent parmi les plus utilisées en imagerie RADAR. Le choix de la bande dépend fortement de l'application visée. Par exemple, la bande X (utilisée par le satellite allemand TerraSAR-X) est particulièrement bien adaptée à des applications en cartographie ou en détection d'objets dans le cadre de l'imagerie RADAR haute résolution.

1.2.2 Formation des images

Équation RADAR Dans le cas d'un RADAR monostatique, "l'équation du RADAR", qui permet d'établir le lien entre puissance reçue $P_{\text{reçue}}$ et puissance émise $P_{\text{émise}}$ est donnée par la formule :

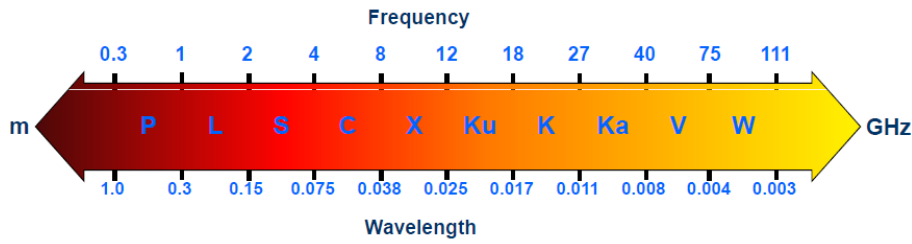


FIGURE 1.1 – Bandes de fréquence utilisées en imagerie RADAR RSO. Source : [Den09].

$$P_{\text{reçue}} = P_{\text{émise}} \frac{G^2 \lambda^2}{(4\pi)^3 D^4 a} \sigma \quad \text{dans le cas d'une cible ponctuelle} \quad (1.1)$$

ou par la formule :

$$P_{\text{reçue}} = P_{\text{émise}} \frac{G^2 \lambda^3}{(4\pi)^3 D^3 a} \frac{\sigma^0 \rho}{L \sin(\theta_i)} \quad \text{dans le cas d'une cible diffuse} \quad (1.2)$$

avec les notations suivantes :

- G correspond au gain de l'antenne ;
- σ correspond à la surface équivalente RADAR (SER) ;
- σ^0 correspond au coefficient de rétro-diffusion ;
- a correspond au facteur de pertes relatif à la traversée des milieux ;
- ρ correspond à la taille de la cellule de résolution en "géométrie radar" suivant l'axe des distances (voir le paragraphe intitulé **Typologie des données** présenté plus loin dans cette section) ;
- θ_i correspond à l'angle d'incidence du capteur RADAR ;
- L correspond à la longueur de l'antenne ;
- D correspond à la distance cible-capteur.

Géométrie capteur La géométrie du système RADAR est caractérisée par deux axes formant un repère cylindrique (voir figure 1.2) :

- l'axe des distances, défini par la direction de visée de l'antenne ;
- l'axe des azimuths, défini par la trajectoire du porteur supposée rectiligne durant l'acquisition.

La trace du porteur correspond à la projection orthogonale au sol de sa trajectoire.

Comme le système imageur utilise une vision en distance, l'acquisition a lieu en visée latérale, c'est-à-dire suivant un angle d'incidence non nul, noté θ_i .

On note \vec{v} le vecteur vitesse de déplacement du porteur (parallèle à sa trajectoire) et \vec{u} le vecteur unitaire orthogonal à \vec{v} .

On définit l'angle de dépointage de l'antenne ou l'angle de squint, noté θ_s , par l'angle qui existe entre la direction de visée du capteur et le vecteur unitaire \vec{u} . Sur la figure 1.2, la direction de visée du capteur coïncide avec \vec{u} ; le dépointage est donc nul. Notons que, suivant le mode d'acquisition choisi, cet angle de squint peut varier au cours de la trajectoire.

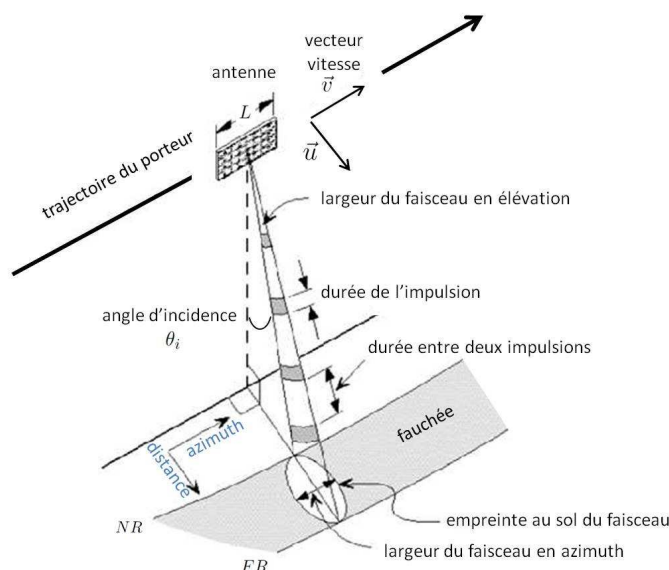


FIGURE 1.2 – Géométrie du système d'acquisition RADAR. Source : [Den09].

Acquisition des données En mode basique (mode StripMap), l'antenne est fixe et l'acquisition se fait récurrence par récurrence de la façon suivante. Le capteur émet une impulsion et enregistre le signal rétrodiffusé à partir du temps t_0 et pendant une durée Δt . Ces mesures correspondent aux réponses de l'ensemble des points présents dans l'empreinte au sol de l'antenne et situés à une distance du capteur comprise entre la distance minimale NR ("near range") et la distance maximale FR ("far range"). En supposant constante la vitesse de la lumière c , les distances NR et FR sont respectivement égales à ct_0 et $c(t_0 + \Delta t)$.

La fauchée, qui définit la largeur de l'image, correspond à la distance au sol séparant les points situés à distance NR du capteur de ceux situés à distance FR du capteur (voir figure 1.2).

On appelle fréquence de répétition des impulsions (FRI) l'inverse de l'intervalle de temps séparant deux impulsions.

D'autres modes d'acquisition spécifiques et plus complexes existent sur les capteurs satellitaires récents :

- le mode ScanSAR, où la direction de visée de l'antenne varie en élévation au cours de l'acquisition (θ_i non constant) afin de générer une combinaison de faisceaux couvrant des fauchées adjacentes ; ceci permet d'améliorer la couverture imagée au détriment de la résolution spatiale ;
- le mode SpotLight (ou SlipSAR), où la direction de visée de l'antenne varie en azimuth au cours de l'acquisition (θ_s non constant) afin de maximiser le temps d'illumination de la cible ; ceci permet d'améliorer la résolution spatiale au détriment de la couverture imagée.

La figure 1.3 présente les trois modes d'acquisition classiquement utilisés en imagerie RADAR satellitaire.

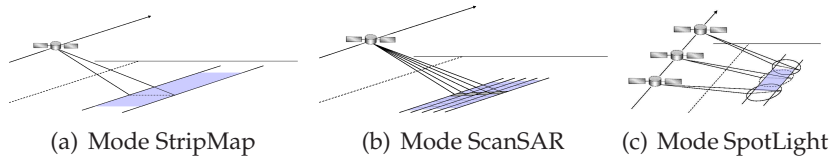


FIGURE 1.3 – Modes d'acquisition utilisés en imagerie RADAR satellitaire. Source : [Den09].

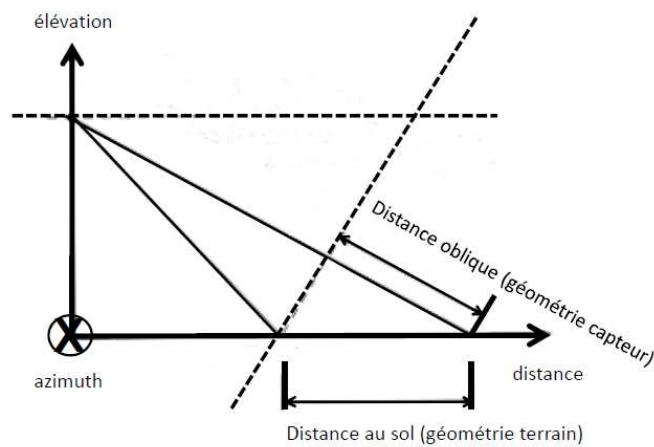


FIGURE 1.4 – Géométrie image et géométrie terrain.

Typologie des données Les données RADAR directement issues de la mesure (en amplitude et en phase) du signal réfléchi par la scène imagée sont appelées données "Raw".

Celles-ci sont converties en données "Single Look Complex (SLC)" suite à différents traitements, détaillés dans la section 1.2.4 (filtrage adapté en distance, synthèse d'ouverture en azimuth et échantillonnage).

Les données SLC sont représentées en "géométrie capteur" - ici "géométrie radar" - ("slant geometry"), c'est-à-dire dans le référentiel naturel d'acquisition des données composé des deux axes : azimuth et distance oblique (slant range) (voir figure 1.4).

Elles peuvent ensuite être projetées en "géométrie terrain" ("ground geometry") et corrigées géométriquement et radiométriquement (filtrage multi-vues). On parle alors de données "projetées" ("projected products"), c'est-à-dire représentées dans le référentiel terrain composé des deux axes : azimuth et distance au sol ou distance projetée (ground range) (voir figure 1.4).

Cette projection est réalisée suivant un ellipsoïde de référence (par exemple, celui issu du système WGS 84). Elle intègre éventuellement la correction de certaines distorsions géométriques dues aux effets de relief et prises en compte grâce à l'utilisation de modèles numériques de terrain (MNT) ou d'élévation (MNE).

1.2.3 Caractéristiques radiométriques

Les données RADAR SLC sont classiquement codées sur 32 bits (16 bits pour la partie réelle et 16 bits pour la partie imaginaire). L'image d'amplitude présente donc une large dynamique s'étalant sur une plage de 65536 valeurs et permettant de considérer de très fortes variations d'amplitude, caractéristiques des données RADAR (par exemple, entre les zones d'ombres et les zones d'échos forts).

1.2.4 Caractéristiques spatiales

Résolution en distance avant traitement La résolution en distance spécifie la capacité du système imageur à pouvoir séparer deux cibles selon l'axe des distances.

De façon générale, la résolution en distance est inversement proportionnelle à la largeur de bande du signal émis.

Dans le cas particulier d'une impulsion contenant une onde monochromatique (sinusoïde pure), elle est directement proportionnelle à la durée, notée τ , de l'impulsion émise.

Afin que deux cibles soient correctement résolues, la distance les séparant doit être suffisamment grande pour que le capteur reçoive des signaux rétrodiffusés distincts.

La résolution en distance, notée r_d , est alors égale à [Tis04, Cel06] :

$$r_d = \frac{c\tau}{2} \quad (1.3)$$

Résolution en azimuth avant traitement On rappelle que le diagramme de rayonnement d'une antenne de longueur L se caractérise par son angle d'ouverture à 3 dB, noté α_{3dB} et relatif à la largeur du lobe principal.

Lorsque α_{3dB} est faible (approximation en champ lointain), on obtient la formule suivante [Tis04, Cel06] :

$$\alpha_{3dB} \approx 0.88 \frac{\lambda}{L} \quad (1.4)$$

La résolution en azimuth spécifie la capacité du système imageur à pouvoir séparer deux cibles selon l'axe des azimuths.

Elle est directement proportionnelle à l'angle d'ouverture en azimuth de l'antenne.

Afin que deux cibles soient correctement résolues, elles doivent être suffisamment distantes selon l'axe des azimuths pour ne pas se trouver à l'intérieur d'un même cône d'angle α_{3dB} .

La résolution en azimuth, notée r_a , d'une cible située à la distance D du capteur, est égale à [Tis04, Cel06] :

$$r_a = \alpha_{3dB} D \approx 0.88 \frac{\lambda}{L} D \quad (1.5)$$

Application numérique

Cas du satellite TerraSAR-X en bande X

- fréquence porteuse : 9.65 GHz (soit longueur d'onde λ : 3.1 cm)
- hauteur nominale au-dessus de l'équateur h : 514 km
- longueur d'antenne L : 4.8 m

- durée d’impulsion τ : 30 à 50 μs

On en déduit que, sans traitement :

- la résolution en distance r_d serait de 4500 à 7500 m ;
- la résolution en azimuth r_a , pour un angle d’incidence θ_i égal à 30° et une distance capteur-cible D de 551 km, serait de 3132 m.

On fait ainsi appel à deux techniques de traitement du signal afin de pouvoir améliorer ces résolutions : la compression en distance et la synthèse d’ouverture en azimuth.

Compression en distance Il s’agit d’effectuer une compression de l’onde émise en la modulant linéairement en fréquence par un chirp. Celui-ci est défini par sa largeur de bande, appelée bande utile, notée B_{utile} et égale à :

$$B_{\text{utile}} = K\tau \quad (1.6)$$

où K , appelé taux de modulation, correspond à la pente du chirp utilisé.

Le signal rétrodiffusé par la cible doit ensuite être démodulé par application d’un filtrage adapté (qui maximise le rapport signal à bruit).

Après compression, la résolution ultime obtenue en distance, notée r_d^* , ne dépend plus que de B_{utile} selon la formule [Tis04, Cel06] :

$$r_d^* \approx \frac{0.88c}{2B_{\text{utile}}} \quad (1.7)$$

Application numérique

Cas du satellite TerraSAR-X en bande X

- largeur de bande utile B_{utile} maximale : 150 MHz

On déduit que, après traitement :

- la résolution en distance théorique maximale r_d^* est de 0.88 m ; toutefois, en pratique, une marge de 10 % étant incluse dans les performances reportées pour les produits commercialisés, une résolution en distance de 1 m est annoncée dans les spécifications produits [EFM⁺09] (résolution donnée en géométrie RADAR pour des données SLC).

Synthèse d’ouverture Une démarche similaire est adoptée pour la synthèse d’ouverture en azimuth, mais, dans ce cas, la modulation fréquentielle se fait grâce au mouvement du capteur le long de sa trajectoire (avec génération d’un effet Doppler).

On cherche à artificiellement synthétiser une antenne plus grande que l’antenne réelle, en utilisant le fait qu’un même point de la scène est éclairé plusieurs fois, depuis différentes positions en azimuth du satellite, grâce à la largeur du lobe d’antenne.

La bande passante du système utilisée lors de la synthèse est appelée bande Doppler et notée B_{Doppler} .

Après synthèse, la résolution ultime obtenue en azimuth, notée r_a^* , ne dépend plus que de λ , θ_s , v (vitesse de déplacement du porteur) v_{sol} (vitesse de déplacement de l’empreinte au sol) et α (ouverture angulaire sur laquelle est effectuée la synthèse) selon la formule [Tis04, Cel06] :

$$r_a^* \approx \frac{0.88\lambda v_{\text{sol}}}{2v \cos^3(\theta_s) (\tan(\frac{\alpha}{2} - \theta_s) + \tan(\frac{\alpha}{2} + \theta_s))} \quad (1.8)$$

Dans le cas d'un angle de squint nul ($\theta_s = 0$) et d'une faible ouverture angulaire, l'équation 1.8 se simplifie :

$$r_a^* \approx \frac{0.88\lambda v_{sol}}{2v\alpha} \quad (1.9)$$

Puis, dans le cas où la synthèse est effectuée sur l'ouverture complète de l'antenne ($\alpha = \alpha_{3dB}$), on obtient :

$$r_a^* \approx \frac{L v_{sol}}{2 v} \quad (1.10)$$

Enfin, si la vitesse de déplacement de l'empreinte au sol est assimilée à celle du porteur (cas vérifié en imagerie aéroportée), on obtient la formule suivante :

$$r_a^* \approx \frac{L}{2} \quad (1.11)$$

Application numérique

Cas du satellite TerraSAR-X en bande X

On déduit que, après traitement :

- en mode StripMap, la résolution en azimuth théorique maximale r_a^* est de 2.4 m ; toutefois, en pratique, suite à un traitement opéré afin de réduire les ambiguïtés causées par le phénomène de repliement spectral ("aliasing"), on obtient finalement une résolution en azimuth de 3 m, annoncée dans les spécifications produits [EFM⁺09] (résolution donnée pour des données SLC) ;
- en mode SpotLight, avec un angle de dépointage variant entre -0.75° et 0.75° , on obtient une résolution en azimuth de 1 m en mode HS (High Resolution SpotLight Mode) et de 2 m en mode SL (SpotLight Mode), annoncées dans les spécifications produits [EFM⁺09] (résolution donnée pour des données SLC).

Échantillonnage en distance et en azimuth Afin de respecter le théorème de Shannon, le signal doit être échantillonné :

- en distance, avec une fréquence d'échantillonnage F_d supérieure à la bande utile B_{utile} ; ceci implique un pas d'échantillonnage en distance P_d - définissant la taille des pixels en distance - inférieur à la résolution en distance r_d^* ;
- en azimuth, avec une fréquence d'échantillonnage F_a (conditionnant le choix de la fréquence de répétition des impulsions FRI) supérieure à la bande Doppler $B_{Doppler}$; ceci implique un pas d'échantillonnage en azimuth P_a - définissant la taille des pixels en azimuth - inférieur à la résolution en azimuth r_a^* .

1.2.5 Spécificités géométriques

La vision en distance du système RADAR induit des particularités quant à la géométrie de l'image suivant l'axe des distances. Les phénomènes de distorsions géométriques observés sont susceptibles de rendre plus difficiles l'interprétation visuelle et l'exploitation des données RADAR.

On recense trois types de déformations géométriques en imagerie RADAR à haute résolution en milieu urbain [Tis04, Den09] (voir figure 1.5) :

- la "contraction" ("foreshortening"), qui a pour effet de raccourcir sur l'image RADAR les pentes orientées vers le capteur et d'étirer celles orientées en sens inverse ;

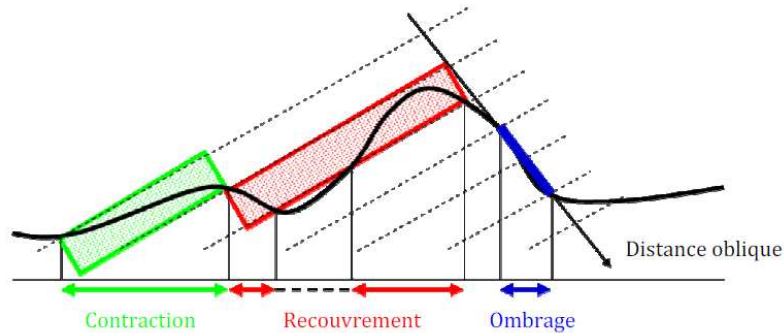


FIGURE 1.5 – Déformations géométriques observées en imagerie RADAR. Source : [Den09].

- le "recouvrement" ("layover"), qui a pour effet de mélanger les signaux issus de différents objets mais qui, situés à une même distance de l'antenne, répondent simultanément ;
- l'"ombrage" ("shadow"), où une partie de la scène est susceptible de ne pas être éclairée (et donc de ne pas répondre), à cause d'objets masquants situés devant elle.

Ces phénomènes interviennent pour certaines configurations géométriques de la scène déterminées par la taille des objets (hauteur, ...), leur position relative et les pentes locales du terrain.

1.2.6 Propriétés statistiques et filtrage des images

Les images RADAR se caractérisent par un bruit multiplicatif bien particulier appelé chatolement ou bruit de "speckle" et dû à la nature cohérente des ondes électromagnétiques rétrodiffusées.

De nombreux modèles statistiques ont été proposés dans la littérature pour décrire les statistiques des images RSO.

Certains modèles se sont appuyés sur des considérations physiques régissant les phénomènes de propagation des ondes électro-magnétiques à l'origine de la formation des images.

En présence de speckle dit pleinement développé et dans le cas de zone homogène, une loi Gamma généralisée à deux paramètres (cas multi-vues) est classiquement utilisée pour modéliser la distribution en intensité de l'image RADAR. Ce modèle s'est naturellement avéré être bien adapté dans le cas de scènes naturelles en imagerie basse résolution mais moins satisfaisant dans le cas de milieux urbains en imagerie haute résolution (zones hétérogènes).

En présence de speckle pleinement développé et dans le cas de zone hétérogène avec texture, plusieurs densités de probabilités telles que la loi K, les distributions U, B et W du système de Pearson, ou encore la loi gaussienne inverse ont été proposées pour la modélisation en intensité de l'image RADAR.

Enfin, en présence de speckle non pleinement développé (cas de cellules de résolution de petite taille) et pour des zones texturées, d'autres lois en intensité ont été étudiées, parmi lesquelles la distribution Log-Normale, la distribution de Weibull et la distribution de Nakagami-Rice.

D'autres modèles, plus récents, relèvent davantage de considérations mathématiques et de constatations empiriques. On peut citer notamment la loi de Fisher, particulièrement bien adaptée pour la modélisation des statistiques présentes sur les images monovues à haute résolution en milieu urbain. Il est montré dans [Tis04], que la loi de Fisher, décrite par trois paramètres (un facteur d'échelle et deux facteurs de forme), présente une souplesse lui permettant de modéliser aussi bien la distribution en intensité des surfaces naturelles (végétation, ...) que celle des surfaces non-naturelles (toits, ...) en imagerie RSO haute résolution. Elle permet d'unifier l'ensemble des modèles précédents sous la forme de différentes déclinaisons issues d'un formalisme générique commun. Plusieurs méthodes d'estimation de ses paramètres ont été proposées, parmi lesquelles la méthode des log-moments développée par Nicolas [Nic06] et basée sur des statistiques de seconde espèce.

Plusieurs types de filtres ont été développés afin de réduire au maximum ce bruit de speckle, tout en conservant au mieux le signal (en veillant notamment à préserver les discontinuités de l'image).

Certains filtres font appel à des considérations spatiales et fréquentielles simples ; c'est le cas des filtres multi-vues (multi-looks).

D'autres filtres, plus adaptés aux données RADAR, se fondent sur des raisonnements statistiques plus complexes. On peut citer quelques uns des filtres adaptés les plus célèbres en imagerie RADAR : le filtre de Frost, le filtre de Lee, le filtre de Kuan, les filtres MAP (maximum a posteriori) tels que le Gamma-MAP ou le Fisher MAP, le filtre EAP (espérance a posteriori) ou encore le filtre de Wu et Maître [Tis04, Maï01].

1.3 Nouveaux capteurs satellitaires optiques et RADAR à haute résolution

Dans cette section, nous nous intéressons aux caractéristiques des capteurs satellitaires récents, utilisés en imagerie optique et en imagerie RSO.

Les informations relatives aux capteurs exposés sont issues des guides de produits suivants : [EFM⁺09, Ima07, Spo09a, Spo09d, Spo09b, Spo09c, Spo03, Geo10, Orb06, Rap10, Dig09, Dig09a, Dig09b, MDA10, OHB09, Ita07, ADHB05, ABE⁺07].

Les tables 1.1 et 1.2 présentent certaines spécificités techniques (bande spectrale, mode d'acquisition, résolution spatiale, taille de la scène, fauchée) de quelques capteurs, lancés au cours des dix dernières années ou prochainement mis en orbite.

En imagerie satellitaire optique, ces dix dernières années sont marquées par le passage de la moyenne résolution (MR) (cas de Spot-5, lancé en 2002 et présentant une résolution de quelques mètres en mode panchromatique) à la haute résolution (HR) (cas de WorldView-2, lancé en 2009 et présentant une résolution submétrique en mode panchromatique). Par comparaison avec l'imagerie aéroportée optique, atteignant aujourd'hui la gamme des très hautes résolutions (THR) (de quelques centimètres à quelques dizaines de centimètres), l'imagerie satellitaire optique se situe actuellement dans un contexte intermédiaire, celui de la résolution métrique et submétrique.

En imagerie satellitaire RADAR, les grands changements des dernières années se caractérisent par l'émergence de nouveaux modes d'acquisition permettant l'obtention de données avec une meilleure résolution (cas de TerraSAR-X, lancé en 2007 et présentant une résolution inférieure au mètre en mode SpotLight) ou avec une plus grande surface imagée (par l'utilisation du mode ScanSAR).

1.4 Comparatif imagerie optique - imagerie RADAR : avantages et inconvénients

L'ensemble des points forts et des points faibles qui caractérisent les données satellitaires optiques et RADAR rend particulièrement intéressant leur utilisation conjointe dans le cadre de la fusion d'informations.

En ce qui concerne les avantages des données optiques, on note d'abord une facilité d'interprétation des scènes imagées lors d'un examen par un opérateur ; ceci n'est pas le cas pour les données RADAR, qui présentent un fort bruit de speckle et de nombreuses distorsions géométriques, rendant moins intuitive une interprétation visuelle directe (réservée alors aux experts).

On note également sur les données satellitaires optiques :

- un fort potentiel spatial avec l'obtention d'images atteignant aujourd'hui des résolutions de l'ordre de quelques dizaines de centimètres ;
- un fort potentiel spectral, particulièrement utile dans le cadre d'applications en classification, permettant une discrimination rapide et simple de certains types d'objets présentant une signature spectrale caractéristique (exemple de la végétation aisément visible dans l'infra-rouge).

En ce qui concerne les avantages des données RADAR, on souligne d'abord deux particularités du système d'acquisition :

- le capteur RADAR est un capteur actif, ce qui lui confère une capacité à imager de jour comme de nuit ;
- il s'agit d'un système d'imagerie cohérente, ce qui induit une information de phase exploitée en synthèse d'ouverture pour obtenir de meilleures résolutions ou en interférométrie pour obtenir des estimations d'altitudes ou de mouvements très précises.

On remarque également que les données satellitaires RADAR peuvent s'avérer particulièrement intéressantes en situation de crise par leur rapidité d'obtention (notamment grâce au fait qu'elles soient très peu dépendantes des conditions météorologiques) ; ceci n'est pas le cas pour les données optiques, qui peuvent être partiellement non exploitables à cause du couvert nuageux.

Enfin, il convient de souligner sur les données satellitaires RADAR :

- un fort potentiel "3D" ; ce potentiel relève premièrement de déformations géométriques spécifiques (recouvrement, ombres) permettant d'extraire des informations de hauteurs en monoscopie RADAR ; il se rapporte deuxièmement aux deux techniques différentes que sont la radargrammétrie (équivalent RADAR de la photogrammétrie optique) et l'interférométrie RADAR, qui, appliquées sur des couples de données RADAR, produisent la troisième dimension ;
- un fort potentiel polarimétrique, qui rend possible l'extraction de nombreuses informations et caractéristiques, à partir des quatre images RADAR polarimétriques disponibles (même si cet aspect ne sera pas abordé dans nos travaux) ;

TABLE 1.1 – Quelques nouveaux capteurs satellitaires optiques.

Nom	Pays	Lancement	Bande spectrale	Résolution	Taille de la scène
EROS-B	Israël	2006	-	70 cm (PAN)	14 km x 14 km ("basique") et 26 km x 28 km ("mosaïque")
FORMOSAT-2	Taiwan	2004	B, G, R, NIR	2 m (PAN) 8 m (MUL)	24 km x 24 km
GEOEYE-1	USA	2008	-	41 cm (PAN) 1.65 m (MUL)	15 km x 15 km ("scène seule") 300 km x 50 km ("région large") ou 100 km x 100 km ("région large")
KOMPSAT-2	Corée	2006	B, G, R, NIR	1 m (PAN) 4 m (MUL)	15 km x 15 km
ORBVIEW-3	USA	2003	B, G, R, NIR	1 m (PAN) 4 m (MUL)	8 km x 8 km
PLEIADES (HR1, HR2)	France	2010 -2011	-	50 cm (PAN) 2 m (MUL)	20 km x 20 km ("scène seule") 100 km x 100 km ("mosaïque")
QUICKBIRD	USA	2001	B, G, R, NIR	61/72 cm (PAN) 2.44/2.88 m (MUL)	16.5 km x 16.5 km ("scène seule") et 115 km x 16.5 km ("bandeau")
RAPIDEYE	Allemagne	2008	B, G, R, RE, NIR	6.5 m	77 km x 50 à 300 km
SPOT-5	France	2002	G, R, NIR, SWIR	2.5 ou 5 m (PAN) 10 m (MUL)	60 km x 60 à 80 km
SPOT (6, 7)	France	2012 -2013	-	2 m (PAN) 8 m (MUL)	60 km x 60 km
WORLDVIEW-1	USA	2007	-	50/59 cm (PAN)	17.6 km x 17.6 km ("scène seule") 330 km x 17.6 km ("bandeau") 110 km x 60 km ("région large")
WORLDVIEW-2	USA	2009	B, G, R, NIR RE, C, Y, NIR 2	46/52 cm (PAN) 1.84 m (MUL)	16.4 km x 16.4 km ("scène seule") 250 km x 16.4 km ("bandeau") 110 km x 65.6 km ("région large")

Légende :

PAN : mode panchromatique - MUL : mode multi-spectral - B : blue (canal bleu) - G : green (canal vert) - R : red (canal rouge) - C : coastal - Y : yellow - RE : red edge
 NIR : near infra red (proche infra-rouge) - NIR 2 : near infra red 2 (proche infra-rouge 2) - SWIR : short wave infra red (infra-rouge à ondes courtes)

TABLE 1.2 – Quelques nouveaux capteurs satellitaires RADAR.

Nom	Pays	Lancement	Bande spectrale	Résolution	Fauchée
COSMOSKYMED (1 à 4)	Italie	2007 à 2010	X	1 m (SL) 3 m (SM) 30 m (SSW) 100 m (SSH)	11 km (SL) 40 km (SM) 100 km (SSW) 200 km (SSH)
ENVISAT ASAR	Europe	2002	C	30 m (IM) 150 m (WSM) 1000 m (GMM)	56 à 100 km (IM) 400 km (WSM) 400 km (GMM)
RADARSAT-2	Canada	2007	C	0.74 m (SL) 7.7 m (F) 7.7 m (STD) 100 m (SSW)	18 km (SL) 50 km (F) 100 km (STD) 500 km (SSW)
SARLUPE	Allemagne	2006 à 2008	X	≈ 50 cm (SL) ≈ 1 m (SM)	5.5 km (SL)
SENTINEL-1	Europe	2011	C	5 m (SM) 20 m (WSM)	80 km (SM) 250 km (WSM)
TANDEM-X	Allemagne	2010	X	1 m (SL) 3 m (SM) 16 m (SC)	10 km (SL) 30 km (SM) 100 km (SC)
TERRASAR-X	Allemagne	2007	X	1.1 m (HS) 1.7 m (SL) 3.3 m (SM) 18.5 m (SC)	10 km (HS) 10 km (SL) 30 km (SM) 100 km (SC)

Légende :
SL : StripMap - SC : ScansAR - HS : High SpotLight - SSW : ScansAR Wide - SSH : ScansAR Huge
- F : Fine - STD : Standard - IM : Image Mode - WSM : Wide Swath Mode - GMM : Global Monitoring Mode

- la capacité à mettre facilement en avant certains objets d'intérêt qui réfléchissent fortement les ondes RADAR (grâce à leur configuration géométrique et à leurs propriétés électro-magnétiques) et qui sont susceptibles d'être non imagés dans le visible ou confondus avec le fond ; c'est notamment le cas de certaines constructions en matériau métallique tels que les toits de bâtiments. En contrepartie, la forme générale des objets est souvent plus difficile à retrouver, avec une forte variabilité des objets en fonction de leur géométrie par rapport à l'angle d'incidence.

On conclut ainsi que les deux types de données optiques-RADAR, présentant de nombreuses complémentarités, favorisent une exploitation jointe des images particulièrement bien adaptée à l'étude des milieux urbains en imagerie satellitaire haute résolution.

Chapitre 2

État de l'art en reconstruction 3D du bâtiment à partir d'images optiques et RADAR à haute résolution

Nous proposons dans ce chapitre un état de l'art relatif au domaine de la reconstruction tridimensionnelle des bâtiments à partir d'images optiques et RADAR à résolution métrique et submétrique. On précise que l'on ne s'intéressera, côté RADAR, qu'à l'imagerie RSO (RADAR à synthèse d'ouverture) et pas au LIDAR.

Les méthodes recensées sont regroupées selon trois familles qui feront l'objet de trois sections dédiées : les méthodes basées sur l'utilisation de données optiques uniquement, celles basées sur l'utilisation de données RADAR uniquement et celles basées sur l'utilisation conjointe de données optiques et RADAR.

Suivant la méthodologie développée, une ou plusieurs des problématiques suivantes, relatives à la reconstruction tridimensionnelle des bâtiments, seront traitées : pré-traitements, extraction de primitives caractéristiques 2D-3D, estimation des hauteurs sol / terrain ou des hauteurs sursol / bâti, estimation des paramètres d'un modèle 3D, validation de la présence d'un bâtiment.

2.1 Généralités

Modèles 3D de bâtiments La démarche suivie et la modélisation tridimensionnelle adoptée pour la reconstruction 3D dépendent directement du type de données disponibles en entrée (images optiques et/ou RADAR), de leur nombre et de leur configuration (données monoscopiques, stéréoscopiques, radargrammétriques, interférométriques ou multi-aspects) et de la gamme de résolution à laquelle on travaille (métrique ou submétrique).

Dans le cas des données RADAR à résolution métrique, relativement difficiles à traiter en raison de leur complexité (bruit de speckle, distorsions géométriques, fortes discontinuités en milieu urbain), des modèles 3D simples de bâtiments sont souvent adoptés (parallélépipède 3D).

Au contraire, dans le cas des données optiques à résolution submétrique, présentant une aisance à l'interprétation, des modèles 3D plus sophistiqués sont proposés : différents types d'emprises (rectangulaires, en "L", en "T", polygonales ...) et de toits (plats, galbés, avec pentes) sont envisagés et il est également possible d'intégrer certaines superstruc-

tures (cheminées, balcons ...), qui, la plupart du temps, ne peuvent pas être considérées en imagerie RADAR.

Primitives et signatures On désigne par "primitive" tout élément 2D ou 3D (point, coin, segment, région, plan) caractéristique de la présence d'un bâtiment dans les données. Le niveau de détail des primitives recherchées, leur géométrie et leur position dépendent directement du modèle 3D de bâtiment choisi. La "signature" d'un bâtiment est définie comme l'ensemble des primitives caractéristiques s'y référant.

Parmi les primitives optiques figurent entre autres :

- comme primitives ponctuelles : les sommets de toits (points ou coins) ;
- comme primitives linéaires : les arêtes des toits (segments) ;
- comme primitives surfaciques : les zones de toits (plans), les zones de façades (plans) et les zones d'ombres optiques (régions).

Parmi les primitives RADAR figurent entre autres :

- comme primitives linéaires : les échos doubles mur-sol (segments) ;
- comme primitives surfaciques : les zones de toits avec ou sans recouvrement (régions) et les zones d'ombres RADAR (régions).

2.2 Reconstruction des bâtiments en imagerie optique

On recense dans la littérature de nombreux travaux récents portant sur la reconstruction tridimensionnelle des bâtiments à partir d'images optiques à résolution métrique et submétrique.

Nous nous intéressons dans la suite aux trois enjeux majeurs étudiés dans ces travaux :

1. La détection de bâtiments et l'extraction de leurs contours.
2. L'estimation des hauteurs de bâtiments.
3. La génération de MNS ou de MNE et la reconstruction 3D des bâtiments.

Les méthodologies proposées utilisent en entrée d'algorithme des données images seules (une image [LN98, KP08], un couple d'images [Ori03], trois images ou plus) ou des données auxiliaires additionnelles (MNS ou MNE [ODZ07b, LDZPD06a, KST06], limites cadastrales [Jib02, Tai05, FMJ03] et bases de données [KST06]).

2.2.1 Détection des bâtiments et extraction de leurs contours

Nous traitons dans cette section la problématique de détection de bâtiments, ayant comme objectif l'extraction d'informations relatives à la présence, à la localisation et au dimensionnement des bâtiments.

La plupart des méthodes étudiées se basent sur l'accumulation d'indices identifiant la présence d'une construction et relevant de différents types : indices géométriques et morphologiques (induisant des informations a priori sur la forme ou structure des bâtiments) [LN98, MZ05, BPA03, CBF06, PB01, SDW04, JD05, LW07], indices radiométriques ou photométriques (utilisant par exemple l'hypothèse de bâtiments radiométriquement homogènes et d'une gamme définie de couleurs) [MZ05], indices contextuels (vérifiant par exemple la juxtaposition entre un bâtiment et son ombre) [MZ05, SDW04, JD05, WZS04, Bac09, AA10] et indices spectraux dans le cas d'images multi-spectrales (se basant par

exemple sur la valeur du NDVI afin de distinguer les régions de label "végétation" et les régions de label "bâtiment") [KST06, Bac09].

Selon que les images sont acquises avec une visée verticale ou éloignée du nadir, l'étape d'extraction des contours de bâtis se réfère à l'extraction des emprises au sol et/ou des bords de toits et peut inclure ou non la gestion des façades.

Nous proposons de décomposer les approches de détection de bâtiments suivant cinq classes :

- les approches "basées données auxiliaires" ;
- les approches "basées images", qui exploitent directement l'information au niveau du pixel ;
- les approches "basées primitives 2D-3D", qui induisent une étape intermédiaire d'identification de primitives caractéristiques du bâtiment ;
- les approches "basées modèles 2D", qui se réfèrent à une librairie a priori de modèles bidimensionnels de bâtiments ;
- les approches "mixtes", qui combinent deux ou plus des trois classes citées précédemment.

Chaque classe de méthodes présente naturellement ses avantages et ses inconvénients en termes de robustesse, précision et généralité : les méthodes "basées images", offrent un fort degré de généralité mais demeurent souvent peu robustes ; les méthodes "basées modèles 2D" intègrent des connaissances géométriques a priori et ne peuvent donc être appliquées que pour un nombre limité de types géométriques de bâtiments mais bénéficient d'une grande robustesse ; enfin, les méthodes "basées primitives 2D-3D" ainsi que les méthodes "mixtes" se situent à un niveau intermédiaire.

Approches "basées données auxiliaires" Les contours polygonaux des bâtiments peuvent être fournis à partir de données auxiliaires :

- de type vectoriel, telles que des cartes issues de plans cadastraux [Jib02, Tai05, FMJ03], directement superposables à des images ortho-rectifiées, qui indiquent l'emplacement des limites planimétriques des bâtis recensés ;
- de type raster, telles que des MNE, qui fournissent une description altimétrique des scènes urbaines et facilitent l'identification des bâtiments. Dans [ODZ07b, ODZ07a], Ortner et al. présentent une méthode basée sur un processus ponctuel marqué permettant d'extraire, à partir d'un MNE approximatif, les emprises de bâtiments sous forme de quadrilatères composés de rectangles connectés entre eux.

Approches "basées images" Parmi les approches "basées images", on distingue les algorithmes "orientés contours" qui se focalisent sur l'extraction directe des contours (limite extérieure des bâtis) [WZS04] et les algorithmes "orientés pixels" ou "orientés régions" (contenu intérieur des bâtis) qui cherchent à établir une classification (label "bâti" ou "non bâti") des pixels ou des régions issues d'une pré-segmentation [MZ05, CBF06].

Dans [WZS04], Wei et al. proposent d'utiliser la localisation des ombres optiques pour extraire des contours polygonaux de bâtiments à partir d'un détecteur de bords (opérateur de Canny) et de transformées de Hough.

Dans [MZ05, CBF06], il s'agit de classifier par apprentissage respectivement les pixels et les régions d'une image à partir d'un vecteur de caractéristiques. Dans [MZ05], Müller et al. utilisent un ensemble de considérations géométriques, photométriques et structurales pour caractériser ce vecteur de caractéristiques. Dans [CBF06], Chanussot et al.

introduisent des profils morphologiques différentiels (fournissant une information de forme, de taille et de contraste) afin de définir le vecteur caractéristique d'un pixel.

Approches "basées primitives 2D-3D" Deux sortes de primitives 2D-3D caractéristiques du bâtiment sont essentiellement utilisées : on trouve d'une part celles "orientées contours" avec les segments 2D-3D (arêtes de toit ou d'emprise) [LN98, Che05, Tai04] et les coins 2D-3D (jonctions orthogonales entre des segments 2D-3D), et, d'autre part, celles "orientées régions" avec les facettes ou plans 2D-3D (associées aux intérieurs de toits) [Che05].

En monoscopie optique, Lin et al. [LN98] développent un procédé de génération - sélection - validation d'hypothèses pour retrouver des bordures de toits rectilinéaires à partir de segments 2D regroupés.

En stéréoscopie optique, des stratégies de mise en correspondance sont classiquement utilisées pour sélectionner des primitives fiables préalablement extraites. Dans [Che05, Tai04], les segments 2D sont issus d'un procédé en trois étapes basé sur une détection de contours (filtre de Canny-Deriche et seuillage à hystérésis), un chaînage et une polygonalisation. Les facettes 2D quant à elles sont extraites soit par recherche de demi-plans autour des segments 2D appariés sur les deux images (et formant ainsi des segments 3D) soit par mise en correspondance sous contraintes (contrainte de taille, contrainte épipolaire, contrainte altimétrique, contrainte de similarité et contrainte de recouvrement) de régions issues d'une pré-segmentation multi-échelle et hiérarchique de l'image. Les facettes appariées fiables sont sélectionnées comme toits des bâtiments et les contours des régions associées indiquent alors l'emplacement et les dimensions des bords de toits.

Approches "basées modèle 2D" Les méthodes "basées modèle 2D" introduisent un a priori complet quant à la forme bidimensionnelle des emprises recherchées. Elles permettent ainsi de résoudre les problèmes de robustesse qui affectent les approches "basées images" et qui sont notamment dûs aux ombres et aux occlusions présentes sur les images optiques.

En monoscopie optique, Karantzalos et al. [KP08] proposent, dans un cadre variationnel, un procédé de détection des zones de bâtiments par segmentation sous contrainte, à travers la minimisation d'une fonctionnelle intégrant un terme d'attache aux données et un terme d'a priori relatif aux différentes formes d'emprises envisagées (rectangulaires, en croix, en "L", en "U", ...).

Approches "mixtes" Ces approches visent à combiner les points forts des approches "basées images", "basées primitives 2D-3D" et "basées modèle 2D".

Dans [KST06], San et al. décrivent une méthode de détection de bâti à partir d'un couple stéréoscopique (fournissant un MNS) et d'une image multispectrale (fournissant une mesure de l'indice NDVI), en tenant compte d'aspects "images" et d'aspects "modèles 2D" : il s'agit d'abord de détecter les régions labellisées "bâtis" de l'image par classification supervisée selon un maximum de vraisemblance et ensuite de délimiter les contours de bâtis en associant à chaque région un modèle polygonal parmi une librairie de modèles a priori, issue d'une base de données.

De même, dans [Ori03], Oriot met en œuvre une approche par contours actifs statistiques ("snakes") pour l'extraction de bâtis à partir d'un couple stéréoscopique d'images optiques aériennes. Pour chaque bâti, le polygone optimal est estimé en minimisant une

énergie définie dans l'espace du coefficient de corrélation et tenant compte de deux termes : un terme d'attache aux données calculé à partir d'informations (intervalles de disparité, zones d'occlusions, intensité) issues du couple d'images stéréoscopiques et un terme de régularisation introduisant un a priori géométrique sur la forme du bâti recherché.

Enfin, dans [KHN01], Kim et al. développent un procédé de détection des bâtiments complexes (toit plats ou toits avec pentes) à partir d'images multi-aspects (c'est-à-dire vues selon des traces d'acquisition multi-directionnelles), en intégrant des considérations "basées images", "basées primitives 2D-3D" et "basées modèles 2D". Premièrement, des primitives 2D-3D (lignes et jonctions) sont extraites, filtrées, appariées et sélectionnées, à partir des données images et d'un MNE, selon une approche probabiliste. Deuxièmement, des hypothèses visant à modéliser les toits de chaque bâti sont générées, vérifiées puis sélectionnées par analyse des zones de toits (pentes et recouvrement). Une description 3D du modèle de bâti est finalement obtenue et permet l'extraction des contours 2D des bâtiments par projection sur les données images.

2.2.2 Estimation des hauteurs de bâtiments

En ce qui concerne la problématique isolée d'estimation des hauteurs de bâtiments (hauteur à la base du toit, hauteur au sommet du toit ou hauteur le long d'arêtes sur le toit), différentes méthodes peuvent être mises en œuvre en fonction des données disponibles en entrée :

- en monoscopie optique, par mesure de la parallaxe (si les façades sont visibles) ou par mesure des ombres portées / projetées [WL06, SS98] ;
- en stéréoscopie et multiscopie optique, par mesure des disparités après mise en correspondance de points ou de primitives 2D-3D [CDJS02, Che05, Tai04, BZ00] ;
- à partir de données auxiliaires telles qu'un MNE, en utilisant localement les informations d'élévation fournies ([LDZPD06a, VC02]) (il s'agit alors d'identifier en planimétrie les zones de bâti puis de restreindre à ces emplacements d'intérêt l'information altimétrique disponible) ; par exemple, dans [VC02], Vinson et al. estiment la hauteur des bâtiments en moyennant les hauteurs MNE des pixels se trouvant à l'intérieur d'emprises extraites.

2.2.3 Génération de MNS ou de MNE et reconstruction 3D des bâtiments

Les différentes méthodes recensées dans la littérature pour la reconstruction du bâtiment en imagerie optique à haute résolution cherchent à établir une reconstruction 3D des structures toits. Celles-ci, une fois combinées aux emprises planimétriques détectées, permettent finalement d'obtenir une reconstruction complète des bâtis composant la scène urbaine (sous l'hypothèse de façades verticales).

Ces méthodes de reconstruction du bâti 3D sont :

- indirectes, lorsqu'elles se fondent sur l'usage de primitives 2D-3D intermédiaires ; elles sont décomposables alors en deux catégories : les approches "basées primitives 3D sans modèle 3D" (dans ce cas, aucun modèle 3D de bâtiment n'est introduit a priori) et les approches "basées primitives 3D avec modèles 3D" (dans ce cas, un modèle ou une bibliothèque de modèles 3D de bâtiments sont considérés) ;
- directes, lorsque l'objet 3D est reconstruit sans l'usage de telles primitives (approches "basées objets 3D").

Approches indirectes "basées primitives 3D sans modèle 3D" Il s'agit d'approches dites "ascendantes", génériques mais peu robustes, visant à reconstruire le bâti en s'appuyant sur des primitives caractéristiques précédemment extraites et sans information a priori sur la structure 3D à reconstruire.

Dans [Jib02, BZ00], deux méthodes de reconstruction sont développées à travers l'utilisation de plans 3D et de segments 3D issus d'images aériennes.

Dans [Jib02], Jibrini définit les surfaces recherchées associées aux toits comme des surfaces 2,5 D, continues, faites de plans et recouvrant les régions délimitées par des contours polygonaux 2D issus de cartes cadastrales. La méthodologie proposée se décompose en trois étapes : une étape de détection de plans 3D à partir d'un accumulateur de Hough 3D contraint par les orientations des façades, une étape de sélection heuristique parmi différentes hypothèses de plans et une étape de recherche de la surface 3D optimale à l'intérieur d'un graphe 3D se référant aux plans sélectionnés.

Dans [BZ00], Baillard et al. reconstruisent précisément la scène urbaine à partir de six images avec recouvrement. Dans un premier temps, des demi-plans dits fiables sont obtenus par balayage de plans en rotation autour de segments 3D extraits. Dans un second temps, les segments 3D extraits sont regroupés ou complétés selon des règles de coplanéité et colinéarité. Dans un dernier temps, les contours de plans (définissant des facettes de toit) sont délimités par intersection de segments 3D et les objets bâtis 3D sont enfin fermés.

Approches indirectes "basées primitives 3D avec modèles 3D" Il s'agit de reconstruire le bâti à partir d'approches dites :

- "descendantes", robustes mais peu génériques car basées sur des bibliothèques de modèles selon lesquelles on cherche à apparier au mieux les primitives ;
- ou "mixtes", qui intègrent une phase "ascendante" et une phase "descendante".

Dans [Tai05], Taillandier développe une méthode "descendante" de reconstruction du bâti à partir de multiples images aériennes (fournissant un MNE de référence par corrélation) et de cartes cadastrales (délimitant les bords de toit). Premièrement, pour chaque segment en bord de toit, un plan initial (avec une pente et une hauteur arbitrairement fixées) est inféré et donne lieu à un ensemble exhaustif de différents arrangements possibles de facettes 3D délimitées par les intersections de tous les plans (graphe 3D de facettes). Deuxièmement, des surfaces dites "admissibles" sont sélectionnées sur ce graphe et génèrent différentes hypothèses de modèles simples de toits (par la méthode des graphes compatibles ou des cliques maximales). Un tri est alors effectué selon des critères géométriques afin de réduire le nombre de modèles potentiels parmi une famille de modèles. Enfin, le meilleur modèle est sélectionné par maximisation de la corrélation centrée entre le MNE de référence et le MNE de synthèse relatif à ce modèle. Les deux paramètres (pente et hauteur des gouttières) caractérisant ce modèle sont obtenus par minimisation d'une distance entre les deux MNE.

Dans [Che05], Chehata propose une méthode "mixte" de reconstruction du bâti à partir d'un couple optique d'images stéréoscopiques à haute résolution (fournissant un MNE initial). Une première phase ("ascendante") permet d'extraire des primitives 3D fiables (segments 3D et facettes 3D) à partir du couple d'images et fournit ainsi des hypothèses de modèles. Une seconde phase ("descendante") permet de résoudre le problème de modélisation de surface 3D en validant les hypothèses de modèles : une méthode de flots de graphes est utilisée afin d'obtenir un MNE dense en contraignant le problème d'optimisation à partir des segments 3D et du MNE initial.

Approches directes "basées objets 3D" On cherche ici à reconstruire directement l'objet bâti 3D en introduisant des informations a priori sur le modèle 3D de bâtiment mais sans utiliser de primitives 3D intermédiaires.

Dans [LDZPD06a, LDZPD06b], Lafarge et al. mettent en œuvre une approche structurale dans un cadre bayésien pour la reconstruction des bâtiments à partir d'un MNE de référence (issu d'images satellitaires optiques à haute résolution) et d'emprises quadrilatérales de bâtiments. La méthode vise à assembler des structures urbaines simples, appartenant à une grammaire de modèles 3D paramétriques adaptés au paysage urbain, en maximisant a posteriori une densité de probabilité formulée selon une énergie de Gibbs. Cette énergie intègre un terme d'a priori modélisant les interactions entre bâtiments voisins en désignant des objets dits "joignables" (selon des critères altimétriques et morphologiques) et un terme d'attache aux données (vraisemblance). Deux types d'algorithmes d'optimisation sont cités : l'un basé sur une méthode stochastique nommée RJMCMC (Reversible Jump Markov Chain Monte Carlo) et l'autre sur une méthode déterministe (Vitterbi). L'approche développée présente l'avantage d'être automatique, adaptative et relativement robuste et générique (conservation de la forme des bâtiments reconstruits).

Dans [FMJ03], Flammand et al. développent une approche opérationnelle pour la reconstruction de modèles 3D de villes à partir d'images aériennes optiques (fournissant un MNS précis) et de plans cadastraux. Il s'agit de déduire la structure 3D des toits de bâtis à partir de la forme 2D des bâtiments de la façon suivante : on génère d'abord un ensemble d'hypothèses recouvrant une large variété de toits possibles à partir d'arêtes centrales de taille modulable (déduites par des considérations de symétrie bidimensionnelle). On recherche ensuite parmi ces hypothèses celles qui collent au mieux avec le MNS généré afin d'obtenir la reconstruction finale.

2.2.4 Conclusions : Performances et limites

Un large nombre de critères de qualité ont été proposés dans la littérature afin d'évaluer les performances atteintes par les méthodes de reconstruction 3D des bâtiments d'une scène urbaine [BPJ⁺06, SW03, Pap98, DJ01] :

- précision de localisation des contours planimétriques extraits,
- précision de localisation des primitives 3D extraites (arêtes 3D et coins 3D),
- précision altimétrique, planimétrique et précision des pentes du MNS ou du MNE de sortie,
- exactitude / dissimilitude des formes 3D reconstruites,
- taux de détection (ou taux d'exhaustivité),
- taux de fausse alarme (ou taux de divergence ou de non existence),
- taux de qualité,
- cohérence géométrique, syntaxique et sémantique des objets 3D (avec tests de continuité, connexité, planéité, fermeture et intersection) ...

Dans [FBG02], Fraser et al. mènent une étude intéressante sur les potentiels et limites de la reconstruction 3D des bâtis en multiscopie optique satellitaire (résolution métrique). Les performances atteintes en matière de précision de localisation des primitives extraites et en matière d'identification de bâtiments sont comparées au cas aérien (résolution sub-métrique). Il est montré que, sur la scène d'étude considérée, environ 15 % des bâtis détectés à partir d'images aériennes ne peuvent pas être modélisés à partir d'images satellitaires. Les problèmes majeurs rencontrés en imagerie satellitaire métrique dans le cadre de la modélisation 3D de scènes urbaines sont cités en conclusion : variabilité de la

qualité des images, identification difficile des bâtiments trop petits, reconstruction précise de la forme des bâtis délicate à réaliser sans généralisation excessive (structures de toit souvent simplifiées).

2.3 Reconstruction des bâtiments en imagerie RADAR

On recense dans la littérature plusieurs travaux récents portant sur la reconstruction tridimensionnelle des bâtiments à partir d'images RADAR à résolution métrique et sub-métrique.

Nous nous intéressons dans la suite aux quatre enjeux majeurs étudiés dans ces travaux :

1. L'extraction de primitives caractéristiques et la génération d'emprises de bâtiments.
2. L'estimation des hauteurs de bâtiments.
3. La génération de MNT en milieu urbain.
4. La génération de MNS ou de MNE et la reconstruction 3D des bâtiments.

Les méthodologies proposées utilisent en entrée d'algorithme des données de différents types (données monoscopiques, radargrammétriques, interférométriques ou multi-aspects), susceptibles de donner lieu à des résultats plus ou moins denses et précis.

On précise que les méthodes basées sur des données RADAR polarimétriques (PolSAR et PolInSAR) ne seront pas présentées dans cette section.

2.3.1 Extraction de primitives caractéristiques et génération d'emprises de bâtiments

En imagerie RADAR, les primitives caractéristiques usuelles (ponctuelles, linéaires ou surfaciques) sont souvent extraites pour décrire partiellement ou totalement les emprises de bâtiments [Cel06, CON06, Sim02, SOG05, SOGLC03].

Les primitives fortement brillantes se réfèrent aux parties des objets présentant une configuration géométrique et des propriétés électro-magnétiques favorables à une forte rétrodiffusion du signal RADAR [FIR02, BBFL09].

Les primitives ponctuelles sont dues à la présence de cibles objets telles que les coins triédriques de bâtis, constitués de matériaux métalliques.

Les primitives linéaires simples existent majoritairement sous trois formes (citées par ordre d'intensité décroissante) :

- les échos doubles ou diédriques qui se manifestent par des lignes très brillantes ; ils sont la signature des coins réflecteurs sol-mur orientés vers le capteur (il s'agit de la somme d'échos apparaissant dans une même cellule de résolution) ;
- les échos simples qui se manifestent par des lignes relativement brillantes ; ils sont souvent la signature des bords de toit (gouttières) ou d'éléments métalliques appartenant au toit (jonction entre deux pans) ;
- les échos triples qui se manifestent avec une amplitude plus faible ; ils sont la signature des échos sol-mur-sol ou mur-sol-mur (contrairement aux échos doubles, ils sont localisés plus loin du capteur et n'apparaissent pas tous dans une même cellule de résolution).

Les primitives surfaciques existent majoritairement sous deux formes :

-
- les zones d’ombres RADAR qui se manifestent par des régions de très faible amplitude composées uniquement de bruit thermique ;
 - les zones de toits (toit seul, toit avec recouvrement) et de façades ; ces primitives apparaissent si la zone imagée est suffisamment rugueuse par rapport à la longueur d’onde utilisée.

Ces primitives peuvent être délicates à extraire et leur apparence dépend du capteur (importance de l’angle d’incidence) et de la scène (matériau de construction des bâtis, agencement des objets environnants, orientation capteur - bâti).

Nous proposons dans la suite de cette section une liste de détecteurs proposés dans la littérature pour extraire ces primitives. Certaines méthodes sont mises en œuvre sur les données brutes et d’autres sur les données filtrées (exemple : utilisation d’un filtrage Mean-Shift [Cel06, CON06]).

Extraction des primitives ponctuelles en monoscopie RADAR Un certain nombre de méthodes ont été proposées dans la littérature pour l’extraction de primitives ponctuelles en monoscopie RADAR, parmi lesquelles :

- le détecteur de cibles brillantes utilisé par Michaelsen et al. dans [MSS02] et basé sur l’utilisation d’un filtre défini par un pixel central et 3 couronnes ;
- le détecteur de cibles brillantes de Petit et al. [PAD01], développé dans le cadre d’une classification floue et permettant d’obtenir en sortie la classe des points brillants et celle des points très brillants (coins réflecteurs) ;
- le détecteur de cibles brillantes de Lopes et al. [LTN90] ; il s’agit d’un détecteur à taux constant de fausse alarme, basé sur le rapport de l’intensité moyenne de deux régions (la région cible et la région fond environnant).

Extraction des primitives linéaires simples en monoscopie RADAR De nombreux algorithmes ont été proposés dans la littérature pour l’extraction de lignes brillantes en monoscopie RADAR, parmi lesquels :

- le détecteur de bords de Touzi et al. [TLB88], basé sur une approche statistique et géométrique ;
 - le détecteur de lignes brillantes de Tupin et al. [Tup97, TMM⁺98, THD02], issu de la fusion de 2 détecteurs à taux constant de fausse alarme ; ce détecteur est développé sous l’hypothèse d’une modélisation de l’amplitude RADAR par une loi de Nakagami et combine, selon une somme associative symétrique, un premier détecteur basé sur un rapport de bords et un second détecteur basé sur une corrélation centrée normée entre deux populations de pixels ;
 - le détecteur de lignes brillantes de Tison [Tis04] et de Simonetto et al. [Sim02, SOG05, SOGLC03], combinant les résultats du détecteur de Tupin et al. et les résultats issus d’une classification markovienne adaptée aux statistiques RADAR (classe des "échos forts" associée à la loi Log-Normale ou à la loi de Fisher) selon un système de vote majoritaire par voisinage ;
 - le détecteur de lignes brillantes de Cellier et al. [Cel06, CON06], basé d’une part sur la détection d’orientations privilégiées et de points d’initialisation à partir d’un maillage binaire (obtenu par transformées de Hough) et d’autre part sur l’algorithme de croissance de régions AROCS proposé par Simonetto et al. [Sim02, SOG05, SOGLC03] et mis ici en œuvre séquentiellement pour la génération de segments ;
 - le détecteur de lignes brillantes de Soergel et al. [STS03], basé sur un seuillage des
-

images en intensité et sur un amincissement des contours allongés, connectés et brillants ;

- le détecteur de lignes brillantes de Quartulli et al. [QD04], développé sous l'hypothèse d'une modélisation de la statistique RADAR par une loi inverse gaussienne généralisée.

Extraction des primitives linéaires complexes en imagerie radargrammétrique, interférométrique ou multi-aspects Des primitives plus complexes, à l'origine d'emprises plus sophistiquées (dites en "L", "T", "U" ou de forme rectangulaire) ont pu être extraites à partir de données plus complètes.

C'est par exemple le cas de :

- Simonetto et al. [Sim02, SOG05, SOGLC03], qui proposent une méthode d'extraction d'emprises brillantes sous forme de rectangles, à partir d'un couple d'images radargrammétiques ; une approche reposant sur un procédé en quatre étapes est proposée : premièrement, la classification des pixels en trois catégories selon leur intensité ; deuxièmement, l'extraction des lignes de forte intensité par application du détecteur de Tupin et al. [Tup97, TMM⁺98, THD02] afin d'obtenir un masque binaire de lignes candidates ; troisièmement, la détection sur ce masque de structures orthogonales (en "L" ou "T") marquées par des croix via l'utilisation de transformées de Hough ; dernièrement, la génération de rectangles candidats par la mise en œuvre de l'algorithme AROCS appliqué uniquement sur les croix validées par une mise en correspondance entre les deux images radargrammétiques ;
- Thiele et al. [TCS⁺07a], qui proposent une méthode d'extraction d'emprises brillantes sous forme de parallélogrammes et de quadrangles à partir de quatre couples d'images InSAR multi-aspects acquises avec des directions de visée orthogonales. Le double intérêt des données disponibles (information interférométrique d'élévation et information multi-aspects à l'origine de résultats denses et complets) est exploité dans l'approche développée. Deux cas sont considérés suivant la taille des bâtiments recherchés (premier cas pour les petits bâtiments et second cas pour les immeubles étendus). L'approche repose sur un procédé en quatre étapes. Premièrement, on extrait des lignes brillantes en "L" (dans le premier cas) et des objets supplémentaires (dans le second cas) tels que les bords brillants face au capteur, les bords de côté et les bords dans l'ombre, à partir des images de magnitude fusionnées et en utilisant des opérateurs de Steger, de Canny et de seuillage. Deuxièmement, on discrimine les lignes brillantes causées par les coins diédriques en utilisant les profils de gradient en magnitude, en cohérence et en hauteur (les échos doubles diédriques se caractérisent par une forte amplitude et une forte cohérence). Troisièmement, on assemble les structures complémentaires détectées par les deux vues orthogonales suivant des contraintes en angle et en distance, afin de générer des parallélogrammes et des quadrangles, appelés candidats potentiels pour les emprises des bâtiments. Dernièrement, on discrimine les emprises effectives parmi les emprises candidates suivant des considérations surfaciques (étendue minimale) et altimétriques (variance maximale de la hauteur interférométrique) à partir d'un MNT et d'un interférogramme ;
- Soergel [Soe06], qui propose un détecteur de lignes brillantes en "L" à partir d'un couple d'images RADAR multi-aspects acquises avec des directions de visée orthogonales ; le détecteur utilise les résultats issus d'un filtre gradient moyen carré [Foe94] et les complète en remplissant les faibles espaces entre primitives coli-

néaires.

Extraction des zones d'ombres RADAR en monoscopie RADAR Plusieurs algorithmes classiques ont été proposés dans la littérature pour l'extraction des ombres en monoscopie RADAR, parmi lesquels :

- le détecteur de Hoepfner [Hoe99], basé sur un filtre binaire rotatif suivant 36 directions et intégrant un critère de force sur l'orientation du contour ; il permet d'extraire la limite entre le bâtiment et son ombre ;
- le détecteur de Bennett et al. [BB03b], basé sur les contours actifs statistiques ("snakes"). Il s'agit de délimiter le bâtiment et son ombre de l'environnement extérieur en faisant croître des régions délimitées par un contour actif (formé de noeuds reliés entre eux) à partir d'une grille initiale. On cherche simultanément à minimiser la longueur du périmètre du contour actif et à maximiser la probabilité que les deux régions (intérieure et extérieure) suivent des distributions de probabilité exponentielles négatives différentes ;
- les deux détecteurs de Bolter [Bol00, Bol01]. Le premier est basé sur une stratégie simple combinant des outils radiométriques (seuillage sur les niveaux de gris après application d'un filtre passe-bas sur les données en amplitude), des outils géométriques (suppression des régions de trop petite taille) et des outils morphologiques. Le second est basé sur un filtre binaire rotatif suivant 16 directions et intègre un critère de force de contour.

Extraction des zones d'ombres RADAR en imagerie interférométrique Des stratégies plus élaborées ont pu être mises en œuvre sur des données interférométriques.

C'est par exemple le cas de :

- Cellier et al. [Cel06, CON06], qui proposent une méthode d'extraction des ombres en cinq étapes, basée sur des considérations géométriques, radiométriques et utilisant la nature cohérente du signal. Premièrement, un filtre Mean Shift est appliqué sur l'image d'amplitude et de cohérence (les ombres s'y caractérisent par une faible amplitude et une faible cohérence). Deuxièmement, une image fusionnée est obtenue par intersection des images précédentes filtrées et seuillées. Troisièmement, les contours des ombres sont détectés par un algorithme de croissance de régions et sont polygonalisés par un algorithme classique de la corde afin d'obtenir un premier maillage des ombres. Quatrièmement, un second maillage des ombres est obtenu par une transformée de Hough mettant en évidence les trois directions privilégiées des ombres. Dernièrement, les deux maillages sont fusionnés par minimisation d'une distance et sont régularisés au niveau de leur orientation et de leur largeur ;
- Tison [Tis04], qui propose une méthode d'extraction des ombres en trois étapes, basée sur des considérations statistiques, radiométriques, morphologiques et utilisant la nature cohérente du signal. Un processus de classification markovienne est mis en œuvre conjointement sur l'image d'amplitude (où la classe d'ombres se caractérise par la plus faible amplitude) et sur l'image de cohérence (où la classe d'ombres se caractérise par la plus faible cohérence).

Extraction des zones de toits et de façades en monoscopie RADAR Plusieurs algorithmes ont été proposés dans la littérature pour l'extraction en monoscopie RADAR de

zones homogènes (indépendamment du phénomène de speckle), telles que les zones de toits (avec ou sans recouvrement) et les zones de façades, parmi lesquels :

- ceux basés sur des outils géométriques et radiométriques ; c'est par exemple le cas de Cellier et al. [Cel06, CON06], qui proposent une méthode d'extraction des zones de toits en quatre étapes, similaire à celle proposée pour l'extraction des ombres. Deux maillages des zones de toits sont combinés. Le premier est issu d'une segmentation par croissance de régions et d'une polygonalisation. Le second est issu d'une transformée de Hough mettant en évidence les deux directions privilégiées des régions extraites ;
- ceux se référant à des procédés statistiques (segmentation et classification avec apprentissage), sous l'hypothèse de zones homogènes suivant des distributions de probabilités adaptées à la statistique RADAR en imagerie urbaine haute résolution (exemples : contours actifs appliqués avec des lois de Fisher [LGNT06], "grille active" [GR01] appliquée avec des lois Gamma [PDR09] ou des lois de Fisher [GNS⁺09]).

Extraction des zones de toits et de façades en imagerie interférométrique Des stratégies plus élaborées ont pu être mises en œuvre sur des données interférométriques.

C'est par exemple le cas de :

- Tison [Tis04], qui propose une méthode d'extraction des toits en trois étapes, similaire à celle proposée pour l'extraction des ombres, dans le cadre d'une classification markovienne. Cette fois-ci, on utilise les trois classes de labels "toit sombre", "toit moyen" et "toit clair-coin réflecteur", caractérisées par une très forte cohérence.

2.3.2 Estimation des hauteurs de bâtiments

En ce qui concerne la problématique isolée d'estimation des hauteurs de bâtiments (hauteur à la base du toit, hauteur au sommet du toit ou hauteur le long d'arêtes sur le toit), différentes méthodes peuvent être mises en œuvre en fonction des données disponibles en entrée :

- en monoscopie RADAR, par mesure de la "zone de recouvrement", par mesure de la "zone d'ombre", par mise en œuvre d'approches électro-magnétiques, ou encore par mise en œuvre d'approches itératives en simulation RADAR ;
- en radargrammétrie, en exploitant le principe de la stéréoscopie ;
- en interférométrie, en exploitant l'information de phase ;
- à partir de données RADAR multi-aspects, qui permettent d'améliorer l'estimation des hauteurs de bâtis par fusion d'informations.

Mesure de la "zone de recouvrement" en monoscopie RADAR (et extension en multi-scopie RADAR) Lorsqu'une seule image RADAR est disponible, la hauteur des bâtiments peut être extraite en utilisant la formule suivante, dans le cas d'un bâti parallélépipédique à toit plat :

$$h_{bâti} = \frac{L_{recouvrement}}{\cos(\theta_i)} \quad \text{si } h_{bâti} < h_{limite} \quad (\text{configuration 1}) \quad (2.1)$$

avec les notations suivantes (voir figure 2.1) :

- $h_{bâti}$ correspond à la hauteur du bâtiment ;
- θ_i correspond à l'angle d'incidence du capteur ;

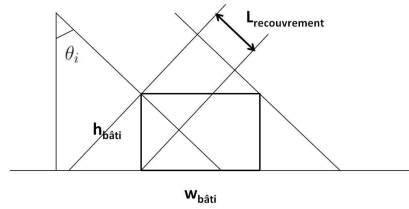
(a) Configuration 1 : $h_{b\hat{a}ti} < h_{limite}$

FIGURE 2.1 – Estimation de la hauteur des bâtis à partir de la "zone de recouvrement".

- $L_{recouvrement}$ correspond à l'extension de la "zone de recouvrement" projetée en "géométrie RADAR" suivant l'axe des distances RADAR.

L'équation 2.1 est valable lorsque la hauteur du bâtiment est suffisamment grande devant son "épaisseur vue" en distance, c'est-à-dire lorsque la hauteur du bâti $h_{b\hat{a}ti}$ est inférieure à une valeur limite h_{limite} définie par :

$$h_{limite} = w_{b\hat{a}ti} \tan(\theta_i) \quad (2.2)$$

avec la notation suivante :

- $w_{b\hat{a}ti}$ correspond à "l'épaisseur vue" du bâtiment en "géométrie terrain" suivant l'axe des distances RADAR.

La bande de recouvrement imagée avec une forte intensité se définit par un premier bord (en hauteur) situé à la limite avec la "zone de fond" (et potentiellement marqué par des échos simples issus de structures en bordure de toit) et par un second bord (au sol) situé à la limite avec la "zone d'écho double" (au niveau de la jonction sol-mur)

C'est cette relation qui est exploitée par :

- Tupin [Tup03], qui propose une méthode de sélection automatique des bords de "zones recouvrement" parmi un ensemble de segments brillants extraits, en intégrant des critères d'alignement et de proximité et en estimant la hauteur des bâtis via l'utilisation d'un histogramme de hauteurs construit selon un système de vote ;
- Simonetto et al. [Sim02, SOG05, SOGLC03], qui fusionnent plusieurs méthodes d'estimation de la hauteur de structures industrielles, dont l'une d'entre elles est basée sur la mesure de la "zone de recouvrement", pour des bâtis avec une façade face au capteur et sous l'hypothèse que les primitives de type écho simple (bords de toits) et de type écho double (coins mur-sol) soient toutes deux visibles.

Mesure de la "zone d'ombre" en monoscopie RADAR (et extension en multiscopie RADAR) Lorsqu'une seule image RADAR est disponible, la hauteur des bâtiments peut être extraite en utilisant l'une des formules suivantes, dans le cas d'un bâti parallélépipédique à toit plat :

$$h_{b\hat{a}ti} = L_{ombre} \cos(\theta_i) \quad \text{si } h_{b\hat{a}ti} < h_{limite} \quad (\text{configuration 1}) \quad (2.3)$$

et

$$h_{b\hat{a}ti} = \frac{1}{\tan(\theta_i)} \left(\frac{L_{ombre}}{\sin(\theta_i)} - w_{b\hat{a}ti} \right) \quad \text{sinon} \quad (\text{configuration 2}) \quad (2.4)$$

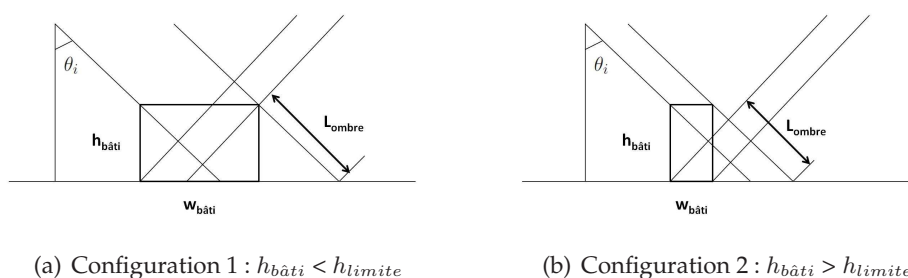


FIGURE 2.2 – Estimation de la hauteur des bâtis à partir de la "zone d'ombre".

avec la notation suivante (voir figure 2.2) :

- L_{ombre} correspond à l'extension de la "zone d'ombre" projetée en "géométrie RADAR" suivant l'axe des distances RADAR.

C'est cette relation qui est exploitée par :

- Bennett et al. [BB03b, BB03a], qui utilisent les ombres pour la détermination des hauteurs et des largeurs de bâtis en contexte multiscopique. Bennett et al. étendent les équations précédentes au cas de différents types de bâtis (toit plat, toit bi-pentes) et précisent que les résultats fournis ne sont satisfaisants que pour des bâtis de dimensions suffisamment grandes (suivant l'axe des élévations et suivant celui des distances RADAR) ;
- Bolter [Bol00], qui utilise également la largeur des ombres dans les images afin d'estimer conjointement les hauteurs et les positions des bâtis en contexte multiscopique. La méthodologie mise en œuvre se base sur la fusion de résultats d'estimation de hauteurs issus d'un jeu de quatre images acquises selon les quatre points cardinaux ;
- Tison et al. [TTM04], qui proposent, dans un contexte interférométrique, de déduire et de régulariser la forme et la hauteur des bâtis à partir d'ombres extraites.

Remarque sur les méthodes d'estimation des hauteurs de bâtiments à partir de mesures monoscopiques (sur la "zone d'ombre" et sur la "zone de recouvrement") Il convient de souligner que les méthodes basées sur une unique mesure d'ombre ou de recouvrement requièrent souvent l'intervention manuelle d'un utilisateur et demeurent peu précises et peu fiables en contexte monoscopique. Elles sont donc difficilement exploitables en milieu opérationnel : fortes hypothèses de visibilité rarement vérifiées en milieu urbain et dépendant des propriétés électro-magnétiques des bâtiments, mesures délicates à cause des distorsions géométriques et de la présence d'échos perturbateurs (échos triples).

Toutefois, en contexte multiscopique, elles sont susceptibles d'apporter une information complémentaire exploitable et s'avèrent beaucoup plus robustes.

Mise en œuvre d'approches électro-magnétiques en monoscopie RADAR (et extension en multiscopie RADAR) Lorsqu'une seule image RADAR est disponible, la hauteur des bâtiments peut également être estimée par mise en œuvre d'approches électro-magnétiques, en supposant connue la localisation des emprises planimétriques des bâtiments.

C'est par exemple le cas de Franceschetti et al. [FGI⁺07], qui proposent une approche déterministe pour l'extraction des hauteurs de bâtis en monoscopie RADAR. Il s'agit d'utiliser, après extraction des échos linéiques diédriques, l'équation reliant la hauteur d'un bâti à l'ensemble de paramètres suivants : la section RADAR efficace, l'incidence de l'onde RADAR, la longueur et l'orientation du bâti, les paramètres de rugosité du sol environnant et la matrice de rétro-diffusion. L'hypothèse de toit plat et l'approximation géométrique de Kirchhoff sont supposées valides. Cette approche permet d'estimer avec précision des hauteurs de bâtis (à quelques centimètres près) mais nécessite de nombreuses connaissances a priori. De plus, elle requiert une phase de calibration, permettant de mettre en correspondance la hauteur d'un bâti et sa contribution en intensité au sein de la section RADAR efficace.

Mise en œuvre d'approches itératives par simulation en monoscopie RADAR (et extension en multiscopie RADAR) À nouveau, lorsqu'une seule image RADAR est disponible, la hauteur des bâtiments peut également être estimée par mise en œuvre d'approches itératives basées sur des stratégies de type "génération d'hypothèses de hauteurs - simulation - vérification", en supposant là encore connue la localisation des emprises planimétriques des bâtiments.

C'est par exemple le cas de Brunner et al. [BLB08, BLBG10], qui développent un algorithme d'estimation des hauteurs, applicable en monoscopie RADAR (ou dans le cadre de la fusion optique-RADAR) et fournissant des résultats très intéressants. Cet algorithme, nommé "génération d'hypothèses - restitution - appariement", sera détaillé en section 2.4.2.2.

Exploitation du principe de la stéréoscopie en radargrammétrie Il s'agit d'exploiter le principe de la stéréoscopie consistant à remonter à une information de hauteur, en utilisant les propriétés de transformation des distances par projection, lorsqu'une même scène est acquise par deux capteurs différents.

Après ré-échantillonnage et mise en géométrie épipolaire, des cartes appelées cartes de disparité sont calculées par mise en correspondance de pixels ou primitives homologues. L'information 3D est ensuite estimée en utilisant la relation entre hauteur et disparité, dérivant des équations fondamentales du système d'acquisition RADAR (équation du "cône doppler", équation de la "sphère distance" et "équation de surface").

Lors de la phase d'appariement, différents types de corrélateurs peuvent être mis en œuvre. Il peut s'agir de corrélateurs classiques (tel que le coefficient de corrélation centré normé) utilisés après filtrage des données RADAR ou de corrélateurs adaptés utilisés sur les données brutes, parmi lesquels : les corrélateurs adaptés de Tupin et al. [TN02], basés sur un critère de rapport et sur un critère lié au coefficient de variation ; le corrélateur spécifique de Nocera [Noc96], déduit d'une approche probabiliste ; les corrélateurs spécifiques d'Hervet [Her00], basés sur les intensités et sur les réflectivités sous-jacentes des pixels avoisinants ; ou encore le corrélateur adapté de Deledalle et al. [DDT09], s'appuyant sur une mesure de similarité probabiliste.

Dans [Sim02, SOG05, SOGLC03], Simonetto et al. proposent d'estimer la hauteur de bâtiments industriels rectangulaires ou circulaires à toits plats, à partir de couples d'images RAMSES aéroportées en configuration radargrammétrique. L'extraction des hauteurs des structures se fait par fusion d'informations, en combinant deux mesures stéréoscopiques : premièrement, la mesure de la hauteur moyenne des pixels situés à l'intérieur d'emprises de bâtis dilatées et deuxièmement, la mesure de la hauteur moyenne

le long de fines primitives toits appariées sur les deux images. On obtient une estimation des hauteurs qui se révèle très précise pour les bâtis de taille grande ou moyenne. La principale limitation de cette méthode réside dans l'absence, pour certains bâtis, d'éléments en hauteur susceptibles de répondre fortement dans les deux images.

Dans [WE10], Weydahl et al. proposent d'estimer les hauteurs de très grands bâtis (avec une élévation de plusieurs centaines de mètres) à partir d'un couple satellitaire stéréoscopique TerraSAR-X en mode SpotLight. L'information de hauteur est issue de la mise en correspondance de points d'intérêt situés au niveau des toits (cibles brillantes ponctuelles associées à de petites fenêtres) sur les deux images. D'après les auteurs, on obtient une estimation très précise des hauteurs.

Exploitation de l'information de phase en interférométrie Il s'agit d'exploiter l'information de phase contenue dans le signal RADAR complexe afin de remonter à une information de hauteur.

Différentes étapes sont impliquées dans le procédé d'interférométrie RADAR. Premièrement, après recalage du couple d'images, le produit interférométrique est calculé en intégrant un filtrage multi-look et la correction des franges topographiques. On rappelle que l'argument du produit interférométrique correspond à "l'image de phase" ou à "l'interférogramme" et que son module correspond à "l'image de cohérence". Deuxièmement, les "zones d'ombres", sources de bruit dans l'interférogramme, sont supprimées. Troisièmement, les franges de l'interférogramme sont déroulées et l'information de phase est convertie en information d'élévation. On obtient ainsi une information de hauteur relative pour chaque point de la scène par rapport à une surface de référence.

Comme souligné dans [SST03], la fabrication d'interférogrammes pour l'estimation de la hauteur de bâtiments reste très délicate en milieu urbain : présence de fortes discontinuités, présence de zones de recouvrement avec mélange des phases et de zones d'ombres avec très faible cohérence, existence d'occlusions aboutissant à des reconstructions partielles, phénomène perturbateur des propagations multi-rebonds.

Dans [Cel06, CON06], l'estimation de la hauteur des bâtiments est issue d'une combinaison de mesures : une mesure monoscopique et une mesure interférométrique. Cette dernière utilise un modèle original de mélange altimétrique au niveau des "zones de repliement" entre échos forts afin d'améliorer les hauteurs estimées.

Dans [Tis04, TTM07], une approche jointe de classification markovienne et d'estimation des hauteurs (avec régularisation d'interférogramme) est mise en œuvre sur un graphe de régions à partir d'un couple interférométrique. On cherche à optimiser conjointement la hauteur d'une région et son label via l'utilisation de relations contextuelles. En sortie du processus d'optimisation, on obtient un MNS-MNE très satisfaisant.

Récemment, des méthodologies plus complexes, développées pour la construction d'interférogrammes en milieu urbain et dédiées au traitement de données à résolution métrique et à leur mode d'acquisition spécifique (cas notamment du mode SpotLight) ont été étudiées et intégrées au sein de chaînes interférométriques [EAB⁺09].

Estimation des hauteurs de bâtis à partir de données multi-aspects Il est également possible d'améliorer l'estimation des hauteurs de bâtis en fusionnant l'information issue de données multi-aspects (les techniques utilisées étant alors inapplicables aux données satellitaires).

En contexte radargrammétrique, Goel et al. [GA10] développent une méthode d'estimation de hauteurs à travers une approche bayésienne mise en œuvre sur un jeu de trois

images TerraSAR-X radargrammétriques acquises pour différents angles d'incidence. L'estimation de la hauteur de certains points d'intérêt stables et fortement réflecteurs (appelés "permanent scatterers"), situés sur la structure urbaine d'intérêt, est réalisée dans un cadre bayésien à partir d'un modèle de graphe dirigé.

En contexte interférométrique, Bolter et al. [Bol00] estiment une carte des hauteurs interférométriques par fusion d'interférogrammes, calculés pour quatre directions de visée orthogonales, en retenant pour chaque point la hauteur issue du couple présentant un coefficient de corrélation maximal. Cela permet notamment de supprimer des zones de repliement.

En contexte multi-aspects, Thiele et al. [TCS⁺07a] estiment directement la hauteur des bâtiments à partir d'une mesure interférométrique. L'image des phases est soumise à un recalage sous-pixelaire basé sur la maximisation de la cohérence. Afin de compenser les problèmes d'ambiguïté de phase pour les grands bâtiments et de gérer les discontinuités de hauteurs, les phases dont la valeur est significativement en dessous de la moyenne locale sont rehaussées de 2π .

2.3.3 Génération de MNT en milieu urbain

Un MNT fournit l'information d'élévation au niveau du sol (sans prise en considération des objets du sursol).

En contexte radargrammétrique, Simonetto et al. [Sim02, SOG05, SOGLC03] génèrent un MNT en milieu urbain à partir d'un couple unique d'images RADAR aéroportées à résolution submétrique et suite à la détection préalable des objets du sursol. Les cartes de disparités sont estimées par une mise en correspondance dense selon un algorithme pyramidal à 3 niveaux et suivant une stratégie d'aller-retour assurant la contrainte d'unicité. L'appariement est effectué sur les pixels n'appartenant pas à la classe des pixels dits "brillants". Les régions correspondant aux emprises dilatées des objets détectés sont également ôtées de la carte des disparités. Le MNT ainsi obtenu est soumis à deux post-traitements : la suppression des hauteurs jugées aberrantes et une interpolation du MNT sur les zones non couvertes après triangulation de Delaunay de la scène. En outre, localement, l'altitude au sol est estimée à proximité des bâtiments par stéréoscopie appliquée aux échos doubles.

En contexte interférométrique, Cellier et al. [Cel06] déterminent l'altitude au sol par une méthode de croissance de régions appliquée sur un interférogramme, issu d'un couple d'images RADAR aéroportées.

En contexte multi-aspects, Thiele et al. [TCS⁺07a] développent une approche pour la génération de MNT en milieu urbain à partir de données interférométriques relatives à quatre couples d'images. La scène imagée ne présentant pas de brutales descentes en profondeur au niveau du sol, le déroulement des franges est réalisé aisément. Toutefois, un lissage des hauteurs interférométriques est mis en œuvre dans les régions présentant une faible cohérence et de fortes variances de hauteurs. Dans ce type de régions, la hauteur est fixée selon la hauteur locale moyenne de terrain.

2.3.4 Génération de MNS ou de MNE et reconstruction 3D des bâtiments

Le MNE précis d'une scène urbaine résulte de la combinaison du MNT généré avec les structures 3D des bâtis détectées et reconstruites.

Les différentes méthodes recensées dans la littérature pour la reconstruction du bâti 3D en imagerie RADAR à haute résolution sont :

- indirectes, lorsqu'elles se fondent sur l'usage de primitives 2D-3D intermédiaires ; ces méthodes sont largement répandues et se décomposent en trois catégories (approches "basées primitives 2D + hauteurs", approches "basées primitives 3D sans modèle 3D" et approches "basées primitives 3D avec modèles 3D") ;
- ou directes, lorsque l'objet 3D est reconstruit sans l'usage de telles primitives (approches "basées objets 3D") ; ces méthodes sont moins largement répandues et font appel à une stratégie de reconstruction 3D basée sur une unique étape d'estimation jointe des paramètres planimétriques et altimétriques décrivant l'objet 3D.

Approches indirectes "basées primitives 2D + hauteurs" En contexte radargrammétrique, Simonetto et al. [Sim02, SOG05, SOGLC03] proposent une méthode de reconstruction de bâtiments industriels (cuves et parallélépipèdes 3D), en caractérisant simplement les structures 3D bâties par leur emprise 2D (circulaire ou rectangulaire) et leur hauteur (chaque bâti n'étant associé qu'à une seule hauteur). Les hauteurs sursol / sol sont issues de différentes méthodes d'estimation (mesure par monoscopie ou par stéréoscopie). Celles-ci sont validées comme correspondant à la hauteur du toit ou du terrain selon une technique de seuillage et sont ensuite fusionnées selon un estimateur du maximum de vraisemblance. Le MNE est finalement construit par simple association des informations planimétriques relatives aux primitives 2D extraites (localisation et dimensions des emprises affinées) et des informations altimétriques estimées (hauteurs finales des bâtis).

En contexte interférométrique multi-aspects, Thiele et al. [TCS⁺07a] génèrent le MNE d'une scène urbaine en associant les informations planimétriques d'emprises sélectionnées (et modélisées par leur boîte englobante) avec les informations d'élévation issues d'un interférogramme, construit à partir de paires d'images acquises le long de deux trajectoires orthogonales.

Approches indirectes "basées primitives 3D sans modèle 3D" En contexte radargrammétrique multi-aspects, Soergel et al. [Soe06] développent une méthode de reconstruction des bâtiments basée sur l'extraction d'angles 3D, à partir d'un couple d'images RADAR acquises suivant des trajectoires orthogonales. L'idée consiste à apparier, suivant une sorte de contrainte épipolaire adaptée à 45°, des couples d'angles 2D, extraits indépendamment sur chaque image par assemblage de formes brillantes en "L" adjacentes et non colinéaires. La mise en correspondance d'angles 2D issus d'une même structure est réalisée sous hypothèse de hauteurs croissantes et en utilisant la relation entre la hauteur d'un bâti et la position associée de la "zone de recouvrement". Les jeux d'angles 3D (primitives 3D) ainsi générés sont alors utilisés pour la reconstruction tridimensionnelle des toits de bâtiments.

En contexte interférométrique multi-aspects, Bolter [Bol00] propose de reconstruire le MNE d'une scène urbaine à partir de quatre images acquises suivant des trajectoires orthogonales. La technique présentée se base sur l'utilisation des ombres (primitives 2D) pour la détection des quatre murs orthogonaux décrivant un bâti, ainsi que sur une technique d'interpolation permettant de retrouver les formes globales des toits (primitives 3D). Les informations de localisation et de forme des structures 3D sont ensuite combinées aux informations d'élévation issues de la carte des hauteurs interférométriques afin de générer le MNE de la scène.

Approches indirectes "basées primitives 3D avec modèles 3D" En contexte interférométrique, Cellier et al. [Cel06, CON06] proposent une approche fondée sur la gestion d'hypothèses à partir d'un ensemble incomplet de primitives pour la reconstruction de bâtiments à toits plats et à emprises rectangulaires. Après extraction des primitives 2D de niveau 0 (échos, ombres et toits), celles-ci sont interprétées et reliées à des éléments ou primitives 3D du bâti de niveau 1 (arêtes, toits ou façades), en utilisant des informations de hauteurs issues de mesures monoscopiques et interférométriques. La gestion d'hypothèses (validation, infirmation ou classification des hypothèses) permet alors d'accéder aux primitives 3D de niveau 2 (sous-parties des bâtiments). Les façades sont localisées grâce aux segments de plus fort rapport signal à bruit. Des séries de règles géométriques sur des paires de segments consécutifs permettent de formuler, par analyse de contraintes et d'interactions, une ou plusieurs hypothèses sur chacune des sous-parties. Enfin, une combinaison de ces hypothèses suivant un ensemble de modèles 3D envisagés permet d'atteindre le niveau 3, c'est-à-dire la reconstruction 3D du bâtiment complet.

Approches directes "basées objet 3D" En contexte monoscopique RADAR, une approche stochastique est développée dans [QD04] pour la reconstruction 3D de bâtis à partir d'une unique image RADAR aéroportée à résolution métrique. Quartulli et al. se placent dans un cadre bayésien et définissent une hiérarchie des éléments de la scène : la scène est composée d'objets, constitués de plusieurs facettes, issues elles-mêmes d'un regroupement de pixels et possédant certaines propriétés géométriques et radiométriques. Un processus ponctuel marqué permettant la description paramétrique d'objets (caractérisés par leur position et leurs attributs) est mis en œuvre afin de maximiser une probabilité a posteriori, composée d'un terme de régularisation et d'un terme de vraisemblance. Ce dernier fait l'hypothèse d'une modélisation de la répartition des centres des objets par une loi de Poisson. Chaque facette est détectée suivant une décision du rapport de vraisemblance. Une procédure de recuit simulé permet enfin de converger vers différentes formes de bâtiments. Toutefois, les résultats de reconstruction 3D fournis restent très limités (beaucoup de bâtiments sont perdus).

Dans [LT07], Lemoigne et al. décrivent une approche générique par contours actifs pour la reconstruction 3D de bâtiments à emprises polygonales. La méthodologie présentée traite le cas où les données d'entrée sont composées d'un couple InSAR et d'une image optique (voir section 2.4.3.3). Toutefois, une méthode similaire peut être appliquée dans le cas où seul un couple interférométrique est disponible. Le principe consiste à optimiser, dans un espace 3D cartographique, les positions des noeuds définissant les sommets de bâtis potentiels, en maximisant un critère de Log-Vraisemblance RADAR, calculé conjointement sur les images d'amplitude, de phase et de cohérence.

2.3.5 Conclusions : Performances et limites

Les résultats de reconstruction 3D de bâtiments obtenus en imagerie RADAR sont souvent moins précis que ceux obtenus en imagerie optique pour les deux raisons suivantes.

Premièrement, la résolution des données RADAR satellitaires (métrique) est généralement moins bonne que celle des données optiques satellitaires (submétrique). Cela induit nécessairement une détection des bâtiments moins fine et empêche la reconstruction de certains éléments de petite taille (cheminées ...).

Deuxièmement, la difficulté inhérente aux images RADAR (due au bruit de speckle et aux

déformations géométriques) rend fréquemment nécessaire l'usage de primitives 2D-3D intermédiaires (approches indirectes). Cela génère des résultats souvent trop simplifiés, erronés ou incomplets (primitives partiellement vues ou non vues à l'origine de non-détections ou d'erreurs de dimensionnement).

De façon générale, les résultats de reconstruction obtenus sont souvent bruités sur le plan altimétrique (comme c'est le cas de l'information d'élévation issue d'un interférogramme non régularisé) et éparses sur le plan planimétrique (comme c'est le cas lorsque la détection des bâtiments se fait en monoscopie RADAR ou par radargrammétrie).

Toutefois, l'imagerie RADAR présente certains avantages majeurs susceptibles d'apporter une contribution notable pour la reconstruction 3D des bâtiments en milieu urbain. Premièrement, la composante en amplitude du signal RADAR met en évidence un certain nombre de primitives linéaires brillantes caractéristiques du bâti, qui permettent une focalisation directe vers les zones d'intérêt (c'est notamment le cas des primitives issues de matériaux métalliques non nécessairement rapidement identifiables dans l'optique). En outre, l'information qu'elles apportent en terme d'orientations privilégiées apparaît comme très fiable. Par ailleurs, la présence d'un ensemble complet et large de régions caractéristiques différentes ("zone de recouvrement", "zone d'ombre" ...) permet l'obtention de résultats plus robustes.

Deuxièmement, la composante en phase fournit une information d'élévation tout à fait intéressante à exploiter pour estimer la hauteur des bâtiments.

Une première reconstruction 3D globale et approximative des scènes urbaines à partir d'images RADAR uniquement peut ainsi être envisagée. Des prévisions optimistes, quant à l'exploitation des données issues des futurs capteurs RADAR pour la reconstruction 3D précise de zones urbaines, peuvent être émises.

Enfin, la complémentarité des informations issues des deux types d'images optiques-RADAR nous oriente naturellement vers leur utilisation jointe dans le cadre de la fusion de données.

2.4 Reconstruction des bâtiments par fusion d'informations optiques-RADAR

Dans le cadre de la fusion de données, plusieurs pistes, combinant des informations optiques et RADAR ont été récemment explorées en vue de la détection et de la reconstruction de bâtiments en milieu urbain.

Nous nous intéressons dans la suite aux trois problématiques majeures investies dans ces travaux :

1. Le recalage de données optiques et RADAR à résolution métrique.
2. La reconstruction 3D à partir d'une image optique et d'une image RADAR.
3. La reconstruction 3D à partir d'une image optique et d'un couple RADAR.

2.4.1 Recalage de données optiques et RADAR à résolution métrique

La problématique de recalage est étroitement liée au contexte de la fusion de données pour la reconstruction 3D de structures urbaines. De nouvelles approches dédiées à l'appariement ou à la superposition de primitives homologues spécifiques aux données à résolution métrique ont été développées au cours des dernières années [SR10, TR03, HS05, IG04].

Parmi les approches de recalage, on distingue :

- les approches "basées modèles de projection" ; dans ce cas, on dispose soit de la connaissance d'un modèle physique (avec paramètres du capteur et de la plateforme d'acquisition) soit d'un modèle polynômial (ou sous la forme d'un rapport de deux polynômes) ; un référentiel terrestre de géoréférencement (cartographique, géographique ou cartésien) est alors utilisé comme référentiel intermédiaire entre l'image de départ et celle d'arrivée ;
- les approches "basées données images" ; dans ce cas, aucune donnée auxiliaire n'est nécessaire et le recalage se fonde sur des mesures de similarité directement calculées sur les données images.

Approches "basées modèles de projection" Dans le cas des approches "basées modèles de projection", on utilise les fonctions dites fonctions de localisation directe et inverse pour projeter des primitives d'une image vers une autre. Lorsque les modèles physiques sont disponibles, ces transformations utilisent directement les équations fondamentales optiques et RADAR relatives aux systèmes imageurs [TP94]. Ces dernières mettent en relation les coordonnées pixels (2D) d'un point dans un référentiel image avec ses coordonnées physiques (3D) dans un référentiel terrestre. Ces fonctions utilisent la connaissance des paramètres des systèmes d'acquisition et nécessitent l'introduction d'une information de hauteur (en pratique, celle-ci est souvent fournie par un MNT donnant la hauteur des points au sol ou peut faire l'objet d'hypothèses quant à l'élévation des objets du sur-sol).

Dans [TR03, TR04, Tup04, TR05], Tupin et al. mettent en œuvre une telle méthode de projection sur des données aéroportées afin d'apparier des primitives homologues optiques et RADAR, extraites à partir de détecteurs de cibles linéaires ou ponctuelles sur l'image RADAR et à partir d'un détecteur de contours sur l'image optique.

Étant données les erreurs résiduelles restant après projection (majoritairement dues aux imprécisions du MNT utilisé), cette étape de projection est souvent suivie d'une étape d'affinage de recalage "basée données images", employant les stratégies présentées dans le paragraphe suivant.

Approches "basées données images" Dans le cas des approches "basées données images", deux types de méthodes sont proposées dans la littérature :

- les méthodes classiques de recalage global, où l'on cherche à estimer sur l'ensemble de l'image un modèle analytique ou paramétrique de déformation ;
- les méthodes de recalage local, visant à estimer une grille de déformation, c'est-à-dire à fournir un échantillonnage de la carte de déformation.

Dans [TD07, LTT⁺08, Leh10], une méthode de recalage global basée sur un procédé hiérarchique est présentée. Dans un premier temps, des primitives sont extraites sur les données optiques (détection des contours de Canny-Deriche) et sur les données RADAR (détection des échos linéaires avec le détecteur de Tupin et al. [Tup97, TMM⁺98]). Dans un deuxième temps, les paramètres d'un modèle rigide de déformation sont estimés suivant une approche multi-échelles : les paramètres de translation sont estimés via une mesure de similarité et les paramètres de rotation et d'homothétie sont estimés via une transformation de Fourier-Mellin. Dans un dernier temps, le recalage est affiné par l'estimation d'une transformation polynômiale par mise en correspondance de points d'intérêt (issus d'un détecteur de Harris).

Un procédé similaire et robuste, permettant l'obtention de résultats de recalage satisfaisants, est proposé par Hong et al. dans [HS05], sous l'hypothèse de sol localement plat. Dans un premier temps, un simple décalage est estimé par appariement de points de contours en utilisant successivement un coefficient de corrélation (fournissant des décalages candidats) et une fonction de coût géométrique (sélectionnant le décalage optimal). Dans un second temps, un affinage polynômial est estimé par utilisation de points de contours pré-appariés.

Dans [IG04], une méthode de recalage fin et local est présentée et appliquée sur des images satellitaires SPOT-5 et ERS-2. À travers l'utilisation de mesures de similarité locales, Inglada et al. estiment une grille dense de déformations de façon totalement automatique. Le recalage sub-pixelique obtenu en fin de traitement s'avère très intéressant pour certaines applications en fusion de données à haute résolution (reconstruction 3D, détection de changements, ...).

Des mesures de similarité locales, visant à appairer des primitives homologues optiques et RADAR, sont utilisées pour les deux types de méthodes "basées données images".

On peut citer, parmi les mesures d'appariement étudiées et comparées dans [STC07, IG04], les critères suivants :

- le coefficient ou le rapport de corrélation,
- le critère de Woods,
- les mesures de similarité appartenant à la famille des f-divergences (le rapport de vraisemblance, l'information mutuelle, les distances de Kullback et de Kolmogorov, la divergence du χ^2),
- l'indice "Cluster Reward Algorithm" (CRA).

Certaines de ces mesures se révèlent particulièrement bien adaptées au recalage de données multi-modales ; c'est le cas de l'information mutuelle et de l'indice CRA.

2.4.2 Reconstruction 3D à partir d'une image optique et d'une image RADAR

La configuration d'entrée "une image optique - une image RADAR" fait l'objet de peu de travaux présentés dans la littérature.

Il s'agit d'un cas difficile à exploiter, puisque aucune des méthodes connues (stéréoscopie, radargrammétrie et interférométrie) permettant de remonter à une information de hauteur à partir d'un couple d'images ne peut être mise en œuvre.

Pour beaucoup d'approches, l'intervention partielle d'un opérateur et l'addition de certaines données auxiliaires sont souvent requises.

En outre, les différentes composantes impliquées dans la problématique de reconstruction 3D de bâtiments par fusion optique-RADAR ne sont pas systématiquement toutes traitées dans les chaînes proposées.

Dans plusieurs travaux, les indices caractéristiques du bâti sont indépendamment extraits sur chacune des images (optique-RADAR) par l'intermédiaire de primitives. Celles-ci recouvrent un aspect sémantique précis grâce à la résolution des données satellitaires aujourd'hui disponibles. Les informations complémentaires 2D-3D ainsi obtenues sont ensuite fusionnées afin de reconstruire la scène urbaine.

C'est dans ce cadre de travail en plein essor que s'inscrivent nos travaux de thèse.

2.4.2.1 Détection de bâtiments et extraction de leurs contours

Nous proposons de décomposer les approches de détection de bâtiments suivant trois classes :

- les approches "basées contours optiques-RADAR" ;
- les approches "basées régions optiques-RADAR" ;
- les approches "basées contours et régions optiques-RADAR".

Approches "basées contours optiques-RADAR" Dans [TR03], Tupin et al. développent une approche de reconnaissance des contours de bâtis (rectangles ou composition de rectangles) dans une image optique, à partir de données RADAR qui en contraignent la recherche. Dans un premier temps, deux types de primitives sont extraites : les échos doubles linéiques sur l'image RADAR et les contours de Canny-Deriche sur l'image optique. Dans un deuxième temps, ces dernières sont épaissies, polygonalisées et filtrées suivant des critères de distance et d'orientation vis-à-vis des échos RADAR. Dans un dernier temps, la génération des contours de bâtis est réalisée sur l'image optique : les segments et les coins candidats sont utilisés comme éléments d'initialisation et les formes identifiées (de type simple ou complexe) sont optimisées par maximisation d'un score (réponse moyenne à un détecteur de bords).

Approches "basées régions optiques-RADAR" Dans [PIS08, PIS⁺09], Poulain et al. proposent une nouvelle méthode de détection de bâtiments pour la mise à jour de bases de données vectorielles. Suite à une pré-segmentation de l'image optique, les régions susceptibles d'être des bâtiments sont classifiées comme telles grâce à l'accumulation d'indices issus des données optiques et RADAR (vérification de la présence d'ombres optiques, recherche de lignes et d'angles droits sur les bords optiques, utilisation de l'indice NDVI pour la végétation, recherche d'échos doubles linéiques RADAR). Ces différents indices sont fusionnés dans le cadre de la théorie des croyances de Dempster-Shafer [PIS⁺09] afin d'exploiter la redondance des informations et de réduire les incertitudes.

Dans [Bac09], Baccouche propose une méthode de détection des bâtiments à partir d'une image optique panchromatique (de résolution submétrique), d'une image optique multi-spectrale et d'une image RADAR (de résolution métrique). Suite à une sur-segmentation par Mean-Shift de l'image optique panchromatique, une stratégie, similaire à celle proposée par Poulain et al. [PIS08, PIS⁺09] et basée sur une accumulation d'évidences issues des données optiques et RADAR (vérification de la présence d'ombres optiques et RADAR suivant des directions données, utilisation de l'indice NDVI pour la végétation, recherche de zones brillantes RADAR linéiques ou surfaciques) permet d'abord de créer des cartes de scores et de déduire un masque binaire de régions labellisées "bâtiments". Un algorithme de recherche de rectangles est ensuite mis en œuvre sur ce masque afin de générer les contours de bâtiments rectangulaires.

Approches "basées contours et régions optiques-RADAR" Dans [JCHM06], Junjie et al. développent une stratégie double (détection approximative et détection fine), basée sur l'extraction de contours et sur l'extraction de régions caractéristiques. Dans un premier temps, les "zones de recouvrement" sont extraites afin d'estimer les hauteurs de bâtis. Dans un deuxième temps, les localisations des emprises au sol sont approximativement déduites dans l'image RADAR par translation des "zones de recouvrement",

sous l'hypothèse suivante : "les dimensions planimétriques du bâti sont supposées suffisamment faibles devant sa hauteur pour pouvoir assimiler les dimensions de la "zone de recouvrement" à celles de l'emprise du bâti". Dans un dernier temps, les bordures de toits sont précisément extraites sur l'image optique présentant une meilleure résolution de la façon suivante : une classification des régions (par méthode ISODATA) et une polygonalisation des contours (par transformée de Hough) sont successivement réalisées autour des "fenêtres de recouvrement" projetées dans l'image optique (à partir de la hauteur estimée du bâti).

2.4.2.2 Estimation des hauteurs de bâtiments

En ce qui concerne la problématique isolée d'estimation des hauteurs de bâtiments, différentes méthodes peuvent être mises en œuvre dans le cadre de la configuration d'entrée "une image optique - une image RADAR" :

- en monoscopie (optique ou RADAR), par mesure manuelle des ombres (optiques ou RADAR), de la parallaxe optique ou de la "zone de recouvrement" RADAR ;
- en utilisant conjointement l'image optique et l'image RADAR, par mise en œuvre d'approches exploitant l'information d'élévation issue d'un appariement d'éléments optiques-RADAR ;
- en monoscopie RADAR, par mise en œuvre d'approches itératives en simulation RADAR.

Mesures manuelles en monoscopie optique ou RADAR (et extension en multiscopie)

C'est par exemple le cas de Junjie et al. [JCHM06], qui, dans le cadre de la fusion optique-RADAR, utilisent uniquement l'image RADAR pour estimer la hauteur des bâtis : après avoir extrait les "zones de recouvrement" (supposées nettes et régulières), les hauteurs de bâtis sont simplement déduites à partir de l'équation reliant l'extension de la "zone de recouvrement" à la hauteur du bâtiment. Il convient toutefois de souligner que cette méthode d'estimation des hauteurs s'avère peu précise et peu robuste.

Exploitation de l'information d'élévation issue d'un appariement d'éléments optiques-RADAR en fusion de données

Dans [TR04], Tupin et al. proposent une nouvelle méthode, basée sur la projection et sur l'appariement de primitives optiques-RADAR, pour la construction d'un MNT local et d'un MNE relativement épars (estimation de la hauteur de certains bâtiments) sans hypothèse de forme a priori. L'algorithme développé calcule, pour un jeu de hauteurs tests (utilisées lors de la projection) et pour un ensemble de primitives extraites, un score d'appariement optique-RADAR. Celui-ci combine des informations sur la magnitude du gradient optique et sur la similarité directionnelle entre primitives optiques et primitives RADAR. L'optimisation de ce score permet alors de conserver un ensemble de hauteurs candidates. Un filtrage par noyau gaussien permet finalement de différencier et de regrouper les primitives au sol et les primitives avec élévation (situées à l'intérieur des emprises de bâtis), puis de générer un MNT et un MNE approximatifs.

Mise en œuvre d'approches itératives par simulation en monoscopie RADAR (et extension en multiscopie RADAR)

Dans [BLB08, BLBG10], Brunner et al. décrivent une procédure originale nommée "génération d'hypothèses - restitution - appariement" ("hypothesis generation - rendering - matching") pour l'estimation de la hauteur des bâtis,

en supposant connue la localisation des emprises planimétriques rectangulaires. Les paramètres définissant ces dernières sont issus d'une base de données ou d'une extraction manuelle sur l'image optique à disposition. Un critère d'appariement (exemple : information mutuelle), testant la similarité entre une image RADAR de bâtiment simulée (sous différentes hypothèses de hauteur) et l'image RADAR réelle, est utilisé pour estimer la hauteur du bâti en monoscopie RADAR.

2.4.2.3 Génération de MNS ou de MNE et reconstruction 3D des bâtiments

Étant donnée la difficulté du cadre de travail (fusion d'informations avec des données d'entrée très restreintes), les algorithmes de reconstruction 3D proposés dans ce contexte combinent généralement de façon très simple les informations (localisations, dimensions et hauteurs) issues indépendamment à partir des signatures optiques et RADAR.

Il s'agit la plupart du temps d'approches de reconstruction indirectes "basées primitives 2D + hauteurs".

Approches indirectes "basées primitives 2D + hauteurs" Dans les travaux de Junjie et al. [JCHM06], un modèle 3D simple (parallélépipède défini par une unique hauteur et une emprise rectangulaire) est utilisé pour effectuer la reconstruction 3D des bâtiments d'une scène urbaine. L'information altimétrique (hauteur des bâtis) est issue de la donnée RADAR à partir des "zones de recouvrement". L'information planimétrique (emplacement des emprises rectangulaires) est issue des deux images à partir d'une approche "basée contours et régions optiques-RADAR". L'association de ces deux types d'informations permet enfin de reconstruire en 3D la scène d'étude.

2.4.3 Reconstruction 3D à partir d'une image optique et d'un couple RADAR

On souligne le fait que, lorsqu'un couple RADAR (radargrammétrique ou interférométrique) associé à une image optique est disponible, la configuration est particulièrement bien adaptée à l'extraction de hauteurs 3D des bâtis à partir des données RADAR et à la déduction de la forme et de la position des emprises à partir de celles optiques.

2.4.3.1 Détection de bâtiments et extraction de leurs contours

Approches "basées classification jointe optique-RADAR" Une nouvelle méthode d'extraction du bâti, à partir de données optiques et interférométriques, est proposée dans [WST09]. Le bénéfice apporté par les données RADAR, utilisées en complément de l'image optique, est illustré sur des données réelles aéroportées. Dans un premier temps, les données RADAR interférométriques sont utilisées pour fournir une carte des échos linéiques dûs aux coins mur-sol, grâce à l'utilisation des images d'amplitude et de cohérence et de l'interférogramme. Ces échos servent à établir des vecteurs caractéristiques RADAR utilisés dans la dernière étape du procédé. Dans un deuxième temps, les données optiques sont traitées : un ensemble de "régions candidates" (toits candidats) est d'abord généré sur l'image optique par utilisation d'un algorithme de croissance de régions et d'outils morphologiques ; un ensemble de vecteurs caractéristiques optiques (basés sur la géométrie, la forme, la radiométrie et la structure des "régions candidates") est ensuite extrait puis utilisé en entrée d'une classification optique afin de générer un ensemble de "régions toits potentiels". Dans un troisième temps, une seconde passe de classification, utilisant

cette fois conjointement les vecteurs caractéristiques optiques et RADAR, est mise en œuvre sur les "régions toits potentiels". La carte des "régions toits validés" est alors établie par seuillage sur des mesures de qualité optiques-RADAR associées aux différentes régions. Wegner et al. démontrent ainsi l'amélioration réalisée (augmentation du taux de détection et réduction des fausses alarmes) par comparaison à une méthode typique de classification en monoscopie optique.

2.4.3.2 Estimation des hauteurs de bâtiments

Exploitation de l'information d'élévation issue d'une optimisation jointe optique-RADAR en fusion de données Dans [Tup04, TR05], Tupin et al. développent une méthode d'estimation de hauteurs sur un graphe de régions adjacentes, au sein d'un cadre markovien permettant la prise en compte de relations contextuelles entre primitives. Le problème posé est celui de l'estimation d'un champ de hauteurs (information d'élévation 3D) à partir de données radargrammétriques, sur un graphe de régions adjacentes obtenu par sur-segmentation de l'image optique (information de forme et de position). Il s'agit d'identifier, sur le couple radargrammétrique, les primitives RADAR avec élévation et de projeter celles-ci dans l'image des régions optiques afin d'optimiser un critère énergétique en fonction de leurs hauteurs. Une approche d'optimisation énergétique, basée sur le calcul de la coupe de capacité minimale dans un graphe (technique des "graph cuts"), est mise en œuvre. Les hauteurs radargrammétriques estimées permettent de projeter des points d'intérêt RADAR vers l'image optique afin de calculer le champ des observations, auquel se réfère le terme d'attache aux données. Celui-ci tient compte en effet de la présence de primitives RADAR à l'intérieur des régions optiques. Le terme d'a priori, quant à lui, permet de préserver les contours tout en régularisant correctement les hauteurs. Il prend en compte les discontinuités radiométriques optiques grâce à l'introduction d'un coefficient de contraste et les discontinuités de hauteurs à l'intérieur de régions homogènes grâce à l'utilisation d'un processus bord implicite. Les résultats obtenus sont satisfaisants et les principales limitations sont dues aux mauvais appariements ou aux non appariements radargrammétriques (zones d'ombres et présence de petits bâtis sans éléments répondant en hauteur). Un procédé similaire peut être employé à partir des primitives RADAR avec élévation issues d'un couple interférométrique [TD07].

2.4.3.3 Génération de MNS ou de MNE et reconstruction 3D des bâtiments

Approches directes "basées objets 3D" Dans [Lem07], une nouvelle approche, basée sur des contours actifs polygonaux, est brièvement présentée pour la reconstruction 3D de bâtiments relativement complexes, à partir de données optiques et InSAR. Il s'agit d'optimiser, dans un espace 3D cartographique, les positions des noeuds de bâtis potentiels en minimisant un coût énergétique global, conjointement calculé sur les données optiques et RADAR (phase, amplitude, cohérence). Lemoigne suggère de combiner au sein de ce critère global un critère statistique RADAR (exemple : Log-Vraisemblance d'une image calculée pour un partitionnement en régions caractéristiques) et un critère géométrique optique (exemple : critère de gradient optique).

2.4.4 Conclusions : Performances et limites

L'utilisation combinée des données optiques et RADAR permet l'obtention de résultats très satisfaisants pour la reconstruction 3D des bâtiments.

Dans de nombreuses méthodes, les algorithmes proposés tirent profit :

- du gradient optique, pour l'obtention d'informations précises en termes de localisation et de dimensions des bâtiments ;
- de l'image optique pour la détection des formes de bâtis ;
- des primitives RADAR linéaires pour l'obtention d'informations robustes en termes d'orientation ;
- du potentiel "3D" de la donnée RADAR pour l'estimation des hauteurs de bâtiments (en monoscopie par mesure des "zones d'ombres" ou des "zones de recouvrement", en radargrammétrie par appariement d'éléments en hauteur homologues et en interférométrie grâce à la nature cohérente du signal) ;
- du potentiel "régularisateur" de la donnée optique (exemple : régularisation d'interférogrammes RADAR) ;
- du potentiel "focalisateur" de la donnée RADAR, qui permet de rapidement converger vers les zones d'intérêt ;
- des potentiels de "validation croisée" et "d'amplification" des deux types de données : en effet, celles-ci, utilisées conjointement ou seules (si l'une des deux images n'est pas exploitable), permettent de valider et/ou de compléter les résultats de reconstruction.

On souligne ainsi la complémentarité manifeste des deux types d'images, qui rend leur usage joint tout à fait pertinent dans un tel contexte.

2.5 Applications de la reconstruction 3D des bâtiments

On recense de nombreux champs d'application de la reconstruction 3D des bâtiments en télédétection optique et RADAR, parmi lesquels :

- la cartographie avec la représentation (simple ou complexe) du paysage urbain et semi-urbain [WZS10, GHS00, BH03, FBB09, AA10, BM99, DJ01, Pap98, Jib02, Tai05] ;
- la détection de changements avec l'étude de l'évolution des aménagements et des structures urbaines dans les applications de surveillance et l'évaluation des dommages causés en situation de crise [Ben10, MS10, Mer10, Bal04, TPCI09, BLB10] (gestion du cas classique où une image optique pré-événement ou une base de données pré-événement et une image RADAR post-événement sont disponibles) ;
- la simulation 3D, avec certaines applications dans le domaine civil (exemple : élaboration de jeux de plateforme grand public) et militaire (exemple : préparation de missions).

Chapitre 3

Présentation de la chaîne de traitements proposée

Dans ce chapitre, nous décrivons dans sa globalité la chaîne de traitements proposée. Les fonctionnalités du système, les choix stratégiques associés et l'organisation de la chaîne sont d'abord exposés. Les hypothèses de travail adoptées pour réaliser la reconstruction des bâtiments et les limites inhérentes à notre approche de reconstruction sont ensuite explicitées. Enfin, nous présentons les jeux de données réelles sur lesquels la chaîne sera évaluée.

3.1 Description de la chaîne complète

3.1.1 Objectifs, fonctionnalités et contraintes

Le but de ces travaux est de développer une chaîne complète de traitements semi-automatiques, capable de fournir en sortie une reconstruction simple des bâtiments présents dans une scène urbaine ou semi-urbaine, à partir d'images satellitaires optiques et RADAR à haute résolution.

Le système développé prendra uniquement en entrée un couple composé d'une seule image optique et d'une seule image RADAR d'une même scène, ainsi qu'un MNT associé.

Les contraintes suivantes devront être satisfaites :

1. La chaîne doit être capable de fonctionner en conditions opérationnelles, ce qui implique la capacité à pouvoir traiter de larges zones en un temps relativement court et avec une intervention de l'opérateur aussi limitée que possible (traitements semi-automatiques).
2. Les différentes étapes composant la chaîne doivent être en mesure de fournir l'intégralité des informations planimétriques et altimétriques, requises pour une reconstruction simple, précise et fiable des bâtiments.
3. Un indice de confiance pertinent, témoignant de la qualité de la reconstruction, doit être proposé en fin de chaîne pour chaque bâtiment reconstruit.

Il s'agira donc d'explorer et d'intégrer, à travers une séquence de méthodes innovantes, l'ensemble des informations complémentaires susceptibles d'être issues des données optiques et RADAR, en vue d'une reconstruction de la scène urbaine.

Remarquons enfin que la configuration spécifique des données d'entrée (un unique duo optique-RADAR) correspond à un cas pratique susceptible d'être observé en milieu opérationnel et nécessite donc pleinement d'être étudiée. On peut citer : la vérification de bâtiments dans un contexte militaire, la cartographie rapide dans les situations de catastrophes majeures ou de surveillance du territoire (inondations par exemple). Fournir une méthodologie de reconstruction 3D des bâtis dédiée à cette configuration s'avère ainsi particulièrement intéressant en termes de potentiel d'intégration au sein de certains systèmes ou applications en télédétection.

Par ailleurs, il convient de souligner, qu'étant donnée cette configuration d'entrée, aucune des méthodes classiquement développées dans la littérature et basées sur un couple de deux images optiques ou RADAR (stéréoscopie optique, radargrammétrie et interférométrie RADAR) ne peut ici être mise en œuvre pour la reconstruction des bâtiments.

3.1.2 Choix stratégiques

Nous optons pour deux choix stratégiques majeurs, pour lesquels nous détaillons nos motivations ci-dessous :

- premièrement, étant donnée la complexité qui caractérise les scènes urbaines en imagerie RADAR à haute résolution, il s'avère très délicat de pouvoir détecter la présence de bâtiments directement à partir d'une seule image RADAR ; nous proposons donc d'extraire l'information concernant la présence de bâtis potentiels à partir de l'image optique et ensuite de valider cette information à partir des données RADAR ;
- deuxièmement, la visée latérale des capteurs RADAR induit un certain nombre de déformations géométriques, telles que les "zones de recouvrement" et les "zones d'ombre", qui peuvent être utilisées pour retrouver une information de hauteur ; nous proposons donc d'estimer la hauteur des bâtis à partir de l'image RADAR (ce choix étant d'autant plus justifié, que les images optiques, souvent acquises avec un angle de visée proche du nadir, ne révèlent que peu de façades de bâtis et présentent ainsi un faible potentiel "3D").

Nous voyons que la stratégie développée repose sur des rôles distinctement répartis entre la donnée optique et la donnée RADAR : l'information planimétrique est apportée par la source optique, quand celle altimétrique provient du capteur RADAR. Une amélioration de la chaîne de traitements proposée pourrait reposer sur une meilleure symétrisation de l'apport des deux sources. Des perspectives en ce sens sont présentées dans le chapitre 8 ainsi que des résultats préliminaires.

3.1.3 Organisation de la chaîne proposée

La figure 3.1 présente un schéma synoptique de la chaîne complète, que nous développerons dans ce manuscrit.

Celle-ci peut être décomposée en quatre étapes successives qui correspondent aux quatre contributions majeures de nos travaux :

1. La détection de bâtiments potentiels et l'extraction précises de leurs emprises en monoscopique optique.
 2. La projection et le recalage de ces emprises potentielles depuis la donnée optique vers la donnée RADAR.
-

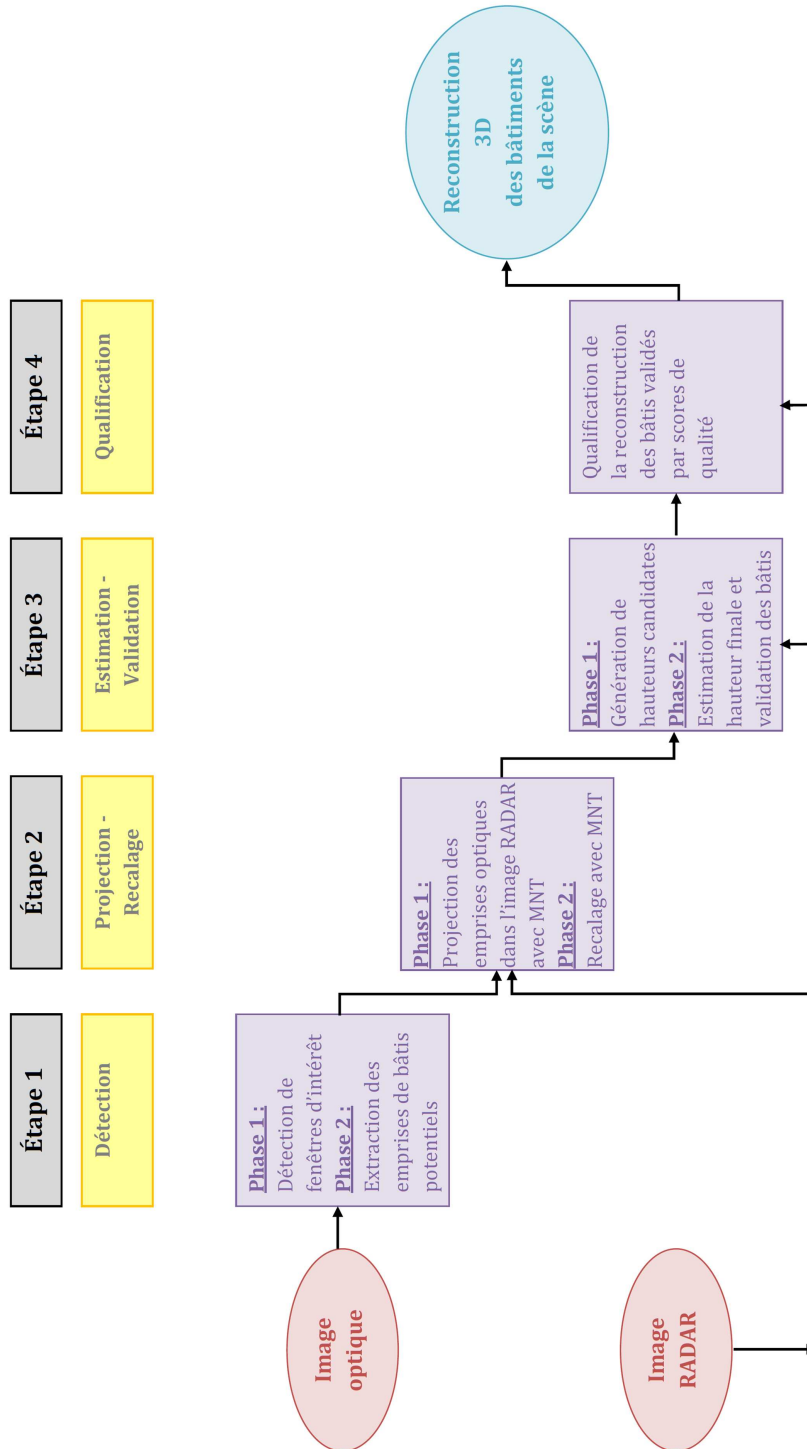


FIGURE 3.1 – Schéma synoptique de la chaîne complète proposée.

3. L'estimation de la hauteur des bâtis potentiels et la validation simultanée de leur présence à partir de l'image RADAR.
4. L'évaluation de la qualité des bâtiments reconstruits par l'introduction de scores de confiance RADAR.

Chacune de ces étapes est décrite en détails et mise en application sur nos scènes d'étude au sein d'un chapitre dédié de ce manuscrit.

Premièrement, nous montrons dans le chapitre 4 que l'information planimétrique relative à la localisation et aux dimensions des emprises de bâtis peut être extraite de l'image optique disponible, à travers la mise en œuvre d'un procédé basé sur des critères morphologiques, géométriques et radiométriques. Nous obtenons en fin d'étape une carte binaire fournissant, dans le référentiel de l'image optique, l'emplacement précis des emprises potentielles, considérées dans la suite comme les primitives optiques auxquelles nous nous référerons.

Deuxièmement, nous expliquons dans le chapitre 5 la méthodologie utilisée lors de l'étape de projection de ces primitives optiques vers la donnée RADAR. Nous proposons notamment un procédé original de recalage fin visant à apparier les primitives optiques projetées avec leurs homologues RADAR. Nous obtenons en fin d'étape un double produit, composé d'une part, d'une carte binaire fournissant, dans le référentiel image RADAR, l'emplacement précis des emprises optiques potentielles, projetées et recalées et, d'autre part, un MNT affiné localement au niveau de ces emprises.

Troisièmement, nous démontrons dans le chapitre 6 que l'introduction de données RADAR permet l'obtention d'une information altimétrique. Une nouvelle approche, combinant un critère statistique RADAR et un critère de contraste RADAR, est mise en œuvre pour estimer la hauteur des bâtis et valider conjointement leur présence. Nous obtenons en fin d'étape une carte en niveaux de gris fournissant, dans le référentiel image RADAR, la localisation et les hauteurs estimées des bâtis validés. La reconstruction des bâtis de la scène est alors réalisée en intégrant l'ensemble des informations altimétriques et planimétriques issues du jeu de cartes généré tout au long des traitements.

Dernièrement, nous présentons dans le chapitre 7 une nouvelle méthode de qualification de la reconstruction des bâtiments en définissant, pour chaque bâti reconstruit, un score de confiance global. Ce dernier tient compte d'un ensemble de mesures individuelles effectuées à partir des critères RADAR pré-cités. Il permet d'assortir chaque bâti d'un niveau de fiabilité, susceptible d'être fort utile en conditions opérationnelles dans le cadre d'applications semi-supervisées.

3.2 Hypothèses de travail

Dans ces travaux, deux hypothèses majeures sont faites : la première hypothèse est liée à la modélisation géométrique des bâtis que l'on souhaite reconstruire, tandis que la seconde se réfère aux conditions d'acquisition des données optiques.

3.2.1 Modélisation des bâtiments

Nous limitons notre cadre de travail à celui de la reconstruction tridimensionnelle de bâtiments simples présentant les caractéristiques suivantes : emprise rectangulaire, façades verticales et toit plat horizontal (donc associé à une seule hauteur).

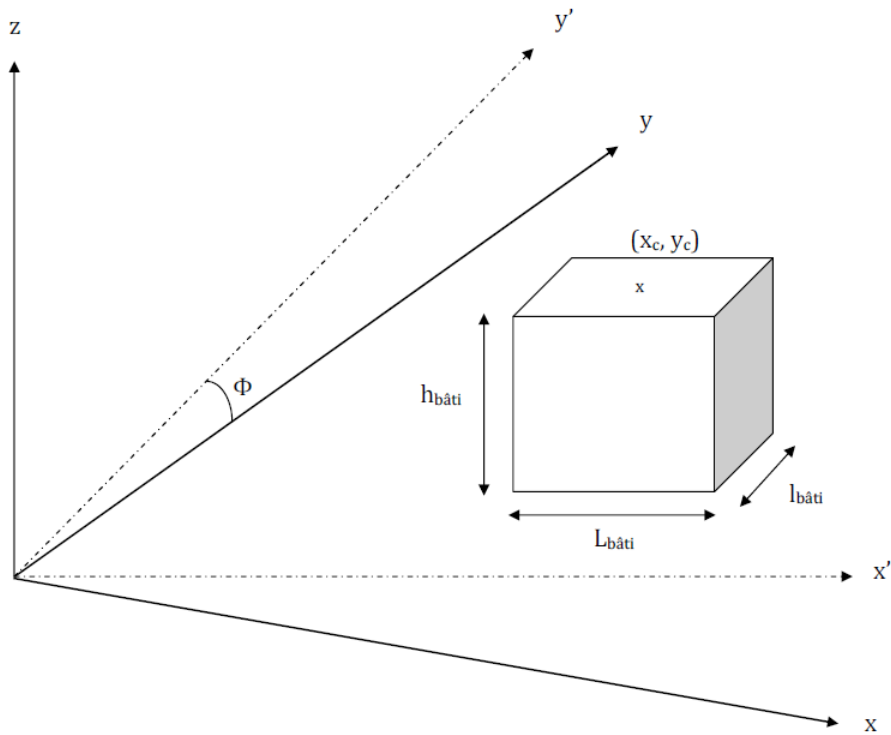


FIGURE 3.2 – Modèle 3D simple de bâtiment adopté pour la reconstruction, représenté dans le référentiel terrestre $(O, \vec{x}, \vec{y}, \vec{z})$. Ce modèle est décrit par 5 paramètres planimétriques et 1 paramètre altimétrique.

Nous modélisons ainsi notre bâti comme un simple parallélépipède, défini dans un référentiel terrestre $(O, \vec{x}, \vec{y}, \vec{z})$ par cinq paramètres planimétriques décrivant son empreinte rectangulaire (largeur $l_{\text{bâti}}$, longueur $L_{\text{bâti}}$, orientation ϕ et coordonnées (x_c, y_c) de son centre) et un paramètre altimétrique (hauteur du bâti $h_{\text{bâti}}$) (voir figure 3.2).

Le choix d'un modèle 3D simple (restreignant nécessairement le type de bâti susceptible d'être reconstruit) est principalement motivé par le cadre de travail particulièrement délicat et difficile à aborder, qu'est celui de la fusion de données optiques et RADAR à résolution métrique en milieu urbain. Nous proposerons donc de concentrer davantage nos efforts sur l'aspect "fusion d'informations" à partir d'un unique couple d'images optique-RADAR que sur l'aspect "généricité du modèle 3D employé". Par ailleurs, l'extension des algorithmes développés le long de la chaîne de traitement au cas de modèles 3D plus sophistiqués sera discutée au chapitre 8 et pourra faire l'objet de travaux ultérieurs, les bases théoriques et pratiques ayant été posées.

3.2.2 Conditions d'acquisition

Nous supposons que l'image optique disponible en entrée de chaîne est acquise avec un angle de visée au nadir (ou très proche du nadir). Ceci facilite en effet l'étape de détection des bâtis sur la donnée optique puisque nous n'avons pas à gérer le problème des façades, qui ne sont pas visibles (voire très légèrement visibles pour les bâtis extrêmement hauts).

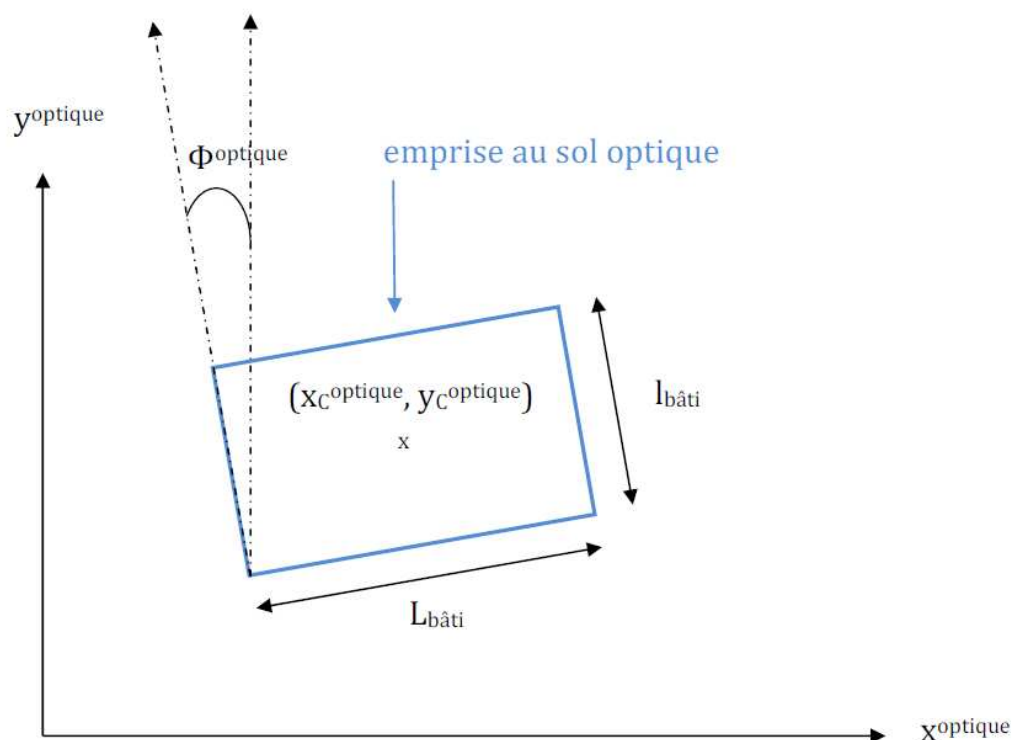


FIGURE 3.3 – Modélisation de la signature optique d'un bâtiment adoptée pour la reconstruction, représentée dans le référentiel image optique.

En conséquence, la localisation des emprises au sol des bâtis et la localisation des bordures de toits (extraites directement sur l'image optique) sont confondues dans le référentiel planimétrique optique ; dans la suite de ce travail, nous ne nous référerons plus qu'aux emprises au sol des bâtis.

3.2.3 Conséquences sur la modélisation des signatures de bâtiments

Signature optique des bâtis à reconstruire Étant donné le modèle 3D parallélépipédique adopté et sous l'hypothèse d'une image optique acquise selon une visée au nadir, nous déduisons que la signature optique d'un bâti se caractérise dans nos travaux par un simple rectangle. Celui-ci, appelé primitive optique, identifie le contour du toit et indique l'emplacement de l'emprise au sol du bâti (voir figure 3.3 représentée dans le référentiel image optique $(O, \vec{x}^{optique}, \vec{y}^{optique})$).

Signature RADAR des bâtis à reconstruire Étant donné le modèle 3D parallélépipédique adopté, nous proposons dans nos travaux deux configurations géométriques possibles pour la signature RADAR d'un bâti et susceptibles d'intervenir en fonction de la hauteur et des dimensions au sol du bâti.

Celles-ci sont présentées sur la figure 3.4 dans le référentiel image RADAR $(O, \vec{x}^{RADAR}, \vec{y}^{RADAR})$.

La configuration 1 a lieu si $(h_{bâti} < h_{limite})$ et la configuration 2 a lieu si $(h_{bâti} > h_{limite})$.

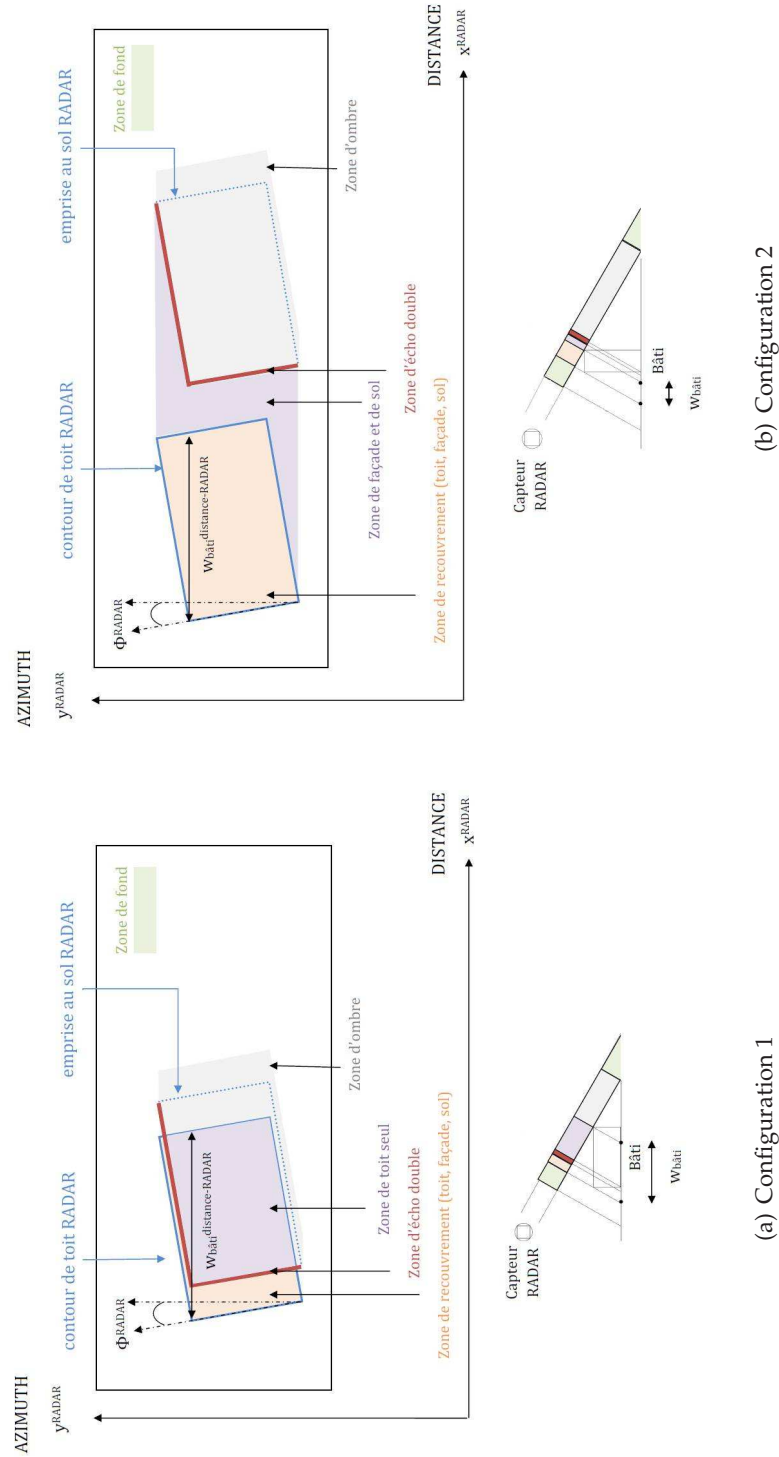


FIGURE 3.4 – Modélisation de la signature RADAR d'un bâtiment adoptée pour la reconstruction, représentée dans le référentiel image RADAR.

La limite entre les deux configurations correspond au cas où le coin au sol situé à l'avant du bâti (sur la face proche du capteur) et le coin en hauteur situé à l'arrière du bâti (sur la face éloignée du capteur) se trouvent sur la même iso-distance.

On rappelle que la hauteur h_{limite} est définie par :

$$h_{limite} = w_{bâti} \tan(\theta_i) \quad (3.1)$$

avec les notations suivantes :

- θ_i correspond à l'angle d'incidence du capteur RADAR ;
- $w_{bâti}$ correspond à "l'épaisseur vue" du bâtiment en "géométrie terrain" suivant l'axe des distances RADAR.

Remarque sur le calcul de h_{limite} Étant donnée la relation :

$$w_{bâti}^{distance-RADAR} = w_{bâti} \sin(\theta_i) \quad (3.2)$$

h_{limite} s'exprime aussi sous la forme :

$$h_{limite} = \frac{w_{bâti}^{distance-RADAR}}{\cos(\theta_i)} \quad (3.3)$$

avec la notation suivante :

- $w_{bâti}^{distance-RADAR}$ correspond à "l'épaisseur vue" du bâtiment projetée en "géométrie RADAR" suivant l'axe des distances RADAR (mesurée dans l'image).

Les différentes régions composant la signature RADAR du bâti, appelées primitives RADAR, sont listées ci-dessous :

- pour la configuration 1 : la "zone de fond" ; la "zone de recouvrement" correspondant à la rétrodiffusion simultanée du toit, de la façade et du fond ; la "zone d'écho double" ; la "zone de toit seul" et la "zone d'ombre" ;
- pour la configuration 2 : la "zone de fond" ; la "zone de recouvrement" correspondant à la rétrodiffusion simultanée du toit, de la façade et du fond ; la "zone de façade" où seuls la façade et le sol répondent ; la "zone d'écho double" et la "zone d'ombre".

La localisation et l'étendue de chaque région peut être calculée pour une hauteur de bâtiment donnée $h_{bâti}$, à partir de l'emplacement de l'emprise rectangulaire au sol dans l'image RADAR et en fonction de certains paramètres du système d'acquisition RADAR (angle d'incidence et pas d'échantillonnage en distance oblique) [BB03b, BB03a, BLB08, BLBG10, BLB10].

Mise en correspondance optique-RADAR Seules certaines primitives optiques-RADAR peuvent être appariées dans notre cadre de travail.

Il convient d'ailleurs de noter les points suivants se rapportant à cette mise en correspondance :

- l'emprise au sol optique peut être mise en correspondance partielle avec les échos doubles RADAR ; seuls les deux segments qui sont face au capteur RADAR parmi les quatre segments décrivant le rectangle apparaîtront en tant qu'échos doubles RADAR dûs aux coins diédriques mur-sol ;
- les façades optiques n'étant quasiment pas visibles sur nos images, il nous est impossible de les mettre en correspondance avec leurs homologues RADAR ;

TABLE 3.1 – Caractéristiques du jeu de données optique-RADAR.

Caractéristiques	Image optique	Image RADAR
Capteur	Quickbird	TerraSAR-X
Mode	Panchromatique (PAN)	High SpotLight (HS)
Polarisation	-	HH
Niveau de traitement	2A - Standard	1B - SSC Slant Range
Format du fichier image	GeoTIFF (.TIF)	COSAR (.COS)
Année d'acquisition	2003	2007
Angle au nadir	19.2°offNadir	-
Angle d'incidence	-	31.8°au centre de l'image
Localisation du coin gauche (longitude, latitude)	(5.29, 43.29)	(5.43, 43.25)
Région imagée	Marseille et ses environs	Marseille et ses environs
Taille de la scène	11.6 km x 21.8 km	9.3 km x 5.3 km
Nombre de lignes	17548	6025
Nombre de colonnes	31604	10784
Pas d'échantillonnage	0.66 m au sol en ligne et 0.69 m au sol en colonne	0.45 m en distance oblique (0.86 m en distance au sol) et 0.87 m en azimuth
Résolution au centre de l'image RADAR	- - -	0.59 m en distance oblique (1.1 m en distance au sol) et 1.1 m en azimuth

- les ombres optiques (dues au non-éclairage d'une zone par la lumière en provenance du soleil) et les ombres RADAR (dues au non-éclairage d'une zone par les ondes électro-magnétiques en provenance du capteur RADAR) proviennent de phénomènes physiques totalement différents et se situent donc à des endroits différents dans un référentiel terrestre ; elles ne peuvent donc pas être mises en correspondance l'une avec l'autre et ne trouvent pas d'équivalent respectif dans l'image RADAR et dans l'image optique.

3.3 Présentation des données

3.3.1 Jeu de données complet

Données images Nous disposons d'un couple de données métriques satellitaires optique-RADAR, composé d'une unique image Quickbird (en mode panchromatique et acquise selon un angle de visée quasiment au nadir) et d'une unique image TerraSAR-X (en mode High SpotLight, avec une résolution de 1.1 m en distance au sol et en azimuth et acquise avec incidence de 31.8°), toutes deux situées dans la région de Marseille (France). Notons que nous nous sommes placés dans un cadre particulièrement difficile car nous n'avons pas considéré de données multi-spectrales pour l'image optique.

Le tableau 3.1 répertorie les caractéristiques des deux images présentées sur les figures 3.5, 3.6 et 3.7.



FIGURE 3.5 – Image optique disponible : capteur Quickbird ©DigitalGlobe, mode pan-chromatique, région de Marseille - France.

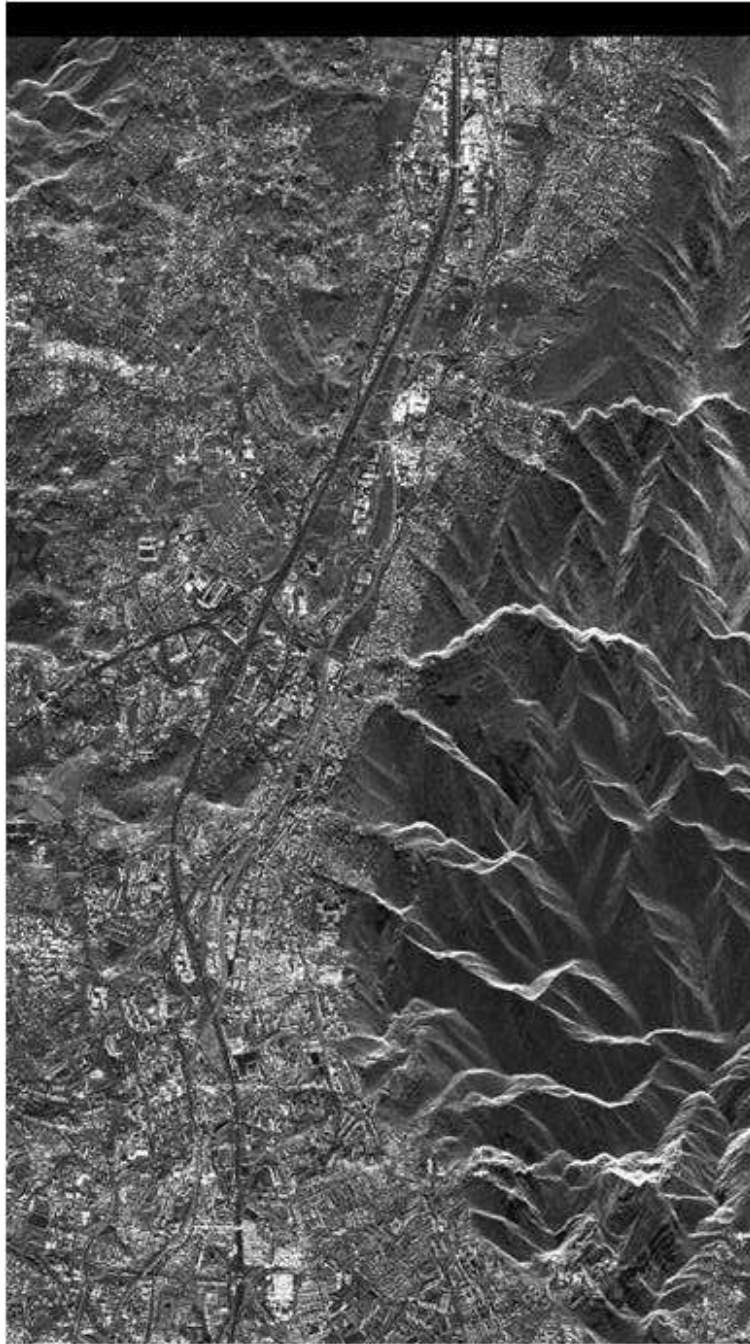


FIGURE 3.6 – Image RADAR disponible : capteur TerraSAR-X ©Infoterra, mode High SpotLight, région de Marseille - France.

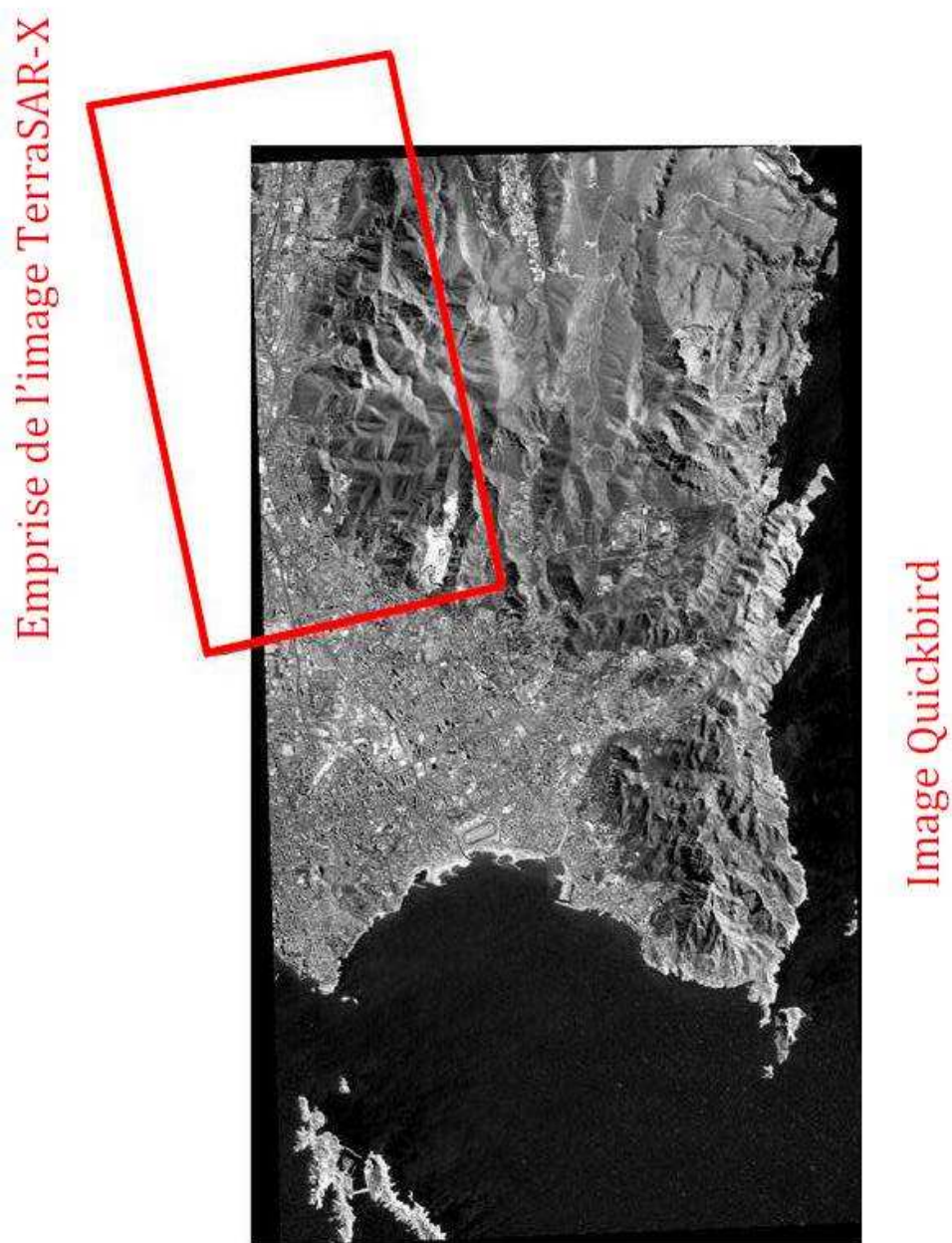


FIGURE 3.7 – Recouvrement entre l'image optique ©DigitalGlobe et l'image RADAR. L'emprise de l'image RADAR sur l'image optique est représentée en rouge.



FIGURE 3.8 – Modèle numérique de terrain (MNT) disponible : MNT DIMAP ©SpotImage, précision altimétrique à 90 % de 10-20 m, région de Marseille - France.

Modèle numérique de terrain (MNT) Un MNT DIMAP produit par la société SPOT est disponible sur la zone imagée. Celui-ci, proposé en figure 3.8, présente une résolution altimétrique à 90 % variant entre 10 et 20 m, c'est-à-dire comparable à ce qui peut être fourni par un MNT de type DTED 2. Il sera utilisé lors de l'étape de projection et de recalage décrite au chapitre 5, afin de déterminer un intervalle de recherche.

Précisons que la nécessité de disposer en entrée d'un MNT d'assez bonne qualité (de type DTED 2) nous a empêché de tester la chaîne complète de traitement développée sur un second jeu de données optiques-RADAR, disponible sur la région de Toulouse (France) mais pour lequel aucun MNT de qualité suffisante n'a pu être obtenu.

Données auxiliaires Les données auxiliaires, permettant notamment de géo-référencer les données, sont également disponibles pour chacune des images du couple optique-RADAR :

- pour l'image Quickbird, on dispose d'un ensemble de fichiers auxiliaires (.JPG, .TIL, .RPB, .IMD, .TXT), parmi lesquels le fichier .RPB, fournissant un modèle de projection dit RPC (Quickbird Rational Polynomial Coefficients), qui sera utilisé

lors de l'étape de projection et de recalage ;

- pour l'image TerraSAR-X, on dispose d'un ensemble de fichiers auxiliaires (.TIF, .PNG, .XML, .HTML), parmi lesquels un fichier .XML, fournissant un modèle de projection physique (c'est-à-dire répertoriant l'ensemble des paramètres capteur et plateforme du système d'acquisition), qui sera également utilisé lors de l'étape de projection et de recalage.

3.3.2 Scènes d'intérêt

Nous proposons d'illustrer les différents résultats, obtenus à chaque niveau intermédiaire de la chaîne de traitements, sur deux scènes d'intérêt, composées d'une vingtaine de bâtiments et sélectionnées pour leur variété.

La figure 3.9 présente les deux duos optiques-RADAR des scènes étudiées le long de ce manuscrit.

La première scène d'étude (scène A), située en milieu semi-urbain, est essentiellement composée de larges bâtiments rectangulaires industriels comparables à des hangars, assez bas et relativement éloignés les uns des autres.

La seconde scène d'étude (scène B), située en milieu urbain, est essentiellement composée de bâtiments résidentiels rectangulaires, assez hauts et relativement proches les uns des autres. Cette seconde scène apparaît sur l'image RADAR comme plus difficile à traiter en raison de la densité des primitives caractéristiques des bâtis qui s'y trouvent et qui sont susceptibles de se recouvrir.

Ces deux scènes sont relativement petites mais vont être utilisées pour analyser en détails les performances de la chaîne proposée, grâce à la vérité terrain qui leur est associée (et que nous n'aurions pu établir de façon exhaustive pour de larges zones).

Par ailleurs, précisons que le faible recouvrement existant entre l'image optique et l'image RADAR a limité le nombre de zones de tests disponibles (voir figure 3.7). Ceci a été, en outre, amplifié par le fait que seule une partie de l'image RADAR est située en milieu urbain et semi-urbain, la grande majorité correspondant à une zone montagneuse (voir figure 3.6).

3.3.3 Vérité terrain

La scène A correspondant à une zone industrielle à accès limité (entourée d'une voie de chemin de fer et de routes à accès contrôlé), il nous a été impossible de photographier les bâtiments de cette scène pour établir une vérité terrain. Par contre, la scène B correspondant à une zone résidentielle, il nous a été possible de réaliser des photos vérité terrain pour tous les bâtiments composant cette scène (voir figure 3.10).

Afin d'évaluer la qualité de la reconstruction obtenue, nous prendrons en considération différentes sources d'informations pour créer une vérité terrain complète :

- quant à la présence et à la localisation des emprises réelles des bâtis : par combinaison des informations issues des photographies disponibles avec les informations déduites d'une carte d'emprises manuelle ; cette dernière est réalisée manuellement sur l'image optique d'après une appréciation visuelle et avec l'aide d'images complémentaires disponibles sous Google Map ;
- quant aux hauteurs réelles des bâtis : par combinaison des hauteurs calculées sur les photographies disponibles avec les hauteurs déduites manuellement par mesure sur l'image optique des ombres projetées par les bâtiments.

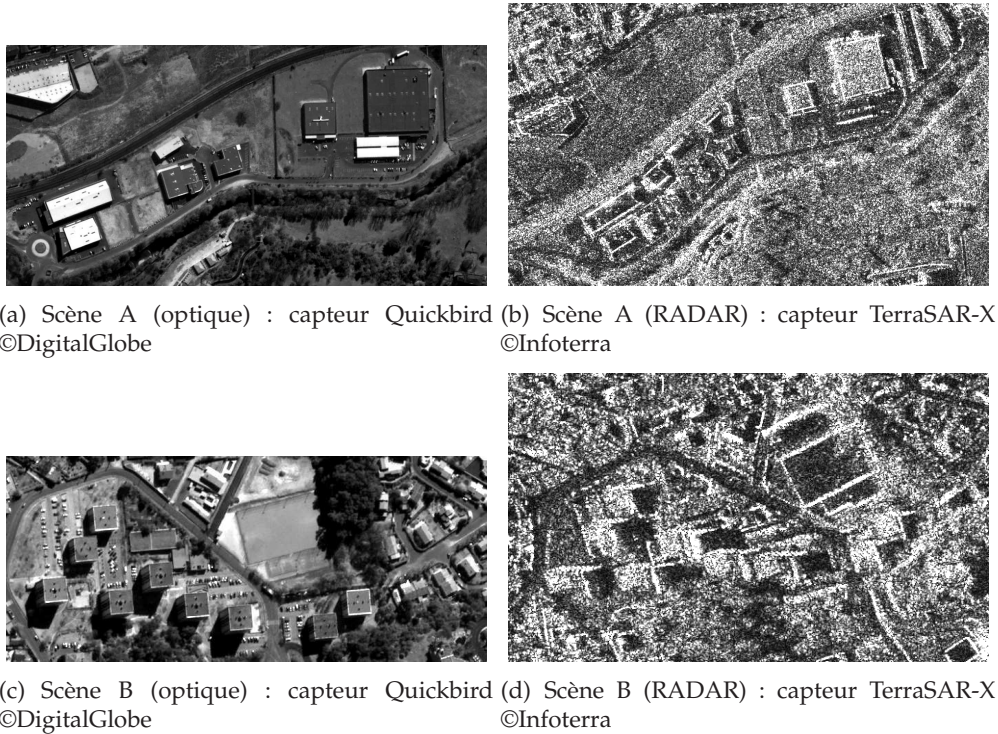


FIGURE 3.9 – Scènes d'étude A et B.

La figure 3.11 présente, pour chaque scène d'étude, une carte binaire vérité terrain indiquant la localisation des emprises des bâtiments dans le référentiel image optique et référençant les labels utilisés.

Notons que seuls les bâtis de taille suffisamment importante ont été répertoriés. En particulier, sur la scène B, les petites maisons individuelles à toits bi-pentes, situées essentiellement sur la partie en haut à droite de l'image, ne sont pas prises en compte car il ne nous paraît pas raisonnable de les modéliser à résolution métrique.

Les tableaux 3.2 et 3.3 présentent, pour chaque scène d'étude, les hauteurs réelles relatives à ces bâtiments.



(a) Bâtiment F

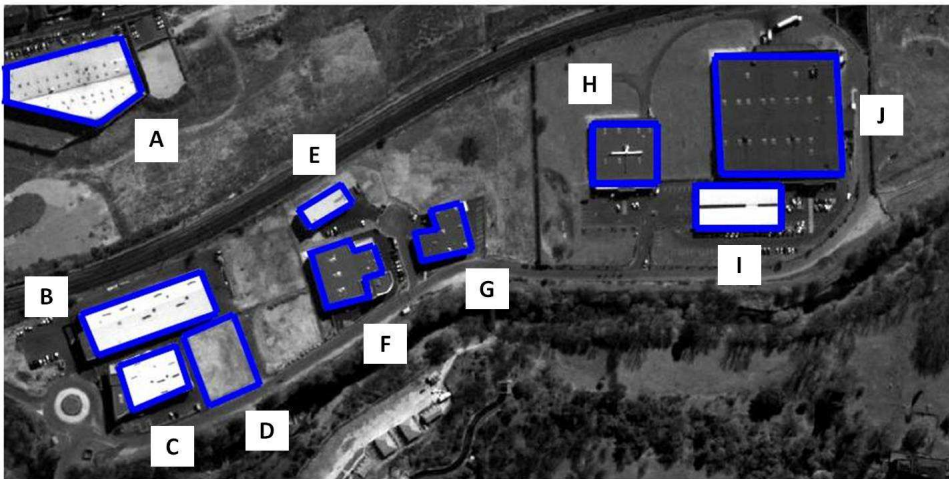
(b) Bâtiment E

(c) Bâtiment D

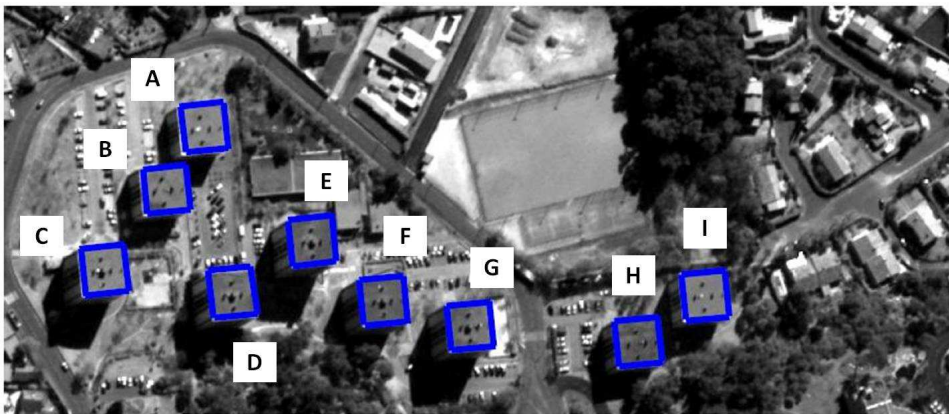
(d) Bâtiment H

(e) Bâtiment A

FIGURE 3.10 – Photos vérité terrain de quelques bâtiments composant la scène d'étude B.



(a) Scène A (optique) : extraction manuelle des emprises (en bleu) des bâtiments de la vérité terrain (10 emprises au total, dont 7 rectangulaires) ©DigitalGlobe



(b) Scène B (optique) : extraction manuelle des emprises (en bleu) des bâtiments de la vérité terrain (9 emprises au total, toutes rectangulaires) ©DigitalGlobe

FIGURE 3.11 – Cartes de la vérité terrain pour les scènes d'étude A et B.

TABLE 3.2 – Hauteurs réelles des bâtiments de la vérité terrain pour la scène d'étude A.

Label issu de la vérité terrain	Hauteur estimée par ombre portée sur l'image optique h_{ombre}	Hauteur réelle moyenne h_{reel}
A	10.3 m	10.3 m
B	9.1 m	9.1 m
C	7.6 m	7.6 m
D	-	-
E	6.8 m	6.8 m
F	6.9 m	6.9 m
G	7 m	7 m
H	5.4 m	5.4 m
I	4.9 m	4.9 m
J	9.3 m	9.3 m

TABLE 3.3 – Hauteurs réelles des bâtiments de la vérité terrain pour la scène d'étude B.

Label issu de la vérité terrain	Hauteur estimée par longueur d'ombre sur l'image optique h_{ombre}	Hauteur estimée sur photographie h_{photo}	Hauteur réelle moyenne h_{reel}
A	10.4 m	11.9 m	11.2 m
B	13.2 m	13.8 m	13.5 m
C	19.5 m	20.3 m	19.9 m
D	20.3 m	20.3 m	20.3 m
E	22.8 m	22.8 m	22.8 m
F	19 m	19.7 m	19.4 m
G	22.3 m	18.9 m	20.6 m
H	15.4 m	14.5 m	15 m
I	14.5 m	14.3 m	14.4 m

Chapitre 4

Extraction des bâtiments en monoscopie optique

Nous présentons dans ce chapitre une nouvelle méthode de détection de bâtiments en monoscopie optique.

Dans un premier temps, une nouvelle méthodologie, permettant l'identification des bâtiments et l'extraction de leur contour, est développée à partir d'outils morphologiques, géométriques et radiométriques.

Dans un second temps, le procédé présenté est mis en œuvre sur les scènes d'études et fournit les "cartes de localisation des contours optiques", issues de la première étape de la chaîne de traitements proposée.

4.1 Méthode proposée pour la détection de bâtiments

4.1.1 Stratégie globale

L'objectif est de développer une nouvelle méthode opérationnelle pour l'extraction des contours rectangulaires de bâtiments à partir d'une seule image optique panchromatique (acquise avec une visée proche du nadir) et sans donnée auxiliaire supplémentaire (de type emprises cadastrales ou base de données de l'occupation des sols).

C'est sur cette faible quantité d'informations en entrée et sur l'aspect semi-automatique sollicité pour le procédé, que reposent les principales difficultés du problème.

La méthodologie présentée fait appel à la combinaison d'outils morphologiques (profil différentiel), géométriques (coefficient de recouvrement, direction principale, contours rectangulaires et taille minimale) et radiométriques (score de radiométrie moyenne).

Elle peut être décomposée en deux phases :

- la première phase, globale, s'intitule "détection des fenêtres d'intérêt". Elle est fondée sur une approche "orientée régions" et fournit une carte approximative de zones rectangulaires susceptibles de contenir chacune un bâtiment de la scène urbaine ; ces zones à fort potentiel, appelées "fenêtres d'intérêt", sont utilisées en entrée de la seconde phase afin d'orienter localement l'algorithme d'extraction de contours ;
 - la seconde phase, locale, s'intitule "extraction fine des emprises potentielles". Elle est fondée sur une approche "orientée contours" et produit une carte affinée des contours de bâtiments, extraits à l'intérieur des fenêtres d'intérêt.
-

Rappelons, qu'étant données les hypothèses émises dans ces travaux (visée optique proche du nadir et modélisation parallélépipédique du bâtiment), les contours optiques extraits sont de forme rectangulaire et fournissent indifféremment la localisation planimétrique des emprises au sol ou la localisation planimétrique des bordures de toit.

4.1.2 Phase 1 : détection des fenêtres d'intérêt

4.1.2.1 Définition des profils morphologiques différentiels

Le concept de profil morphologique différentiel (P. M. D.), introduit par Pesaresi et al. [PB01], a déjà été utilisé pour certaines applications en milieu urbain, telles que la classification [BPA03, CBF06] et la détection de changements [DMBBB07] à partir d'images optiques satellitaires.

Le P. M. D. fournit une information à propos de la taille, de la forme et du contraste d'un objet dans une image. Il est construit à partir d'opérateurs d'ouverture et de fermeture par reconstruction, faisant appel à une métrique géodésique afin de conserver la forme des objets.

Une approche multi-échelle, basée sur un jeu de différents éléments structurants, est adoptée afin d'explorer une large gamme de tailles d'objets.

Notons $\mathbf{PMD}(x, y)$, $\mathbf{PMD}_\phi(x, y)$ et $\mathbf{PMD}_\gamma(x, y)$ les vecteurs différentiels appelés profil morphologique différentiel, profil morphologique différentiel de fermeture et profil morphologique différentiel d'ouverture, respectivement définis au pixel (x, y) de l'image I par leur i -ème élément :

$$\mathbf{PMD}^i(x, y) \hat{=} \begin{cases} \mathbf{PMD}_\phi^i(x, y), & \forall i \in [1, p] \\ \mathbf{PMD}_\gamma^i(x, y), & \forall i \in [-p, -1] \end{cases} \quad (4.1)$$

$$\mathbf{PMD}_\phi^i(x, y) \hat{=} \mathbf{PM}_\phi^i(x, y) - \mathbf{PM}_\phi^{i-1}(x, y) \quad (4.2)$$

$$\mathbf{PMD}_\gamma^i(x, y) \hat{=} \mathbf{PM}_\gamma^i(x, y) - \mathbf{PM}_\gamma^{i-1}(x, y) \quad (4.3)$$

où les vecteurs $\mathbf{PM}_\phi(x, y)$ et $\mathbf{PM}_\gamma(x, y)$ désignent le profil morphologique de fermeture et le profil morphologique d'ouverture, respectivement définis au pixel (x, y) de l'image I par leur i -ème élément :

$$\mathbf{PM}_\phi^i(x, y) \hat{=} \phi_{\epsilon_i} [I(x, y)] \quad (4.4)$$

$$\mathbf{PM}_\gamma^i(x, y) \hat{=} \gamma_{\epsilon_i} [I(x, y)] \quad (4.5)$$

en notant :

- ϕ l'opérateur de fermeture par reconstruction ;
- γ l'opérateur d'ouverture par reconstruction ;
- i l'index des éléments des vecteurs précédemment définis (chaque valeur de i correspond à un niveau du profil c'est-à-dire à une itération du procédé de construction du profil $\mathbf{PMD}(x, y)$) ;
- $N_{\mathbf{PMD}} = (2p + 1)$ la taille ou dimension du profil $\mathbf{PMD}(x, y)$, obtenu par concaténation du profil $\mathbf{PMD}_\phi(x, y)$ et du profil $\mathbf{PMD}_\gamma(x, y)$;
- $\epsilon_i = i \times \epsilon_1$ la taille de l'élément structurant ES utilisé au niveau i , où ϵ_1 correspond à la taille de l'élément structurant élémentaire utilisé au premier niveau.



FIGURE 4.1 – Exemple partiel de profil morphologique différentiel associé à un bâtiment ©DigitalGlobe. De gauche à droite : l'extrait panchromatique initial puis quelques images composant son profil et associées à un élément structurant de taille croissante.

Les ouvertures et les fermetures affectent respectivement les structures plus claires ou plus foncées que leur région environnante.

Lorsque l'élément structurant ES atteint la taille d'un objet homogène dans le profil, chacun des pixels le composant reçoit une valeur commune (celle de la région environnante la plus claire ou la plus sombre), ce qui induit un pic dans le profil (correspondant au contraste local) et provoque la "disparition" de l'objet (fusionné avec son environnement extérieur). Les objets sont ainsi progressivement "retirés" aux différents niveaux du profil (selon les tailles croissantes) et une pile d'images dites "simplifiées" à différentes échelles peut être contruite à partir d'une image panchromatique initiale.

Dans l'algorithme présenté, nous tirons profit des propriétés de simplification de la scène à différentes échelles et de préservation des formes du P. M. D. afin d'introduire un critère géométrique relatif à la forme a priori des bâtiments. En effet, ces derniers sont supposés être imagés sur la donnée optique comme des régions rectangulaires homogènes de différentes tailles, susceptibles d'être plus aisément identifiables aux différents niveaux du profil. Par exemple, dans le cas d'un bâtiment présentant un toit plat avec des petites superstructures (exemple : cheminées), il existe un niveau du P. M. D. pour lequel l'ensemble de ces petites superstructures disparaissent (car la taille de l'élément structurant a dépassé leur taille propre). En conséquence, on dispose, à un niveau du profil, d'une image représentant le toit comme un seul rectangle homogène (plus facile à traiter) qui favorise la mise en évidence d'une fenêtre d'intérêt grâce au test d'adéquation géométrique décrit dans la section suivante. Un tel phénomène est illustré sur la figure 4.1, où est représentée une partie du P. M. D. relatif au bâtiment imagé.

4.1.2.2 Présentation de l'algorithme

L'algorithme proposé pour l'identification globale des fenêtres d'intérêt à partir d'une image optique panchromatique se base sur un critère d'adéquation géométrique, testé hiérarchiquement sur les objets du P. M. D. (voir figure 4.2).

Algorithme Il se décompose en quatre étapes :

1. **Pré-traitements** Premièrement, les ombres (facilement visibles sur l'image d'entrée) sont extraites par un simple seuillage radiométrique puis masquées sur l'image optique afin de faciliter la suite du traitement¹. Deuxièmement, le P. M. D. de l'image est construit.

1. Les ombres ne sont pas prises en compte dans l'algorithme considéré. Toutefois, il aurait été possible de les intégrer à la méthode développée afin d'améliorer les performances d'extraction des contours de bâtis. Ceci fera l'objet d'une discussion en section 4.3 et 4.5.

2. **Première passe** Le procédé suivant est appliqué itérativement sur les différentes images simplifiées du P. M. D., en parcourant le profil depuis les images obtenues avec l'élément structurant de taille maximale vers celles obtenues avec l'élément structurant de taille minimale. Pour chaque région objet présente dans l'image courante et de radiométrie non nulle, une boîte englobante de référence, de forme rectangulaire et associée à un bâti potentiel, est générée en utilisant les axes d'inertie de la région objet testée. Un coefficient de recouvrement, testant l'adéquation géométrique entre la région rectangulaire de référence et la région objet considérée, est calculé. Celui-ci est défini comme le rapport de surface commune entre les deux régions. Si ce dernier est supérieur à un seuil, la région de référence est validée comme fenêtre d'intérêt et la région objet correspondante est masquée pour l'itération suivante. Un profil appelé "profil masqué" est ainsi obtenu en fin de première passe.
3. **Seconde passe** Cette passe est nécessaire pour traiter en partie le cas de bâtiments définis par des régions objet imparfaites (faible contraste, forme rectangulaire peu marquée, contours irréguliers ou discontinus). Premièrement, de façon à "dessiner" plus nettement les limites sous-jacentes entre de telles régions objet et leur environnement, des opérations classiques de morphologie mathématique (érosion et dilatation) sont appliquées avec un élément structurant élémentaire sur les différentes images composant le "profil masqué" issu de la première phase (exemple : mise en œuvre d'une ouverture morphologique pour supprimer les petits détails en bordure des objets). Un profil appelé "profil amélioré" est ainsi obtenu. Deuxièmement, le même procédé hiérarchique que celui précédemment décrit est mis en œuvre sur ce "profil amélioré" et permet d'augmenter le taux de détection de fenêtres d'intérêt validées².
4. **Post-traitement** Un seuillage sur la taille minimale des fenêtres validées est réalisé afin de supprimer celles de trop petite taille (en pratique, il s'agit souvent de fausses alarmes).

4.1.3 Phase 2 : extraction fine des emprises potentielles

4.1.3.1 Présentation de l'algorithme

Étant donné que les paramètres géométriques définissant les fenêtres d'intérêt issues de la première phase ne sont souvent pas assez précis, une seconde phase est nécessaire pour affiner la localisation et les dimensions des emprises rectangulaires.

Dans cette phase, une approche locale complémentaire est adoptée afin d'obtenir une fine superposition entre les contours de bâtiments et les discontinuités radiométriques de l'image.

L'algorithme proposé est mis en œuvre localement autour de chaque fenêtre d'intérêt afin d'affiner en orientation, dimension et position l'emprise rectangulaire associée (voir figure 4.3).

Algorithme Il se décompose en quatre étapes :

2. Les opérations morphologiques, décrites au début de cette seconde passe et appliquées dans l'intention d'obtenir un "profil amélioré", auraient également pu être intégrées dans une étape de pré-traitement général, de façon à appliquer directement, lors d'un seul passage, le procédé hiérarchique présenté.

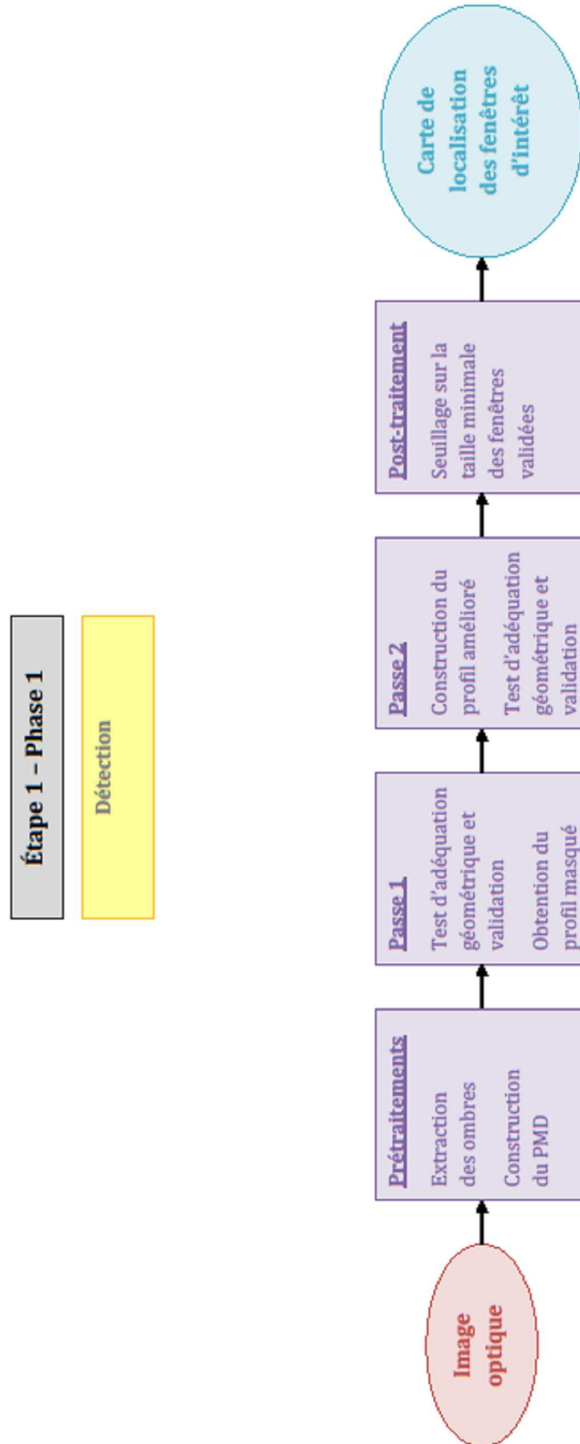


FIGURE 4.2 – Schéma synoptique de la phase 1 de l'étape 1.

1. **Initialisation** Le rectangle recherché est initialisé à partir du contour R^0 de la fenêtre d'intérêt considérée.
2. **Ré-orientation suivant la direction principale** Un accumulateur de Hough est localement calculé pour l'ensemble des pixels contenus à l'intérieur de la boîte englobante de la fenêtre d'intérêt considérée à partir de l'image seuillée de gradient de Sobel. La direction privilégiée du bâtiment est obtenue en sélectionnant dans l'accumulateur le couple polaire (ρ, θ) correspondant à la plus grande valeur d'accumulation. Une rotation est alors appliquée sur le rectangle initial R^0 de façon à ce que son orientation principale "colle" avec la direction privilégiée trouvée. On note R^1 le rectangle issu de cette rotation.
3. **Optimisation en position et dimensions** De faibles variations en position et en dimensions sont envisagées sur R^1 . Une double boucle, testant simultanément sur le rectangle R^1 un jeu de modifications possibles en termes de translation du centre et elongations des bords, est mise en place. Elle permet de générer un ensemble exhaustif de rectangles tests R^k , définis à chaque itération k par :

$$R^k = \left(T_{(n1, n2)}^{translation} \circ H_{n3}^{longueur} \circ H_{n4}^{largeur} \right) (R^1) \quad (4.6)$$

où $T_{(n1, n2)}^{translation}$ correspond à l'opérateur de translation suivant le vecteur $(n1, n2)$, $H_{n3}^{longueur}$ correspond à l'opérateur d'homothétie de rapport $n3$ appliqué sur les deux bords longs du rectangle, et $H_{n4}^{largeur}$ correspond à l'opérateur d'homothétie de rapport $n4$ appliqué sur les deux bords courts du rectangle.

On définit alors un score $S_{extraction}(R^k)$ par :

$$S_{extraction}(R^k) = \sum_{(i,j) \in C_{R^k}} \frac{\delta_{(i,j)}}{P_{R^k}} \quad (4.7)$$

où $\delta_{(i,j)}$ correspond à la radiométrie (0 ou 1) du pixel (i, j) dans l'image seuillée de gradient de Sobel, C_{R^k} désigne le contour de R^k et P_{R^k} désigne le périmètre de R^k . Ce score correspond à la radiométrie moyenne calculée le long du contour du rectangle testé R^k dans l'image seuillée de gradient de Sobel et prend donc des valeurs entre 0 et 1.

L'optimisation de ce score permet enfin d'extraire le contour rectangulaire optimal R^* défini par :

$$R^* = \arg \max_{R^k} \left(S_{extraction}(R^k) \right) \quad (4.8)$$

4. **Validation** L'emprise rectangulaire sélectionnée est finalement conservée si le score optimal associé $S_{extraction}(R^*)$ est supérieur à un seuil (éventuellement modifiable par l'opérateur), assurant une valeur minimale de radiométrie moyenne pour le contour rectangulaire extrait.

Les "cartes de localisation des contours optiques", indiquant l'emplacement des emprises potentielles détectées, sont obtenues en sortie de la phase 2.

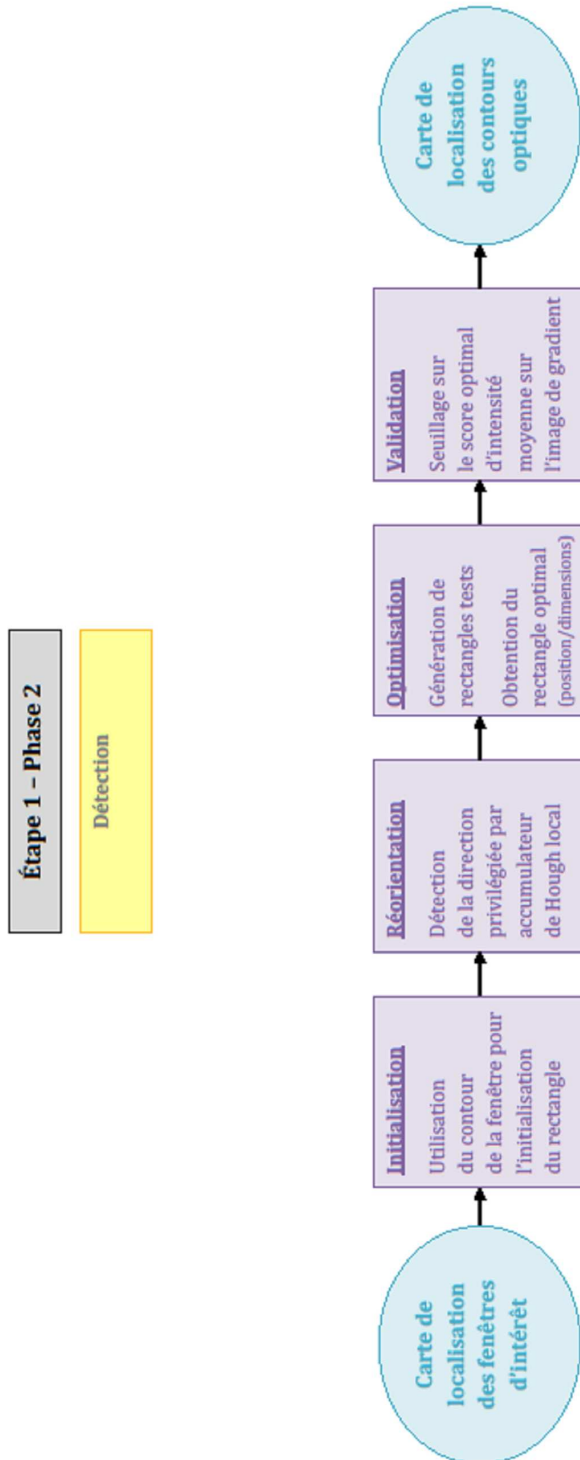


FIGURE 4.3 – Schéma synoptique de la phase 2 de l'étape 1.

4.1.4 Discussion sur la gestion des paramètres

Paramètres de la phase 1 Les différents paramètres utilisés dans la phase de détection des fenêtres d'intérêt sont les suivants :

- le seuil S_{ombres} utilisé pour le masquage des ombres pendant l'étape des pré-traitements ; celui-ci peut être réglé par apprentissage sur les zones d'ombres pour un capteur optique donné et pour des conditions locales de luminosité fixées ;
- les paramètres utilisés pour la construction du P. M. D. ; il s'agit de la dimension N_{PMD} du profil (c'est-à-dire du nombre d'images composant le profil) et de la taille minimale ϵ_1 et maximale $\epsilon_{N_{PMD}} = N_{PMD} \times \epsilon_1$ des éléments structurants utilisés ; ces paramètres dépendent du compromis adopté entre le niveau de simplification souhaité dans le profil et le temps de calcul relatif à la construction de celui-ci ;
- le seuil $S_{recouvrement}$ relatif à la valeur minimale du coefficient de recouvrement testant l'adéquation géométrique entre la région rectangulaire de référence et la région objet considérée ; ce seuil, défini comme un rapport de surface, est empiriquement fixé à 0.75 ;
- les paramètres de surface minimale S^{min} , longueur minimale L^{min} et largeur minimale l^{min} utilisés pendant l'étape de post-traitement ; ceux-ci peuvent être déduits à une résolution donnée des dimensions minimales et maximales des bâtiments recherchés.

Paramètres de la phase 2 Les différents paramètres utilisés dans la phase d'extraction fine des emprises potentielles sont les suivants :

- le seuil S_{Sobel} utilisé pour le calcul de l'image optique de gradient seuillée ; celui-ci peut être réglé par apprentissage sur les zones de discontinuité pour un capteur optique donné et pour un type de région donné (ici : milieu urbain, milieu semi-urbain) ;
- les valeurs maximales $n1^{max}$, $n2^{max}$, $n3^{max}$ et $n4^{max}$ utilisées pour réaliser les translations de centre et les élongations de bords du rectangle à optimiser ; ces valeurs maximales sont empiriquement fixées à quelques pixels ;
- le seuil S_{score} utilisé pour confirmer la fiabilité de l'emprise rectangulaire extraite à partir du calcul de son score de radiométrie moyenne ; ce seuil, relatif à un score variant entre 0 et 1, est empiriquement fixé à 0.5 (en effet, ce score est supposé prendre une valeur proche de 1 lorsque la majorité du contour rectangulaire généré se situe effectivement au niveau des discontinuités déduites de l'image de gradient seuillée et proche de 0 lorsque la superposition observée entre contour généré et discontinuités reste faible).

4.2 Application sur les scènes d'études

La première étape de la chaîne de traitements proposée est mise en œuvre sur les données réelles pour les deux scènes d'étude considérées.

4.2.1 Paramètres utilisés

Les paramètres utilisés pour la première phase et pour la seconde phase de l'étape d'extraction des bâtiments en monoscopie optique sont indiqués en table 4.1.

TABLE 4.1 – Paramètres utilisés pour l'étape d'extraction des emprises potentielles.

Noms des paramètres	Scène A	Scène B
S_{ombres}	210	150
N_{PMD}	24	24
$\epsilon_1(\text{pixels})$	5	5
$\epsilon_{N_{PMD}}(\text{pixels})$	120	120
$S_{recouvrement}$	0.75	0.75
$S^{min}(\text{pixels}^2)$	200	500
$L^{min}(\text{pixels})$	20	-
$l^{min}(\text{pixels})$	20	-
S_{Sobel}	10	15
$n1^{max}(\text{pixels})$	5	5
$n2^{max}(\text{pixels})$	5	5
$n3^{max}(\text{pixels})$	7	7
$n4^{max}(\text{pixels})$	7	7
S_{score}	0.5	0.5

Les valeurs des seuils sont données pour une image optique panchromatique Quick-bird codée sur 16 bits (c'est-à-dire sur 65536 niveaux de gris).

L'ensemble des paramètres utilisés sont sensiblement égaux pour les deux scènes à quelques exceptions près. Par exemple, la scène B (en milieu urbain) présentait une multitude de tout petits bâtis (sous forme de maisons individuelles rectangulaires à toit non plat) que nous ne souhaitions pas intégrer dans l'ensemble des bâtiments tests caractéristiques. Il a donc été nécessaire de fixer une valeur de surface minimale S^{min} supérieure pour la scène B à celle utilisée pour la scène A. De même, la scène A (en milieu semi-urbain) présentait par ailleurs un ensemble de petites surfaces homogènes rectangulaires très fines et très longues (du type morceaux de routes) que nous ne souhaitions pas non plus intégrer dans l'ensemble des bâtiments tests. Il a donc été nécessaire, pour cette scène, de fixer des valeurs non nulles de largeur minimale l^{min} et de longueur minimale L^{min} afin de supprimer ces fenêtres d'intérêt inutiles. De façon générale, il aurait été nécessaire de réaliser un test sur la compacité minimale des objets à conserver, afin de supprimer ce type de surfaces non souhaitées.

4.2.2 Résultats obtenus

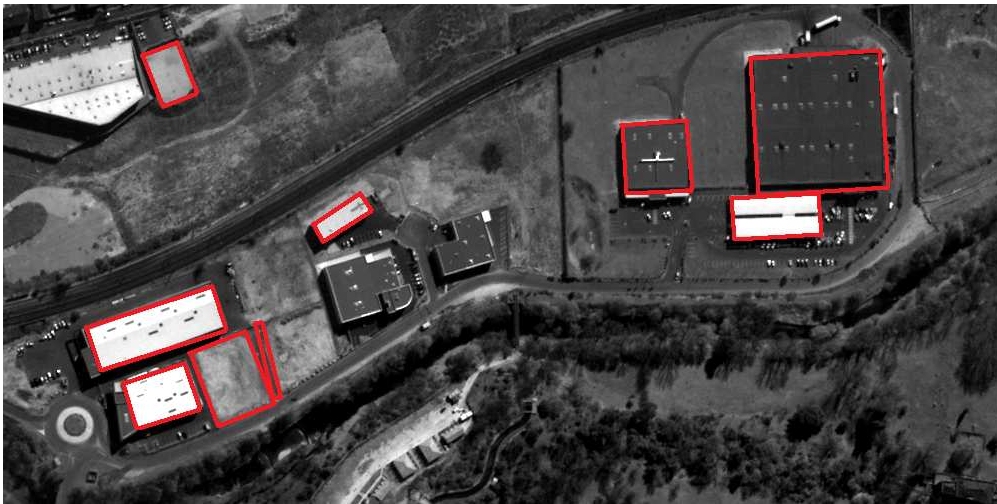
La figure 4.4 présente les deux "cartes de localisation des contours optiques", obtenues à la fin de la première étape et indiquant l'emplacement précis (dans le référentiel image optique) des emprises potentielles extraites.

Les emprises rectangulaires détectées sont représentées en rouge et sont superposées sur l'image optique panchromatique.

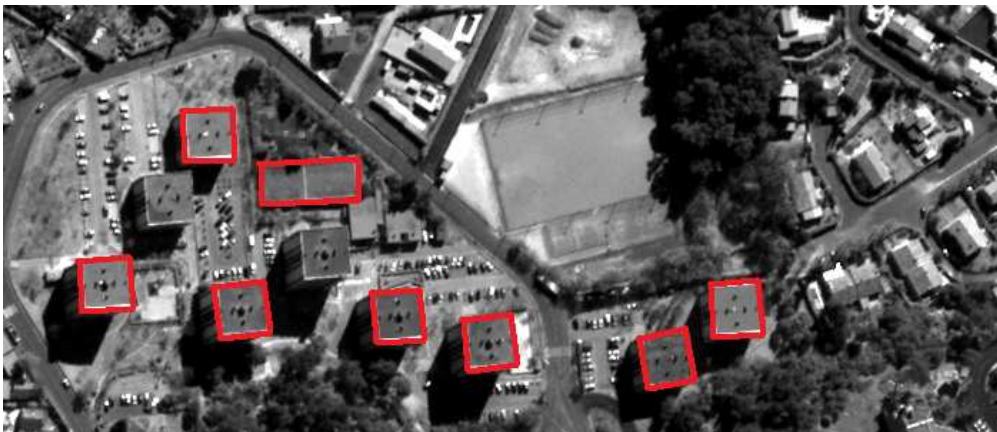
4.2.3 Analyse qualitative et quantitative

Après la première étape de la chaîne, 9 emprises potentielles sont détectées sur la scène A et 8 sur la scène B.

Parmi celles-ci, on recense 7 bonnes détections et 2 fausses alarmes sur la scène A et 7 bonnes détections et une fausse alarme sur la scène B.



(a) Scène A (optique) : emprises extraites en fin d'étape 1 (en rouge) ©DigitalGlobe



(b) Scène B (optique) : emprises extraites en fin d'étape 1 (en rouge) ©DigitalGlobe

FIGURE 4.4 – Carte de localisation des contours optiques extraits pour les scènes d'étude A et B.

En ce qui concerne les bonnes détections, des résultats d'extraction obtenus sont très précis : les coordonnées planimétriques des coins de bâtiments ainsi déduites pourront être utilisées comme un bon point de départ pour la fusion avec les données RADAR.

En ce qui concerne les fausses détections, un taux de fausses alarmes de 18 % est globalement obtenu sur les deux scènes d'intérêt. Ces fausses alarmes sont majoritairement dues à la présence de larges zones rectangulaires homogènes présentes sur les terrains d'étude. Celles-ci pourront être corrigées en grande majorité lors des étapes ultérieures du traitement.

En ce qui concerne les non détections, on dénombre, par comparaison avec la vérité terrain, 3 emprises non détectées sur la scène A et 2 emprises non détectées sur la scène B.

Celles-ci s'expliquent par la présence sur les zones d'intérêt de bâtiments non rectangulaires ou de bâtiments présentant un très faible contraste ou des contours trop imparfaits. En ne prenant en compte que les bâtiments à emprises rectangulaires, on obtient de façon satisfaisante pour les deux scènes d'étude un taux moyen de détection de 88 %.

Les tables 4.2 et 4.3 résument les résultats obtenus individuellement pour chacun des bâtiments composant les deux scènes d'étude.

4.3 Étude détaillée de quelques bâtiments d'intérêt

Fenêtres d'intérêt et contours affinés La figure 4.5 présente deux exemples de fenêtres d'intérêt associées à des détections correctes de bâtiments. On constate, à l'issue de la première phase, que les fenêtres obtenues (représentées en jaune) ne se superposent pas précisément aux bords des bâtiments, ce qui justifie une seconde phase d'affinage. À l'issue de cette seconde phase, les contours rectangulaires extraits (représentés en rouge) sont alors satisfaisants en termes de localisation du centre, d'orientation et de dimensionnement.

Bonnes détections La figure 4.6 présente deux exemples de bonnes détections. Ces deux bâtis, conformes à l'hypothèse de forme rectangulaire, présentent un contraste suffisamment marqué par rapport au fond et des contours suffisamment nets pour faciliter leur détection.

Fausses alarmes La figure 4.7 présente deux exemples de fausses alarmes. Il s'agit dans le premier cas d'une parcelle sans construction et dans le second cas d'une cour. À ce niveau de la chaîne, il nous est encore impossible de les distinguer des éléments du sursol tels que les bâtiments (qui présentent une élévation propre), puisque nous ne disposons pour l'instant que d'informations planimétriques et que nous n'avons pas exploité la détection des ombres. La suppression de telles erreurs sera donc traitée ultérieurement au sein d'autres étapes du procédé (étape d'estimation-validation et étape de qualification) en utilisant les données RADAR.

Non détections La figure 4.8 présente deux exemples de non détections. Le premier cas s'explique par la présence d'un bâti de forme en "L", qui ne colle pas à l'hypothèse de modélisation faite quant à la forme a priori rectangulaire des contours de bâtiments. Le second cas est dû à la présence d'un bâti qui présente des contours imparfaits (limite du bord en bas à droite ne se détachant pas assez nettement du fond) : ce bâti n'apparaît pas

TABLE 4.2 – Résultats individuels de détection des bâtiments sur la scène d'étude A.

Label issu de l'étape d'extraction	Label issu de la vérité terrain	Type de détection
1	-	fausse alarme
2	I	bonne détection
3	E	bonne détection
4	B	bonne détection
5	D	bonne détection
6	C	bonne détection
7	J	bonne détection
8	H	bonne détection
9	-	fausse alarme
-	A	non détection (non rectangulaire)
-	F	non détection (non rectangulaire)
-	G	non détection (non rectangulaire)

TABLE 4.3 – Résultats individuels de détection des bâtiments sur la scène d'étude B.

Label issu de l'étape d'extraction	Label issu de la vérité terrain	Type de détection
1	A	bonne détection
2	C	bonne détection
3	F	bonne détection
4	D	bonne détection
5	I	bonne détection
6	G	bonne détection
7	-	fausse alarme
8	H	bonne détection
-	B	non détection (rectangulaire)
-	E	non détection (rectangulaire)

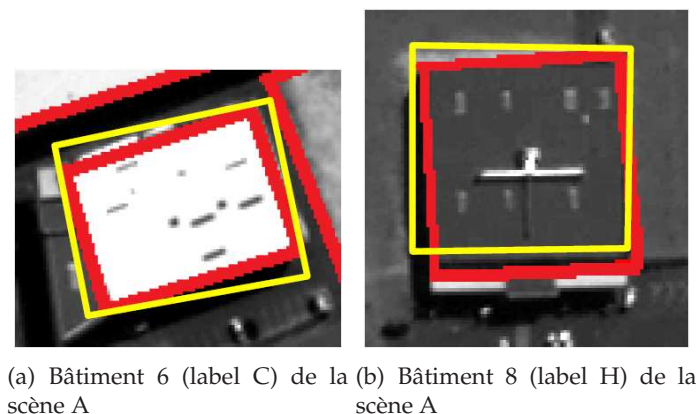


FIGURE 4.5 – Exemples de fenêtres d’intérêt associées à des bonnes détections sur la scène d’étude A. En jaune : la fenêtre d’intérêt détectée (à l’issue de la phase 1). En rouge : l’emprise finale extraite après affinage (à l’issue de la phase 2).

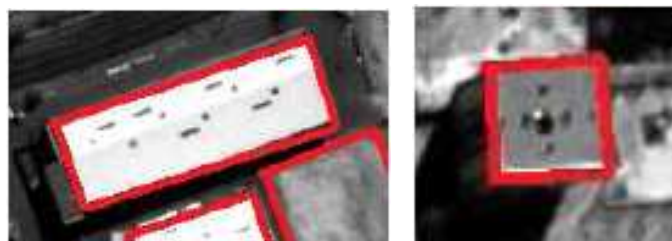
clairement dans le profil comme un objet rectangulaire distinct et ne donne donc pas lieu à une fenêtre de recherche associée lors de la première phase du procédé. Un premier type d’intervention, permettant d’ajouter certaines fenêtres d’intérêt (non automatiquement détectées mais approximativement délimitées par l’opérateur), permettrait de résoudre ce type de problème en orientant l’algorithme localement vers les zones d’intérêt souhaitées lors de la seconde phase. Un second type d’intervention consisterait à ajouter à la liste des bâtis automatiquement extraits, la liste des bâtis manquants, précisément extraits par l’opérateur. Une autre approche consisterait à exploiter l’information apportée par les ombres et la position du soleil, comme cela a déjà été réalisé dans les travaux de Poulain et al. [PIS⁺09] et de Akçay et al. [AA10]. Plus précisément, nous pourrions adapter localement la valeur du seuil $S_{recouvrement}$, utilisé lors de la phase de détection des fenêtres d’intérêt, en admettant une valeur de recouvrement plus faible lorsqu’une ombre adjacente est détectée dans le voisinage de la région testée (indiquant alors la présence probable d’un bâtiment).

4.4 Application sur des scènes complémentaires

Cette première étape d’extraction des emprises de bâtiments potentiels a été validée de façon complémentaire sur d’autres scènes optiques à résolution métrique ou submétrique, situées dans les régions de Toulouse, Saint-Mandrier et Marignane (France).

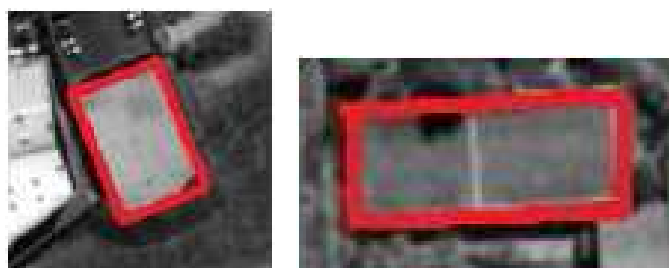
Précisons que, ne disposant pas sur ces zones d’un MNT suffisamment précis ou d’une image RADAR à résolution métrique recouvrant la scène imagée, les étapes suivantes de la chaîne de traitements proposée ne pourront pas être appliquées sur ces données tests.

Les résultats de détection des contours optiques obtenus sur ces données complémentaires (emprises représentées en rouge et superposées sur l’image optique panchromatique) sont proposés en figure 4.9 (Toulouse), en figure 4.10 (Saint-Mandrier) et en figure 4.11 (Marignane), pour des bâtiments individuels.



(a) Bâtiment 4 (label B) de la scène A (b) Bâtiment 2 (label C) de la scène B

FIGURE 4.6 – Exemples de bonnes détections sur les scènes d'étude A et B.



(a) Bâtiment 1 de la scène A (b) Bâtiment 7 de la scène B

FIGURE 4.7 – Exemples de fausses alarmes sur les scènes d'étude A et B.



(a) Bâtiment de label G sur la scène A (b) Bâtiment de label B sur la scène B

FIGURE 4.8 – Exemples de non détections sur les scènes d'étude A et B.

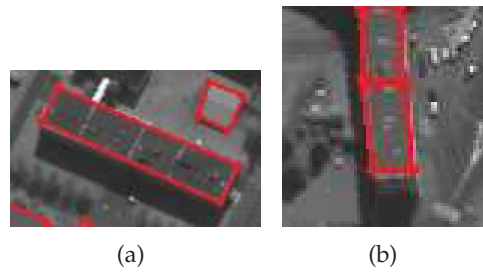


FIGURE 4.9 – Exemples de détections sur la zone de Toulouse (capteur satellitaire)
©DigitalGlobe.

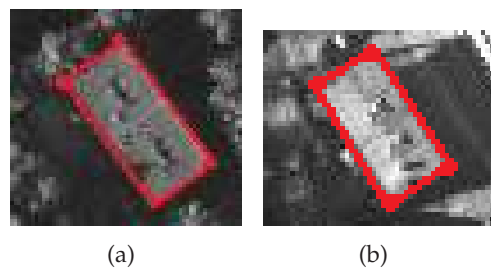


FIGURE 4.10 – Exemples de détections sur la zone de Saint-Mandrier (capteur aéroporté).

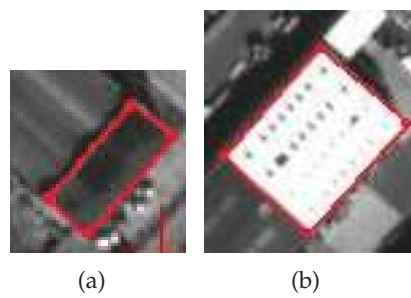


FIGURE 4.11 – Exemples de détections sur la zone de Marignane (capteur satellitaire)
©DigitalGlobe.

4.5 Conclusions et améliorations

Une méthode originale de détection des bâtiments en monoscopie optique panchromatique a été proposée dans ce chapitre. Celle-ci, basée sur la combinaison d'une approche "orientée régions" et d'une approche "orientée contours", s'est révélée fiable lorsque les conditions de contraste et de forme sont remplies et a permis l'obtention de résultats satisfaisants sur les données réelles testées.

La principale limitation de l'algorithme d'extraction proposé repose sur le modèle bidimensionnel a priori (simple rectangle) utilisé pour la recherche des emprises. Celui-ci pourrait être complexifié afin de gérer des emprises plus sophistiquées telles que des emprises circulaires ou régulières en forme de "L", de "T" ou de "U", obtenues par combinaison de rectangles.

De même, la prise en compte de plusieurs rectangles homogènes - mais de radiométrie distincte - juxtaposés et présentant des similarités en termes de direction principale et de dimensions, nous permettrait sans doute de traiter le cas des bâtiments simples et réguliers à toits non plats (exemple : bâtiments à toits bi-pentes où chaque pan, recevant son éclairage solaire propre, est imagé comme un rectangle homogène).

En outre, l'intégration de relations contextuelles (telles que l'adjacence avec les ombres dans un voisinage proche et suivant une direction connue) permettrait sans doute d'améliorer encore les résultats de détection du bâtiment.

Par ailleurs, nous avons pu voir que la méthode proposée utilisait seulement l'image optique pour détecter des bâtiments. Il serait judicieux d'étudier l'amélioration des performances obtenue en utilisant conjointement les deux sources de données par fusion d'indices optiques-RADAR. Cet axe de recherche fera d'ailleurs l'objet d'un approfondissement dans le chapitre 8.

Il conviendrait également de comparer quantitativement la méthode proposée à d'autres approches existant dans la littérature (même si cela n'a pas été réalisé dans le cadre de notre thèse, l'objet de nos travaux portant principalement sur la chaîne de traitements complète et non spécifiquement sur la seule problématique de détection des bâtis en monoscopie optique) et de substituer l'une d'entre elles à l'approche proposée dans la première étape de la chaîne, si celle-ci présente de meilleures performances.

De même, dans le cadre d'une démarche hiérarchique, nous pourrions envisager, dans un premier temps, de traiter les bâtiments les plus faciles à détecter, en appliquant la méthodologie simple proposée, puis, dans un second temps, de s'intéresser aux bâtiments plus délicats à extraire, grâce à la mise en œuvre de procédés plus élaborés.

Soulignons enfin le fait que nous avons traité la problématique d'extraction des bâtiments dans le cadre précis où seule une image optique de type panchromatique était disponible en entrée. Dans le cas où l'utilisateur disposerait de données multi-spectrales, les approches développées par Poulain et al. [PIS⁺09] (en fusion optique-RADAR) et par Akçay et al. [AA10] (en optique seulement), basées sur des stratégies d'accumulation d'évidences (ombre, végétation, couleur des toits, contours linéaires, ...), permettraient d'obtenir des résultats de détection satisfaisants, qui pourraient être utilisés en entrée de la seconde étape de la chaîne.

Chapitre 5

Recalage au sol de primitives caractéristiques optiques et RADAR

Afin de combiner les informations issues des images optiques et RADAR à disposition dans le cadre de la fusion de données, il est nécessaire de projeter et de recaler dans un même référentiel des éléments homologues, caractéristiques des bâtiments à reconstruire.

Nous présentons dans ce chapitre une nouvelle méthode de projection et de recalage fin des emprises optiques extraites dans l'image RADAR.

Dans un premier temps, l'ensemble des notations spécifiques utilisées dans ce chapitre est introduit.

Dans un deuxième temps, on rappelle les équations fondamentales optiques et RADAR, mises en jeu lors de la projection de points (fonctions de localisation directes et inverses) depuis une image optique vers une image RADAR.

Dans un troisième temps, une nouvelle méthodologie, permettant l'obtention d'un recalage fin (à l'échelle métrique) de primitives au sol, est développée à partir d'outils spatiaux, radiométriques et contextuels.

Dans un dernier temps, le procédé présenté est mis en œuvre sur les scènes d'études et fournit les "cartes de localisation des contours RADAR projetés" puis les "cartes de localisation des contours RADAR recalés", issues de la deuxième étape de la chaîne de traitements proposée.

5.1 Notations spécifiques

Les notations suivantes sont employées dans ce chapitre :

- $(i_P, j_P)_{optique}$ désignent les coordonnées image (ligne, colonne) du point P dans le référentiel image optique noté O ;
- $(i_P, j_P)_{radar}$ désignent les coordonnées image (ligne, colonne) du point P dans le référentiel image RADAR noté R ;
- $(X_P, Y_P, Z_P)_{terrestre}$ désignent les coordonnées terrain du point P dans le référentiel terrestre noté T ; ce référentiel a pour origine le centre de la Terre O_T ; $(X_P, Y_P, Z_P)_{terrestre}$ correspondent donc aussi aux coordonnées du vecteur position $\overrightarrow{O_T P}$ relatif au point P dans le référentiel T ;
- $P_{X \rightarrow Y}$ désigne l'opérateur de projection (fonction de localisation directe ou inverse)

permettant de passer d'un référentiel image X à un référentiel terrestre Y ou vice-versa (avec $X = O, R$ ou T et $Y = O, R$ ou T);

- P -joint $_{X \rightarrow Y}$ désigne l'opérateur joint de projection (concaténation de deux opérateurs de projection) permettant de passer du référentiel image X au référentiel image Y (avec $X = O$ ou R et $Y = O$ ou R).

5.2 Rappels : fonctions de localisation directes et inverses

Les rappels présentés dans cette section sont issus des ouvrages suivants : [KE01, TP94, TR05].

5.2.1 Système d'équations optique

Le modèle géométrique utilisé pour décrire le système d'acquisition optique est défini de la façon suivante : chaque point de l'image optique correspond à l'intersection entre le plan image et la ligne reliant le centre optique C du capteur au point 3D imagé P .

Localisation inverse Le système à deux équations suivant permet de déduire les coordonnées image $(i_P, j_P)_{optique}$ du point P à partir de ses coordonnées terrain $(X_P, Y_P, Z_P)_{terrestre}$ (fonction de localisation inverse) :

$$\left\{ \begin{array}{l} i_{P,optique} = \frac{a_1 X_P + b_1 Y_P + c_1 Z_P + d_1}{a_3 X_P + b_3 Y_P + c_3 Z_P + 1} \\ j_{P,optique} = \frac{a_2 X_P + b_2 Y_P + c_2 Z_P + d_2}{a_3 X_P + b_3 Y_P + c_3 Z_P + 1} \end{array} \right\} \quad (5.1)$$

où les coefficients a_i, b_i, c_i et d_i sont relatifs au calibrage interne et externe du système d'acquisition. Il s'agit donc d'un modèle physique de localisation.

Le système d'équation 5.1 est valable en imagerie aéroportée (sans prise en compte du phénomène de distorsion).

En imagerie satellitaire, avec des acquisitions de type "pushbroom" (cas du satellite Quickbird), les coefficients utilisés dans la fonction de localisation inverse ne sont plus constants (ils varient en fonction du temps).

Notons que lorsqu'un modèle physique de localisation n'est pas disponible (paramètres capteurs non fournis), les images optiques peuvent être livrées avec des modèles de type polynômial. C'est par exemple le cas du modèle RPC (Rational Polynomial Coefficients) souvent livré avec les images Quickbird.

Localisation directe La localisation directe, encore appelée géo-référencement (passage des coordonnées image aux coordonnées terrain), nécessite de connaître la hauteur sursol h_P du point P . Elle se fait classiquement à partir d'un schéma d'inversion numérique utilisant le système d'équations 5.1.

5.2.2 Système d'équations RADAR

Localisation directe La localisation directe en imagerie RADAR est donnée par la résolution du système 5.2 fondé sur les équations fondamentales RADAR (après filtrage adapté en distance et synthèse d'ouverture). Ces équations sont respectivement appelées

équation de la "sphère distance", équation du "cône Doppler" et "équation de surface" (présentée ici dans le cas d'un modèle ellipsoïde augmenté).

$$\left\{ \begin{array}{l} (X_P - X_S(t))^2 + (Y_P - Y_S(t))^2 + (Z_P - Z_S(t))^2 = r(t)^2 \\ \frac{2}{\lambda r(t)} (X_P V_{X_S}(t) + Y_P V_{Y_S}(t) + Z_P V_{Z_S}(t)) = f_{DD} \\ \frac{X_P^2 + Y_P^2}{(A_T + h_P)^2} + \frac{Z_P^2}{(B_T + h_P)^2} = 1 \end{array} \right\} \quad (5.2)$$

avec les notations complémentaires suivantes :

- $(X_S(t), Y_S(t), Z_S(t))_{terrestre}$ désignent les coordonnées du vecteur position $\overrightarrow{O_T S(t)}$ du satellite S à l'instant t ;
- $(V_{X_S}(t), V_{Y_S}(t), V_{Z_S}(t))_{terrestre}$ désignent les coordonnées du vecteur vitesse du satellite S à l'instant t ;
- $r(t)$ correspond à la distance point P - satellite S à l'instant t , c'est-à-dire à la norme du vecteur $\overrightarrow{PS(t)}$;
- λ correspond à la longueur d'onde du capteur RADAR ;
- f_{DD} correspond à la fréquence Doppler de datation, indiquée comme donnée auxiliaire (aujourd'hui, toutes les images RADAR sont datées à Doppler nul) ;
- A_T (demi grand-axe) et B_T (demi petit-axe) définissent l'ellipsoïde de référence utilisé (on se place typiquement dans le système WGS 84, pour lequel on a : $A_T = 6378$ km et $B_T = 6357$ km) ;
- h_P donne l'altitude supposée connue du point P au dessus de la surface de référence (hauteur sursol).

Les relations de correspondance, permettant de passer des coordonnées image du point P à l'instant t auquel il est imagé et à la distance point - satellite $r(t)$ à cet instant, sont données par :

$$\left\{ \begin{array}{l} t = \frac{1}{FRI} i_{P,radar} + t_0 \\ r(t) = \frac{c}{2F_d} j_{P,radar} + r_0 \end{array} \right\} \quad (5.3)$$

avec les notations complémentaires suivantes :

- t_0 et r_0 sont des constantes données dans les paramètres système (t_0 correspond à la "date de référence" à l'acquisition de la première ligne et r_0 correspond à la "distance au plus proche") ;
- FRI est la fréquence de répétition des impulsions ;
- c est la vitesse de la lumière ;
- F_d est la fréquence d'échantillonnage en distance.

Localisation inverse La localisation inverse se fait à partir des deux premières équations du système 5.2 (équation de la "sphère distance" et équation du "cône Doppler"). Les coordonnées $(X_P, Y_P, Z_P)_{terrestre}$ étant données, on a un système de deux équations à deux inconnues (t et $r(t)$), que l'on résout numériquement. Une fois connu l'instant t_P auquel le point P est imagé et la distance point - satellite $r(t_P)$ à cet instant, il suffit d'inverser les relations de correspondance présentées en 5.3 pour en déduire la position du point P sur l'image RADAR.

5.2.3 Projection depuis une image optique vers une image RADAR

La projection d'un point d'intérêt P depuis une image optique vers une image RADAR se fait par composition de transformations de la façon suivante :

- premièrement, en utilisant la fonction de localisation directe liée au capteur optique, permettant de passer, sous une hypothèse de hauteur sursol h_P , des coordonnées image $(i_P, j_P)_{optique}$ du point P dans le référentiel image optique à ses coordonnées terrain $(X_P, Y_P, Z_P)_{terrestre}$ dans le référentiel terrestre ;
- deuxièmement, en utilisant la fonction de localisation inverse liée au capteur RADAR, permettant de passer des coordonnées terrain $(X_P, Y_P, Z_P)_{terrestre}$ à ses coordonnées image $(i_P, j_P)_{radar}$ dans le référentiel image RADAR.

Dans la suite, la concaténation des deux modèles optique et RADAR, utilisés pour projeter des points d'intérêt depuis l'image optique vers l'image RADAR, sera appelée "modèle joint de projection".

5.3 Méthode proposée pour le recalage

5.3.1 Stratégie globale

L'objectif est d'obtenir une fine superposition au sol entre éléments homologues optiques-RADAR, afin de tirer profit conjointement des deux sources d'informations complémentaires disponibles.

Pour cela, on propose un procédé de recalage basé sur les primitives caractéristiques suivantes :

- côté optique : les arêtes au sol des bâtiments, issues des contours rectangulaires extraits sur l'image optique ;
- côté RADAR : leurs primitives homologues RADAR correspondantes (lorsque celles-ci sont visibles), c'est-à-dire les échos linéaires brillants issus de doubles réflexions mur-sol.

La méthodologie de recalage présentée se base sur une approche plus élaborée qu'une simple corrélation. Elle fait appel à la combinaison d'outils spatiaux (contrainte de visibilité), radiométriques (score RADAR de recalage et contrainte de radiométrie) et contextuels (contrainte de tendance) intégrant des informations a priori.

Elle utilise également un certain nombre d'informations issues de données auxiliaires (modèles de projection à disposition et hauteurs terrain approximatives issues du MNT).

Elle peut être décomposée en deux phases :

- la première phase, relative à la projection des emprises, fournit une "carte de localisation des contours RADAR projetés" ;
- la seconde phase, relative à l'affinage du recalage entre arêtes optiques et échos doubles RADAR, fournit une "carte de localisation des contours RADAR recalés".

5.3.2 Phase 1 : projection des primitives optiques

5.3.2.1 Modèles de projection disponibles pour nos données

On rappelle que :

- pour l'image optique Quickbird, nous disposons d'un modèle polynômial de projection (modèle physique non disponible) appelé modèle RPC (Quickbird Rational
-

Polynomial Coefficients) ; celui-ci donne directement l'ensemble des coefficients nécessaires pour coder les fonctions de localisation optiques. Étant donné que seuls les coefficients de l'équation de localisation inverse sont fournis dans le fichier .RPB, un schéma itératif d'optimisation (basé sur une stratégie d'aller-retour image-terrain) est employé pour coder la fonction de localisation directe optique ici utilisée ;

- pour l'image RADAR TerraSAR-X, nous disposons d'un vrai modèle physique de projection, grâce au fichier .XML livré avec l'image ; celui-ci fournit l'ensemble des paramètres décrivant le modèle physique de prise de vue : paramètres d'orbite (précisément connus grâce à la stabilité de la trajectoire du satellite), paramètres système et paramètres de traitement. Ceux-ci sont utilisés pour coder les fonctions de localisation RADAR et en particulier la fonction de localisation inverse RADAR ici utilisée.

5.3.2.2 Affinage optionnel des paramètres du "modèle joint de projection"

En imagerie satellitaire, le "modèle joint de projection" peut éventuellement être affiné grâce à l'introduction de quelques points homologues au sol, saisis manuellement par l'opérateur sur les deux images (exemple : intersection de routes, coins de bâtiments).

Il s'agit de ré-estimer les valeurs des composantes du vecteur $\vec{\theta}$, relatif aux paramètres du "modèle joint de projection" (ici correspondant uniquement à ceux du modèle physique RADAR), par minimisation d'une erreur quadratique moyenne $\epsilon_{EQM}(\vec{\theta})$ définie par :

$$\epsilon_{EQM}(\vec{\theta}) = \sum_{n=1}^{N_{points-homologues}} \left\| P\text{-joint}_{O \rightarrow R}^{\vec{\theta}} [U_{n,optique}] - V_{n,radar} \right\|^2 \quad (5.4)$$

avec les notations complémentaires suivantes :

- $N_{points-homologues}$ est le nombre de points homologues utilisés ;
- $U_{n,optique}$ désigne un point saisi par l'opérateur dans l'image optique et $V_{n,radar}$ son point homologue dans l'image RADAR ;
- $\| \quad \|$ désigne la norme euclidienne dans l'espace bidimensionnel des coordonnées image.

5.3.2.3 Mise en œuvre de la projection

Pour chaque emprise rectangulaire potentielle détectée sur l'image optique, l'étape de projection (depuis l'image optique vers l'image RADAR) est appliquée aux quatre points au sol définissant les sommets du rectangle considéré.

Dans la suite, ces points d'intérêt seront appelés "points sommets".

L'information de hauteur utilisée pour mettre en œuvre cette projection est issue du MNT, qui fournit, pour tout "point sommet" de l'emprise, une information de hauteur au sol approximative.

Après la phase de projection, on constate de légères erreurs de recalage optique-RADAR sur les "cartes de localisation des contours RADAR projetés" (de l'ordre de quelques pixels). Les causes à l'origine de telles erreurs sont décrites en section 5.3.3 et les décalages observés sont précisément quantifiés en section 5.4.3.

5.3.3 Phase 2 : affinage du recalage sur l'image RADAR

5.3.3.1 Motivations

Parmi les différentes causes à l'origine des problèmes de recalage observés après projection, on peut citer les sources d'erreurs suivantes :

- premièrement, les hauteurs au sol fournies par le MNT utilisé ne présentent pas une précision altimétrique suffisante (on rappelle que cette précision est de 10 à 20 m à 90 %). Des erreurs altimétriques de géo-référencement sont alors générées dans le référentiel terrestre ;
- deuxièmement, le modèle RPC optique utilisé peut induire de légères erreurs (altimétriques et planimétriques) de géo-référencement lors de l'étape de projection depuis le référentiel image optique vers le référentiel terrestre intermédiaire ;
- troisièmement, certaines erreurs planimétriques peuvent provenir de l'étape d'extraction des contours de bâtiments en monoscopie optique, notamment en termes de dimensionnement et d'orientation ;
- quatrièmement, de petits décalages en termes de positionnement planimétrique sont dus à l'hypothèse approximativement vérifiée de visée optique au nadir (on rappelle qu'en visée oblique, le phénomène de parallaxe implique une translation entre la localisation des contours de toits et la localisation des emprises au sol).

Les erreurs altimétriques et planimétriques de géo-référencement ainsi causées sont à l'origine d'erreurs le long des axes RADAR distance et azimuth, obtenues après l'étape de projection des "points sommets" depuis le référentiel terrestre intermédiaire vers le référentiel image RADAR.

De façon à obtenir un recalage au sol fin, il est donc nécessaire d'estimer localement, pour les "points sommets", des hauteurs de terrain très précises.

Nous proposons ainsi de mettre en place une phase d'affinage du recalage permettant d'obtenir conjointement un MNT plus précis et un appariement des "points sommets" optiques-RADAR satisfaisant.

5.3.3.2 Hypothèses et principe

Ce procédé utilise la correspondance qui existe entre les arêtes au sol de bâtiments extraites sur l'image optique et leurs échos doubles RADAR homologues.

On suppose que ces échos sont clairement visibles sur la donnée RADAR et qu'ils représentent la majorité des structures linéaires brillantes.

L'idée consiste à modifier les coordonnées intermédiaires des "points sommets" dans le référentiel terrestre de façon à assurer une fine superposition entre ces primitives caractéristiques, une fois réalisée l'étape de projection depuis le référentiel terrestre intermédiaire vers le référentiel image RADAR.

Le procédé de recalage présenté se décompose en deux sous-phases : une optimisation des déplacements et une régularisation des déplacements.

5.3.3.3 Optimisation des déplacements

Étant donné qu'après la phase de projection, les erreurs de recalage observées sont légères, seuls de petits déplacements seront envisagés dans le repère terrestre intermédiaire.

On cherche à trouver le jeu de déplacements optimal $(\delta_x, \delta_y, \delta_z)^*$ à appliquer aux "points sommets" dans le référentiel terrestre (une fois projetés depuis l'image optique), de façon à définir de nouveaux "points sommets" qui optimisent un score RADAR de recalage $S_{recalage}$, ce dernier étant calculé après la projection des nouveaux points depuis le référentiel terrestre vers le référentiel image RADAR (voir figure 5.1).

On rappelle que chaque jeu de déplacements 3D $(\delta_x, \delta_y, \delta_z)$ appliqué dans le référentiel terrestre trouve un jeu équivalent de translations 2D $(t_{distance}, t_{azimuth})$ appliqué dans le référentiel image RADAR.

On suppose un sol localement plat pour chaque bâti individuel, ce qui implique que les "points sommets" (situés au sol) d'une même emprise soient situés à une même altitude terrain.

On considère les notations complémentaires suivantes :

- T_k définit l'opérateur de translation 3D, appliqué dans le référentiel terrestre avec le jeu de déplacements 3D $(\delta_x, \delta_y, \delta_z)^k$, à la k-ième itération du procédé d'optimisation ;
- $(t_{distance}, t_{azimuth})^k$ correspond au jeu de translations 2D dans le référentiel image RADAR induit par le jeu de déplacements 3D $(\delta_x, \delta_y, \delta_z)^k$ dans le référentiel terrestre, après projection des points depuis le référentiel terrestre vers le référentiel image RADAR, à la k-ième itération du procédé d'optimisation.

Algorithme Pour chaque contour de bâtiment potentiellement extrait sur l'image optique, on procède de la façon suivante :

1. Initialisation

- Les deux segments d'arêtes au sol optiques $(Sab, Sac)_{optique}$ susceptibles d'être imagés en tant qu'échos doubles mur-sol sont désignés dans le référentiel image optique étant donnés leur emplacement et la position du capteur RADAR.
- Les trois points $(Pa, Pb, Pc)_{optique}$ situés à l'extrémité des deux segments $(Sab, Sac)_{optique}$ sont déduits dans le référentiel image optique.
- Ces points $(Pa, Pb, Pc)_{optique}$ sont projetés dans le référentiel terrestre et fournissent les points notés $(Pa, Pb, Pc)_{terrestre}^0$ (qui serviront de points d'initialisation au procédé itératif) :

$$(Pa, Pb, Pc)_{terrestre}^0 = P_{O \rightarrow T} [(Pa, Pb, Pc)_{optique}] \quad (5.5)$$

où la projection s'effectue en utilisant pour chaque point une valeur de hauteur identique z_{moyen} issue du MNT disponible et estimée en moyennant les trois hauteurs des trois points $(Pa, Pb, Pc)_{optique}$.

- ### 2. Optimisation de la superposition entre primitives homologues
- Le schéma itératif d'optimisation est mis en œuvre à travers un algorithme "gourmand", qui teste exhaustivement un ensemble de déplacements possibles $(\delta_x, \delta_y, \delta_z)^k$ (appliqués dans le référentiel terrestre) de la façon suivante :

- définition des points tests $(Pa, Pb, Pc)_{terrestre}^k$ à l'itération k dans le référentiel terrestre par :

$$(Pa, Pb, Pc)_{terrestre}^k = T_k [(Pa, Pb, Pc)_{terrestre}^0] \quad (5.6)$$

- obtention de leur points tests équivalents $(Pa, Pb, Pc)_{radar}^k$ dans le référentiel image RADAR par :

$$(Pa, Pb, Pc)_{radar}^k = P_{T \rightarrow R} \left[(Pa, Pb, Pc)_{terrestre}^k \right] \quad (5.7)$$

où la projection s'effectue en utilisant pour chaque point la valeur de hauteur identique ($z_{moyen} + (\delta_z)^k$);

- calcul et enregistrement du score de recalage $S_{recalage}^k$ relatif aux deux segments $(Sab, Sac)_{radar}^k$ (associé aux trois points $(Pa, Pb, Pc)_{radar}^k$) par :

$$S_{recalage}^k = \sum_{(i,j) \in (Sab \cup Sac)_{radar}^k} \frac{r_{(i,j)}}{L_{(Sab \cup Sac)_{radar}^k}} \quad (5.8)$$

où $r_{(i,j)}$ correspond à la radiométrie du pixel (i, j) dans l'image RADAR et $L_{(Sab \cup Sac)_{radar}^k}$ correspond à la longueur cumulée des deux segments $(Sab, Sac)_{radar}^k$;

- estimation du jeu de déplacement optimal $(\delta_x, \delta_y, \delta_z)^*$ (qui assure le meilleur appariement entre arêtes au sol optiques et échos doubles RADAR) par maximisation de score :

$$(\delta_x, \delta_y, \delta_z)^* = \arg \max_{(\delta_x, \delta_y, \delta_z)^k} \left(S_{recalage}^k \right) \quad (5.9)$$

- déduction de la localisation optimale des trois points $(Pa, Pb, Pc)_{terrestre}^*$ dans le référentiel terrestre en appliquant l'équation (5.6) avec le jeu $(\delta_x, \delta_y, \delta_z)^*$;
- déduction de la localisation optimale des trois points projetés équivalents $(Pa, Pb, Pc)_{radar}^*$ dans le référentiel image RADAR en appliquant l'équation (5.7) avec la hauteur optimale $z^* = z_{moyen} + (\delta_z)^*$.

3. **Affinage local du MNT** Le MNT affiné est généré à partir des valeurs z^* localement obtenues pour chaque emprise de bâtiment.
4. **Conversion des jeux optimaux** Le jeu de déplacements 3D optimal $(\delta_x, \delta_y, \delta_z)^*$ dans le référentiel terrestre est converti en un jeu équivalent de translations 2D optimales $(t_{distance}, t_{azimuth})^*$ dans le référentiel image RADAR pour la seconde sous-phase.

5.3.3.4 Régularisation des déplacements

Dans l'étape d'optimisation précédente, nous supposons que l'hypothèse relative à la présence d'échos doubles visibles et majoritaires est vérifiée pour tout bâti et nous appliquons ainsi de façon indépendante le même traitement à chaque bâtiment individuellement.

Or, sur des données RADAR réelles, cette hypothèse n'est pas toujours vraie.

En effet, il existe en pratique de nombreuses raisons empêchant la "zone d'écho double" d'apparaître sur l'image RADAR : présence d'objets environnants (bosquets de végétation, bâti voisin) susceptibles de masquer le coin mur-sol, existence d'un sol irrégulier à l'origine d'une onde de rétrodiffusion envoyée dans une direction différente de celle du capteur...

En outre, rien n'assure dans l'étape précédente que les primitives linéaires de forte radiométrie auxquelles se sont "accrochées" les arêtes optiques (lors de l'optimisation du score de recalage) soient effectivement les primitives issues d'échos doubles mur-sol. Il pourrait en effet également s'agir de primitives linéaires brillantes issues de la réponse

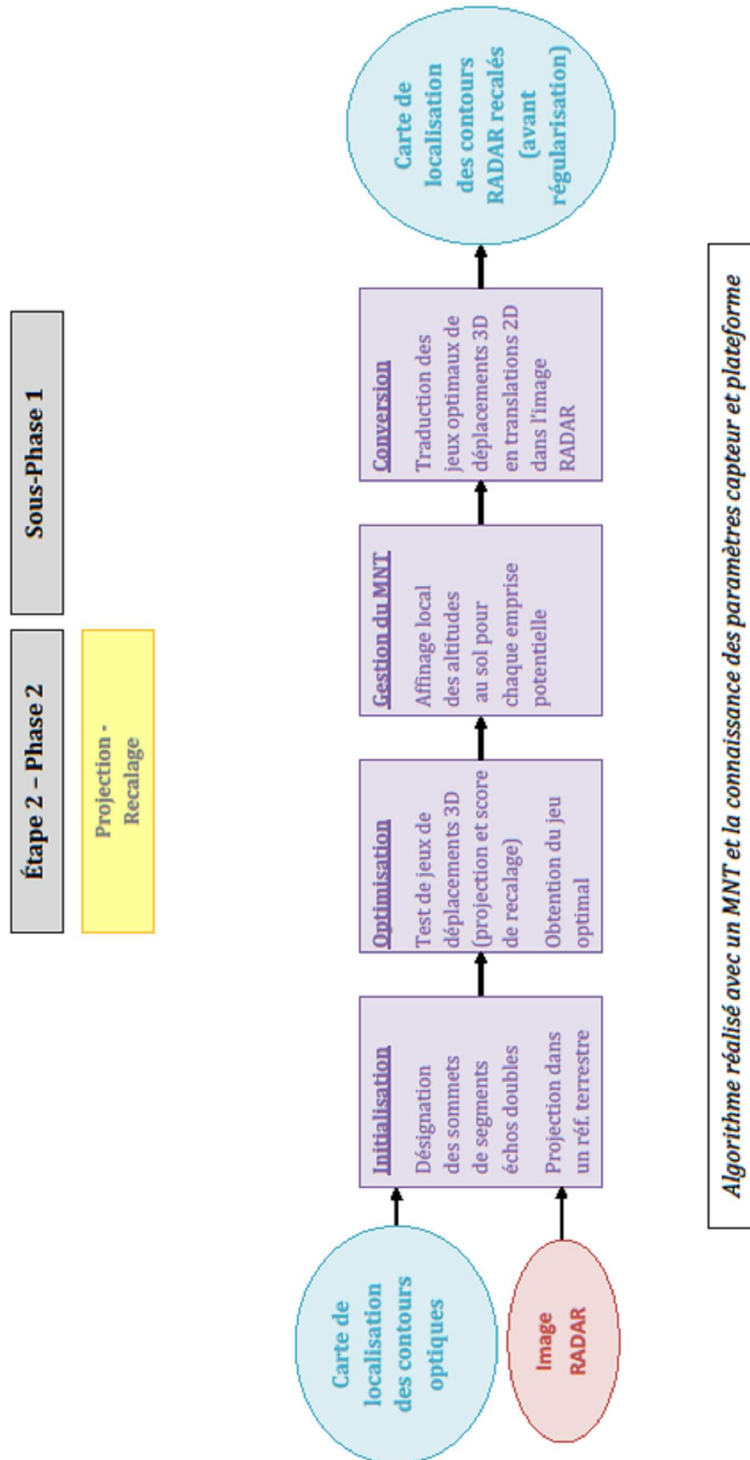


FIGURE 5.1 – Schéma synoptique de la phase 2 de l'étape 2 (sous-phase 1).

d'autres éléments ou parties du bâti (au sol ou en hauteur) qui auraient fortement répondu (exemples : échos linéaires simples issus de gouttières métalliques ou échos linéaires doubles issus de parapets sur le toit). Or, il nous est bien sûr impossible, à ce niveau de la chaîne, de distinguer les échos doubles murs-sol effectifs des autres échos, juste à partir d'une information de radiométrie.

Nous suggérons ainsi de réaliser une étape de régularisation sous contraintes afin d'obtenir des résultats de recalage plus robustes.

L'idée consiste à sélectionner, parmi les translations optimales $(t_{distance}, t_{azimuth})^*$ précédemment calculées, celles qui se révèlent être les plus fiables et de les propager localement (voir figure 5.2).

Dans ce but, nous considérons que le MNT présente de faibles variations et que les erreurs de géo-référencement sont localement similaires.

Algorithme Pour chaque contour de bâtiment projeté sur l'image RADAR et recalé suivant la translation optimale $(t_{distance}, t_{azimuth})^*$, on procède de la façon suivante :

1. **Évaluation de contraintes** Trois contraintes nommées $C_{visibilite}$, $C_{radiometrie}$ et $C_{tendance}$ sont introduites et examinées. Elles s'appuient sur des relations spatiales, radiométriques et contextuelles, qui intègrent des connaissances a priori.

Ces contraintes sont définies par :

- $C_{visibilite}$ se rapporte à la possibilité pour les deux segments arêtes désignés $(Sab, Sac)_{RADAR}^*$ d'être vus comme échos doubles sur l'image RADAR, étant donné leur environnement spatial. Cette contrainte $C_{visibilite}$ est vérifiée si, à proximité des segments considérés, aucun autre bâti ne se trouve suffisamment proche pour cacher les coins mur-sol associés ;
- $C_{radiometrie}$ se réfère au score de recalage $S_{recalage}^*$ obtenu après optimisation pour l'emprise considérée. Celui-ci doit être suffisamment élevé pour garantir un recalage effectif sur les échos doubles très fortement brillants. Cette contrainte $C_{radiometrie}$ est vérifiée si $S_{recalage}^*$ est supérieur à un seuil $S_{echo-double}$;
- $C_{tendance}$ se réfère aux valeurs de translations optimales $(t_{distance}, t_{azimuth})^*$ obtenues pour l'emprise considérée. Celles-ci doivent être localement relativement similaires les unes aux autres afin de pas être jugées comme aberrantes. Cette contrainte $C_{tendance}$ est vérifiée si les valeurs $(t_{distance}, t_{azimuth})^*$ sont suffisamment proches d'une valeur, appelée valeur tendance $(t_{distance}, t_{azimuth})_{tendance}^*$ issue d'un moyennage local sur les valeurs de translations relatives aux bâtis voisins.

2. **Sélection des translations fiables** La décision quant à la fiabilité des translations optimales $(t_{distance}, t_{azimuth})^*$ obtenues est alors prise. Lorsque les trois contraintes sont simultanément vérifiées, la translation optimale considérée est jugée fiable et validée. Le cas contraire (au moins une des trois contraintes n'est pas vérifiée) signifie probablement que les arêtes optiques du bâti considérées ont été recalées dans l'image RADAR en s'accrochant à des structures linéaires différentes des échos doubles souhaités : la translation optimale considérée doit donc être rejetée.

3. **Propagation des translations fiables** Dans le cas où la translation optimale considérée a été rejetée, celle-ci est remplacée par la translation optimale la plus proche qui a été validée pour un bâti voisin.

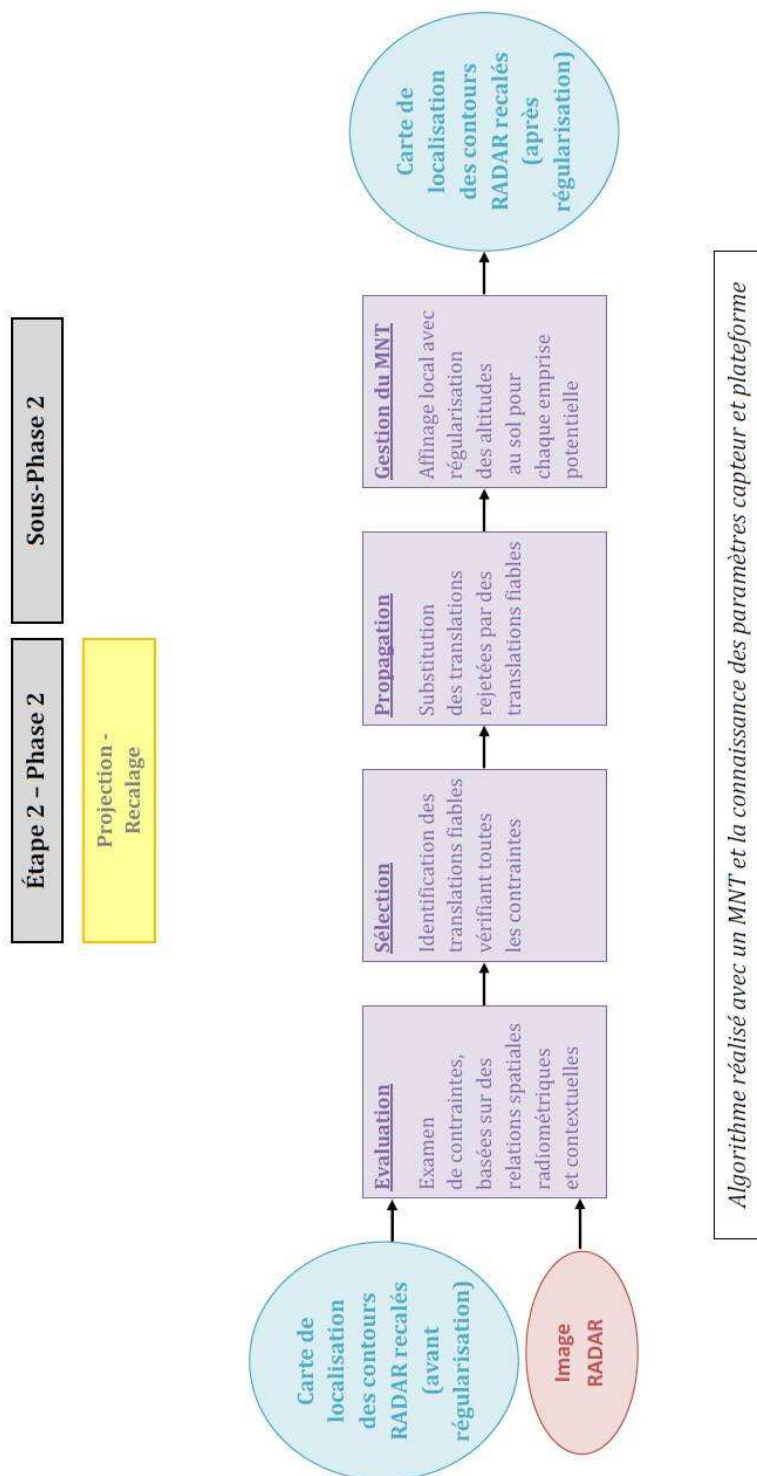


FIGURE 5.2 – Schéma synoptique de la phase 2 de l'étape 2 (sous-phase 2).

4. **Affinage local du MNT avec régularisation** L'altitude au sol z^* estimée au niveau de l'emprise considérée est également régularisée en utilisant une stratégie similaire. Si z^* induit une valeur de translation optimale $(t_{distance}, t_{azimuth})^*$ validée, elle est conservée telle quelle. Si z^* correspond au contraire à une valeur rejetée, elle est remplacée par la moyenne des altitudes au sol z^* qui ont été conservées pour les bâtis voisins.

À la fin du procédé de recalage, on obtient des "cartes de localisation des contours RADAR recalés", indiquant l'emplacement précis (dans le référentiel image RADAR) des emprises potentielles recalées.

5.3.4 Discussion sur la gestion des paramètres

Paramètre de la phase 1 Le seul paramètre utilisé dans la phase de projection est le suivant :

- le nombre $N_{points-homologues}$ de points homologues au sol, manuellement extraits sur l'image optique et sur l'image RADAR et susceptibles d'être utilisés afin d'affiner le "modèle joint de projection" optique-RADAR (il s'agit d'un paramètre optionnel, d'ailleurs non spécifique à notre démarche).

Paramètres de la phase 2 Les différents paramètres utilisés dans la phase d'affinage du recalage sont les suivants :

- les valeurs de translations maximales $t_{distance}^{max}$ et $t_{azimuth}^{max}$ envisagées dans le référentiel image RADAR lors de l'étape d'optimisation des déplacements ; ces valeurs sont empiriquement fixées à quelques pixels ;
- les valeurs des pas $p_{distance}$ et $p_{azimuth}$ utilisés entre les différentes translations testées ; ces paramètres dépendent du compromis adopté entre la précision de recalage souhaitée et le temps de calcul dédié ;
- le diamètre $d_{voisinage}$ utilisé pour définir la taille du voisinage à considérer autour de chaque bâtiment pour valider la contrainte $C_{visibilite}$; ce diamètre est empiriquement fixé à quelques pixels ;
- le seuil $S_{echo-double}$ utilisé pour valider la contrainte $C_{radiometrie}$; celui-ci peut être réglé par apprentissage sur les zones des doubles échos pour un capteur RADAR donné ;
- les deux seuils $t_{distance}^{tolerance}$ et $t_{azimuth}^{tolerance}$ utilisés pour spécifier l'écart maximal autorisé entre les valeurs $(t_{distance}, t_{azimuth})^*$ et $(t_{distance}, t_{azimuth})_{tendance}^*$, lors de la validation de la contrainte $C_{tendance}$.

5.4 Application sur les scènes d'études

La seconde étape de la chaîne de traitements proposée est mise en œuvre sur les données réelles pour les deux scènes d'étude considérées.

5.4.1 Paramètres utilisés

Les paramètres utilisés pour la première phase et pour la seconde phase de l'étape d'extraction des bâtiments en monoscopie optique sont indiqués en table 5.1.

TABLE 5.1 – Paramètres utilisés pour l'étape de projection et de recalage des emprises potentielles.

Noms des paramètres	Scène A	Scène B
$N_{points-homologues}$	7	7
$t_{distance}^{max} (pixels)$	+/- 3	+/- 6
$t_{azimuth}^{max} (pixels)$	+/- 6	+/- 12
$p_{distance} (pixels)$	0.25	0.25
$p_{azimuth} (pixels)$	0.25	0.25
$d_{voisinage} (pixels)$	5	5
$S_{echo-double}$	450	450
$t_{distance}^{tolerance} (pixels)$	+/- 3	+/- 3
$t_{azimuth}^{tolerance} (pixels)$	+/- 6	+/- 6

Les valeurs des seuils sont données pour une image RADAR TerraSAR-X codée sur 32 bits (16 bits pour la partie réelle et 16 bits pour la partie imaginaire).

Toutes les valeurs sont identiques pour les deux scènes, exceptées celles prises pour $t_{distance}^{max}$ et $t_{azimuth}^{max}$. Des translations d'amplitude maximale supérieure sont autorisées en distance et en azimuth pour la scène B. Ceci s'explique par le fait que, l'hypothèse de sol localement plat n'étant pas vérifiée partout sur cette scène, des erreurs de recalage plus importantes sont susceptibles d'être présentes après la phase de projection. De façon générale, ces valeurs peuvent être fixées à une vingtaine de pixels afin d'assurer un recalage correct, même si cela implique un temps de calcul légèrement supérieur mais qui reste toutefois opérationnel.

5.4.2 Résultats obtenus

Les figures 5.3 et 5.4 présentent les deux "cartes de localisation des contours RADAR projetés", obtenues après la phase de projection, suite à l'utilisation du "modèle joint de projection" initial et affiné (grâce à l'introduction de quelques points homologues).

La figure 5.5 présente les deux "cartes de localisation des contours RADAR recalés", obtenues après la phase d'affinage du recalage à la fin de la deuxième étape.

Pour la phase de projection, on utilise, individuellement pour chaque point, sa hauteur propre fournie par le MNT disponible.

Pour la phase d'affinage du recalage, on utilise, pour chaque point appartenant à une même emprise, une valeur de hauteur identique z^* estimée lors du processus d'optimisation - régularisation (hypothèse de sol localement plat introduite en section 5.3.3.3).

Les emprises rectangulaires projetées et recalées sont représentées en rouge et sont superposées sur l'image RADAR en niveaux de gris.

5.4.3 Analyse qualitative et quantitative

On constate visuellement, qu'après la première phase (projection), quelques légères erreurs de recalage sont présentes. Celles-ci sont inférieures en moyenne à une vingtaine de pixels lorsque le "modèle joint de projection" utilisé est le modèle initial et sont inférieures en moyenne à une dizaine de pixels lorsque le "modèle joint de projection" utilisé est le modèle affiné (grâce à la saisie manuelle de quelques points homologues). Ces estimations moyennes sont cohérentes étant données la résolution spatiale à laquelle on



(a) Scène A (RADAR) : emprises projetées (en rouge) ©Infoterra

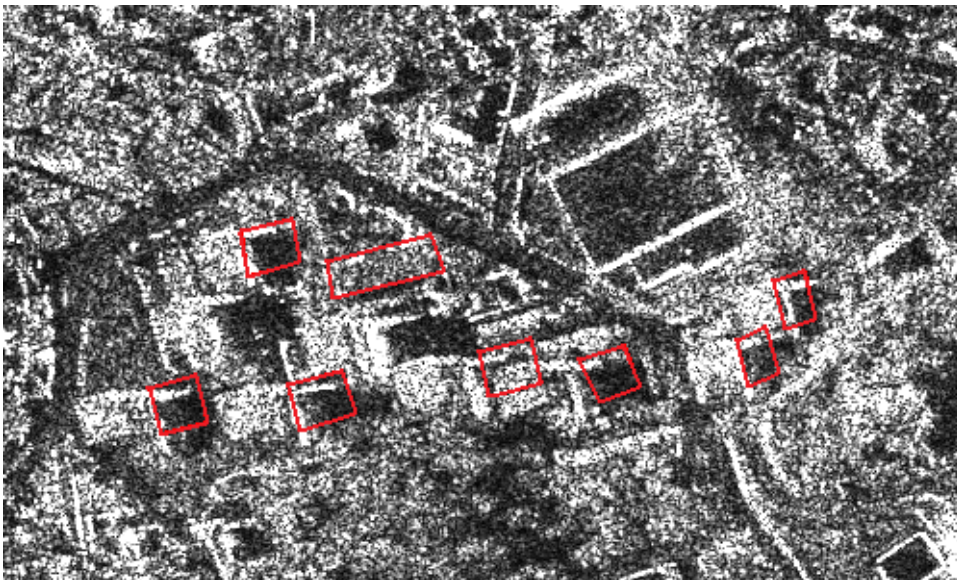


(b) Scène B (RADAR) : emprises projetées (en rouge) ©Infoterra

FIGURE 5.3 – Carte de localisation des contours RADAR projetés pour les scènes d'étude A et B. Le "modèle joint de projection" utilisé ici est le modèle initial, basé sur la seule connaissance du modèle RPC optique et du modèle physique RADAR (utilisant les paramètres capteurs).



(a) Scène A (RADAR) : emprises projetées (en rouge) ©Infoterra

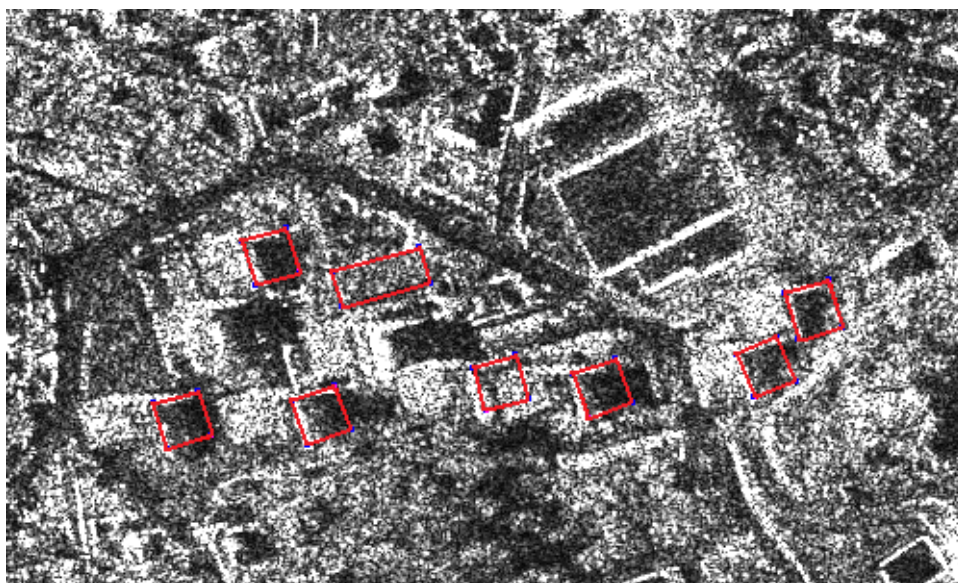


(b) Scène B (RADAR) : emprises projetées (en rouge) ©Infoterra

FIGURE 5.4 – Carte de localisation des contours RADAR projetés pour les scènes d'étude A et B. Le "modèle joint de projection" utilisé ici est le modèle joint affiné, ré-estimé grâce à l'introduction de quelques points homologues.



(a) Scène A (RADAR) : emprises recalées en fin d'étape 2 (en rouge) ©Infoterra



(b) Scène B (RADAR) : emprises recalées en fin d'étape 2 (en rouge) ©Infoterra

FIGURE 5.5 – Carte de localisation des contours RADAR recalés pour les scènes d'étude A et B.

travaille et la précision du MNT DIMAP utilisé.

Après la seconde phase (affinage du recalage), le recalage obtenu est globalement plus précis et plus robuste. Suite aux sous-phases d'optimisation et de régularisation, les arêtes au sol extraites de l'image optique se trouvent précisément positionnées au milieu des échos doubles RADAR. Les erreurs observées sont inférieures en moyenne à trois pixels, ce qui est tout à fait satisfaisant.

Les tables 5.2 et 5.3 résument les résultats obtenus individuellement pour chacun des bâtiments composant les deux scènes d'étude.

On y indique :

- les valeurs des scores maximaux de recalage $S_{recalage}^*$ obtenus en fin d'optimisation ;
- les statuts (V : vérifiée - NV : non vérifiée) des contraintes $C_{visibilite}$, $C_{radiometrie}$ et $C_{tendance}$, obtenues lors de la régularisation ;
- le statut (fiable ou régularisée) des translations optimales $(t_{distance}, t_{azimuth})^*$, déduit de la vérification ou non des contraintes.

Le statut de la contrainte $C_{radiometrie}$ est examiné en premier lieu. Si celle-ci est vérifiée (V), nous examinons alors le statut de la contrainte $C_{tendance}$. Si celle-ci est vérifiée (V), nous examinons enfin le statut de la contrainte $C_{visibilite}$. Dès lors qu'une contrainte est non vérifiée (NV) pour un bâti, il devient en fait inutile, pour ce dernier, de s'intéresser aux statuts des contraintes suivantes.

Durant la régularisation, seules les translations relatives à 4 bâtiments dans la scène A et à 6 bâtiments dans la scène B ont été jugées fiables et ont été utilisées pour régulariser l'ensemble des autres translations.

5.5 Étude détaillée de quelques bâtiments d'intérêt

Emprises recalées après optimisation sans régularisation La figure 5.6 présente trois exemples de bâtiments, dont les emprises ont été recalées par optimisation sans régularisation.

Pour une facilité d'interprétation, les arêtes visualisées (en rouge) sont représentées avec un trait plus fin que celui utilisé pour la figure 5.5.

Le capteur RADAR est situé à la gauche de l'image et il image sur sa droite. Une ligne de l'image est donc parallèle à l'axe des distances et une colonne de l'image est parallèle à celui des azimuths.

En reprenant les notations de la section 5.3.3.3, les trois points $(Pa, Pb, Pc)_{RADAR}^*$ (définissant au sol les deux échos doubles RADAR) correspondent dans chaque cas aux trois points situés le plus à gauche dans l'image.

Les translations optimales individuellement associées à chacun des bâtiments sont jugées fiables (toutes contraintes satisfaites) et sont directement utilisées pour recalculer les emprises au sol.

La superposition précise ainsi obtenue entre arêtes optiques (fines) et échos doubles RADAR (plus larges et s'étalant sur 1 à 4 pixels) est mise en évidence.

Emprises recalées après optimisation et régularisation La figure 5.7 présente trois exemples de bâtiments, dont les emprises ont été recalées par optimisation et régularisation.

TABLE 5.2 – Résultats individuels de recalage des bâtiments sur la scène d'étude A.

Label issu de l'étape d'extraction	$S_{recalage}^*$	$C_{radiometrie}$	$C_{tendance}$	$C_{visibilite}$	Statut de la translation optimale
1	195	NV	-	-	régularisée
2	180	NV	-	-	régularisée
3	849	V	V	V	fiable
4	1551	V	V	V	fiable
5	378	NV	-	-	régularisée
6	409	NV	-	-	régularisée
7	682	V	V	V	fiable
8	725	V	V	V	fiable
9	340	NV	-	-	régularisée

TABLE 5.3 – Résultats individuels de recalage des bâtiments sur la scène d'étude B.

Label issu de l'étape d'extraction	$S_{recalage}^*$	$C_{radiometrie}$	$C_{tendance}$	$C_{visibilite}$	Statut de la translation optimale
1	615	V	V	V	fiable
2	523	V	V	V	fiable
3	561	V	V	V	fiable
4	588	V	V	V	fiable
5	491	V	V	V	fiable
6	363	NV	-	-	régularisée
7	324	NV	-	-	régularisée
8	509	V	V	V	fiable

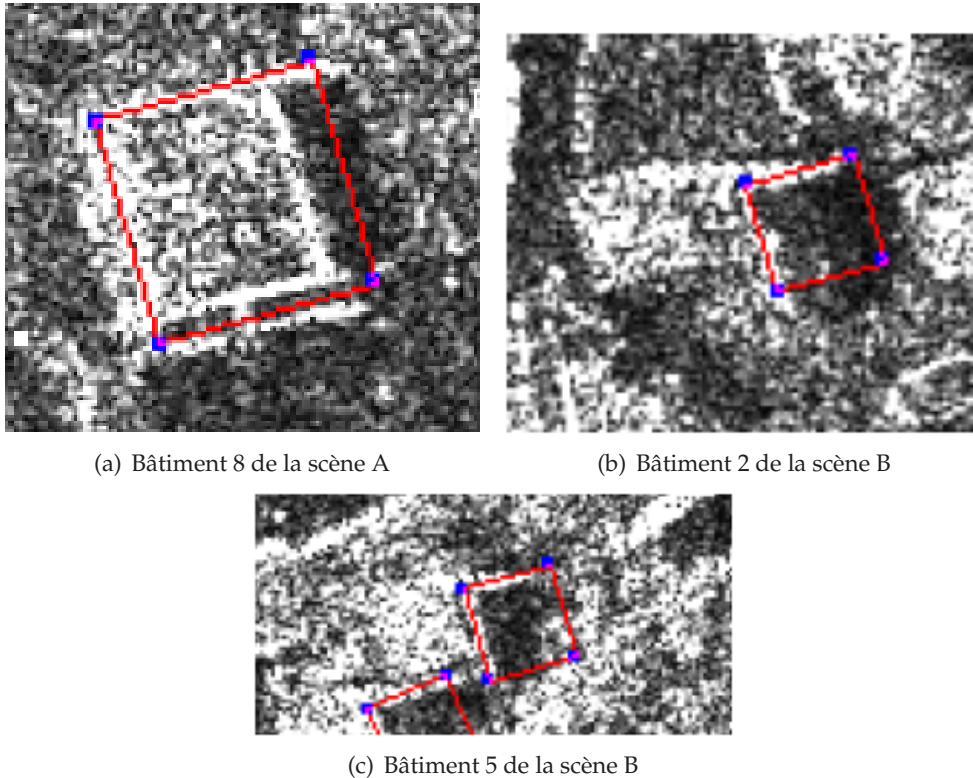


FIGURE 5.6 – Exemples de bâtiments recalés après optimisation sans régularisation sur les scènes d'étude A et B.

Les translations optimales individuellement associées à chacun des bâtiments sont rejetées (au moins une contrainte non satisfaite) et sont remplacées par les translations fiables les plus proches.

Pour le bâtiment 2 de la scène A, la contrainte de visibilité n'est pas vérifiée. Celui-ci est trop proche du bâtiment 7 (son coin diédrique est donc probablement partiellement caché).

Pour les bâtiments 6 de la scène A et 7 de la scène B, la contrainte de radiométrie n'est pas vérifiée. Les scores maximaux de recalage $S_{recalage}^*$, obtenus en fin d'optimisation (avant régularisation) et respectivement égaux à 409 et 324, restent inférieurs au seuil fixé ($S_{echo-double} = 450$). Ceci s'explique :

- pour le bâtiment 6 de la scène A, par le fait que la translation optimale (avant régularisation) se soit "accrochée" à un élément brillant présent en bord de toit et non à l'écho double souhaité (partiellement caché par le bâtiment 4) ;
- pour le bâtiment 7 de la scène B, par le fait qu'il s'agisse d'une fausse alarme issue de l'étape de détection.

5.6 Conclusions et améliorations

Une nouvelle méthode de projection et de recalage, basée sur un procédé d'optimisation - régularisation, a été proposée dans ce chapitre et a permis l'obtention de résultats précis sur les données réelles testées (fine superposition des primitives linéaires homo-

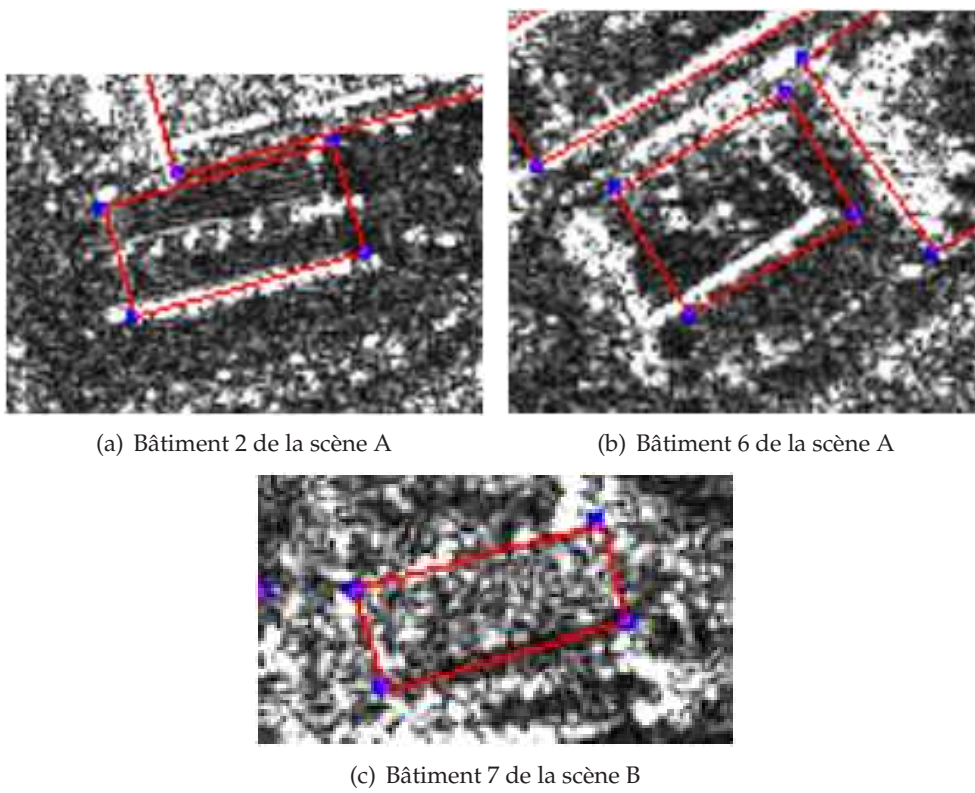


FIGURE 5.7 – Exemples de bâtiments recalés après optimisation avec régularisation sur les scènes d'étude A et B.

logues à l'échelle du mètre).

Les principales limitations de la méthode proposée reposent sur les fortes hypothèses considérées (mais non nécessairement vérifiées en pratique) et relatives à :

- la présence d'échos doubles supposés visibles et majoritaires sur l'image RADAR (hypothèse utilisée lors de l'optimisation pour le calcul des scores de recalage) ;
- la similarité locale entre les erreurs de géo-référencement présentes (hypothèse utilisée lors de la régularisation pour la propagation des translations optimales fiables).

D'un autre côté, il convient de noter que ces hypothèses fortes assurent un recalage très fiable (une simple méthode par corrélation fournirait sans doute des résultats beaucoup plus "élastiques").

Les performances de recalage à l'issue de cette deuxième étape de traitement pourraient être améliorées.

D'une part, la mise en œuvre d'une approche davantage orientée "fusion des trois contraintes", lors de la sélection des translations fiables (sous-phase de régularisation), serait peut être préférable afin de donner moins d'importance aux seuils (en effet, ici, la vérification individuelle de trois tests "en dur" peut rendre parfois la régularisation trop "sévère"). L'utilisation d'opérateurs de fusion adaptés permettrait ainsi de globaliser ou de pondérer les trois contraintes introduites.

D'autre part, une adaptation de certaines méthodes plus globales de régularisation de champs vectoriels avec optimisation de critères, proposées dans la littérature (exemple : méthode du "Gradient Vector Flow" [XP98]) pourrait être envisagée pour la phase d'affinage du recalage avec la recherche de translations optimales.

Chapitre 6

Estimation des hauteurs et validation des bâtiments

La méthode d'estimation des hauteurs et de validation des bâtiments en monoscopie RADAR, présentée dans ce chapitre, est une des contributions principales de nos travaux.

Dans un premier temps, une nouvelle méthodologie, permettant d'estimer les hauteurs des bâtiments potentiels détectés et de valider leur présence, est développée à travers l'utilisation d'un critère RADAR statistique et d'un critère RADAR de contraste. La méthodologie présentée est appliquée et analysée sur des données RADAR simulées, afin de valider la pertinence de l'approche. L'évolution des critères RADAR utilisés est étudiée dans le cas des deux configurations géométriques envisagées pour la signature RADAR des bâtiments. Des "cartes des hauteurs estimées des bâtiments validés" sont fournies à l'issue de cette troisième étape de la chaîne de traitements.

Dans un deuxième temps, un procédé simple, basé sur la combinaison des informations planimétriques et altimétriques issues des différentes cartes obtenues tout au long de la chaîne de traitements, est proposé pour réaliser la reconstruction tridimensionnelle des bâtiments de la scène urbaine.

Dans un troisième temps, la méthode d'estimation des hauteurs et de validation des bâtiments ainsi que le procédé de reconstruction finale de la scène sont mis en œuvre sur les scènes d'études.

6.1 Méthode proposée pour l'estimation des hauteurs et pour la validation des bâtiments sur l'image RADAR

6.1.1 Stratégie globale

L'introduction des données RADAR dans cette étape nous permet de bénéficier du cadre de travail choisi (celui de la fusion d'informations).

L'objectif visé est de conjointement confirmer (ou non) la présence d'un bâtiment potentiel (validation) et d'estimer sa hauteur (extraction de l'information d'élévation).

Une approche semi-automatique, basée sur l'utilisation de deux critères RADAR dépendants de la hauteur du bâti, est présentée.

La méthodologie exposée peut être décomposée en deux phases :

- la première phase, basée sur un critère statistique global, permet de générer un ensemble de hauteurs candidates ;
-

- la seconde phase, basée sur un critère de contraste local permet de sélectionner les hauteurs candidates les plus appropriées puis d'estimer les hauteurs finales potentielles et de valider la présence des bâtiments ; cette phase fournit une "carte des hauteurs estimées des bâtiments validés".

6.1.2 Génération de partitionnements tests sur une imagerie RADAR

6.1.2.1 Rappels : modélisation des signatures RADAR de bâtiments

On rappelle que deux configurations géométriques sont adoptées pour modéliser la signature RADAR d'un bâti.

Celles-ci découlent directement de l'hypothèse de modèle 3D choisi pour le bâtiment (modèle parallélépipédique) et ont été décrites en détails à la section 3.2.3 du chapitre 3.

On rappelle que la configuration 1 a lieu si la hauteur du bâti $h_{bâti}$ vérifie ($h_{bâti} < h_{limite}$) et que la configuration 2 a lieu si celle-ci vérifie ($h_{bâti} > h_{limite}$) avec h_{limite} définie par :

$$h_{limite} = w_{bâti} \tan(\theta_i) \quad (6.1)$$

avec les notations suivantes :

- θ_i correspond à l'angle d'incidence du capteur RADAR ;
- $w_{bâti}$ correspond à "l'épaisseur vue" du bâtiment en "géométrie terrain" suivant l'axe des distances RADAR.

6.1.2.2 Technique utilisée pour la génération de partitionnements tests

Nous décrivons dans cette section la technique de génération employée pour délimiter les contours de chaque région caractéristique d'un bâtiment, à partir de l'emprise du bâti dans l'image RADAR et de la connaissance des paramètres capteurs.

Cette technique se base sur la projection, suivant l'axe des distances obliques (référentiel image RADAR), d'un modèle de parallélépipède 3D présent dans un référentiel terrestre. Elle utilise les relations connues qui régissent la dépendance de l'extension des régions caractéristiques (projetées en "géométrie RADAR" suivant l'axe des distances RADAR) vis-à-vis des dimensions du bâti (voir figure 6.1 et figure 6.2, pour lesquelles on suppose que le capteur est situé à gauche de l'image et qu'il image sur sa droite).

Génération de la "zone de recouvrement" et de la "zone de toit seul" pour la configuration 1 Dans le cas de la configuration 1, la "zone de recouvrement" et la "zone de toit seul" sont définies de façon complémentaire à l'intérieur de la position du contour du toit. Cette dernière est obtenue par une simple translation vers la gauche (sens des distances décroissantes) de l'emprise au sol rectangulaire suivant le vecteur translation $\vec{t}_{recouvrement}$ défini par :

$$\left\| \vec{t}_{recouvrement} \right\| = \frac{h_{bâti} \cos(\theta_i)}{P_d} \quad (\text{en pixels}) \quad (6.2)$$

où P_d correspond au pas d'échantillonnage en distance oblique.

La "zone de recouvrement" est alors située à gauche des deux segments face au capteur (échos doubles), tandis que la "zone de toit seul" est située sur leur droite (voir figure 6.1(a) et figure 6.1(b)).

Génération de la "zone de recouvrement" et de la "zone de façade" pour la configuration 2 Dans le cas de la configuration 2, la "zone de recouvrement" est également définie à l'intérieur de la position du contour du toit. Mais, dans ce cas, la translation $\overrightarrow{t_{recouvrement}}$ effectuée est tellement grande que l'ensemble des pixels situés à l'intérieur du contour du toit constituent la "zone de recouvrement". Le contour du toit ne chevauchant pas du tout la "zone d'emprise au sol", la "zone de recouvrement" est donc ici parfaitement rectangulaire (voir figure 6.1(c)).

La "zone de façade", quant à elle, est définie en joignant les sommets situés sur la partie droite de la "zone de recouvrement" et les sommets situés sur la partie gauche de la "zone d'emprise au sol" (voir figure 6.1(d)).

Génération de la "zone d'ombre" pour la configuration 1 Dans la configuration 1, le bâti est considéré comme un objet totalement occultant (voir figure 6.2(a)). La largeur de l'ombre induite est alors indépendante des dimensions de l'objet ; elle est fonction de la hauteur du bâti $h_{bâti}$ et commence à partir des sommets du contour du toit les plus éloignés du capteur RADAR. L'extension de la "zone d'ombre" correspond à la norme du vecteur translation $\overrightarrow{t_{ombre-1}}$ défini par :

$$\left\| \overrightarrow{t_{ombre-1}} \right\| = \frac{h_{bâti}}{\cos(\theta_i) P_d} \quad (\text{en pixels}) \quad (6.3)$$

Toutefois, pour les points situés aux deux extrémités (basse et haute) de l'emprise au sol, l'approximation d'objet totalement occultant peut être considérée comme non valide. On peut alors envisager pour eux qu'ils répondent au cas d'objet non occultant (objet très fin suivant l'axe des distances RADAR tel qu'un pylône). La largeur de l'ombre induite est alors indépendante des dimensions de l'objet ; elle est fonction de la hauteur du bâti $h_{bâti}$ et commence à partir des sommets de l'emprise au sol les plus éloignés du capteur RADAR. L'extension de l'ombre correspond à la norme du vecteur translation $\overrightarrow{t_{ombre-bord}}$ défini par :

$$\left\| \overrightarrow{t_{ombre-bord}} \right\| = \frac{h_{bâti} \sin(\theta_i) \tan(\theta_i)}{P_d} \quad (\text{en pixels}) \quad (6.4)$$

Génération de la "zone d'ombre" pour la configuration 2 Dans la configuration 2, le bâti est considéré comme un objet semi-occultant (voir figure 6.2(b)). La largeur de l'ombre induite est alors fonction des dimensions de l'objet et de la hauteur du bâti $h_{bâti}$ et commence à partir des sommets de l'emprise au sol les plus proches du capteur RADAR. L'extension de la "zone d'ombre" correspond à la norme du vecteur translation $\overrightarrow{t_{ombre-2}}$ défini par :

$$\left\| \overrightarrow{t_{ombre-2}} \right\| = \frac{w_{bâti} \sin(\theta_i)}{P_d} + \frac{h_{bâti} \sin(\theta_i) \tan(\theta_i)}{P_d} \quad (\text{en pixels}) \quad (6.5)$$

où $w_{bâti}$ correspond à "l'épaisseur vue" du bâtiment en "géométrie terrain" suivant l'axe des distances RADAR.

Toutefois, pour les points situés aux deux extrémités (basse et haute) de l'emprise au sol, la translation appliquée reste identique à la translation de vecteur $\overrightarrow{t_{ombre-bord}}$ définie pour la configuration 1.

Génération de la "zone d'écho double" pour les deux configurations La "zone d'écho double" est générée en considérant les deux segments de l'emprise rectangulaire du bâti qui font face au capteur RADAR et en les dilatant suivant l'axe des distances RADAR d'une épaisseur $\epsilon_{echo-double}$ (voir figure 6.2(c) : on a la relation $\left\| \overrightarrow{t_{echo-double}} \right\| = \epsilon_{echo-double}$). La position et les dimensions de la "zone d'écho double" sont donc indépendantes de la hauteur du bâti $h_{bâti}$.

Génération de la "zone de fond" pour les deux configurations La "zone de fond" est la plus simple à définir et la dernière zone à être générée. On rappelle que les bords de l'imagette RADAR considérée sont obtenus en élargissant la boîte englobante de l'emprise au sol d'une épaisseur ϵ_{boite} . La "zone de fond" correspond alors à la région complémentaire située à l'intérieur de l'imagette RADAR et à l'extérieur du contour englobant l'ensemble des autres régions caractéristiques précédemment citées. La figure 6.2(d) présente deux exemples de "zones de fond" générées sur des données réelles.

Remarque sur la génération du partitionnement en régions caractéristiques Il est important de noter que les différentes régions considérées doivent être générées de telle sorte qu'elles forment un partitionnement de l'imagette RADAR considérée. L'intersection entre deux régions doit donc systématiquement être nulle et l'union de toutes les régions doit correspondre à la surface de l'imagette entière. Afin que ces points soient parfaitement vérifiés, il est nécessaire, lors de la génération du partitionnement, de soustraire aux régions "recouvrement", "façade", "toit seul" et "ombre" précédemment décrites les quelques pixels, appartenant à la "zone d'écho double", qui se trouveraient inclus à l'intérieur du contour de chacune d'entre elles.

6.1.3 Phase 1 : utilisation d'un critère RADAR global statistique pour la génération de hauteurs candidates

6.1.3.1 Principe

On note $RC(h_{bâti})$ l'ensemble des régions caractéristiques qui composent la signature RADAR d'un bâti et qui dépendent de sa hauteur $h_{bâti}$.

L'idée consiste à exploiter, pour chaque bâti considéré, la localisation et l'étendue des régions appartenant à $RC(h_{bâti})$ afin d'estimer sa hauteur.

Nous proposons d'utiliser comme critère statistique la log-vraisemblance des intensités de l'image RADAR pour les régions définies par le partitionnement induit par $RC(h_{bâti})$ ("zone de recouvrement", "zone de toit seul", "zone d'ombre", ...).

Plus précisément, étant données une imagette RADAR englobant un bâti potentiel et une hauteur de bâti supposée $h_{bâti}$, nous utilisons un critère de log-vraisemblance globale correspondant à la somme des log-vraisemblances individuelles, relatives aux différentes régions caractéristiques composant $RC(h_{bâti})$.

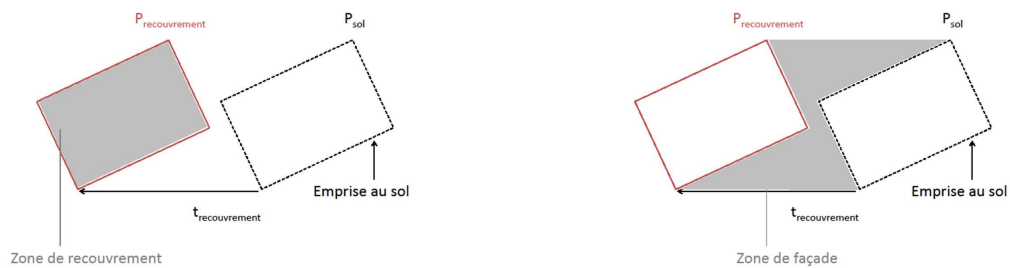
6.1.3.2 Modélisation de la distribution d'intensité dans chaque région caractéristique

On choisit de modéliser la distribution d'intensité de chaque région caractéristique i , supposée individuellement homogène, par une loi Gamma généralisée à deux paramètres, classiquement utilisée en imagerie RADAR.

On note $[\mu_i, L_i]$ le couple de paramètres de la loi Gamma G définie par [Nic06] :

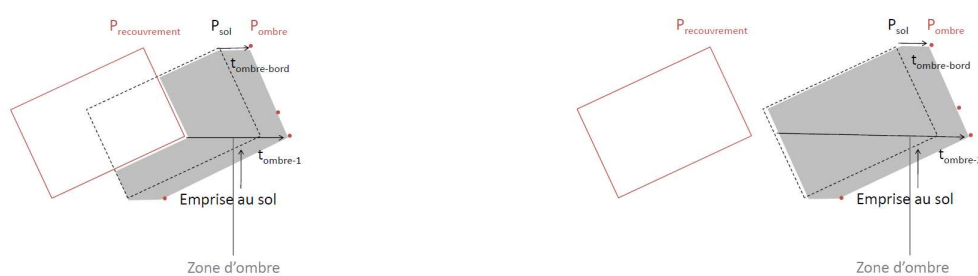


(a) Génération de la "zone de recouvrement" pour la configuration 1 (b) Génération de la "zone de toit seul" pour la configuration 1



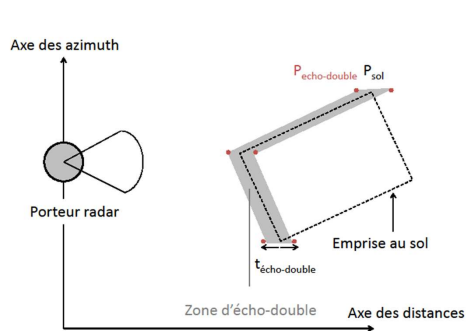
(c) Génération de la "zone de recouvrement" pour la configuration 2 (d) Génération de la "zone de façade" pour la configuration 2

FIGURE 6.1 – Génération des zones caractéristiques de la signature du bâtiment. [Partie 1/2]

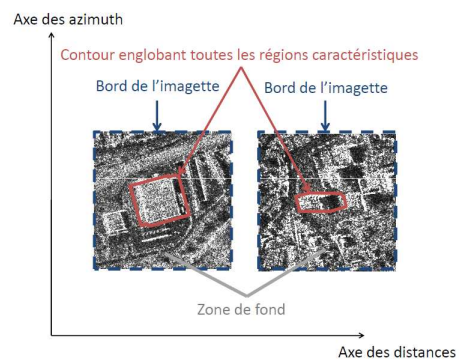


(a) Génération de la "zone d'ombre" pour la configuration 1

(b) Génération de la "zone d'ombre" pour la configuration 2



(c) Génération de la "zone d'écho double"



(d) Génération de la "zone de fond" ©Infoterra

FIGURE 6.2 – Génération des zones caractéristiques de la signature du bâtiment. [Partie 2/2]

$$G_{[\mu_i, L_i]}(x) = \frac{1}{\Gamma(L_i)} \frac{L_i}{\mu_i} \left(\frac{L_i x}{\mu_i} \right)^{L_i-1} \exp \frac{-L_i x}{\mu_i} \quad \forall x \in \mathbb{R}^+ \quad (6.6)$$

avec les notations suivantes :

- μ_i correspond à la valeur moyenne (moment d'ordre 1) ;
- L_i correspond à un facteur de forme ; en imagerie RADAR, il s'agit du "nombre de vues" ; celui-ci est normalement égal à 1 sur les images SLC ; toutefois, ce paramètre sera ré-estimé localement dans chacune des régions considérées (voir section 6.1.3.4), où il fournira des valeurs proches de 1 ;
- Γ désigne la fonction Gamma.

6.1.3.3 Expression analytique de la log-vraisemblance individuelle de chaque région caractéristique

L'expression analytique de la log-vraisemblance généralisée d'une région LVG_{region} , calculée sur N_i échantillons pour une région homogène i qui suit en intensité une loi Gamma $G_{[\mu_i, L_i]}$, est donnée par [Nic06] :

$$\begin{aligned} LVG_{region}[\mu_i, L_i] &= N_i(L_i \log(L_i) - \text{Log}\Gamma(L_i)) \\ &\quad - \frac{L_i}{\mu_i} \sum_{j \in i} x_j - N_i \log(\mu_i) \\ &\quad + (L_i - 1) \left(\sum_{j \in i} \log(x_j) - N_i \log(\mu_i) \right) \end{aligned} \quad (6.7)$$

avec les notations suivantes :

- N_i correspond au nombre de pixels dans la région i ;
- x_j correspond à l'intensité du pixel j ;
- $\text{Log}\Gamma$ désigne la fonction LogGamma, définie comme le logarithme de la fonction Gamma Γ .

Cette formule est valable individuellement pour chaque région caractéristique du bâti envisagée.

6.1.3.4 Estimation des paramètres locaux par maximum de vraisemblance

On propose d'estimer indépendamment pour chaque région caractéristique test i , les paramètres $[\mu_i, L_i]$ de la loi Gamma utilisée pour modéliser la distribution d'intensité des pixels appartenant à cette région.

On rappelle que les estimées $[\hat{\mu}_i, \hat{L}_i]$ issues de la méthode du maximum de vraisemblance vérifient les formules [Nic06] :

$$\hat{\mu}_i = \frac{\sum_{j \in i} x_j}{N_i} \quad (6.8)$$

et

$$\log(\hat{L}_i) - \Psi(\hat{L}_i) = \log(\hat{\mu}_i) - \frac{\sum_{j \in i} \log(x_j)}{N_i} \quad (6.9)$$

où Ψ désigne la fonction DiGamma, définie comme la dérivée logarithmique de la fonction Gamma Γ .

Une approximation numérique du paramètre \hat{L}_i peut être obtenue à partir de l'expression 6.9 en utilisant le développement au premier ordre suivant [Nic06] :

$$\frac{1}{\log(x) - \Psi(x)} \approx 1.98278 x - 0.2668 \quad \text{pour } 1 \leq x \leq 50 \quad (6.10)$$

On déduit alors l'expression explicite suivante [Nic06] :

$$\hat{L}_i \approx 0.1346 + \frac{0.504}{\log(\hat{\mu}_i) - \frac{\sum_{j \in i} \log(x_j)}{N_i}} \quad (6.11)$$

Cette approche d'estimation locale des paramètres nous dispense d'une phase d'apprentissage global, qui comporterait les deux désavantages suivants :

- d'une part, l'intervention de l'utilisateur serait requise pour sélectionner des zones d'apprentissage sur l'image complète (perte d'automatisme de l'algorithme) ;
- d'autre part, les zones sélectionnées devraient être suffisamment représentatives en termes de distribution d'intensité pour permettre une estimation de paramètres généralisable à l'ensemble des zones traitées. Or, en pratique, ceci est vérifié sur les "zones d'ombre" (pour un capteur donné) mais reste nettement moins valable sur les "zones de toit" et sur "les zones de recouvrement" typiquement, où la radiométrie dépend fortement du matériau imagé et de l'incidence du capteur.

6.1.3.5 Formulation analytique du critère statistique global

On rappelle qu'étant donnée une hauteur test de bâtiment h_{test} , on peut générer le partitionnement test en régions caractéristiques $RC(h_{test})$ d'une imagerie RADAR englobant un bâti.

On définit alors notre critère statistique global $C_{statistique}(h_{test})$ comme l'opposé de la log-vraisemblance généralisée globale calculée pour l'imagerie considérée.

Ce critère statistique est calculé pour un nombre total de pixels N_{total} égal à la somme des N_i relatifs à chaque région i . Il est important de noter que N_{total} est constant (c'est-à-dire indépendant de h_{test}). Il correspond au nombre total de pixels contenus dans l'imagerie RADAR considérée (puisque l'intersection entre deux régions caractéristiques est systématiquement nulle et que l'union de toutes les régions caractéristiques correspond à la surface de l'imagerie entière).

Sous l'hypothèse d'indépendance des pixels de l'imagerie, ce critère $C_{statistique}(h_{test})$ vérifie la formule analytique suivante :

$$\begin{aligned} C_{statistique}(h_{test}) &= - LVG_{globale}(RC(h_{test})) \\ &= - \sum_{i \in RC(h_{test})} LVG_{region}[\mu_i, L_i] \end{aligned} \quad (6.12)$$

Ce critère $C_{statistique}(h_{test})$ correspond à un coût énergétique à minimiser en fonction de la hauteur testée h_{test} .

Lorsque ce critère est minimal, il existe une adéquation optimale entre le partitionnement généré $RC(h_{test})$ et le partitionnement réel de l'imagerie RADAR traitée.

Il convient de noter que ce critère se fonde seulement sur l'hypothèse d'homogénéité des régions caractéristiques mais n'intègre pas d'information a priori quant à la radiométrie relative moyenne des différentes régions.

C'est ce second point qui justifiera l'introduction d'un critère de contraste dans la seconde phase du procédé (voir section 6.1.4).

6.1.3.6 Optimisation du critère statistique pour la génération de hauteurs candidates

On propose d'utiliser, pour chaque bâtiment, un schéma algorithmique permettant de générer un ensemble de hauteurs candidates, susceptibles d'approcher la hauteur réelle du bâti considéré (voir figure 6.3).

Algorithme Pour chaque emprise potentielle projetée et recalée dans l'image RADAR, on procède de la façon suivante :

1. **Définition de l'imagette RADAR** L'imagette RADAR, sur laquelle est calculé le critère $C_{statistique}(h_{test})$, est définie en élargissant la boîte englobante de l'emprise considérée. Une couronne rectangulaire de taille fixe, choisie suffisamment grande (voir section 6.1.7), est ajoutée autour de la boîte englobante considérée. L'imagette RADAR rectangulaire est alors définie comme l'union de la boîte englobante considérée et de la couronne associée.
2. **Boucle itérative de type "Proposition d'une hauteur test - Génération d'un partitionnement test - Évaluation du critère statistique"** Le schéma suivant est alors mis en œuvre :
 - sélection d'une hauteur test h_{test} dans l'intervalle $[h_{min}, h_{max}]$ des hauteurs de bâtis possibles ;
 - génération du partitionnement test en régions caractéristiques $RC(h_{test})$ sur l'imagette considérée, en utilisant la technique de génération de partitionnement décrite en section 6.1.2 ;
 - évaluation du critère $C_{statistique}(h_{test})$ sur le partitionnement test, en utilisant les équations 6.8 et 6.11 pour l'estimation des paramètres puis les équations 6.7 et 6.12 pour le calcul du critère.

3. **Génération de hauteurs candidates par optimisation locale du critère statistique** La courbe représentant l'évolution du critère statistique $C_{statistique}(h_{test})$ en fonction de la hauteur testée h_{test} est tracée.

Lorsqu'un minimum apparaît, une forte adéquation existe entre le partitionnement prédit et le partitionnement réel de l'imagette.

Ce minimum peut être local ou global en fonction de la complexité de la scène.

Dans le cas simple d'une imagette pour laquelle toutes les zones caractéristiques sont imagées (environnement dégagé autour du bâtiment d'intérêt) et nettement différenciables les unes des autres (matériaux de construction répondant de façon distincte), et, pour laquelle la modélisation de la distribution en intensité de chaque zone peut être correctement approchée par une loi Gamma (zones homogènes sans cibles perturbatrices), la courbe d'évolution du critère statistique présente un minimum global.

Au contraire, dans le cas plus complexe d'une imagette dont, par exemple, la "zone de recouvrement" serait faiblement marquée par une texture homogène et la "zone

d'ombre" serait fortement perturbée par la présence de cibles ponctuelles brillantes, l'hypothèse de régions caractéristiques homogènes suivant une modélisation d'intensité en loi Gamma n'est que partiellement vérifiée et le critère de log-vraisemblance généralisée globale n'est que partiellement adapté. En conséquence, suivant le "degré de perturbation" des zones supposées homogènes, la courbe d'évolution du critère statistique peut :

- soit présenter un ensemble de minima locaux, susceptibles de correspondre à des partitionnements qui coïncident partiellement avec la signature réelle du bâti ; il nous faut alors pour l'instant tous les conserver, car il est difficile de déterminer, sans information complémentaire, lequel est le plus pertinent ;
- soit ne présenter aucun minimum local, lorsqu'il n'existe aucune valeur de hauteur test h_{test} pour laquelle le partitionnement induit coïncide avec la signature réelle du bâti, sous les hypothèses émises (homogénéité et modélisation en loi Gamma). Ceci témoigne d'une signature du bâti imagé fortement dégradée (cibles perturbatrices entraînant des régions hétérogènes, matériaux peu répondants, régions occultées par l'environnement, ...) ou d'une signature absente (bâtiment non existant dans l'imagette considérée, c'est-à-dire relevant d'une fausse alarme lors de l'étape d'extraction en monoscopie optique).

Les optima locaux sont ainsi sélectionnés afin de fournir, pour la seconde phase du procédé, des hauteurs, appelées hauteurs candidates $h_{candidate}$ et définies par l'équation suivante :

$$h_{candidate} = \arg \min_{h_{test}} C_{statistique}(h_{test}) \quad (6.13)$$

4. **Sélection des hauteurs candidates représentatives** Afin d'améliorer la sélection des hauteurs candidates $h_{candidate}$, deux contraintes de formes, notées $C_{largeur-pic}(h_{candidate})$ et $C_{hauteur-pic}(h_{candidate})$ et calculées sur la courbe d'évolution du critère statistique, sont introduites.

Celles-ci se réfèrent à la représentativité des optima locaux et sont définies de la façon suivante :

- la contrainte $C_{largeur-pic}(h_{candidate})$ est satisfaite si la largeur $L_{pic}(h_{candidate})$ du pic optimal associé sur la courbe d'évolution est supérieure à un seuil minimal noté $S_{largeur-pic}^{min}$;
- la contrainte $C_{hauteur-pic}(h_{candidate})$ est satisfaite si la hauteur relative $H_{pic}(h_{candidate})$ du pic optimal associé sur la courbe d'évolution est supérieure à un seuil minimal noté $S_{hauteur-pic}^{min}$.

Ces contraintes permettent de filtrer les hauteurs candidates induisant des pics de minima (creux) jugés "représentatifs" et de supprimer toutes celles ne correspondant pas à des pics suffisamment "marqués" sur la courbe.

6.1.3.7 Comparaison avec l'approche développée par Brunner et al.

Remarquons que cette approche est similaire à celle proposée par Brunner et al. dans [BLB08, BLBG10].

Cette dernière propose d'estimer la hauteur des bâtiments en monoscopie RADAR à partir d'une stratégie basée sur un schéma nommé "génération d'hypothèses - restitution

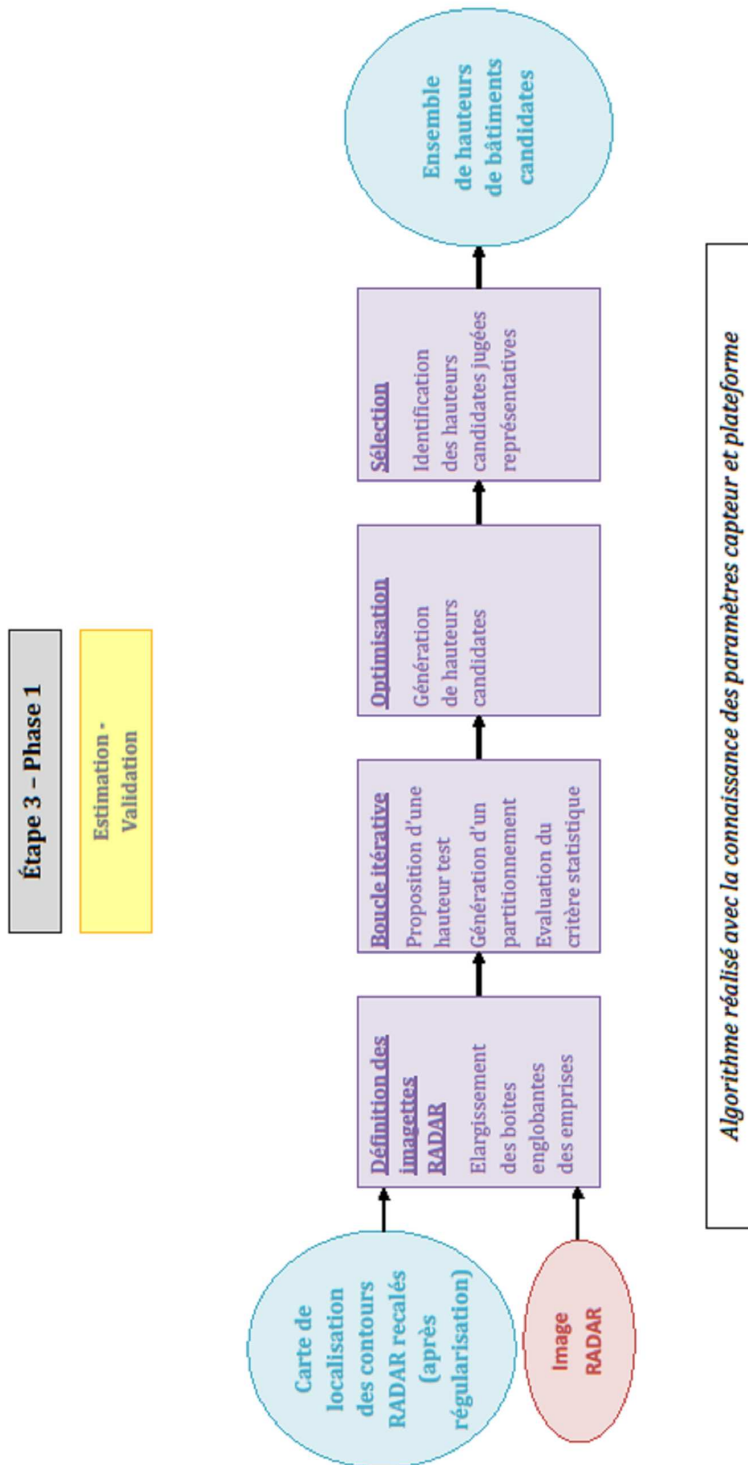


FIGURE 6.3 – Schéma synoptique de la phase 1 de l'étape 3.

- appariement" (sous l'hypothèse d'avoir préalablement estimé les paramètres planimétriques définissant l'emprise au sol du bâti).

Toutefois, les différences suivantes doivent être soulignées :

- au niveau méthodologique : dans l'approche de Brunner et al. , un critère d'appariement (de type corrélation ou information mutuelle) est calculé pour tester la similarité entre une imagerie de bâtiment simulée (dépendant de la hauteur test) et l'imagerie RADAR réelle. Au contraire, dans notre approche, aucune imagerie RADAR n'est simulée. Seul un partitionnement géométrique (dépendant de la hauteur test) l'est et le critère statistique proposé est calculé directement à partir de l'intensité des pixels présents sur l'imagerie RADAR réelle ;
- au niveau statistique : dans l'approche de Brunner et al., le critère d'appariement utilisé pour comparer les deux signatures ne tient pas compte des propriétés statistiques propres à l'imagerie RADAR (bruit de speckle notamment) puisque le critère utilisé compare une imagerie RADAR réelle filtrée et une imagerie de bâti, cette dernière ayant été simulée en tenant compte uniquement des radiométries relatives des différentes régions et des propriétés générales de réflexions (mélange de type spéculaire et lambertien) applicables à la restitution d'une image optique à partir de la technique de "lancer de rayons". Au contraire, dans notre approche, le critère d'adéquation utilisé pour comparer les deux partitionnements tient compte de la statistique RADAR (modélisation des distributions d'intensité par des lois Gamma dans les différentes régions) et est donc calculé sur l'imagerie RADAR réelle non filtrée.

6.1.4 Phase 2 : utilisation d'un critère RADAR de contraste local pour l'estimation des hauteurs finales et pour la validation des bâtiments

6.1.4.1 Principe

Afin de choisir pour chaque bâti la hauteur la plus appropriée parmi les hauteurs candidates issues de la première phase, nous proposons de combiner le critère statistique région avec un critère de contraste local.

Ce dernier intègre une information a priori de type radiométrique qui fait défaut dans le critère statistique précédemment utilisé.

L'idée consiste à définir un critère de contraste local, dépendant de la hauteur h_{test} et testant la présence d'une discontinuité entre deux zones caractéristiques dont le classement relatif des radiométries moyennes est connu a priori.

En l'occurrence, le critère de contraste proposé teste la présence d'une discontinuité entre la "zone de fond" et la "zone de recouvrement" (pour les deux configurations envisagées), dont la localisation peut être déduite en fonction de la hauteur de bâti testée.

6.1.4.2 Formulation du critère de contraste local

On utilise comme critère de contraste local $C_{contraste}(h_{test})$ le contraste local calculé sur l'imagerie considérée entre les deux régions "zone de fond" et "zone de recouvrement".

$C_{contraste}(h_{test})$ est ainsi défini comme la racine du rapport moyen en intensité des pixels appartenant à deux bandes étroites situées de part et d'autre du bord et vérifie la formule analytique suivante :

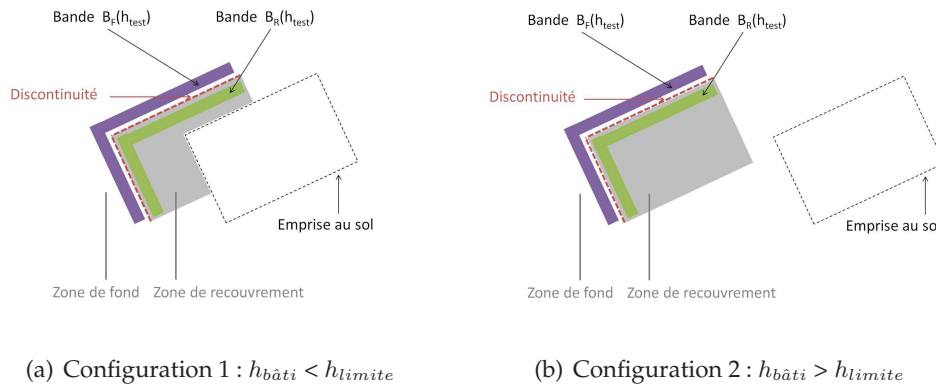


FIGURE 6.4 – Positionnement du contraste au niveau de la discontinuité entre la "zone de fond" et la "zone de recouvrement".

$$C_{\text{contraste}}(h_{\text{test}}) = \sqrt{\frac{1/N_{B_F}(h_{\text{test}}) \sum_{j \in B_F}(h_{\text{test}}) x_j}{1/N_{B_R}(h_{\text{test}}) \sum_{j \in B_R}(h_{\text{test}}) x_j}} \quad (6.14)$$

avec les notations suivantes (voir figure 6.4) :

- $B_F(h_{\text{test}})$ désigne la bande de largeur $\epsilon_{\text{discontinuite}}$ située du côté de la discontinuité donnant sur la "zone de fond" et dont la position dépend de la hauteur testée h_{test} ;
- $B_R(h_{\text{test}})$ désigne la bande de largeur $\epsilon_{\text{discontinuite}}$ située du côté de la discontinuité donnant sur la "zone de recouvrement" et dont la position dépend de la hauteur testée h_{test} ;
- N_{B_F} correspond au nombre de pixels dans la bande $B_F(h_{\text{test}})$;
- N_{B_R} correspond au nombre de pixels dans la bande $B_R(h_{\text{test}})$;
- x_j correspond à l'intensité du pixel j dans l'imagette RADAR réelle.

Étant donné que la "zone de fond" est supposée être plus "foncée" que la "zone de recouvrement", $C_{\text{contraste}}(h_{\text{test}})$ doit être inférieur à 1.

Lorsque le partitionnement de l'imagette "colle" parfaitement bien avec la signature RADAR réelle, $C_{\text{contraste}}(h_{\text{test}})$ est minimal : il s'agit donc d'un coût de contraste à minimiser en fonction de la hauteur testée h_{test} .

Remarque sur le choix des régions sélectionnées pour le calcul du critère de contraste

Nous avons considéré la discontinuité située entre la "zone de fond" et la "zone de recouvrement" pour calculer le critère de contraste local.

Notons cependant que nous aurions pu également nous intéresser à la discontinuité située entre la "zone d'ombre" et la "zone de fond" pour calculer un autre critère de contraste local. Nous aurions alors introduit un second a priori quant à la radiométrie moyenne relative des deux zones (la "zone d'ombre" étant supposée plus foncée que la "zone de fond").

Toutefois, d'après les tests réalisés, ce second type de contraste s'est avéré moins bien visible sur nos images TerraSAR-X. Ceci s'explique d'une part par le fait qu'en imagerie satellitaire, le contraste "ombre"- "fond" est généralement moins marqué qu'en imagerie aéroportée, et, d'autre part, par le fait qu'en milieu urbain dense, les zones d'ombres peuvent être souvent partiellement ou totalement occultées par la réponse de structures

urbaines voisines élevées. L'exploitation de ce second critère de contraste s'est donc révélée plus délicate et nous avons choisi de nous restreindre à l'utilisation du premier critère de contraste ("fond"- "recouvrement") pour le traitement de nos données.

6.1.4.3 Utilisation du critère de contraste pour l'estimation des hauteurs finales et pour la validation des bâtiments

On propose d'utiliser, pour chaque bâtiment, un schéma algorithmique permettant de sélectionner les hauteurs candidates pertinentes, d'estimer les hauteurs finales des bâtis par combinaison des deux critères et de valider leur présence (voir figure 6.5).

Algorithme Pour chaque emprise potentielle projetée et recalée dans l'image RADAR, on procède de la façon suivante :

1. **Boucle itérative de type "Proposition d'une hauteur test - Positionnement de la discontinuité test - Évaluation du critère de contraste"** Le schéma suivant est alors mis en œuvre :
 - sélection d'une hauteur test h_{test} dans l'intervalle $[h_{min}, h_{max}]$ des hauteurs de bâtis possibles ;
 - positionnement de la discontinuité test entre la "zone de fond" et la "zone de recouvrement" en fonction de la hauteur h_{test} ;
 - définition des bandes tests $B_F(h_{test})$ et $B_R(h_{test})$ de part et d'autre de la discontinuité test ;
 - évaluation du critère $C_{contraste}(h_{test})$ sur ces bandes, à partir des intensités des pixels de l'imagette RADAR et en utilisant l'équation 6.14 pour le calcul du critère.
2. **Sélection des hauteurs candidates pertinentes par analyse conjointe des deux critères** La courbe représentant l'évolution du critère de contraste $C_{contraste}(h_{test})$ en fonction de la hauteur testée h_{test} est tracée.

Pour chaque hauteur candidate $h_{candidate}$, on examine localement l'évolution du critère $C_{contraste}(h_{test})$ dans un proche voisinage $[h_{candidate} + \delta h, h_{candidate} - \delta h]$ autour de celle-ci et on retient la valeur minimale du critère $C_{contraste}(h_{test})$ dans ce voisinage.

Ce minimum local, noté $C_{contraste}(\tilde{h}_{candidate})$, peut donc être obtenu pour une valeur de hauteur notée $\tilde{h}_{candidate}$ légèrement différente de $h_{candidate}$.

On sélectionne enfin la hauteur candidate $h_{candidate}$ et la hauteur affinée associée $\tilde{h}_{candidate}$, qui donnent lieu au plus petit minimum local $C_{contraste}(\tilde{h}_{candidate})$ (c'est-à-dire qui correspondent au contraste le plus fort).
3. **Combinaison des deux critères pour l'estimation des hauteurs finales des bâtiments potentiels** La hauteur finale du bâti potentiel considéré $h_{estimee}$ est définie comme la moyenne entre la hauteur candidate sélectionnée $h_{candidate}$ (issue du critère statistique) et la hauteur affinée associée $\tilde{h}_{candidate}$ (issue du critère de contraste).
4. **Validation de la présence des bâtiments par seuillage sur le critère de contraste** Enfin, un seuillage sur la valeur de $C_{contraste}(\tilde{h}_{candidate})$ permet de valider ou non la présence du bâtiment.

À la fin du procédé d'estimation des hauteurs et de validation, on obtient des "cartes des hauteurs estimées des bâtiments validés", qui sont représentées en niveaux de gris (dans le référentiel image RADAR) et qui résultent de la prise en compte des informations altimétriques (hauteur du bâti estimée) et planimétriques (présence du bâti validée) obtenues.

6.1.5 Justifications sur la complémentarité des deux critères

Il convient de remarquer que les principaux atouts et les principales faiblesses de chacun des deux critères sont différents :

- le critère statistique région, qui est un critère global calculé sur toute l'imagette RADAR, présente un avantage en terme de robustesse ; nous verrons, qu'en pratique, il permet de fournir un ensemble de hauteurs candidates fiables mais dont la localisation est légèrement imprécise ;
- le critère de contraste local, quant à lui, reste assez peu robuste car il est calculé de façon locale sur une zone de discontinuité ; il ne pourrait donc suffire à lui seul pour estimer la hauteur finale des bâtis. Toutefois, à partir d'un jeu de hauteurs candidates, son utilisation permet de discriminer les hauteurs les plus fiables et de les ré-estimer avec précision.

L'utilisation conjointe de ces deux critères apparaît donc adaptée puisqu'il s'agit de critères complémentaires, qui, comme nous le verrons en section 6.4, permettent d'atteindre des performances satisfaisantes en termes de robustesse et de précision.

6.1.6 Commentaires sur la propagation des erreurs

Il convient de souligner que les erreurs de dimensionnement planimétrique issues de la première étape (extraction des emprises) se propagent directement dans les étapes suivantes (projection - recalage puis estimation des hauteurs - validation).

L'exactitude des résultats d'une étape est donc directement conditionnée par la précision des résultats fournis au cours des étapes précédentes.

Ce point est particulièrement important pour l'étape d'estimation des hauteurs décrite. En effet, la méthode proposée est fondée sur l'utilisation d'un partitionnement généré à partir des emprises de bâtiments dont la localisation est supposée parfaitement connue. En conséquence, la qualité en termes de dimensionnement (issue de la première étape) et de positionnement (issue de la première étape et de la seconde étape) des résultats précédents a un impact direct sur la capacité de la troisième étape à fournir des hauteurs estimées pertinentes.

6.1.7 Discussion sur la gestion des paramètres

Paramètres de la phase 1 Les différents paramètres utilisés dans la phase de génération des hauteurs candidates sont les suivants :

- la taille $\epsilon_{echo-double}$ utilisée pour définir la largeur de la "zone d'écho double" (lors de la génération de partitionnements tests) et fixée empiriquement à quelques pixels (2 à 3 pixels) ;
 - la taille ϵ_{boite} utilisée pour définir les limites de l'imagette (boîte englobante élargie d'une couronne rectangulaire) autour de l'emprise de chaque bâti ; celle-ci doit être prise suffisamment grande pour assurer que, quelle que soit la valeur de hauteur
-

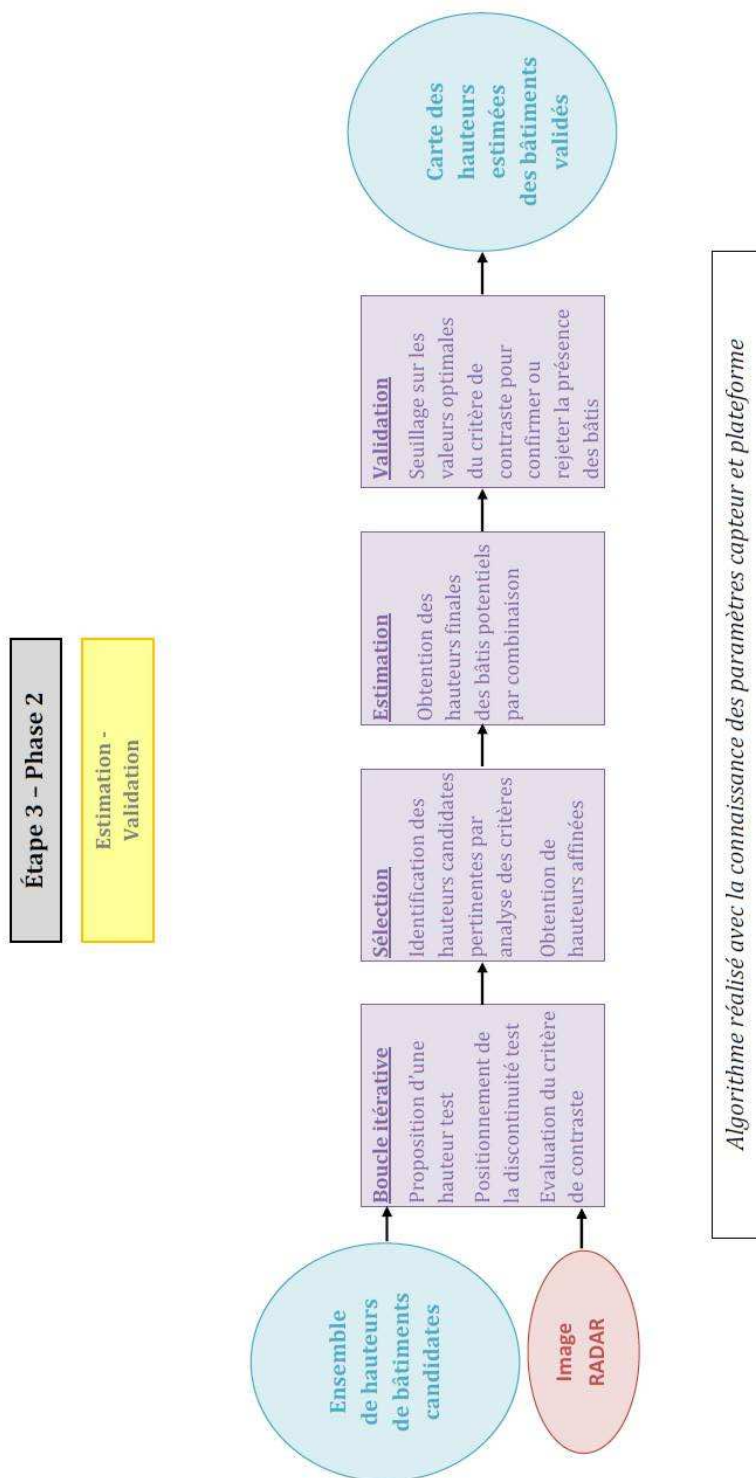


FIGURE 6.5 – Schéma synoptique de la phase 2 de l'étape 3.

- test h_{test} considérée, l'ensemble des régions caractéristiques générées soient toutes contenues dans cette imagerie ;
- les hauteurs h_{min} et h_{max} utilisées pour borner l'intervalle $[h_{min}, h_{max}]$ des hauteurs de bâtis possibles ; celles-ci sont fixées en fonction de la configuration géométrique adoptée ;
 - le pas de quantification h_{pas} utilisé pour parcourir l'intervalle $[h_{min}, h_{max}]$; celui-ci dépend du compromis choisi entre la précision souhaitée pour la hauteur estimée et le temps de calcul ;
 - les seuils $S_{largeur}^{min}$ et $S_{hauteur}^{min}$ utilisés pour valider les contraintes de représentativité des optima locaux sur la courbe d'évolution du critère statistique ; ces seuils sont empiriquement fixés à quelques mètres et à quelques pourcents respectivement.

Paramètres de la phase 2 Les différents paramètres utilisés dans la phase d'estimation des hauteurs finales et de validation des bâtiments sont les suivants :

- la taille $\epsilon_{discontinuite}$ utilisée pour définir la largeur des bandes tests $B_F(h_{test})$ et $B_R(h_{test})$ de part et d'autre de la discontinuité test lors du calcul du critère de contraste ; celle-ci est choisie égale à 3 pixels ;
- la taille δh utilisée pour définir un proche voisinage autour des hauteurs candidates lors de la sélection des hauteurs candidates pertinentes ; celle-ci est choisie égale à 2 m ;
- le seuil $S_{contraste}$ appliqué sur la valeur du plus petit minimum local $C_{contraste}(\tilde{h}_{candidate})$ afin de valider la présence d'un bâti potentiel ; celui-ci, relatif à un contraste a priori inférieur à 1, est empiriquement fixé à 0.8.

6.1.8 Validation de la méthode précédente sur des données RADAR simulées

Afin de valider la méthodologie développée dans l'approche proposée, celle-ci est d'abord testée sur des imageries RADAR de bâtiments simulés.

Les résultats de simulation des signatures RADAR sont présentés, pour les deux configurations étudiées, sur les figures 6.6(a) et 6.6(b).

La table 6.1 indiquent les valeurs des paramètres utilisés pour générer ces données simulées. Le bâtiment S1 a été simulé avec une hauteur $h_{bâti}$ égale à 10 m et inférieure à la hauteur limite h_{limite} calculée (30 m) ; il vérifie donc la configuration géométrique de type 1. Le bâtiment S2 a été simulé avec une hauteur $h_{bâti}$ égale à 27 m et supérieure à la hauteur limite h_{limite} calculée (16 m) ; il vérifie donc la configuration géométrique de type 2.

Des lois Gamma de paramètres $[\mu_i, L_i]$ différents sont utilisées pour simuler l'intensité des pixels à l'intérieur de chaque région caractéristique i , une fois défini le partitionnement de l'imagerie. Afin de faciliter l'interprétation visuelle de la scène simulée, les valeurs données au paramètre μ_i (moyenne) dans les différentes régions caractéristiques satisfont le classement a priori suivant, relatif à la radiométrie moyenne des différentes régions :

$$\mu_{ombre} \leq \mu_{fond} \leq \mu_{toit-seul} \leq \mu_{recouvrement} \leq \mu_{echo-double} \quad (\text{configuration 1}) \quad (6.15)$$

ou

$$\mu_{ombre} \leq \mu_{fond} \leq \mu_{facade} \leq \mu_{recouvrement} \leq \mu_{echo-double} \quad (\text{configuration 2}) \quad (6.16)$$

Les courbes d'évolution du critère statistique $C_{statistique}(h_{test})$ et du critère de contraste $C_{contraste}(h_{test})$ en fonction de la hauteur test h_{test} sont présentées sur les figures 6.6(c), 6.6(d), 6.6(e) et 6.6(f).

Les critères proposés se révèlent être pertinents et efficaces. En effet, pour les deux cas étudiés, un minimum global apparaît clairement sur la courbe d'évolution du critère de log-vraisemblance et sur celle du critère de contraste. Les hauteurs tests optimales permettent de retrouver précisément les hauteurs réelles. Celles-ci sont égales à 10 m pour $h_{candidate}$ et $\tilde{h}_{candidate}$ dans le cas du bâtiment S1 et à 27 m pour $h_{candidate}$ et $\tilde{h}_{candidate}$ dans le cas du bâtiment S2. Les valeurs des plus petits minima locaux $C_{contraste}(\tilde{h}_{candidate})$ sont respectivement égales à 0.48 (voir figure 6.6(e)) et 0.51 (voir figure 6.6(f)) et permettent de valider la présence des bâtis.

Il est à noter que le seul point qui fasse l'objet d'une validation lors de la démarche adoptée est d'ordre méthodologique. Il s'agit de l'équivalence entre minimisation des critères présentés et obtention d'un partitionnement optimal (c'est-à-dire qui "colle" le mieux aux différentes régions de la donnée image) correspondant à la hauteur de bâti recherchée. Cette équivalence a bien été vérifiée en pratique sur nos données simulées.

Toutefois, les aspects suivants n'ont pour l'instant fait l'objet d'aucun test :

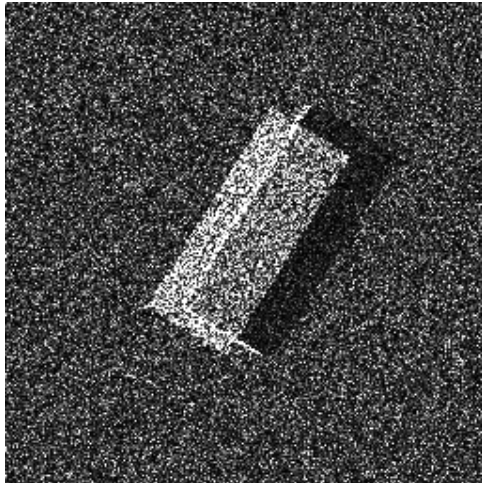
- la technique de génération du partitionnement en régions caractéristiques (en fonction de la hauteur test du bâtiment),
- l'hypothèse relative à l'homogénéité des différentes régions caractéristiques,
- le choix de loi Gamma pour la modélisation des distributions en intensité,
- l'hypothèse de visibilité de l'ensemble des régions caractéristiques dans la signature du bâti (on suppose que l'environnement autour de chaque bâti est suffisamment "dégagé" pour que l'ensemble de ses régions caractéristiques soient correctement imagées).

Ce sont ces points qui seront susceptibles d'être validés ou remis en question lors de la confrontation de l'approche développée avec les données RADAR réelles.

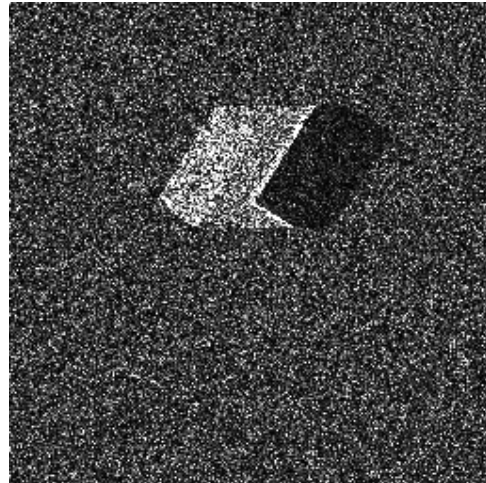
6.2 Reconstruction finale des bâtiments sur la scène

Étant donné le modèle de parallélépipède 3D choisi, la reconstruction finale des bâtiments de la scène se fait de façon simple : il suffit d'introduire les informations (de type planimétrique et altimétrique) relatives à la localisation des emprises des bâtis et à la hauteur de chacun, qui sont fournies par les "cartes des hauteurs estimées des bâtiments validés", au sein d'un visualisateur de scène 3D, puis d'améliorer le rendu visuel en intégrant la composante de texture (de type radiométrique) de l'image utilisée comme référentiel (ici celle de l'image RADAR).

Remarquons qu'il serait tout à fait possible de réaliser cette reconstruction tridimensionnelle dans le référentiel image optique ou dans un référentiel terrestre, en projetant les "cartes des hauteurs estimées des bâtiments validés" dans le référentiel souhaité et en intégrant la composante de texture souhaitée (c'est-à-dire respectivement la texture de l'image optique dans le premier cas ou la texture d'une des deux images projetée en géométrie terrain dans le second cas).



(a) Bâtiment S1 : image de simulation



(b) Bâtiment S2 : image de simulation

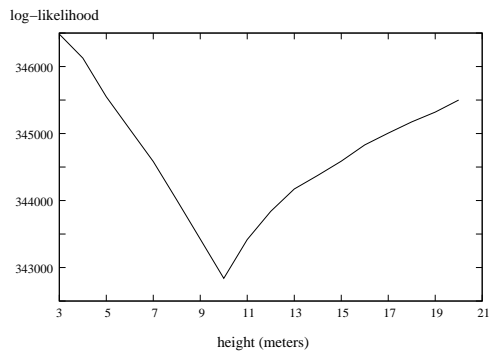
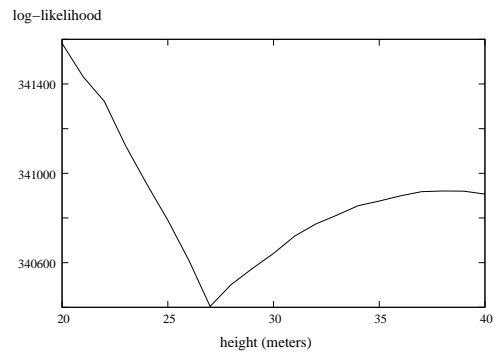
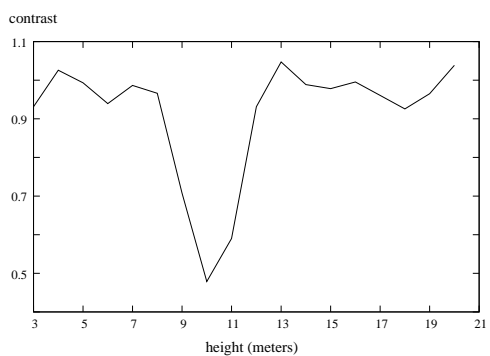
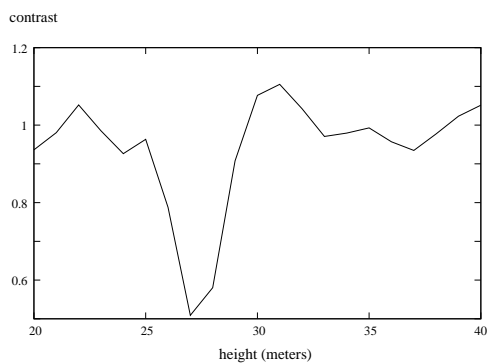
(c) Bâtiment S1 : $C_{statistique}(h_{test})$ (d) Bâtiment S2 : $C_{statistique}(h_{test})$ (e) Bâtiment S1 : $C_{contraste}(h_{test})$ (f) Bâtiment S2 : $C_{contraste}(h_{test})$

FIGURE 6.6 – Résultats obtenus pour les deux bâtiments testés sur données RADAR simulées. (a) et (b) Images issues des simulations RADAR de bâtiments. (c) et (d) Évolution du critère statistique $C_{statistique}(h_{test})$ en fonction de la hauteur. (e) et (f) Évolution du critère de contraste $C_{contraste}(h_{test})$ en fonction de la hauteur.

TABLE 6.1 – Paramètres utilisés pour la simulation de bâtiments RADAR.

	Bâtiment S1	Bâtiment S2
hauteur du bâti $h_{bâti}$	10 m	27 m
hauteur limite calculée h_{limite}	30 m	16 m
configuration géométrique	type 1	type 2
largeur de l'emprise $l_{bâti}$	50 pixels	25 pixels
longueur de l'emprise $L_{bâti}$	125 pixels	60 pixels
orientation de l'emprise ϕ^{RADAR} dans le référentiel image RADAR $(O, \vec{x}^{RADAR}, \vec{y}^{RADAR})$	58°	58°
position du centre $(x_c^{RADAR}, y_c^{RADAR})$ de l'emprise dans le référentiel image RADAR $(O, \vec{x}^{RADAR}, \vec{y}^{RADAR})$	(160, 55)	(160, 55)
incidence du capteur RADAR θ_i	31°	31°
pas d'échantillonnage en distance oblique P_d	0.45 m	0.45 m
couple de paramètres $(\mu_i, L_i)_{fond}$	(64, 1)	(64, 1)
couple de paramètres $(\mu_i, L_i)_{recouvrement}$	(252, 1)	(252, 1)
couple de paramètres $(\mu_i, L_i)_{ombre}$	(30, 1)	(30, 1)
couple de paramètres $(\mu_i, L_i)_{echo-double}$	(500, 1)	(500, 1)
couple de paramètres $(\mu_i, L_i)_{toit-seul}$	(140, 1)	-
couple de paramètres $(\mu_i, L_i)_{facade}$	-	(140, 1)

6.3 Application sur les scènes d'études

6.3.1 Paramètres utilisés

Les paramètres utilisés pour la première phase et pour la seconde phase de l'étape d'estimation des hauteurs et de validation des bâtiments à partir de l'image RADAR sont indiqués en table 6.2.

Les valeurs sont quasiment identiques pour les deux scènes, exceptées celles prises pour h_{min} et h_{max} , celles-ci étant déduites du type de configuration géométrique correspondant aux bâtiments de la scène. En l'occurrence ici, pour les bâtiments bas situés en zone industrielle sur la scène A, la gamme des hauteurs testées varie entre 3 et 15 m alors que pour les bâtiments hauts situés en zone résidentielle sur la scène B, celle-ci varie entre 5 et 30 m.

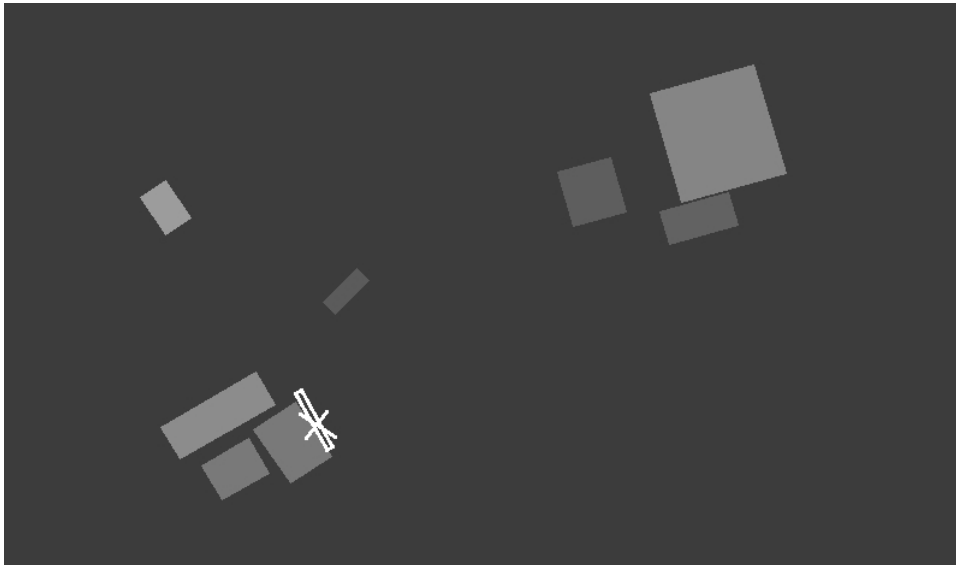
6.3.2 Résultats obtenus

Les figures 6.7(a) et 6.7(b) présentent les "cartes des hauteurs estimées des bâtiments validés", obtenues à la fin de la troisième étape.

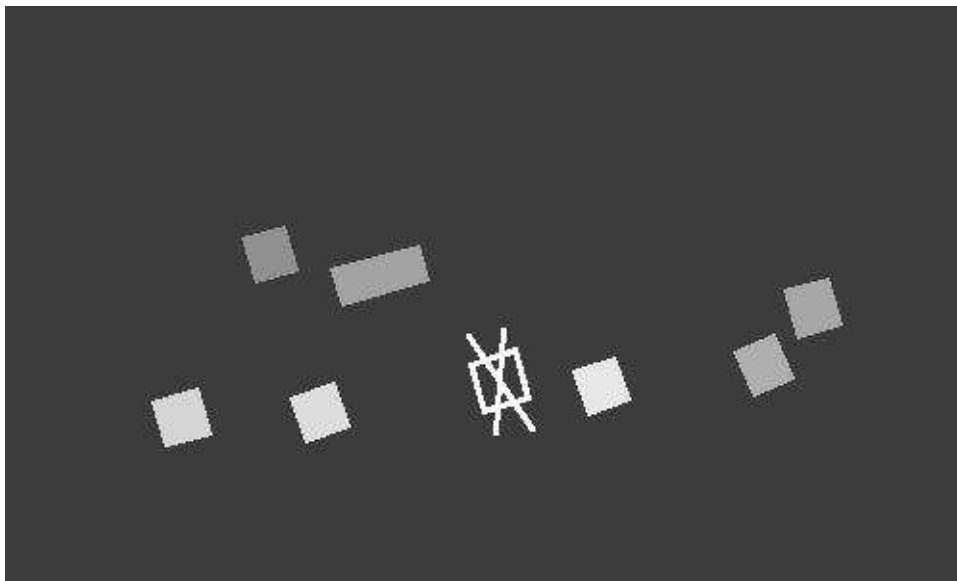
Les vues 3D texturées, représentant la reconstruction finale des bâtiments sur les deux scènes d'étude, sont proposées en figure 6.8(a) et en figure 6.8(b).

6.3.3 Analyse qualitative et quantitative

Les tables 6.3 et 6.4 détaillent, pour chaque bâtiment des deux scènes, les hauteurs ($h_{candidate}$, $\tilde{h}_{candidate}$ et $h_{estimee}$) obtenues à l'issue des différentes phases de l'étape considérée. La quatrième colonne indique le statut (validé ou rejeté) défini en fin de procédé.



(a) Scène A : hauteurs estimées et emprises validées en fin d'étape 3

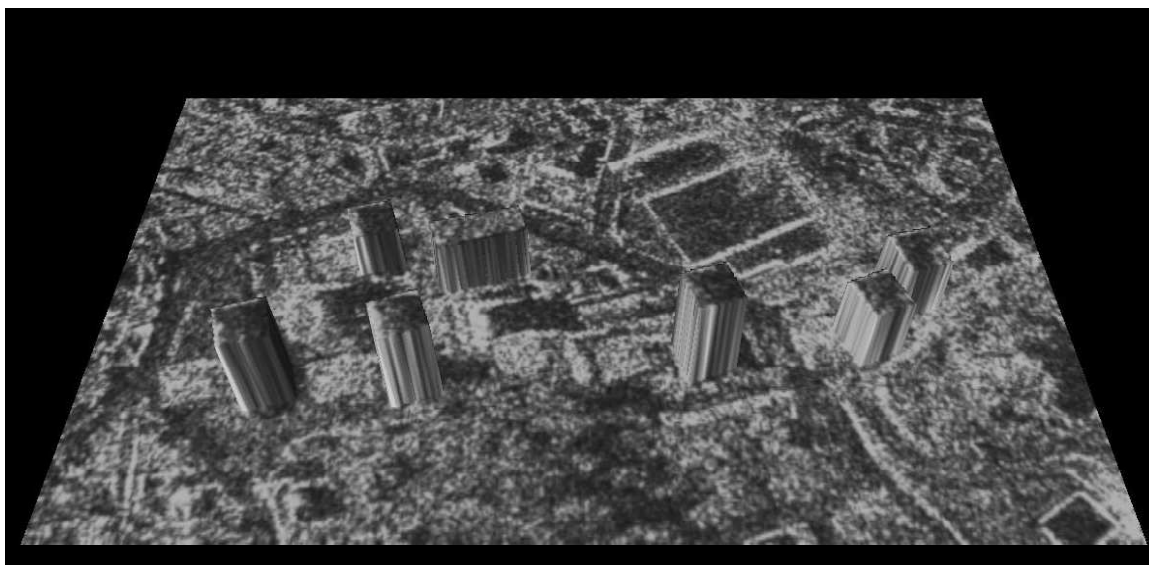


(b) Scène B : hauteurs estimées et emprises validées en fin d'étape 3

FIGURE 6.7 – Cartes des hauteurs estimées des bâtiments validés pour les scènes d'étude A et B. Résultats de l'estimation des hauteurs et de la validation des bâtiments représentés sous forme de cartes en niveaux de gris. La dynamique radiométrique est utilisée pour rendre compte des hauteurs relatives estimées pour les bâtiments validés. Ces cartes sont issues de la fusion des informations altimétriques (hauteur du bâti estimée) et planimétriques (présence du bâti validée) obtenues au cours de la troisième étape de la chaîne développée. Les bâtiments non validés sont désignés par une croix blanche.



(a) Scène A : reconstruction finale obtenue ©Infoterra



(b) Scène B : reconstruction finale obtenue ©Infoterra

FIGURE 6.8 – Résultats de la reconstruction tridimensionnelle des bâtiments sur les scènes d'études A et B (©Infoterra). Les vues 3D sont représentées dans le référentiel image RADAR. La reconstruction est réalisée en combinant les informations planimétriques et altimétriques issues des "cartes des hauteurs estimées des bâtiments validés" avec la composante radiométrique relative à la texture RADAR des objets de la scène.

TABLE 6.2 – Paramètres utilisés pour l'étape d'estimation des hauteurs et de validation.

Noms des paramètres	Scène A	Scène B
$\epsilon_{echo-double}(pixels)$	3	2
$\epsilon_{boite}(pixels)$	100	100
$\epsilon_{discontinuite}(pixels)$	3	3
$h_{min}(mètres)$	3	5
$h_{max}(mètres)$	15	30
$h_{pas}(mètres)$	1	1
$\delta h(mètres)$	2	2
$S_{largeur}^{min}(mètres)$	3	3
$S_{hauteur}^{min}(pourcents)$	2	2
$S_{contraste}$	0.8	0.8

Les tables 6.5 et 6.6 comparent, pour chaque bâti des deux scènes, le statut du bâtiment après validation et son label réel (issu de la vérité terrain) afin de définir le type de validation réalisé :

- validation correcte : lorsqu'un bâtiment existant réellement est validé ;
- validation incorrecte : lorsqu'un bâtiment n'existant pas est validé (cette fausse alarme provient de l'étape d'extraction des bâtis potentiels en monoscopie optique) ;
- rejet correct : lorsqu'un bâtiment n'existant pas (mais provenant d'une fausse alarme de l'étape d'extraction) est rejeté ;
- rejet incorrect : lorsqu'un bâtiment existant réellement est rejeté.

Parmi les 17 emprises potentielles rectangulaires détectées sur les deux scènes, 14 décisions "correctes" ont été prises, à savoir 13 bâtiments existant réellement ont été correctement validés et 1 bâtiment issu d'une fausse alarme a été correctement rejeté.

Les 3 décisions incorrectes correspondent à 2 fausses alarmes validées et au rejet (non-validation) d'un bâtiment existant.

On propose de calculer alors les performances globales obtenues sur les deux scènes d'étude après validation.

On obtient un taux moyen de détection de 81 % (en ne prenant en compte que les bâtiments à emprises rectangulaires) et un taux moyen de fausses alarmes de 13 %.

Par comparaison avec les taux obtenus avant validation à l'issue de l'étape d'extraction optique, qui étaient égaux à 88 % pour le taux moyen de détection et à 18 % pour le taux moyen de fausses alarmes, on constate une diminution intéressante du nombre de fausses alarmes mais accompagnée nécessairement d'une perte en nombre de bâtis détectés.

Le compromis entre ces deux taux dépend directement de la valeur choisie pour le seuil de validation $S_{contraste}$.

Si la démarche opérationnelle choisie est de type "proposition d'une scène 3D reconstruite" ultérieurement soumise à la "vérification manuelle d'un opérateur", susceptible de supprimer certains bâtis, il est préférable d'augmenter la valeur de ce seuil afin d'améliorer le taux de détection, même si cela génère un fort nombre de fausses alarmes.

Si, au contraire, la démarche choisie est de type "proposition d'une scène 3D reconstruite" et "relance locale des algorithmes de reconstruction" sur certaines fenêtres d'intérêt délimitées manuellement par l'opérateur, afin de compléter la reconstruction de certains bâtis de la scène, il est alors préférable de conserver un seuil bas, afin de générer le moins

de fausses alarmes possible.

Les tables 6.7 et 6.8 fournissent, pour chaque bâti des deux scènes, les différences altimétriques entre les hauteurs réelles h_{reel} issues de la vérité terrain et celles estimées $h_{estimee}$ par la méthode présentée.

En ce qui concerne la qualité d'estimation des hauteurs, les valeurs proposées se révèlent être pertinentes. En effet, on obtient des erreurs RMS respectives de 1.34 m et de 0.85 m sur les scènes A et B.

Toutefois, il faut souligner, qu'étant donné le nombre relativement restreint de bâtiments testés (une vingtaine sur les deux scènes), les résultats fournis restent à interpréter avec prudence en termes de représentativité des statistiques. Seul un ordre de grandeur (environ 1 à 3 m d'erreur pour l'estimation de hauteurs de bâtiments correctement validés) est à conserver pour qualifier les performances atteintes par la méthode proposée.

On peut conclure que la reconstruction 3D finale proposée sur les scènes d'étude est globalement satisfaisante lorsque les conditions (taille, contraste, visibilité) précédemment détaillées sont remplies.

6.4 Étude détaillée de quelques bâtiments d'intérêt

Bâtiments faisant l'objet de validations correctes Nous proposons une étude détaillée pour les deux bâtiments d'intérêt suivants :

- le bâtiment 7 de la scène A (voir l'imagette associée en figure 6.9(a)), qui correspond à un exemple de bâtiment respectant la configuration 1 ;
- le bâtiment 2 de la scène B (voir l'imagette associée en figure 6.9(b)), qui correspond à un exemple de bâtiment respectant la configuration 2.

Ces deux bâtiments ont été validés pour des valeurs de contraste optimal respectivement égales à 0.16 et 0.43, obtenues pour des valeurs de $\tilde{h}_{candidate}$ respectivement égales à 9 et 21 m.

Les courbes d'évolution des deux critères $C_{statistique}(h_{test})$ et $C_{contraste}(h_{test})$ sont présentées pour les deux bâtis sur les figures 6.9(c) à 6.9(f).

Le critère de log-vraisemblance apparaît comme suffisamment robuste pour s'affranchir des perturbations (bruit, cibles ponctuelles ...) présentes sur les données RADAR réelles. Les deux minima globaux (respectivement situés à 10 m pour le bâtiment 7 de la scène A et à 19 m pour le bâtiment 2 de la scène B) sont obtenus pour des valeurs de $h_{candidate}$ proches des hauteurs vérité terrain (respectivement égales à 9.3 m et 19.9 m).

Ces valeurs sont ensuite affinées grâce au critère de contraste local. Une estimation précise des hauteurs est finalement obtenue pour chaque bâtiment (hauteur finale $h_{estimee}$ respectivement égale à 9.5 m et à 20 m, ce qui correspond à une erreur altimétrique respective de + 0.2 m et + 0.1 m).

Les figures 6.10 et 6.11 proposent les partitionnements optimaux générés avec les hauteurs finales estimées.

On constate que les régions caractéristiques prédites sont, pour la plupart, assez bien positionnées au niveau de leur localisation visuelle.

Toutefois, on constate certaines imperfections, notamment pour les "zones d'ombres" générées. Ceci s'explique principalement par les erreurs de dimensionnement planimétrique du bâti, issues de l'étape d'extraction des emprises en monoscopie optique, et qui se sont propagées lors de la phase de projection - recalage puis lors de la phase d'estimation des hauteurs - validation. C'est par exemple le cas du bâtiment 7 de la scène

TABLE 6.3 – Résultats individuels d'estimation des hauteurs de bâtiments sur la scène d'étude A.

Label issu de l'étape d'extraction	$h_{candidate}$	$\tilde{h}_{candidate}$	$h_{estimee}$	Statut du bâtiment
1	12 m	13 m	12.5 m	validé
2	6 m	4 m	5 m	validé
3	5 m	3 m	4 m	validé
4	11 m	10 m	10.5 m	validé
5	8 m	8 m	8 m	validé
6	7 m	9 m	8 m	validé
7	10 m	9 m	9.5 m	validé
8	4 m	5 m	4.5 m	validé
9	- m	- m	-	rejeté

TABLE 6.4 – Résultats individuels d'estimation des hauteurs de bâtiments sur la scène d'étude B.

Label issu de l'étape d'extraction	$h_{candidate}$	$\tilde{h}_{candidate}$	$h_{estimee}$	Statut du bâtiment
1	11 m	11 m	11 m	validé
2	19 m	21 m	20 m	validé
3	22 m	21 m	21.5 m	rejeté
4	21 m	21 m	21 m	validé
5	15 m	13 m	14 m	validé
6	22 m	23 m	22.5 m	validé
7	14 m	13 m	13.5 m	validé
8	15 m	15 m	15 m	validé

TABLE 6.5 – Résultats individuels de validation des bâtiments sur la scène d'étude A.

Label issu de l'étape d'extraction	Statut du bâtiment	Label issu de la vérité terrain	Type de validation
1	validé	-	validation incorrecte
2	validé	I	validation correcte
3	validé	E	validation correcte
4	validé	B	validation correcte
5	validé	D	validation correcte
6	validé	C	validation correcte
7	validé	J	validation correcte
8	validé	H	validation correcte
9	rejeté	-	rejet correct
-	-	A	non détection (non rectangulaire)
-	-	F	non détection (non rectangulaire)
-	-	G	non détection (non rectangulaire)

TABLE 6.6 – Résultats individuels de validation des bâtiments sur la scène d'étude B.

Label issu de l'étape d'extraction	Statut du bâtiment	Label issu de la vérité terrain	Type de validation
1	validé	A	validation correcte
2	validé	C	validation correcte
3	rejeté	F	rejet incorrect
4	validé	D	validation correcte
5	validé	I	validation correcte
6	validé	G	validation correcte
7	validé	-	validation incorrecte
8	validé	H	validation correcte
-	-	B	non détection (rectangulaire)
-	-	E	non détection (rectangulaire)

TABLE 6.7 – Différences altimétriques entre les hauteurs réelles vérité terrain et les hauteurs estimées pour la scène d'étude A.

Label issu de l'étape d'extraction	h_{reel}	$h_{estimee}$	Erreur altimétrique
1	non existante	12.5 m	-
2	4.9 m	5 m	+0.1 m
3	6.8 m	4 m	-2.8 m
4	9.1 m	10.5 m	+1.4 m
5	inconnue	8 m	-
6	7.6 m	8 m	+0.4 m
7	9.3 m	9.5 m	+0.2 m
8	5.4 m	4.5 m	-0.9 m
9	non existante	-	-
Erreur RMS pour l'ensemble des bâtiments : 1.34 m.			

TABLE 6.8 – Différences altimétriques entre les hauteurs réelles vérité terrain et les hauteurs estimées pour la scène d'étude B.

Label issu de l'étape d'extraction	h_{reel}	$h_{estimee}$	Erreur altimétrique
1	11.2 m	11 m	-0.2 m
2	19.9 m	20 m	+0.1 m
3	19.4 m	21.5 m mais rejeté	-
4	20.3 m	21 m	+0.7 m
5	14.4 m	14 m	-0.4 m
6	20.6 m	22.5 m	+1.9 m
7	non existante	13.5 m	-
8	15 m	15 m	0 m
Erreur RMS pour l'ensemble des bâtiments : 0.85 m.			

A : l'emprise ayant été légèrement sous-dimensionnée lors de son extraction, la génération de sa "zone d'ombre" suivant l'axe des distances commence en conséquence un peu trop "tôt" ; ceci se manifeste visuellement par un léger empiètement de la partie à l'extrême gauche de la "zone d'ombre" sur la fin du toit du bâti (cette bordure de toit est ici nettement marquée par des structures linéaires brillantes présentes sur sa partie située à l'extrême droite). En revanche, pour le bâtiment 2 de la scène B, un autre phénomène intervient. On constate un chevauchement entre la "zone d'ombre" optimale prédite pour ce bâtiment d'intérêt et la réponse visuelle du toit (région brillante) d'un bâtiment voisin situé sur sa droite (bâtiment 4 de la scène B). Ceci s'explique par le fait que la "zone d'ombre" du bâtiment 2 est à peu près correctement positionnée mais que le bâtiment 4, légèrement plus haut que le bâtiment d'intérêt, voit son sommet répondre avant la fin de la "zone d'ombre" imagée pour le bâtiment d'intérêt.

Bâtiment faisant l'objet d'un rejet incorrect Nous proposons à présent de commenter le cas du bâtiment 3 de la scène B, qui a été rejeté à tort.

Nous avons vu précédemment que, tant que les interférences de rétrodiffusion entre bâtiments proches restaient suffisamment faibles, le procédé permettait de gérer le cas de petites perturbations locales de l'intensité et réussissait à produire des résultats corrects d'estimation de hauteurs et de validation. Toutefois, lorsque ces perturbations deviennent trop importantes (ce qui peut arriver dans le cas de bâtiments grands et très proches), le procédé peut échouer. C'est ce qui s'est produit pour le bâtiment 3 de la scène B.

Ceci s'explique par la combinaison de deux phénomènes : premièrement, l'absence complète de la "zone d'ombre" du bâti considéré (due à la superposition de cette zone avec la "zone de recouvrement" du bâti situé derrière) a empêché le critère statistique de fournir des hauteurs candidates précises ; deuxièmement, la présence de perturbations relativement fortes (dues à des cibles ponctuelles brillantes et à la superposition partielle de la "zone de fond" du bâti considéré avec la "zone d'écho double" d'un bâti voisin) dans la bande étroite B_F , située du côté de la discontinuité donnant sur la "zone de fond", a empêché le critère de contraste d'estimer une hauteur de bâti correctement affinée. Ceci a conduit au rejet du bâtiment, puisque le critère de contraste n'est pas suffisamment bien vérifié (une valeur de contraste optimal égale à 0.87, supérieure au seuil $S_{\text{contraste}}$ fixé, a en effet été obtenue).

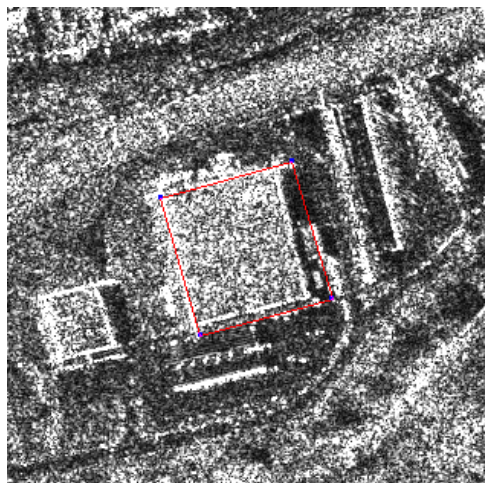
6.5 Conclusions et améliorations

Une nouvelle méthode, basée sur la combinaison de deux critères RADAR, a été développée dans ce chapitre afin de valider la présence des bâtis potentiels extraits de l'image optique et de remonter à leur information d'élévation.

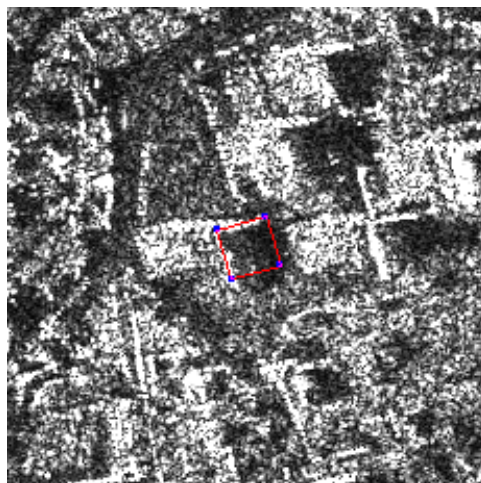
Par comparaison avec une vérité terrain, les résultats obtenus sur les données réelles ont permis d'atteindre de bonnes performances pour la composante de validation (taux de détection et taux de fausses alarmes corrects) et se sont révélés être pertinents quant à la composante d'estimation des hauteurs.

Les scènes urbaines étudiées ont ainsi fait l'objet d'une reconstruction 3D tout à fait satisfaisante à la fin de cette troisième étape.

Les améliorations à envisager sont de deux types.



(a) Bâtiment 7 : emprise



(b) Bâtiment 2 : emprise

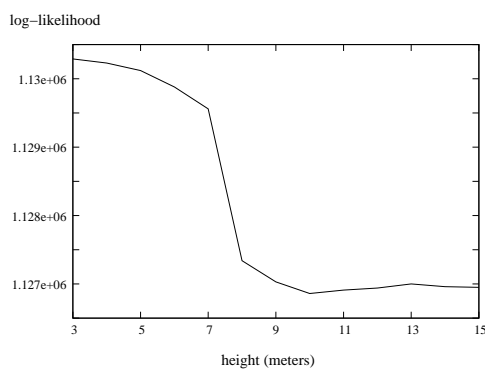
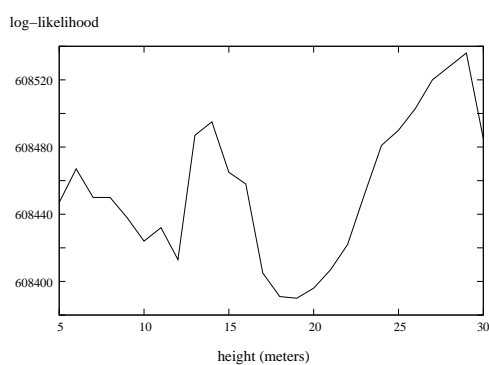
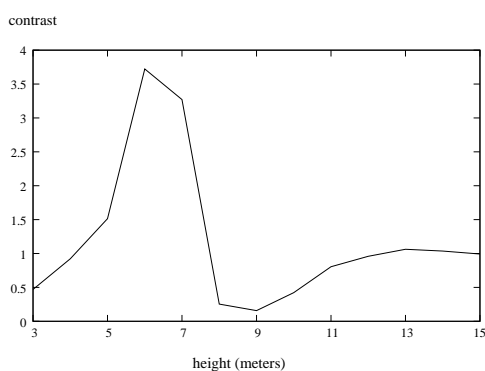
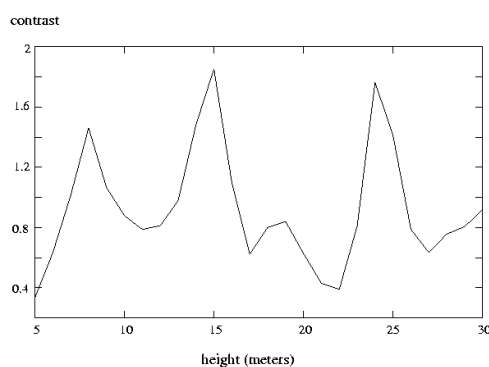
(c) Bâtiment 7 : $C_{statistique}(h_{test})$ (d) Bâtiment 2 : $C_{statistique}(h_{test})$ (e) Bâtiment 7 : $C_{contraste}(h_{test})$ (f) Bâtiment 2 : $C_{contraste}(h_{test})$

FIGURE 6.9 – Résultats obtenus pour deux bâtiments d'intérêt sur données RADAR réelles (©Infoterra). (a) et (b) Limites des emprises de bâtis recalées sur les imagettes RADAR. (c) et (d) Évolution du critère statistique $C_{statistique}(h_{test})$ en fonction de la hauteur. (e) et (f) Évolution du critère de contraste $C_{contraste}(h_{test})$ en fonction de la hauteur.

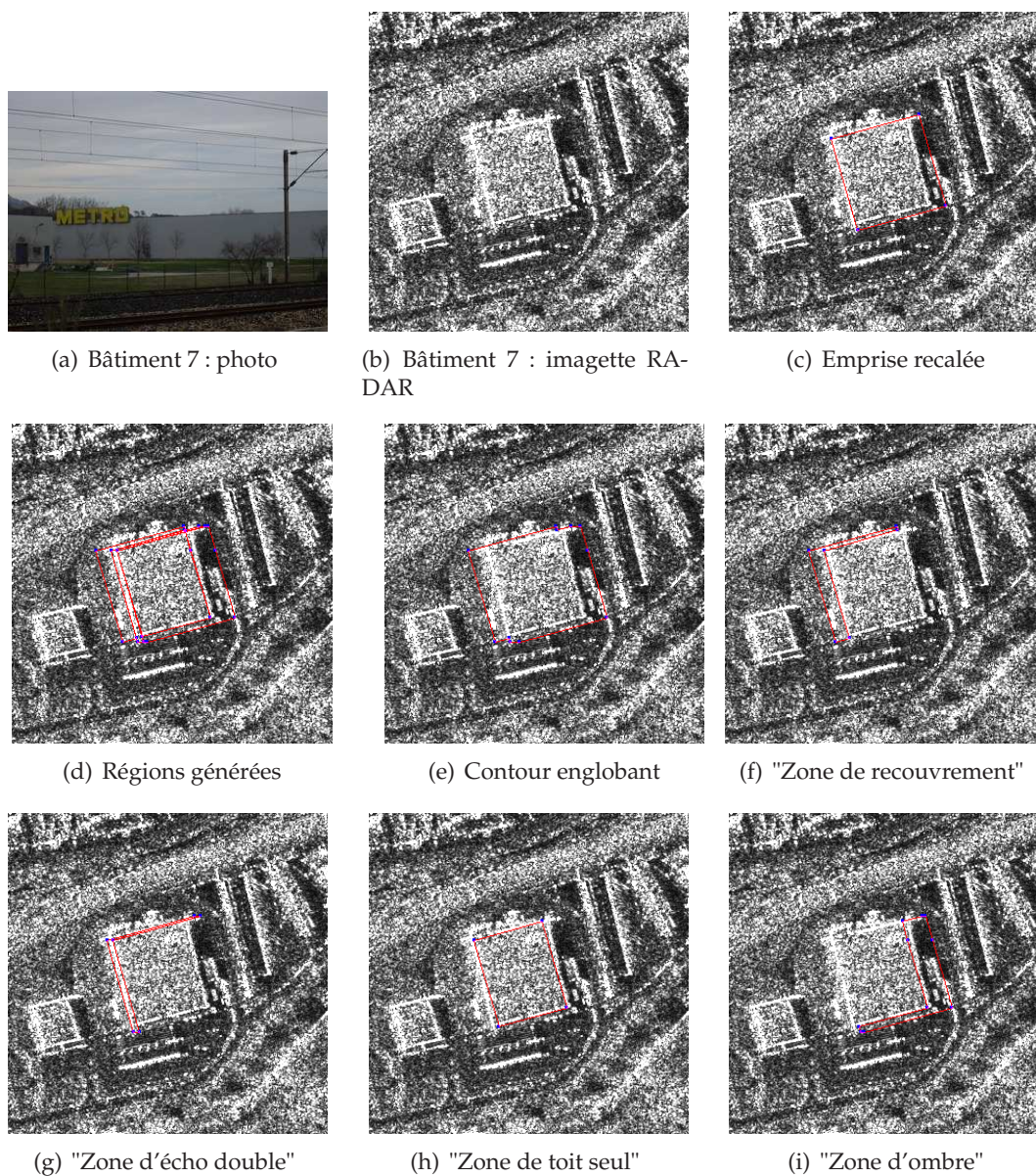


FIGURE 6.10 – Résultats du partitionnement optimal en régions caractéristiques, obtenu pour le bâtiment 7 sur la scène A (©Infoterra). Ce partitionnement a été généré en utilisant la hauteur finale estimée. (d) Toutes les régions générées sont représentées en rouge sur la même imagerie RADAR. (e) La frontière entre la "zone de fond" et le contour englobant, qui contient l'ensemble des autres régions caractéristiques, est représentée en rouge.

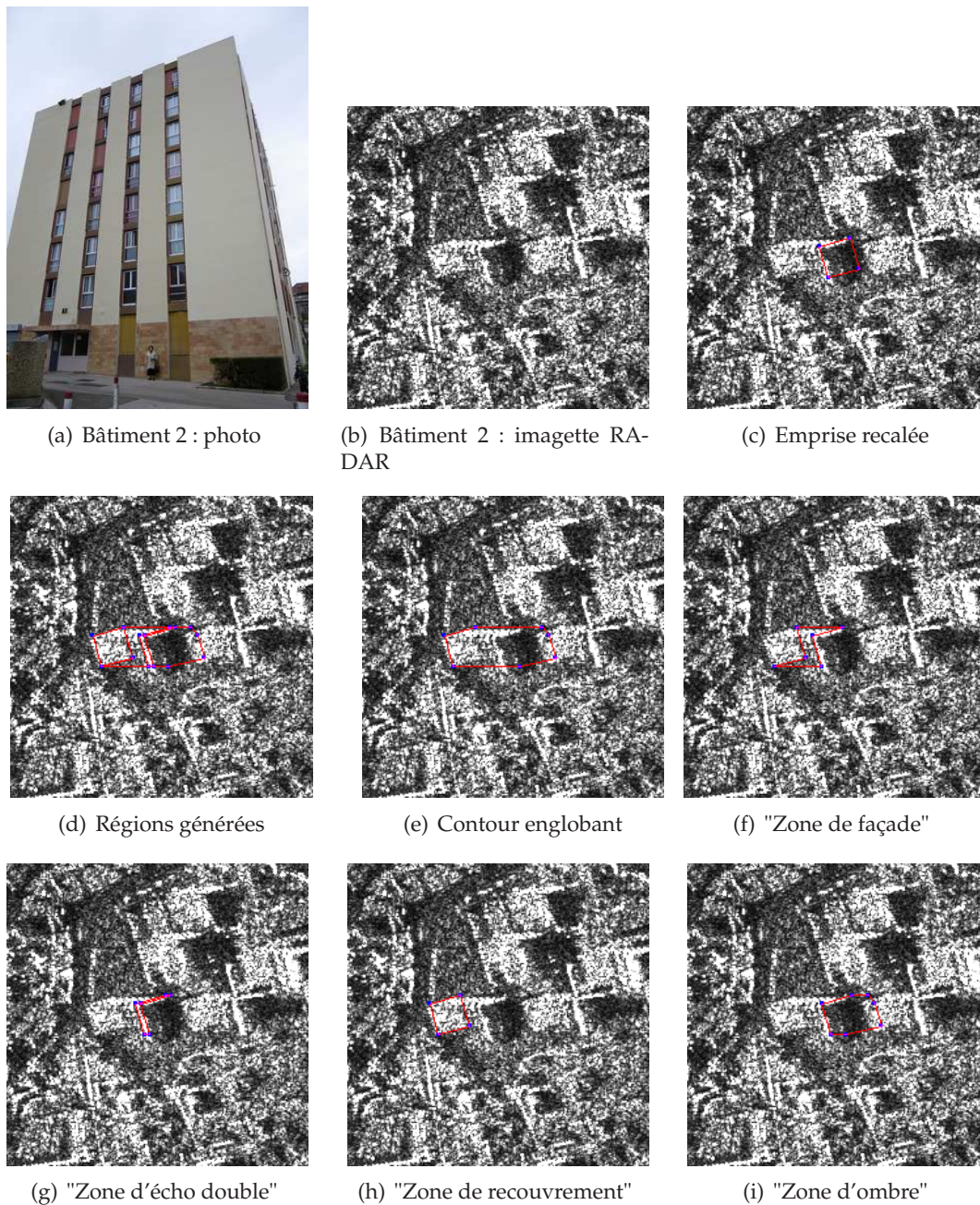


FIGURE 6.11 – Résultats du partitionnement optimal en régions caractéristiques, obtenu pour le bâtiment 2 sur la scène B (©Infoterra). Ce partitionnement a été généré en utilisant la hauteur finale estimée. (d) Toutes les régions générées sont représentées en rouge sur la même imagerie RADAR. (e) La frontière entre la "zone de fond" et le contour englobant, qui contient l'ensemble des autres régions caractéristiques, est représentée en rouge.

Premièrement, les critères utilisés pourraient faire l'objet d'extension ou de généralisation. D'une part, il conviendrait d'adapter le critère statistique aux cas de lois de modélisation des distributions d'intensité plus complexes (exemple : loi de Fisher, à queue lourde) et susceptibles de mieux convenir aux régions caractéristiques bruitées en milieu urbain, pour l'imagerie RADAR à haute résolution. D'autre part, le critère de contraste utilisé pourrait résulter de la combinaison de plusieurs contrastes locaux, calculés sur différentes zones de discontinuités, afin de consolider la phase de sélection des hauteurs candidates (exemple : combinaison du critère de contraste "fond"-"recouvrement" et du critère de contraste "ombre"-"fond").

Deuxièmement, il serait judicieux d'étudier l'amélioration des performances obtenue en procédant à une estimation des hauteurs et à une validation conjointe des bâtiments, à partir de la fusion de critères RADAR et optiques. Cet axe de recherche fera d'ailleurs l'objet d'un approfondissement dans le chapitre 8.

Chapitre 7

Qualification de la reconstruction des bâtiments

Dans ce chapitre, nous présentons une méthode de qualification des bâtiments reconstruits, qui s'appuie sur le chapitre 6.

Cette méthode permet d'évaluer la confiance à accorder à chaque bâti reconstruit à travers la proposition de scores de qualité, calculés à partir des critères RADAR introduits dans le chapitre 6 et susceptibles d'être utilisés comme indicateurs de confiance en conditions opérationnelles. Des "cartes de confiance des bâtiments reconstruits" sont fournies à l'issue de cette quatrième étape de la chaîne de traitements. La méthode proposée est mise en œuvre et commentée sur les scènes d'études.

7.1 Méthode proposée pour la qualification des bâtiments reconstruits

7.1.1 Introduction de scores de qualité à partir de l'image RADAR

Définition de scores partiels Nous introduisons un jeu de scores de qualité partiels, sélectionnés parmi un ensemble de mesures testées et susceptibles de fournir une évaluation de la qualité des bâtiments reconstruits selon différents aspects.

Ces scores prennent en compte des informations relatives à la forme, à la localisation et à la valeur des pics optimaux obtenus dans les courbes d'évolutions des deux critères $C_{statistique}(h_{test})$ et $C_{contraste}(h_{test})$, calculés sur les imagerie RADAR et utilisés pour valider les bâtis et estimer leur hauteur.

Ces scores occupent tous une plage de valeurs s'étalant entre 0 (score attestant une faible confiance) et 1 (score attestant une forte confiance) et sont définis de la façon suivante :

1. Pour le critère $C_{statistique}(h_{test})$, on introduit :
 - le score partiel SP_1 , qui se réfère à la valeur optimale $C_{statistique}(h_{candidate})$ du critère statistique (où $h_{candidate}$ correspond à la hauteur candidate utilisée pour estimer la hauteur du bâti considéré) ; plus celle-ci est faible (log-vraisemblance généralisée globale grande), plus le score SP_1 est grand ;
 - le score partiel SP_2 , qui se réfère à la valeur optimale $L_{pic}(h_{candidate})$ de la largeur du pic optimal ; plus celle-ci est grande (pic avec forte représentativité par rapport aux autres pics candidats), plus le score SP_2 est grand ;
-

- le score partiel SP_3 , qui se réfère à la valeur optimale $H_{pic}(h_{candidate})$ de la hauteur relative du pic optimal ; plus celle-ci est grande (pic profond par rapport aux autres pics candidats), plus le score SP_3 est grand.
2. Pour le critère $C_{contraste}(h_{test})$, on introduit :
- le score partiel SP_4 , qui se réfère à la valeur optimale $C_{contraste}(\tilde{h}_{candidate})$ du critère de contraste ; plus celle-ci est faible (contraste local grand et satisfaisant l'a priori radiométrique au niveau de la discontinuité testée), plus le score SP_4 est grand ;
 - le score partiel SP_5 , qui se réfère à la largeur minimale δh_{min} du voisinage autour de $h_{candidate}$, requise pour obtenir un contraste inférieur au seuil $S_{contraste}$; plus celle-ci est faible, plus le score SP_5 est grand (lorsque $\delta h_{min} > \delta h$, le score SP_5 vaut 0) ;
 - le score partiel SP_6 , qui se réfère à l'écart de localisation suivant l'axe des hauteurs tests entre la hauteur affinée $\tilde{h}_{candidate}$ sélectionnée et la hauteur h_{test} conduisant au minimal local (du critère de contraste) le plus proche ; plus cet écart est petit ($\tilde{h}_{candidate}$ proche du minimal local), plus le score SP_6 est grand (lorsque $\tilde{h}_{candidate}$ coïncide précisément avec le minimal local, le score SP_6 vaut 1).

Définition d'un score total Afin de fournir localement une unique mesure de qualité pour chaque bâti reconstruit, un score total de qualité ST est calculé à partir des scores partiels.

Celui-ci est simplement défini comme la moyenne arithmétique des différents scores de qualité partiels.

L'opérateur de fusion des scores, utilisé dans ces travaux, correspond donc à un opérateur basique, qui donne un poids équivalent à chaque score partiel.

Toutefois, il convient de souligner que les résultats présentés dans ce chapitre ne constituent que des résultats préliminaires. Il serait tout à fait intéressant de mener une étude plus approfondie sur l'apport individuel de chaque score de qualité partiel utilisé dans le calcul de ST , afin de mettre en place une stratégie de fusion des scores plus élaborée, tenant compte de la fiabilité, de la pertinence et de la précision de chaque score partiel considéré (exemple : mise en œuvre de la démarche proposée par Bloch [Blo96] pour la sélection d'opérateurs de fusion adaptés).

7.1.2 Intérêt opérationnel

Ces scores peuvent être utilisés pour guider un opérateur dans le cadre d'une démarche opérationnelle de type "proposition d'une scène 3D reconstruite" et "soumission à une vérification manuelle".

En ne soumettant à une vérification manuelle que les bâtiments identifiés par un score total de qualité ST trop "faible" (c'est-à-dire inférieur ou égal à un seuil de qualité $S_{qualite}$), on parviendrait à la fois à améliorer les performances de détection (suppression de certaines fausses alarmes) et à conserver des temps de traitement acceptables pour la reconstruction 3D de larges zones.

7.2 Application sur les scènes d'études

7.2.1 Résultats obtenus

Les figures 7.1(a) et 7.1(b) présentent les "cartes de confiance des bâtiments reconstruits" obtenues à la fin de la quatrième étape.

Celles-ci se rapportent, pour chaque bâtiment reconstruit, au score total ST attribué.

7.2.2 Analyse qualitative et quantitative

Les tables 7.1 et 7.2 comparent, pour chaque bâti des deux scènes, le type de validation réalisé à l'issue de la troisième étape de la chaîne, l'erreur altimétrique commise et la valeur du score total de qualité ST obtenu.

En analysant ces résultats, on constate que tous les bâtiments ont obtenu un score total de qualité ST strictement supérieur à 0.5, excepté 2 bâtiments : le bâtiment 3 de la scène A et le bâtiment 7 de la scène B.

En conséquence, si le seuil de qualité $S_{qualite}$ était fixé à 0.5, seuls ces 2 bâtiments devraient faire l'objet d'une vérification manuelle par l'opérateur. Ce dernier désignerait le second comme fausse alarme au cours de cette rapide analyse, ce qui permettrait d'augmenter le nombre de décisions "correctes" par rapport à celui issu de l'étape de validation (passage de 14 à 15 décisions "correctes" sur 17, soit un taux moyen de fausses alarmes passant de 13 % à 7 %).

De même, si le seuil de qualité $S_{qualite}$ était fixé à 0.7, 9 bâtiments reconstruits devraient être ré-examinés. L'analyse serait donc plus longue mais l'opérateur serait en mesure de détecter 2 fausses alarmes parmi ces 9 bâtiments : le bâtiment 1 de la scène A et le bâtiment 7 de la scène B. Cela permettrait encore d'augmenter le nombre de décisions "correctes" (passage de 14 à 16 décisions "correctes" sur 17, soit un nombre de fausses alarmes nul). La seule erreur commise serait de type "non-détection" (rejet incorrect du bâtiment 3 de la scène 2).

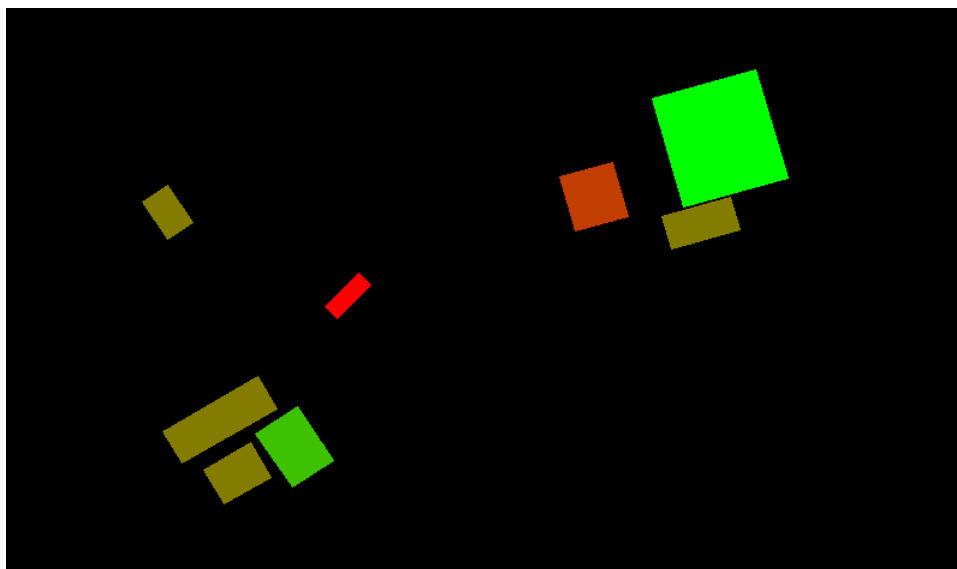
Il serait ainsi opérationnellement possible de laisser à l'opérateur la possibilité de fixer ce seuil $S_{qualite}$ à partir de son appréciation visuelle (par exemple, à travers la mise en place d'un curseur sur la fenêtre de reconstruction 3D, permettant de régler le compromis entre le nombre de bâtis à examiner et le nombre de fausses alarmes potentiellement détectables).

En comparant individuellement les erreurs altimétriques commises et les valeurs des scores de qualité ST obtenues pour les bâtis existants (validations correctes) et correspondants aux cas extrêmes, il est possible d'émettre quelques commentaires quant à la pertinence des mesures de confiance utilisées.

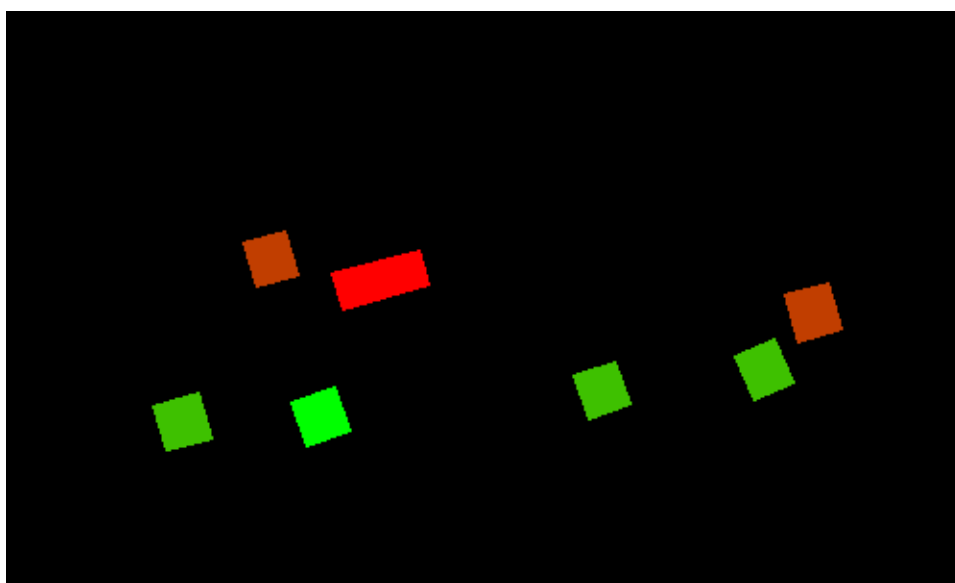
On remarque que la plupart des bâtiments ayant obtenu un "fort" score de qualité (par exemple : le bâtiment 7 de la scène A et les bâtiments 2, 4 et 8 de la scène B dont le score ST est égal à 0.75 ou 0.83) ont donné lieu à des erreurs altimétriques faibles (systématiquement inférieures à 0.7 m pour ceux cités).

Au contraire, le score ST le plus faible a été attribué au bâtiment 3 de la scène A ($ST = 0.5$), pour lequel la plus grande erreur altimétrique a été commise (presque 3 m).

La mesure de confiance fournie par le score ST apparaît donc comme globalement pertinente même si quelques incohérences subsistent dans la qualification des bâtiments reconstruits. C'est notamment le cas pour le bâtiment 6 de la scène B qui présente un score ST relativement élevé (0.75) en dépit d'une erreur altimétrique relativement importante (presque 2 m).



(a) Scène A : scores totaux de qualité obtenus en fin d'étape 4 pour les emprises validées



(b) Scène B : scores totaux de qualité obtenus en fin d'étape 4 pour les emprises validées

FIGURE 7.1 – Cartes de confiance des bâtiments reconstruits pour les scènes d'étude A et B. Résultats de la qualification des bâtiments reconstruits représentés sous forme de cartes en couleurs. La gamme des couleurs rouge - jaune - vert est utilisée pour rendre compte des confiances relatives estimées pour les bâtiments validés. Plus un bâti est proche du rouge, plus son score total de qualité est faible. Plus un bâti est proche du vert, plus son score total de qualité est fort. La dynamique radiométrique utilisée (256 niveaux) s'étale entre les valeurs minimales et maximales observées pour *ST*.

TABLE 7.1 – Résultats individuels de qualification des bâtiments reconstruits sur la scène d'étude A.

Label issu de l'étape de validation	Type de validation	Erreur altimétrique	Score total de qualité <i>ST</i>
1	validation incorrecte	-	0.66
2	validation correcte	+0.1 m	0.66
3	validation correcte	-2.8 m	0.5
4	validation correcte	+1.4 m	0.66
5	validation correcte	-	0.75
6	validation correcte	+0.4 m	0.66
7	validation correcte	+0.2 m	0.83
8	validation correcte	-0.9 m	0.58

TABLE 7.2 – Résultats individuels de qualification des bâtiments reconstruits sur la scène d'étude B.

Label issu de l'étape de validation	Type de validation	Erreur altimétrique	Score total de qualité <i>ST</i>
1	validation correcte	-0.2 m	0.58
2	validation correcte	+0.1 m	0.75
4	validation correcte	+0.7 m	0.83
5	validation correcte	-0.4 m	0.58
6	validation correcte	+1.9 m	0.75
7	validation incorrecte	-	0.5
8	validation correcte	0 m	0.75

7.3 Conclusions et améliorations

Ce chapitre a mis en valeur une propriété de nos travaux importante dans un cadre d'utilisation opérationnelle : la proposition d'indicateurs de qualité pour la scène urbaine reconstruite.

Une méthode, basée sur l'introduction de scores de confiance calculés sur l'image RADAR, a été développée afin d'évaluer globalement la qualité des bâtiments reconstruits.

Sa mise en œuvre sur les deux scènes d'étude a permis de produire des cartes de scores satisfaisantes, reflétant la qualité des bâtiments reconstruits et permettant d'orienter un opérateur dans une démarche de type "vérification manuelle d'une scène de reconstruction 3D soumise".

Les mesures de confiance dont témoignent les valeurs des scores ST sont apparues comme globalement fiables.

Toutefois, deux types d'améliorations peuvent être soulignés.

D'une part, la mise en place d'une stratégie plus élaborée de fusion des scores partiels participant au calcul ST permettrait certainement d'obtenir des scores de qualité plus pertinents (exemples : attribution de poids différents aux différents scores partiels ou sélection des scores partiels les plus adaptés à une situation donnée).

D'autre part, comme cela a été remarqué pour l'étape d'estimation des hauteurs et de validation, il serait judicieux d'étudier l'amélioration des performances obtenue en procédant à une qualification conjointe des bâtiments reconstruits par fusion de scores de qualité issus de critères RADAR et optiques. Cet axe de recherche fera également l'objet d'un approfondissement dans le chapitre 8.

Chapitre 8

Extensions

Nous proposons dans ce chapitre un certain nombre de pistes susceptibles d'être explorées afin d'améliorer les résultats obtenus et d'étendre la portée des travaux présentés dans ce manuscrit.

Les extensions décrites relèvent de deux types : les perspectives relatives aux aspects recherche et les perspectives relatives aux aspects industriels.

Dans un premier temps, nous présentons les quatre points d'investigation suivants, permettant de généraliser et de renforcer la chaîne de traitements développée :

- la complexification du modèle de bâti à considérer afin d'augmenter le type et le nombre de bâtis susceptibles d'être reconstruits ;
- la symétrisation de la chaîne visant à donner un poids équivalent aux composantes optiques et RADAR afin d'améliorer les performances de reconstruction ;
- la proposition d'une variante méthodologique pour la détection des bâtiments (première étape de la chaîne) dans le cadre de la fusion optique-RADAR (ces travaux ont fait l'objet d'une collaboration dans le cadre d'un stage de Master Recherche M2 à Télécom ParisTech) ;
- la gestion d'une nouvelle configuration d'entrée (cas d'une image optique et d'un couple RADAR interférométrique) afin d'étendre le potentiel de la chaîne à un second scénario opérationnel.

Dans un second temps, nous présentons les aspects relatifs à la capitalisation industrielle des travaux présentés. Nous décrivons notamment un projet visant à intégrer l'ensemble des algorithmes développés dans la chaîne de traitements au sein d'un démonstrateur opérationnel pour la reconstruction 3D de scènes urbaines et semi-urbaines.

8.1 Perspectives de recherche

8.1.1 Complexification du modèle de bâtiment

Il serait intéressant de généraliser l'ensemble des algorithmes développés dans le cas du bâtiment 3D parallélépipédique aux cas de bâtiments plus sophistiqués, afin d'élargir le type de bâtis potentiellement traités par la chaîne.

Cette complexification pourrait se faire :

1. Au niveau planimétrique : en considérant des emprises circulaires ou régulières (par exemple, en forme de "L", de "T" ou de "U"). Nous pourrions par exemple
-

utiliser un ensemble de règles de positionnement permettant de combiner des rectangles afin de conclure quant à une première forme géométrique à angles droits. Nous pourrions encore mettre en œuvre la méthodologie développée par Tison [Tis04] pour l'approximation des contours des bâtiments par une forme polygonale (cette méthode est basée sur une adaptation de la méthode des contours actifs au cas de polygones avec un nombre de noeuds limité). Ceci permettrait déjà de couvrir une large gamme parmi les emprises de bâtis les plus répandues.

2. Au niveau altimétrique : en considérant des formes de toits légèrement plus complexes que les toits plats (par exemple, les toits bi-pentes ou les toits galbés). Il faudrait alors considérer plusieurs hauteurs pour caractériser l'élévation d'un bâtiment (hauteur à la base du toit et hauteur au sommet du toit). Notons que ces investigations ont déjà fait l'objet de travaux présentés dans la littérature. Toutefois, ceci a nécessité, d'une part, de travailler sur des images présentant une résolution plus fine (par exemple, dans les travaux de Thiele et al. [TCS⁺07b], deux approches complémentaires sont développées pour l'estimation des hauteurs de bâtiments à toits bi-pentes, à partir de données InSAR aéroportées présentant une résolution submétrique de l'ordre de 10 à 40 cm) et, d'autre part, de fixer un certain nombre de paramètres altimétriques afin de faciliter les étapes d'optimisation (par exemple, dans les travaux de Brunner [Bru09], la reconstruction des bâtiments à toits bi-pentes se fait en fixant initialement leur pente à 35° pour la scène d'étude considérée).

Il conviendrait alors d'adapter les configurations géométriques utilisées pour modéliser les signatures images des bâtiments aux nouveaux modèles 3D envisagés.

Nous pensons que le cadre théorique développé dans cette thèse devrait permettre une telle extension à moindre frais dans le cas des bâtiments à toit plat résultant de la combinaison de rectangles.

8.1.2 Symétrisation de la chaîne de traitements

Une des pistes clefs à explorer dans la suite de ces travaux consisterait à analyser les améliorations qui résulteraient d'une version symétrisée de la chaîne de traitements proposée.

On entend par le mot "symétrisation" le fait de donner aux composantes optiques et RADAR des rôles identiques à l'intérieur des différentes étapes de la reconstruction 3D des bâtiments, de façon à exploiter pleinement les informations complémentaires fournies par les primitives caractéristiques propres à chaque image.

Ceci n'a pas été le cas pour la version de la chaîne présentée au long du manuscrit. En effet, les informations de détection et de localisation des bâtis ont principalement été issues de l'image optique (étape 1) tandis que celles d'élévation-validation (étape 3) et de qualification (étape 4) ont principalement été issues de l'image RADAR (voir figure 3.1). Les composantes optiques et RADAR ont ainsi rempli des rôles complémentaires mais non partagés.

Cette première version de la chaîne de traitements, simple et globalement satisfaisante, pourrait ainsi faire l'objet d'une "symétrisation" afin de gagner en robustesse et en efficacité (augmentation du nombre de bâtis détectés, diminution du nombre de fausses alarmes commises).

L'idée consisterait alors à introduire, au sein de chaque étape, des informations additionnelles issues de l'image non encore exploitée (image RADAR dans l'étape 1 et image

optique dans les étapes 3 et 4) afin d'atteindre de meilleures performances.

Nous proposons ainsi de développer dans cette section une démarche explicite pour la symétrisation de la chaîne de traitements et nous en soulignons les bénéfices.

La figure 8.1 présente un schéma synoptique de la version symétrique de la chaîne complète. Les boîtes et les flèches représentées en pointillés mettent en évidence les boîtes introduites ou modifiées lors de la symétrisation. Idéalement, il conviendrait, à chaque étape, de fusionner au sein d'un même module les différentes boîtes utilisées à ce niveau de la chaîne.

Trois aspects sont traités :

- pour l'étape 1 : d'une part, l'augmentation du taux de détection, à travers la mise en œuvre d'une détection de fenêtres d'intérêt supplémentaires sur l'image RADAR ; et, d'autre part, la ré-estimation affinée de l'orientation principale des contours de bâtiments, grâce à l'introduction de valeurs d'orientation précises issues des images RADAR ;
- pour l'étape 3 : d'une part, l'augmentation de la fiabilité à accorder aux hauteurs de bâtis estimées, à travers l'introduction d'un critère statistique et d'un critère de contraste conjointement calculés sur les données optiques et RADAR, et, d'autre part, la diminution du taux de fausses alarmes à travers la mise en application d'une validation croisée ;
- pour l'étape 4 : l'amélioration de l'étape de qualification, assurée par la prise en compte de scores de confiance conjointement calculés sur les données optiques et RADAR.

8.1.2.1 Pour l'étape 1 : détection de fenêtres d'intérêt supplémentaires et introduction de valeurs d'orientation précises à partir de l'image RADAR

Augmentation du taux de détection Un nouveau schéma pour l'extraction des contours de bâtiments à partir d'un couple optique-RADAR est proposé.

L'idée consiste, dans un premier temps, à extraire une "carte des fenêtres d'intérêt RADAR" à partir de l'image RADAR et dans un second temps, à utiliser celle-ci, pour conditionner la génération de la "carte des fenêtres d'intérêt optiques", décrite au chapitre 4 et utilisée en entrée de la seconde phase du procédé (extraction fine des emprises potentielles).

La "carte des fenêtres d'intérêt RADAR" serait générée de la façon suivante.

Premièrement, certaines primitives RADAR caractéristiques seraient extraites en utilisant par exemple :

- un détecteur de structures linéaires (Tupin et al. [TMM⁺98]) pour les échos doubles brillants ;
- un algorithme de segmentation par grille active statistique (Germain et al. [GR01]) pour les ombres, complété par un seuillage sur la moyenne radiométrique des régions.

Deuxièmement, des couples de primitives écho-ombre associées seraient identifiés à partir de considérations sur leur localisation relative et sur leur proximité.

Troisièmement, la minimisation d'un score énergétique, calculé à partir de l'image RADAR seuillée des structures linéaires détectées, permettrait de générer autour de chaque couple identifié une fenêtre rectangulaire d'intérêt correspondante.

Cette "carte des fenêtres d'intérêt RADAR" serait ensuite projetée dans le référentiel image optique (en utilisant le MNT disponible) afin de conditionner la génération de la

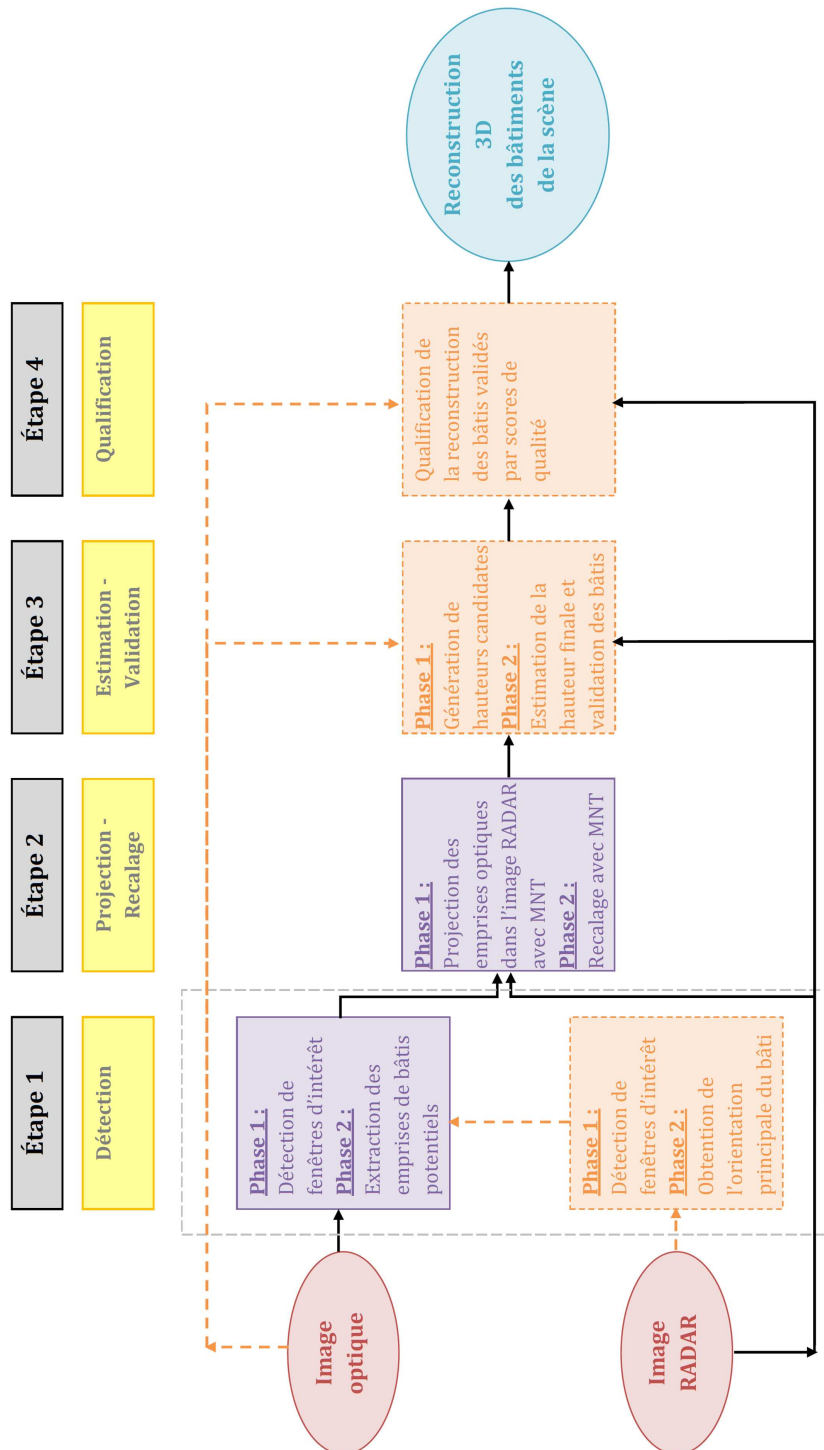


FIGURE 8.1 – Schéma synoptique de la chaîne symétrique proposée. Les boîtes et les flèches faisant l'objet de modifications par rapport à la version initiale de la chaîne de traitements sont représentées en pointillés.

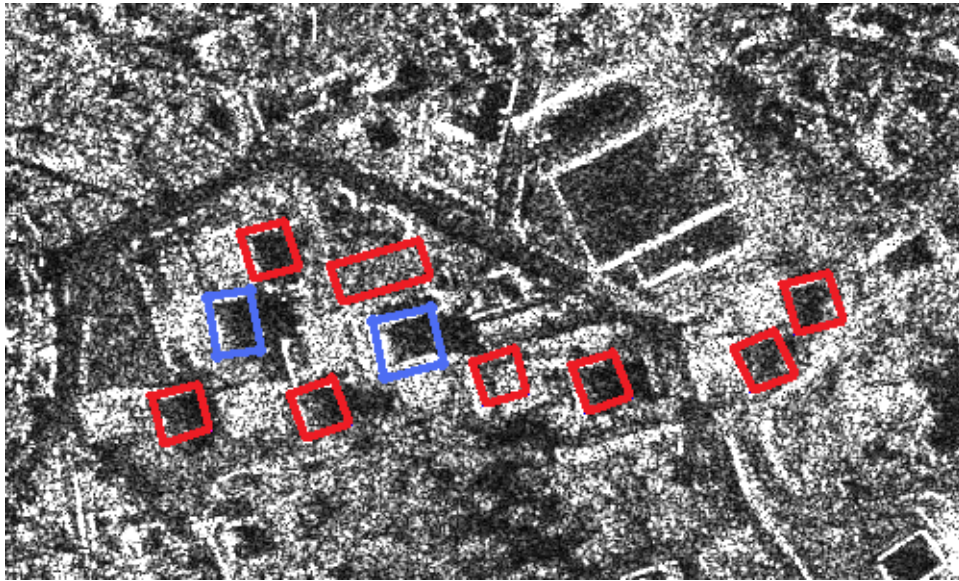


FIGURE 8.2 – Carte de localisation des fenêtres d'intérêt optiques-RADAR détectées pour la scène d'étude B ©Infoterra.

"carte des fenêtres d'intérêt optiques". Par exemple, lors de l'application du test d'adéquation géométrique détaillé en section 4.1.2 et réalisé sur chaque région du P. M. D. optique, nous pourrions localement envisager un seuil de recouvrement minimal plus faible, lorsqu'une fenêtre d'intérêt RADAR apparaît autour de la région testée.

Ceci permettrait de traiter en partie les problèmes de détection relatifs à la forte variabilité des signatures, qui caractérise la visibilité des primitives optiques et RADAR (les contours de bâtis optiques ou les échos RADAR n'étant en effet pas toujours distinguables sur les données images traitées).

Ici, on pourrait espérer pouvoir utiliser les primitives RADAR d'un bâtiment, lorsque ses primitives optiques n'apparaissent pas clairement, afin de conforter l'hypothèse de présence du bâtiment considéré.

Ce schéma symétrique nous permettrait ainsi de compléter les résultats de détection à l'issue de la première étape du traitement.

La figure 8.2 présente, sur la scène d'étude B, un exemple de résultat préliminaire, relatif à cette symétrisation de la première étape du traitement.

Nous avons superposé, dans le référentiel image RADAR, les fenêtres d'intérêt suivantes :

- en rouge, figurent les fenêtres présentes sur la "carte des fenêtres d'intérêt optiques" initialement générée (c'est-à-dire issue d'un traitement en monoscopie optique) ;
- en bleu, figurent les fenêtres supplémentaires présentes uniquement sur la "carte des fenêtres d'intérêt RADAR", générée dans le cadre de la symétrisation.

Nous constatons que deux fenêtres supplémentaires, correspondant effectivement aux deux bâtiments initialement non détectés en monoscopie optique, ont pu être mises en évidence à partir de couples de primitives écho-ombre RADAR. Celles-ci pourront être utilisées pour relancer localement les algorithmes de détection des fenêtres optiques et d'extraction fine des contours, en considérant des seuils de sélection plus faibles.

Ré-estimation affinée de l'orientation principale des contours de bâtiments L'étape de génération d'une "carte des fenêtres d'intérêt RADAR" (précédemment décrite) ne nous permettrait sans doute pas d'extraire directement sur l'image RADAR des informations précises sur la localisation des contours de bâtis (les échos RADAR étant souvent vus partiellement ou sous forme de segments disjoints).

Néanmoins, l'extraction locale d'une direction principale à partir d'une imagerie RADAR est un procédé jugé suffisamment robuste pour nous permettre de considérer comme pertinente l'information d'orientation fournie.

Cette dernière pourrait ainsi être comparée à celle issue de l'image optique (dans la seconde phase de l'étape 1) et, dans le cas de valeurs d'orientation principale proches, une nouvelle valeur affinée pourrait alors être estimée par combinaison.

8.1.2.2 Pour l'étape 3 : estimation jointe des hauteurs des bâtiments et validation croisée

Augmentation de la fiabilité à accorder aux hauteurs de bâtis estimées Le principe serait de combiner les critères RADAR présentés dans le chapitre 6 avec de nouveaux critères optiques, dépendants de la hauteur des bâtiments.

Par exemple, nous pourrions définir un critère combiné de log-vraisemblance généralisée $C_{statistique}^{combine}(h_{test})$ comme la somme pondérée d'un critère statistique RADAR $C_{statistique}^{RADAR}(h_{test})$ et d'un critère statistique optique $C_{statistique}^{optique}(h_{test})$ suivant l'équation :

$$C_{statistique}^{combine}(h_{test}) = C_{statistique}^{RADAR}(h_{test}) + k C_{statistique}^{optique}(h_{test}) \quad (8.1)$$

où k permettrait de pondérer la contribution relative de chacune des composantes RADAR et optique (k pourrait être déduit du degré de "netteté" avec lequel les primitives caractéristiques RADAR et optiques sont imagées).

Pour le calcul de $C_{statistique}^{RADAR}$, nous utiliserions l'équation 6.12, introduite au chapitre 6 et basée sur l'estimation de la log-vraisemblance généralisée globale calculée pour un partitionnement test de l'imagerie RADAR considérée. On rappelle, que, sous l'hypothèse d'une modélisation de la distribution d'intensité de chaque région caractéristique RADAR i par une loi Gamma à deux paramètres $[\mu_i, L_i]$, classiquement utilisée en imagerie RADAR, le critère $C_{statistique}^{RADAR}(h_{test})$ vérifie l'équation suivante :

$$\begin{aligned} C_{statistique}^{RADAR}(h_{test}) &= - LVG_{globale}^{RADAR}(RC^{RADAR}(h_{test})) \\ &= - \sum_{i \in RC^{RADAR}(h_{test})} LVG_{region}^{RADAR}[\mu_i, L_i] \end{aligned} \quad (8.2)$$

où

$$\begin{aligned} LVG_{region}^{RADAR}[\mu_i, L_i] &= N_i(L_i \log(L_i) - \text{Log}\Gamma(L_i)) \\ &\quad - \frac{L_i}{\mu_i} \sum_{j \in i} x_j - N_i \log(\mu_i) \\ &\quad + (L_i - 1) \left(\sum_{j \in i} \log(x_j) - N_i \log(\mu_i) \right) \end{aligned} \quad (8.3)$$

avec les notations suivantes :

- N_i correspond au nombre de pixels dans la région RADAR i ;
- x_j correspond à l'intensité RADAR du pixel j ;
- $\text{Log}\Gamma$ désigne la fonction LogGamma, définie comme le logarithme de la fonction Gamma Γ ;
- μ_i correspond à la valeur moyenne de la loi Gamma pour la région i ;
- L_i correspond au facteur de forme de la loi Gamma pour la région i ;
- $RC^{RADAR}(h_{b\hat{a}ti})$ désigne l'ensemble des régions caractéristiques qui composent la signature RADAR de bâti et qui dépendent de sa hauteur $h_{b\hat{a}ti}$;
- LVG_{region}^{RADAR} correspond à la log-vraisemblance généralisée d'une région homogène i , qui suit en intensité une loi Gamma $[\mu_i, L_i]$;
- $LVG_{globale}^{RADAR}$ correspond à la log-vraisemblance généralisée globale calculée sur l'imagette RADAR considérée pour un partitionnement donné $RC^{RADAR}(h_{b\hat{a}ti})$.

Pour le calcul de $C_{statistique}^{optique}$, nous pourrions estimer, de façon similaire, la log-vraisemblance généralisée globale calculée pour le partitionnement test de l'imagette optique correspondante. Les différentes régions caractéristiques optiques à considérer seraient la "zone de toit optique", la "zone d'ombre optique", éventuellement la "zone de façade" (dans le cas d'une image optique non acquise avec une visée au nadir) et la "zone de fond". Sous l'hypothèse d'une modélisation de la distribution d'amplitude de chaque région caractéristique optique i (supposée individuellement homogène¹) par une loi Gaussienne à deux paramètres $[m_i, \sigma_i]$, classiquement utilisée en imagerie optique, le critère $C_{statistique}^{optique}(h_{test})$ vérifierait l'équation suivante :

$$\begin{aligned} C_{statistique}^{optique}(h_{test}) &= - LVG_{globale}^{optique}(RC^{optique}(h_{test})) \\ &= - \sum_{i \in RC^{optique}(h_{test})} LVG_{region}^{optique}[m_i, \sigma_i] \end{aligned} \quad (8.4)$$

où

$$LVG_{region}^{optique}[m_i, \sigma_i] = -\frac{1}{2\sigma_i^2} \sum_{j \in i} (a_j - m_i)^2 - N_i \log(\sigma_i) - N_i \log(\sqrt{2\pi}) \quad (8.5)$$

avec les notations suivantes :

- N_i correspond au nombre de pixels dans la région optique i ;
- a_j correspond à l'amplitude optique du pixel j ;
- m_i correspond à la valeur moyenne de la loi Gaussienne pour la région i ;
- σ_i correspond à l'écart-type de la loi Gaussienne pour la région i ;
- $RC^{optique}(h_{b\hat{a}ti})$ désigne l'ensemble des régions caractéristiques qui composent la signature optique de bâti et qui dépendent de sa hauteur $h_{b\hat{a}ti}$;
- $LVG_{region}^{optique}$ correspond à la log-vraisemblance généralisée d'une région homogène i , qui suit en amplitude une loi Gaussienne $[m_i, \sigma_i]$;
- $LVG_{globale}^{optique}$ correspond à la log-vraisemblance généralisée globale calculée sur l'imagette optique considérée pour un partitionnement donné $RC^{optique}(h_{b\hat{a}ti})$.

1. On peut imaginer un filtrage préalable de l'image optique (exemples : filtre Mean-Shift, filtres morphologiques) afin de supprimer les petits objets entraînant des inhomogénéités locales (corniches, cheminées).

Ce critère statistique combiné permettrait de générer des hauteurs candidates plus fiables, plus robustes et plus précises : les minima locaux non appropriés ou non représentatifs seraient réduits (voire supprimés) et le pic minimal, associé à la hauteur réelle, apparaîtrait sans doute plus clairement (facilitant ainsi sa sélection).

La nécessité d'introduire ensuite un second critère (critère de contraste combiné $C_{\text{contraste}}^{\text{combine}}(h_{\text{test}})$) devrait alors être ré-évaluée en fonction des performances obtenues grâce à l'introduction du critère statistique combiné $C_{\text{statistique}}^{\text{combine}}(h_{\text{test}})$, lors de la symétrisation de la troisième étape du traitement.

Validation croisée L'étape de validation de la présence des bâtiments pourrait être améliorée en procédant à une vérification croisée, basée sur un ensemble d'indices complémentaires optiques et RADAR évalués le long de la chaîne.

Une stratégie "d'accumulation simultanée d'évidences" permettrait alors de valider la présence d'un bâti avec plus de certitude et de diminuer ainsi le taux de fausses alarmes.

Il s'agirait de vérifier un ensemble de contraintes relatives aux points suivants :

- une contrainte sur la valeur du critère optimal $C_{\text{statistique}}^{\text{combine}}(h_{\text{test}})$ et éventuellement sur la valeur du critère optimal $C_{\text{contraste}}^{\text{combine}}(h_{\text{test}})$, calculés lors de l'étape 3 de la chaîne, sur les imagerie RADAR et optiques ;
- une contrainte sur la valeur du score optimal de radiométrie moyenne, calculé lors de l'étape 1 de la chaîne, le long du contour optique potentiel à partir de l'image de gradient seuillée ;
- la vérification de la présence de primitives discriminantes extraites (écho double RADAR, ombre RADAR, ombre optique) dans un voisinage proche autour des bâtis potentiels.

8.1.2.3 Pour l'étape 4 : qualification jointe des bâtiments reconstruits

Dans la continuité du processus de symétrisation décrit précédemment, il conviendrait de caractériser le degré de confiance des bâtiments reconstruits, en se fondant sur des scores de qualité combinés, issus des données RADAR et optiques.

Ces scores combinés prendraient en compte notamment des informations relatives à la forme, à la localisation et à la valeur du pic optimal obtenu dans la courbe d'évolution du critère $C_{\text{statistique}}^{\text{combine}}(h_{\text{test}})$ et éventuellement dans la courbe d'évolution du critère $C_{\text{contraste}}^{\text{combine}}(h_{\text{test}})$ (ces critères ayant été utilisés pour valider les bâtis et estimer leur hauteur).

Ce schéma symétrique devrait fournir une qualification plus pertinente des scènes reconstruites.

8.1.2.4 Conclusions et perspectives

Nous avons proposé un procédé de symétrisation, basé sur la combinaison systématique d'indices optiques et RADAR appropriés à l'intérieur de chaque étape de la chaîne de traitement, qui permettrait certainement d'obtenir de meilleurs résultats de reconstruction 3D.

Faute de temps, nous n'avons pu effectuer que des tests préliminaires de cette chaîne symétrique. Il conviendrait de la tester sur différentes zones afin d'évaluer quantitativement l'apport de cette nouvelle version vis-à-vis des résultats issus de la version initiale.

8.1.3 Proposition d'une variante méthodologique pour la détection des bâtiments dans le cadre de la fusion optique-RADAR

8.1.3.1 Présentation

Nous proposons dans cette section, une variante méthodologique pour la détection des bâtiments rectangulaires dans le cadre de la fusion optique-RADAR.

Ces travaux ont fait l'objet d'une collaboration menée lors d'un stage de Master Recherche M2 à Telecom ParisTech et ont été exposés dans le rapport de stage suivant [Bac09], intitulé "Recalage et fusion de données optiques et RADAR métriques : Application à la détection de bâtiments".

De façon similaire à la méthode de détection proposée dans la version symétrisée de la chaîne, il s'agit de développer un procédé utilisant à la fois les données optiques et RADAR présentes en entrée.

Les données d'entrée sont ici composées d'une image optique panchromatique (de résolution submétrique), d'une image optique multi-spectrale (B, G, R, NIR) et d'une image RADAR (de résolution métrique) disponibles sur une même scène.

La variante méthodologique proposée est basée sur un ensemble d'indices radiométriques, spectraux, géométriques et contextuels issus des données images.

Elle pourrait être comparée à la version initiale de l'étape 1 de la chaîne de traitements (ainsi qu'à la version symétrisée de cette étape) et pourrait lui être substituée, si ses performances de détection s'avèrent meilleures (dans ce cas, on obtiendrait de fait une symétrisation de la partie détection).

La stratégie mise en œuvre se décompose en deux phases :

- une phase préliminaire de recalage approximatif entre les deux images, nécessaire à la fusion d'informations issues de l'optique et du RADAR ;
- une phase proprement dite de détection des bâtiments.

8.1.3.2 Stratégie

Phase préliminaire de recalage Pour la phase de recalage, il s'agit ici seulement d'obtenir rapidement un recalage simple et global sur toute l'image, de façon à rendre approximativement superposables l'image optique panchromatique initiale et l'image RADAR projetée et recalée dans le référentiel image optique.

La précision de recalage souhaitée ne relève pas de l'ordre du mètre (comme c'est le cas de l'étape 2 de la chaîne de traitements proposée) mais doit simplement permettre d'associer un ensemble d'indices discriminants à un même bâti grâce à leur proximité spatiale (c'est-à-dire grâce à leur appartenance à une région commune, obtenue par segmentation de l'image optique).

Dans un premier temps, l'image optique panchromatique est projetée dans le référentiel image RADAR :

- en appliquant les fonctions de localisation directes et inverses déduites des modèles de projection disponibles et en utilisant un MNT SRTM disponible sur la zone considérée ;
- en introduisant dans ces fonctions des corrections relatives aux décalages moyens (décalage géodésique et décalage pixelaire) observés pour certains points de référence (il s'agit des 4 coins de l'image, pour lesquels la localisation est donnée).

Dans un second temps, le recalage est affiné en appliquant globalement à tous les pixels de l'image la translation et la similitude optimales qui permettent de maximiser

une mesure de similarité (information mutuelle) calculée sur les deux images.

Phase propre de détection des bâtiments La phase de détection des bâtiments est composée des étapes suivantes :

1. L'extraction de cartes d'indices à partir de l'image RADAR filtrée, de l'image optique panchromatique projetée et recalée dans le référentiel image RADAR, et des bandes rouge et infra-rouge de l'image optique multi-spectrale projetées et recalées dans le référentiel image RADAR.
2. La création de cartes de scores par région après segmentation de l'image optique.
3. La déduction d'un masque binaire de régions labellisées "bâtiments".
4. La génération de bâtiments rectangulaires.

Lors de la première étape, l'image RADAR est débruitée en utilisant le filtre adapté proposé par Deledalle et al. [DDT09]. Des cartes d'indices relatifs aux éléments suivants sont alors extraites dans le référentiel image RADAR :

- présence d'une ombre optique (carte binaire extraite par seuillage radiométrique) ;
- présence d'une ombre RADAR (carte binaire extraite par seuillage radiométrique) ;
- présence d'une zone brillante linéique (écho double) ou surfacique (recouvrement RADAR (carte binaire extraite par seuillage radiométrique) ;
- absence de végétation (carte en niveaux de gris utilisant l'indice NDVI).

Lors de la deuxième étape, ces cartes d'indices sont utilisées pour attribuer un score à chacune des régions issue d'une sur-segmentation par Mean-Shift de l'image optique panchromatique. Dans le cas des zones d'ombre optique, des zones d'ombre RADAR et des zones brillantes, ces cartes sont créées en utilisant les outils développés par [Blo99], qui permettent la génération de cartes de distances orientées suivant des directions privilégiées (ces directions étant a priori connues étant données la position des capteurs et du soleil et l'orientation des bâtiments). La combinaison de ces cartes de distances avec les cartes d'indices permet de générer des cartes de scores pour les différents indices. Ces dernières sont alors fusionnées en une carte de score final à partir d'opérateurs, choisis de façon à obtenir une sensibilité et une spécificité de fusion adéquates. Les deux cartes de scores relatives aux ombres optiques et aux ombres RADAR sont d'abord fusionnées suivant un opérateur "somme symétrique". Le résultat de cette première fusion est ensuite combiné avec la carte de scores relative aux zones brillantes RADAR suivant un opérateur "maximum". Enfin, le résultat de cette seconde fusion est combiné avec la carte de scores relative à l'absence de végétation suivant un opérateur "multiplication" permettant une pondération.

Lors de la troisième étape, le masque binaire de régions labellisées "bâtiments" est obtenu en seuillant l'image de score final et en introduisant une contrainte supplémentaire sur la surface minimale des régions "bâtiments".

Lors de la dernière étape, un algorithme de recherche de rectangles, basé sur la maximisation d'une mesure de consistance géométrique, est mis en œuvre sur ce masque binaire afin de générer les contours des bâtiments rectangulaires extraits.

Notons que cette approche ressemble à celle proposée par Poulain et al. [PIS⁺09], où des descripteurs multiples optiques et RADAR sont fusionnés dans le cadre de la théorie des croyances.

8.1.3.3 Exemple de résultat d'application

La figure 8.3 montre les résultats de détection de bâtiments, obtenus à l'issue de la méthode initiale (en monoscopie optique) et ceux obtenus à l'issue de la variante proposée (en fusion optique-RADAR), sur une scène d'étude Quickbird (résolution : 61 cm) représentant le complexe scientifique de Rangueil, situé à Toulouse (France). Dans le cas de la variante proposée, la détection a été réalisée en combinant les données Quickbird à des données RADAR TerraSAR-X (mode SpotLight, résolution : 1 m) disponibles sur cette même zone.

Les bâtiments rectangulaires détectés sont représentés en rouge et sont superposés sur l'image optique panchromatique.

8.1.3.4 Analyse et conclusions

De légères différences peuvent être observées par comparaison des résultats obtenus avec les deux méthodes. Les résultats issus de la variante méthodologique s'avèrent légèrement plus satisfaisants. Le nombre de fausses alarmes a été légèrement réduit (notamment grâce à la prise en compte des ombres optiques et RADAR) et seuls quelques bâtiments ne sont pas détectés. Ces derniers correspondent essentiellement à des bâtiments non rectangulaires ou à des bâtiments fortement texturés qui ont donné lieu à des segmentations incorrectes.

Il convient toutefois de rappeler que cette comparaison reste à interpréter avec prudence, dans la mesure où le cadre de travail des deux méthodes n'est pas exactement le même (données optiques panchromatiques disponibles pour la méthode initiale et données optiques multi-spectrales disponibles pour la variante proposée).

Les principales limitations de la variante proposée reposent sur l'a priori géométrique (forme supposée rectangulaire) et sur les erreurs de segmentation de l'image optique introduites initialement lors de l'application du Mean-Shift.

Le principal avantage de ce procédé relève de sa capacité à intégrer, en plus des indices radiométriques, géométriques et spectraux classiquement utilisés, de multiples relations contextuelles connues a priori, qui se révèlent être pertinentes et tout à fait complémentaires sur les deux types de données. Celles-ci sont notamment relatives à la position des ombres optiques, des ombres RADAR et des zones brillantes suivant des directions privilégiées que l'on peut prédire.

Ces travaux, similaires à ceux développés par Poulain et al. [PIS⁺09] et par Akçay et al. [AA10], pourraient ainsi être utilisés comme alternative à la version symétrisée de l'étape 1 de la chaîne de traitements, pour la détection des bâtiments par fusion des composantes optiques et RADAR.

8.1.4 Gestion d'une nouvelle configuration d'entrée : cas d'une image optique et d'un couple RADAR interférométrique

Nous proposons dans cette section d'envisager un nouveau scénario opérationnel relatif aux données images présentes en entrée de la chaîne de traitements.

Nous nous plaçons dans le cadre où la configuration spécifique disponible se compose cette fois d'une image optique et d'un couple RADAR interférométrique (InSAR) à résolution métrique.

Il serait alors intéressant d'évaluer l'apport de l'information d'élévation issue d'un couple interférométrique (InSAR) pour la reconstruction 3D du bâtiment dans le cadre



(a) Résultat obtenu avec la méthode initiale



(b) Résultat obtenu avec la variante proposée

FIGURE 8.3 – Cartes de localisation des contours optiques pour la scène d'étude sur Toulouse (France) ©DigitalGlobe. Nous proposons une comparaison visuelle des résultats de détection des bâtiments, obtenus à l'issue de la méthode initiale (en monoscopie optique) et à l'issue de la variante proposée (en fusion optique-RADAR).

de la fusion d'informations.

Il conviendrait de comparer les différentes modalités suivant lesquelles cette information de hauteur complémentaire (probablement très bruitée en milieu dense urbain) pourrait être combinée avec les informations altimétriques et planimétriques issues de la chaîne actuelle.

Une idée consisterait à "contraindre" le calcul des hauteurs interférométrique par les informations planimétriques (forme et localisation des emprises de bâtiments) issues de l'image optique.

Par exemple, sous l'hypothèse de toits plats, nous serions amenés à rechercher localement une seule hauteur interférométrique à l'intérieur d'une emprise de bâti préalablement extraite. De même, sous l'hypothèse de toits bi-pentes, nous serions amenés à rechercher localement un jeu de deux hauteurs interférométriques à l'intérieur d'une emprise extraite.

Des techniques d'optimisation avec régularisation devraient alors probablement être mises en œuvre pour estimer cette hauteur interférométrique : le terme d'attache aux données serait relatif à l'information d'élévation issue du couple InSAR et le terme de régularisation intégrerait les a priori (une seule hauteur ou deux hauteurs à l'intérieur d'une forme optique connue).

Ces hauteurs InSAR régularisées obtenues pourraient ensuite être utilisées de façons différentes :

- soit directement pour fournir une estimation de l'élévation des bâtiments ;
- soit indirectement pour conditionner l'étape de sélection des hauteurs candidates et des hauteurs affinées associées lors de l'étape 3 de la chaîne de traitement actuelle.

8.2 Capitalisation industrielle

8.2.1 Projet de développement d'un démonstrateur 3D

8.2.1.1 Descriptif

L'objectif est de capitaliser l'ensemble des algorithmes développés au cours de la thèse en les intégrant au sein d'un démonstrateur 3D opérationnel pour la reconstruction 3D du bâtiment en milieu urbain et semi-urbain par fusion d'informations optique-RADAR.

Ce démonstrateur, développé dans l'environnement logiciel de l'entreprise Thales Communications France (TCF), permettrait de présenter aux clients les fonctionnalités disponibles dans le cadre de la reconstruction 3D du bâtiment dans le cas de scénarios types, d'analyser leur faisabilité industrielle et d'étudier leurs performances.

Les algorithmes validés pourraient ensuite être intégrés au sein de la plateforme d'exploitation des images développée par TCF (appelée PEX SAIM-SSO).

Ce démonstrateur pourrait s'articuler autour de deux modes :

- l'utilisation de modules indépendants, pour répondre à des besoins ponctuels et précis dans des applications semi-supervisées ;
- l'utilisation de la chaîne de traitements complète, pour réaliser la reconstruction 3D des bâtiments sur de grandes zones.

La mise en œuvre de ce démonstrateur serait également l'occasion de traiter, avant intégration, l'aspect d'évaluation détaillée et de minimisation des temps de calcul, à travers l'optimisation de certains algorithmes.

TABLE 8.1 – Temps de calcul relatif à l'application de la chaîne de reconstruction sur un extrait optique (de taille 256 par 256 pixels).

Nom de l'étape	Temps de calcul
(1) Extraction des emprises potentielles optiques	4 minutes et 16 secondes
(2) Projection et recalage dans la donnée RADAR	5 secondes
(3) Estimation des hauteurs et validation	52 secondes
Total	5 minutes et 13 secondes
Les temps de calcul sont donnés en temps CPU utilisateur, pour un processeur AMD Opteron 248, avec une fréquence de 2.2 GHz.	

À ce jour, les étapes 2 (projection-recalage), 3 (estimation des hauteurs et validation) et 4 (qualification) se sont avérées être suffisamment rapides pour être intégrées dans des logiciels opérationnels. Pour l'étape 1 en revanche, la première phase (construction des profils morphologiques différentiels avec des opérateurs par reconstruction) reste relativement coûteuse lorsqu'elle est appliquée sur de larges zones ; il conviendrait ainsi de réfléchir à des pistes de parallélisation en traitant l'image bloc par bloc.

La table 8.1 détaille les temps de calcul des différentes étapes nécessaires pour traiter un extrait optique typique (de taille 256 par 256 pixels) contenant un seul bâtiment et issu d'une sélection manuelle. Environ 5 minutes sont requises pour effectuer l'ensemble des opérations de reconstruction.

8.2.1.2 Utilisation de modules indépendants

Module de détection des bâtiments Ce module prendrait en entrée une zone d'intérêt sélectionnée par l'opérateur sur l'image optique.

Il appliquerait les algorithmes d'extraction du bâtiment en monoscopie optique décrits dans l'étape 1 de la chaîne, afin de produire en sortie une "carte de localisation des contours dans le référentiel image optique".

Cette dernière pourrait être optionnellement soumise à une "validation manuelle" de l'opérateur par appréciation visuelle des bâtiments détectés.

Si la scène d'entrée sélectionnée était composée de plusieurs bâtis, les deux phases de l'étape d'extraction du bâtiment devraient être mises en œuvre (approche globale puis locale). Si, au contraire, la scène d'entrée sélectionnée ne désignait qu'un seul bâti, l'approche locale seule pourrait être mise en œuvre (puisque la zone d'intérêt sélectionnée serait alors directement considérée comme une "fenêtre d'intérêt").

Dans le cadre d'une symétrisation de la chaîne de traitements, les données images à fournir en entrée correspondraient à une même zone d'intérêt sélectionnée sur un couple optique-RADAR. La donnée de sortie correspondrait alors à une "carte de localisation des contours projetée dans le référentiel d'une des deux images".

Module de recalage fin Ce module prendrait en entrée une image RADAR et la "carte de localisation des contours optiques de bâtis à recalier", sélectionnée sur une zone d'intérêt dans le référentiel image optique.

La "carte de localisation des contours optiques de bâtis à recalier" pourrait être issue d'une des deux sources suivantes, projetées dans le référentiel image optique :

- elle pourrait provenir de données extérieures (base de données ou plans cadastraux) ;

- elle pourrait faire l'objet d'une saisie manuelle par l'opérateur dans l'image optique à disposition.

Ce module appliquerait les algorithmes de projection et de recalage décrits dans l'étape 2 de la chaîne, afin de produire en sortie une "carte de localisation des contours optiques recalés dans le référentiel image RADAR".

Module d'estimation des hauteurs de bâtiments Ce module prendrait en entrée une image RADAR et la "carte des emprises de bâtis dont on cherche à estimer la hauteur".

La "carte des emprises de bâtis dont on cherche à estimer la hauteur" pourrait être issue d'une des deux sources suivantes, projetées dans le référentiel image RADAR :

- elle pourrait provenir de données extérieures (base de données ou plans cadastraux);
- elle pourrait faire l'objet d'une saisie manuelle par l'opérateur dans l'image RADAR à disposition.

Ce module appliquerait les algorithmes d'estimation des hauteurs du bâtiment à partir de critères RADAR décrits dans l'étape 3 de la chaîne, afin de produire en sortie une "carte des hauteurs estimées des bâtiments d'entrée, dans le référentiel image RADAR".

Là encore, dans le cadre d'une symétrisation, les données images à fournir en entrée correspondraient à une même zone d'intérêt sélectionnée sur un couple optique-RADAR. La donnée de sortie correspondrait alors à une "carte des hauteurs estimées des bâtiments d'entrée, projetée dans le référentiel d'une des deux images".

8.2.1.3 Utilisation de la chaîne de traitements complète

Il s'agit alors de mettre en œuvre séquentiellement l'ensemble des étapes 1 à 4 composant la chaîne de traitement développée, sur une zone d'intérêt commune optique-RADAR.

Le résultat généré en sortie correspond alors à un double produit composé d'une vue texturée de la reconstruction 3D de la scène traitée accompagnée d'une "carte de confiance des bâtiments reconstruits".

8.2.2 Scénarios types d'utilisation de la chaîne de traitements

Nous présentons ci-dessous quelques scénarios types, susceptibles de se produire en conditions opérationnelles, et qui pourraient alors nécessiter la mise en œuvre de la chaîne complète de reconstruction.

1. **En situation standard** On suppose qu'un ensemble d'images optiques et RADAR est disponible sur une zone, mais une des situations suivantes se présente :
 - aucun couple optique n'est disponible en configuration stéréoscopique ;
 - le couple optique stéréoscopique présente un recouvrement insuffisant au niveau de la zone d'intérêt ;
 - une des deux images du couple optique stéréoscopique disponible n'est que partiellement exploitable (couvert nuageux).

Nous proposons dans ce cas d'appliquer la chaîne de traitements présentée en utilisant en entrée l'image optique présentant la plus grande zone exploitable et le meilleur recouvrement avec la zone d'intérêt, ainsi que l'image RADAR (ou le couple InSAR) disponible présentant la meilleure résolution, le meilleur recouvrement avec la zone d'intérêt et le moins de perturbations possibles.

2. **En situation de crise** On suppose que seule une unique image optique de la zone critique est disponible avant l'évènement, dont on cherche à évaluer les dommages causés.

La commande d'une image RADAR (ou d'une couple InSAR) est alors lancée sur la zone critique, afin d'obtenir rapidement (faible influence des conditions météorologiques sur l'imagerie RADAR) une image de l'état des lieux après l'évènement. Nous proposons dans ce cas d'appliquer la chaîne de traitements présentée en utilisant en entrée le couple composé de l'image optique pré-évènement (utilisée comme vérité terrain, avec une détection manuelle des bâtis existants avant l'évènement) et de l'image RADAR post-évènement, afin de détecter les changements affectant les constructions urbaines.

Conclusion générale

Une chaîne complète et opérationnelle de traitements semi-automatiques, capable de produire une reconstruction fiable des bâtiments, à partir d'un couple d'images satellitaires optiques-RADAR, a été proposée, dans le cadre de la fusion de données à haute résolution, et appliquée sur des données réelles en milieu urbain et semi-urbain.

Les contributions de ces travaux ont essentiellement été d'ordre méthodologique et applicatif.

Au niveau méthodologique, les éléments théoriques et les procédés algorithmiques, mis en jeu dans chacune des quatre étapes principales de la chaîne et constituant les quatre contributions majeures de cette thèse, ont été successivement décrits au sein de chapitres dédiés.

Premièrement, une nouvelle méthode d'extraction des bâtiments en monoscopie optique a été proposée afin de générer des "cartes de localisation des contours optiques potentiels". Certaines propriétés des profils morphologiques différentiels ont d'abord été exploitées de façon spécifique pour produire un ensemble de "fenêtres d'intérêt", grâce à l'introduction d'un critère d'adéquation géométrique, conditionné par la recherche d'emprises rectangulaires et testé hiérarchiquement sur des régions homogènes. Un critère radiométrique de discontinuité, calculé sur l'image seuillée de gradient optique, a ensuite été optimisé pour extraire localement les contours précis des bâtis potentiels.

Deuxièmement, une méthode originale de projection-recalage des emprises optiques potentielles dans la donnée RADAR a été élaborée afin de générer des "cartes de localisation des contours RADAR recalés". La projection de primitives optiques spécifiques (arêtes sélectionnées sur les contours potentiels) vers l'image RADAR a d'abord été réalisée en utilisant les modèles d'acquisition des capteurs disponibles et une information d'élévation approximative issue du MNT. L'affinage du recalage entre les primitives optiques et l'image RADAR a ensuite été effectué en s'appuyant sur la superposition fine des arêtes optiques avec leurs éléments homologues (échos doubles RADAR). L'optimisation d'un score radiométrique a permis de trouver, indépendamment pour chaque bâti, la translation locale assurant le meilleur appariement. Un procédé de régularisation, basé sur la prise en compte d'informations a priori (contraintes spatiales, radiométriques et contextuelles), a alors permis de sélectionner les translations optimales les plus fiables et de les propager dans un voisinage.

Troisièmement, une approche innovante d'estimation des hauteurs et de validation de leur présence, à partir des emprises potentielles recalées et de l'image RADAR, a été développée afin de générer des "cartes des hauteurs estimées des bâtiments validés". Un premier critère de log-vraisemblance a d'abord été utilisé pour générer un ensemble de hauteurs candidates. Sa formulation analytique a été établie sous l'hypothèse de modélisation des distributions en intensité par des lois Gamma et dans le cadre spécifique de

l'analyse du bâtiment en imagerie RADAR à haute résolution (mise en place d'un partitionnement en différentes régions caractéristiques supposées homogènes). Un second critère de contraste a ensuite été introduit afin de sélectionner la hauteur la plus pertinente. La validité de la méthodologie a alors été démontrée sur des données simulées.

Quatrièmement, une méthode de qualification de la reconstruction a été suggérée afin de générer des "cartes de confiance des bâtiments reconstruits". Deux scénarios d'utilisation de ces cartes ont été brièvement envisagés dans le cadre d'une démarche opérationnelle de type "proposition d'une scène reconstruite" et "soumission à une vérification manuelle des bâtis de faibles scores" et ont montré l'apport concret de tels indicateurs de confiance (diminution des fausses alarmes).

Au niveau applicatif, la chaîne a été mise en œuvre séquentiellement sur deux duos de scènes d'étude Quickbird / TerraSAR-X, utilisées pour illustrer, au fil du manuscrit, les résultats intermédiaires et finaux obtenus sur deux types de scènes (cas d'une scène industrielle en milieu semi-urbain et cas d'une scène résidentielle en milieu urbain).

Nous avons montré que l'étape d'extraction des bâtiments, combinée à celle de validation, fournissait des résultats de détection complets et fiables (taux moyen de détection de 81 % et taux moyen de fausses alarmes de 13 %). Par ailleurs, nous avons prouvé, lors de l'étape d'affinage du recalage, qu'il était possible d'obtenir une superposition des primitives linéaires au sol à l'échelle du mètre. Enfin, nous avons mis en évidence que l'information altimétrique déduite des hauteurs estimées permettait de caractériser précisément l'élévation de chaque bâti (erreur altimétrique de l'ordre du mètre).

Finalement, une vue 3D satisfaisante de la scène reconstruite a donc pu être générée en combinant les informations planimétriques, altimétriques et radiométriques (texture de l'image) extraites le long de la chaîne.

Malgré le nombre important de paramètres requis dans les algorithmes, la plupart d'entre eux se sont avérés relativement stables et faciles à fixer (par apprentissage ou de façon empirique pour un capteur donné et une configuration de bâtiment supposée).

La principale limitation de la chaîne présentée repose sur le modèle simple (parallélépipède) de bâti considéré. Celui-ci pourrait être complexifié grâce à l'introduction d'emprises plus sophistiquées (de forme circulaire ou régulière par juxtaposition de rectangles) et à la prise en compte de toits non plats (toits bi-pentes, toits galbés).

Parmi les nombreuses pistes envisageables, deux types d'extensions devraient être prioritairement considérées.

Une première piste prometteuse, axée sur un procédé de symétrisation de la chaîne de traitements, mériterait d'être étudiée afin d'améliorer les performances du système. D'une part, l'étape de détection des bâtiments en monoscopie optique pourrait être conditionnée par la donnée RADAR, en favorisant l'extraction de contours optiques autour de zones pour lesquelles l'accumulation d'indices RADAR (ombres RADAR, échos doubles RADAR) est importante. D'autre part, l'optimisation d'un critère combiné de log-vraisemblance, conjointement calculé sur l'image optique et sur l'image RADAR, devrait fournir une estimation directe, précise et fiable des hauteurs de bâtis. Il serait alors nécessaire de ré-évaluer la nécessité d'introduire le post-traitement jusqu'ici utilisé pour la sélection et l'affinage des hauteurs (test sur le critère de contraste).

Une seconde piste, relevant de contraintes opérationnelles, serait l'adaptation de la chaîne proposée à la gestion d'une autre configuration des données d'entrée. Cette dernière, plus complète, composée d'un couple RADAR interférométrique et d'une image

optique, apporterait une information d'élévation supplémentaire, susceptible de renforcer la qualité de la reconstruction. Le calcul des hauteurs interférométriques pourrait être contraint localement au sein des emprises optiques détectées et régularisé par la prise en compte d'informations a priori (recherche d'un nombre fixé de hauteurs distinctes à l'intérieur d'une emprise pour un modèle donné de bâtiment).

Par ailleurs, ces travaux devraient prochainement être intégrés au sein d'un démonstrateur 3D, visant à capitaliser au niveau industriel les algorithmes développés et à présenter un ensemble de fonctionnalités, disponibles dans le cadre de la reconstruction du bâtiment à partir d'images satellitaires optiques et RADAR à haute résolution. Deux modes, basés sur l'utilisation de modules indépendants ou sur l'utilisation d'un enchaînement séquentiel complet, pourraient être proposés.

Liste des travaux

Publications dans des revues scientifiques avec comité de relecture

[STDb] **H. Sportouche**, F. Tupin et L. Denise, "Extraction and 3D reconstruction of isolated buildings in urban scenes from high-resolution optical and SAR spaceborne images," *IEEE Transactions on Geoscience and Remote Sensing*, accepté pour publication.

[GNS⁺09] F. Galland, J. M. Nicolas, **H. Sportouche**, M. Roche, F. Tupin et P. Réfrégier, "Unsupervised synthetic aperture radar image partitioning using Fisher distributions," *IEEE Transactions on Geoscience and Remote Sensing*, 2009.

Actes de conférence avec comité de relecture

[STD09b] **H. Sportouche**, F. Tupin et L. Denise, "Building extraction and 3D reconstruction in urban areas from high-resolution optical and SAR imagery," in *Urban Remote Sensing Joint Event, 2009. URBAN'09. Proceedings*, 2009.

[STD09a] **H. Sportouche**, F. Tupin et L. Denise, "Building detection by fusion of optical and SAR features in metric resolution data," in *IEEE International Geoscience and Remote Sensing Symposium, 2009. IGARSS'09. Proceedings*, 2009.

[ST10] **H. Sportouche** et F. Tupin, "A processing chain for simple 3D reconstruction of buildings in urban scenes from high resolution optical and SAR images," in *8th European Conference on Synthetic Aperture Radar, 2010. EUSAR 2010. Proceedings*, 2010.

[STD10] **H. Sportouche**, F. Tupin et L. Denise, "Building detection and height retrieval in urban areas in the framework of high resolution optical and SAR data fusion," in *IEEE International Geoscience and Remote Sensing Symposium, 2010. IGARSS'10. Proceedings*, 2010.

[STDa] **H. Sportouche**, F. Tupin et L. Denise, "A symmetric scheme for 3D building reconstruction from a couple of HR optical and SAR data," in *Urban Remote Sensing Joint Event, 2011. URBAN'11. Proceedings*, draft paper accepté.

Interventions libres dans des ateliers

H. Sportouche, F. Tupin et L. Denise "Extraction and 3D building reconstruction in urban areas from high resolution optical and SAR imagery", *Poster, International Summer School on VHR Remote Sensing*, Grenoble, France, 2008.

H. Sportouche, F. Tupin et L. Denise "A morphological method for automatic building extraction on optical HR / VHR images, in the frame of optical and SAR data fusion", *Présentation orale, Workshop CNES - DLR 2008*, Toulouse, France, 2008.

H. Sportouche, F. Tupin et L. Denise "Building extraction and 3D reconstruction in urban areas from high resolution optical and SAR imagery", *Poster, KTD PCC PhD Day 2009*, Palaiseau, France, 2009.

H. Sportouche, F. Tupin et L. Denise "Extraction et reconstruction 3D du bâtiment en milieu urbain et semi-urbain, à partir de données optiques et radar à résolution métrique", *Présentation orale, Journées JIGOT 2009*, Toulouse, France, 2009.

H. Sportouche, F. Tupin et L. Denise "Extraction et reconstruction 3D simple de bâtiments en milieu urbain à partir de données satellitaires optiques et radar à résolution métrique", *Poster, Journée GDR ISIS 2010*, Paris, France, 2010.

Bibliographie

- [AA10] H. G. AKÇAY et S. AKSOY : Building detection using directional spatial constraints. *In IEEE International Geoscience and Remote Sensing Symposium, 2010. IGARSS'10. Proceedings*, 2010.
- [ABE⁺07] E. ATTEMA, P. BARGELLINI, P. EDWARDS, G. LEVRINI, S. LOKAS, L. MOELLER, B. ROSICH-TELL, P. SECCHI, R. TORRES, M. DAVIDSON et P. SNOEIJ : *Sentinel-1 : The radar mission for GMES - Operational land and sea services*. European Space Agency, Noordwijk, The Netherlands, 2007.
- [ADHB05] E. ATTEMA, M. DAVIDSON, E.A HERLAND et S. BRIGGS : *GMES Sentinel-1 mission requirements document*. European Space Agency, Noordwijk, The Netherlands, 2005. Révision 1.4.
- [Amb05] V. AMBERG : *Analyse de scènes péri-urbaines à partir d'images RADAR haute résolution - Application à l'extraction semi-automatique du réseau routier*. Thèse de doctorat, Institut National Polytechnique de Toulouse, Toulouse, France, 2005.
- [Bac09] M. BACCOUCHE : Recalage et fusion de données optiques et radar métriques : Application à la détection de bâtiments. Mémoire de Master 2 Recherche, Institut Télécom, Télécom ParisTech, Paris, France, 2009.
- [Bal04] T. BALZ : SAR simulation based change detection with high-resolution SAR images in urban environments. *In ISPRS Archives*, volume 35, 2004.
- [BB03a] A. J. BENNETT et D. BLACKNELL : Infrastructure analysis from high resolution SAR and InSAR imagery. *In IEEE GRSS/ISPRS Joint Workshop on Data Fusion And Remote Sensing over Urban Areas, 2003. URBAN'03. Proceedings*, pages 230–235, 2003.
- [BB03b] A. J. BENNETT et D. BLACKNELL : The extraction of building dimensions from high resolution SAR imagery. *In International Radar Conference, 2003. Proceedings*, pages 182–187, 2003.
- [BBFL09] D. BRUNNER, L. BRUZZONE, A. FERRO et G. LEMOINE : Analysis of the reliability of the double bounce scattering mechanism for detecting buildings in VHR SAR images. *In IEEE Radar Conference, 2009. RadarCon09. Proceedings*, 2009.
- [Ben10] C. BENEDEK : Efficient building change detection in sparsely populated areas using coupled marked point processes. *In IEEE International Geoscience and Remote Sensing Symposium, 2010. IGARSS'10. Proceedings*, 2010.
- [BFP09] F. BASELICE, G. FERRAIOLI et V. PASCAZIO : DEM reconstruction in layover areas from SAR and auxiliary input data. *IEEE Geoscience and Remote Sensing Letters*, 6(2):253–257, 2009.
-

- [BH03] T. BALZ et N. HAALA : SAR-based 3D-reconstruction of complex urban environments. In *IAPRS Workshop on 3-D reconstruction from airborne laser scanner and InSAR data, 2003. Proceedings*, volume 34, 2003.
- [BLB08] D. BRUNNER, G. LEMOINE et L. BRUZZONE : Extraction of building heights from VHR SAR imagery using an iterative simulation and match procedure. In *IEEE International Geoscience and Remote Sensing Symposium, 2008. IGARSS'08. Proceedings*, volume 4, 2008.
- [BLB10] D. BRUNNER, G. LEMOINE et L. BRUZZONE : Earthquake damage assessment of buildings using VHR optical and SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 48(5):2403–2420, 2010.
- [BLBG10] D. BRUNNER, G. LEMOINE, L. BRUZZONE et H. GREIDANUS : Building height retrieval from VHR SAR imagery based on an iterative simulation and matching technique. *IEEE Transactions on Geoscience and Remote Sensing*, 48(3):1487–1504, 2010.
- [Blo96] I. BLOCH : Information combination operators for data fusion : A comparative review with classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 26(1):52–67, 1996.
- [Blo99] I. BLOCH : Fuzzy relative position between objects in image processing : a morphological approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(7):657–664, 1999.
- [BM99] C. BAILLARD et H. MAÎTRE : 3D reconstruction of urban scenes from aerial stereo imagery : A focusing strategy. *Computer Vision and Image Understanding*, 76(3):244–258, 1999.
- [Bol00] R. BOLTER : Reconstruction of man-made objects from high resolution SAR images. In *IEEE Aerospace Conference, 2000. Proceedings*, volume 3, pages 287–292, 2000.
- [Bol01] R. BOLTER : *Buildings from SAR : detection and reconstruction of buildings from multiple view high resolution interferometric SAR data*. Thèse de doctorat, Université Technique de Graz, Graz, Autriche, 2001.
- [BPA03] J. A. BENEDIKTSSON, M. PESARESI et K. ARNASON : Classification and feature extraction for remote sensing images from urban areas based on morphological transformations. *IEEE Transactions on Geoscience and Remote Sensing*, 41(9):1940–1949, 2003.
- [BPJ⁺06] L. BOUDET, N. PAPANODITIS, F. JUNG, G. MARTINOTY et M. PIERROT-DESEILLIGNY : A supervised classification approach towards quality self-diagnosis of 3D building models using digital aerial imagery. In *ISPRS Archives*, volume 36, 2006.
- [Bru09] D. BRUNNER : *Advanced methods for building information extraction from very high resolution SAR data to support emergency response*. Thèse de doctorat, Université de Trento, Trento, Italie, 2009.
- [BZ00] C. BAILLARD et A. ZISSERMAN : A plane-sweep strategy for the 3D reconstruction of buildings from multiple images. In *ISPRS Archives*, volume 33, pages 56–62, 2000.
- [CBF06] J. CHANUSSOT, J. A. BENEDIKTSSON et M. FAUVEL : Classification of remote sensing images from urban areas using a fuzzy possibilistic model. *IEEE Geoscience and Remote Sensing Letters*, 3(1):40–44, 2006.
-

-
- [CDJS02] N. CHEHATA, M. P. DESSEILLIGNY, F. JUNG et G. STAMON : Extraction of 3D primitives from stereopairs of satellite images for automatic reconstruction of buildings. In *IAPR Conference. Workshop on Machine Vision Applications, 2002. Proceedings*, pages 636–639, 2002.
- [Cel06] F. CELLIER : *Reconstruction 3D de bâtiments en interférométrie RSO haute résolution : approche par gestion d'hypothèses*. Thèse de doctorat, École Nationale Supérieure des Télécommunications, Paris, France, 2006.
- [Che05] N. CHEHATA : *Modélisation 3D de scènes urbaines à partir d'images satellitaires à très haute résolution*. Thèse de doctorat, Université René Descartes, Paris, France, 2005.
- [CJPDS03] N. CHEHATA, F. JUNG, M. PIERROT-DESEILLIGNY et G. STAMON : A region-based matching approach for 3D-roof reconstruction from high-resolution satellite stereo pairs. In *Digital Image Computing Techniques and Applications, 2003. DICTA'03. Proceedings*, volume 2, pages 889–898, 2003.
- [CON06] F. CELLIER, H. ORIOT et J. M. NICOLAS : Hypothesis management for building reconstruction from high resolution InSAR imagery. In *IEEE International Conference on Geoscience and Remote Sensing Symposium, 2006. IGARSS'06. Proceedings*, pages 3639–3642, 2006.
- [DCMG⁺07] J. P. DA COSTA, F. MICHELET, C. GERMAIN, O. LAVIALLE et G. GRENIER : Delineation of vine parcels by segmentation of high resolution remote sensed images. *Precision Agriculture*, 8(1-2):95–110, 2007.
- [DDT09] C. A. DELEDALLE, L. DENIS et F. TUPIN : Iterative weighted maximum likelihood denoising with probabilistic patch-based weights. *IEEE Transactions on Image Processing*, 18(12):2661–2672, 2009.
- [Den09] L. DENISE : SAR image exploitation : introduction to the SAR image functionalities. Rapport technique, Thales Communications, Massy, France, 2009.
- [Dig09] DigitalGlobe. *Quickbird imagery products : Product guide*, 2009. Révision 5.0.
- [Dig09a] DigitalGlobe, Longmont, CO, USA. *DigitalGlobe constellation : WorldView-1 imaging satellite*, 209. Révision 12/09.
- [Dig09b] DigitalGlobe, Longmont, CO, USA. *DigitalGlobe constellation : WorldView-2 imaging satellite*, 209. Révision 07/09.
- [DJ01] O. DE JOINVILLE : *Évaluation de la qualité d'une cartographie urbaine à l'aide d'images aériennes à haute résolution*. Thèse de doctorat, École Nationale Supérieure des Télécommunications, Paris, France, 2001.
- [DJHS08] E. DAI, Y. Q. JIN, T. HAMASAKI et M. SATO : Three-dimensional stereo reconstruction of buildings using polarimetric SAR images acquired in opposite directions. *IEEE Geoscience and Remote Sensing Letters*, 5(2):236–240, 2008.
- [DMBBB07] M. DALLA MURA, J. A. BENEDIKTSSON, L. BRUZZONE et F. BOVOLO : A novel technique based on morphological filters for change detection in optical remote sensing images. In *AIP Conference. 6th International Workshop On Information Optics, 2007. WIO'07. Proceedings*, volume 949, pages 75–82, 2007.
-

-
- [EAB⁺09] M. EINEDER, N. ADAM, R. BAMLER, N. YAGUE-MARTINEZ et H. BREIT : Spaceborne spotlight SAR interferometry with TerraSAR-X. *IEEE Transactions on Geoscience and Remote Sensing*, 47(5):1524–1535, 2009.
- [EFM⁺09] M. EINEDER, T. FRITZ, J. MITTERMAYER, A. ROTH, E. BÖRNER et H. BREIT : *TerraSAR-X ground segment : basic product specification document*. DLR, Germany, 2009. Révision 1.6.
- [FBB09] A. FERRO, D. BRUNNER et L. BRUZZONE : An advanced technique for building detection in VHR SAR images. In *SPIE'09. Proceedings*, volume 7477, 2009.
- [FBG02] C. S. FRASER, E. BALTSAVIAS et A. GRUEN : Processing of Ikonos imagery for submetre 3D positioning and building extraction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 56:177–194, 2002.
- [FGI⁺07] G. FRANCESCHETTI, R. GUIDA, A. IODICE, D. RICCIO, G. RUELLO et U. STILLA : Building feature extraction via a deterministic approach : application to real high resolution SAR images. In *IEEE International Geoscience and Remote Sensing Symposium, 2007. IGARSS'07. Proceedings*, pages 2681–2684, 2007.
- [FIR02] G. FRANCESCHETTI, A. IODICE et D. RICCIO : A canonical problem in electromagnetic backscattering from buildings. *IEEE Transactions on Geoscience and Remote Sensing*, 40(8):1787–1801, 2002.
- [FIRR03] G. FRANCESCHETTI, A. IODICE, D. RICCIO et G. RUELLO : SAR raw signal simulation for urban structures. *IEEE Transactions on Geoscience and Remote Sensing*, 41(9 Part 1):1986–1995, 2003.
- [FMJ03] D. FLAMMANC, G. MAILLET et H. JIBRINI : 3D city models : an operational approach using aerial images and cadastral maps. In *ISPRS Archives*, volume 34, pages 53–58, 2003.
- [Foe94] W. FOERSTNER : A framework for low level feature extraction. In *3rd European Conference on Computer Vision, 1994. Proceedings*, 1994.
- [GA10] K. GOEL et N. ADAM : A bayesian method for very high resolution multi aspect angle radargrammetry. In *8th European Conference on Synthetic Aperture Radar, 2010. EUSAR 2010. Proceedings*, 2010.
- [Geo10] GeoEye, Dulles, VA, USA. *GeoEye 1 Fact Sheet*, 2010. Révision 01/10.
- [GFFRP05] S. GUILLASO, L. FERRO-FAMIL, A. REIGBER et E. POTTIER : Building characterization using L-band polarimetric interferometric SAR data. *IEEE Geoscience and Remote Sensing Letters*, 2(3):347–351, 2005.
- [GHS00] P. GAMBA, B. HOUSHMAND et M. SACCANI : Detection and extraction of buildings from interferometric SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 38(1):611–617, 2000.
- [GNS⁺09] F. GALLAND, J. M. NICOLAS, H. SPORTOUCHE, M. ROCHE, F. TUPIN et P. RÉFRÉGIÉRIER : Unsupervised synthetic aperture radar image partitioning using Fisher distributions. *IEEE Transactions on Geoscience and Remote Sensing*, 47(8):2966–2972, 2009.
- [GR01] O. GERMAIN et P. RÉFRÉGIÉRIER : Statistical active grid for segmentation refinement. *Pattern Recognition Letters*, 22(10), 2001.
-

-
- [HBC⁺08] H. HAMMER, T. BALZ, E. CADARIO, U. SOERGEL, U. THOENNESSEN et U. STILLA : Comparison of SAR simulation concepts for the analysis of high-resolution SAR data. *In 7th European Conference on Synthetic Aperture Radar, 2008. EUSAR 2008. Proceedings, 2008.*
- [Her00] E. HERVET : *Appariement d'images radar à synthèse d'ouverture en conditions radargrammétriques*. Thèse de doctorat, Institut National Polytechnique de Toulouse, Toulouse, France, 2000.
- [Hoe99] K. B. HOEPFNER : Recovery of building structure from IFSAR-derived elevation maps. Rapport technique 99-16 01003-4610, Computer science department, Lederle graduate research center, University of Massachusetts, Massachusetts, Amherst, USA, 1999.
- [HS05] T. D. HONG et R. A. SCHOWENGERDT : A robust technique for precise registration of radar and optical satellite images. *Photogrammetric Engineering and Remote Sensing*, 71(5):585–593, 2005.
- [IG04] J. INGLADA et A. GIROS : On the possibility of automatic multi-sensor image registration. *IEEE Transactions on Geoscience and Remote Sensing*, 42(10):2104–2120, 2004.
- [Ima07] ImageSat International, Tel Aviv, Israël. *Eros imagery products guide*, 2007.
- [Ita07] Italian Space Agency, Italy. *Cosmo-SkyMed Mission : Cosmo-SkyMed system description and user guide*, 2007.
- [JCHM06] Z. JUNJIE, D. CHIBIAO, Y. HONGJIAN et X. MINGHONG : 3D reconstruction of building based on high-resolution SAR and optical images. *In IEEE International Geoscience and Remote Sensing Symposium, 2006. IGARSS'06. Proceedings*, pages 3794–3797, 2006.
- [JD05] X. JIN et C. H. DAVIS : Automated building extraction from high-resolution satellite imagery in urban areas using structural, contextual, and spectral information. *EURASIP Journal on Applied Signal Processing*, 14:2196–2206, 2005.
- [Jib02] H. JIBRINI : *Reconstruction automatique des bâtiments en modèles polyédriques 3D à partir de données cadastrales vectorisées 2D et d'un couple d'images aériennes à haute résolution*. Thèse de doctorat, École Nationale Supérieure des Télécommunications, Paris, France, 2002.
- [KE01] M. KASSER et Y. EGELS : *Photogrammétrie numérique*. Hermès, Paris, France, 2001.
- [KHN01] Z. KIM, A. HUERTAS et R. NEVATIA : A model-based approach for multi-view complex building description. *In ASCONA Conference, 2001. Proceedings*, 2001.
- [KP08] K. KARANTZALOS et N. PARAGIOS : Automatic model-based building detection from single panchromatic high resolution images. *In ISPRS Archives*, 2008.
- [KST06] D. KOC SAN et M. TURKER : Automatic building detection and delineation from high resolution space images using model-based approach. *In Proceedings of the ISPRS Workshop on Topographic Mapping from Space*, 2006.
-

-
- [LDZPD06a] F. LAFARGE, X. DESCOMBES, J. ZERUBIA et M. PIERROT-DESEILLIGNY : A structural approach for 3D building reconstruction. Rapport technique 6048, INRIA, Sophia Antipolis, France, 2006.
- [LDZPD06b] F. LAFARGE, X. DESCOMBES, J. ZERUBIA et M. PIERROT-DESEILLIGNY : An automatic building reconstruction method : a structural approach using high resolution satellite images. In *IEEE International Conference on Image Processing, 2006. ICIP'06. Proceedings*, 2006.
- [Leh10] G. LEHUREAU : *Fusion de données optiques et radar à haute résolution en milieu urbain*. Thèse de doctorat, Institut Télécom, Télécom ParisTech, Paris, France, 2010.
- [Lem07] V. LEMOIGNE : Reconstruction de bâtiments par utilisation de contours actifs sur images SAR et optiques haute résolution. Rapport technique, Institut Télécom, Télécom ParisTech, Paris, France, 2007.
- [LGNT06] V. LEMOIGNE, F. GALLAND, J. M. NICOLAS et F. TUPIN : Statistical polygonal snakes with Fisher distributions. In *6th European Conference on Synthetic Aperture Radar, 2006. EUSAR 2006. Proceedings*, 2006.
- [LN98] C. A. LIN et R. NEVATIA : Building detection and description from a single intensity image. *Computer Vision and Image Understanding*, 72(2):101–121, 1998.
- [LT07] V. LEMOIGNE et F. TUPIN : Statistical polygonal snakes for 3D building reconstruction using high resolution SAR data. In *Urban Remote Sensing Joint Event, 2007. URBAN'07. Proceedings*, 2007.
- [LTN90] A. LOPÈS, R. TOUZI et E. NEZRY : Adaptive speckle filters and scene heterogeneity. *IEEE Transactions on Geoscience and Remote Sensing*, 28(6):992–1000, 1990.
- [LTT⁺08] G. LEHUREAU, F. TUPIN, C. TISON, G. OLLER et D. PETIT : Registration of metric resolution SAR and optical images in urban areas. In *7th European Conference on Synthetic Aperture Radar, 2008. EUSAR 2008. Proceedings*, 2008.
- [LW07] S. LEFÈVRE et J. WEBER : Automatic building extraction in VHR images using advanced morphological operators. In *Urban Remote Sensing Joint Event, 2007. URBAN'07. Proceedings*, 2007.
- [Maî01] H. MAÎTRE : *Traitement des images de RSO*. Hermès, Paris, France, 2001.
- [MBM01] E. M. MIKHAIL, J. S. BETHEL et J. C. MCGLONE : *Introduction to modern photogrammetry*. John Wiley and Sons, Inc., New York, USA, 2001.
- [MDA10] MDA Corporation, Canada. *RadarSat-2 : Product Details*, 2010. Révision 08/10.
- [Mer10] G. MERCIER : Progressive change detection in time series of SAR images. In *IEEE International Geoscience and Remote Sensing Symposium, 2010. IGARSS'10. Proceedings*, 2010.
- [MS08] D. MASSONNET et J. C. SOUYRIS : *Imaging with synthetic aperture radar*. EPFL Press, Lausanne, Switzerland, 2008.
- [MS10] G. MOSER et S. B. SERPICO : Unsupervised change detection with very high-resolution SAR images by multiscale analysis and Markov random fields. In *IEEE International Geoscience and Remote Sensing Symposium, 2010. IGARSS'10. Proceedings*, 2010.
-

-
- [MSS02] E. MICHAELSEN, U. SOERTEL et U. STILLA : Grouping salient scatterers in InSAR data for recognition of industrial buildings. *In 16th International Conference on Pattern Recognition, 2002. ICPR'02. Proceedings*, volume 2, pages 613–616, 2002.
- [MZ05] S. MÜLLER et D. W. ZAUM : Robust building detection in aerial images. *In IAPRS*, volume 36, pages 143–148, 2005.
- [Nic06] J. M. NICOLAS : Application de la transformée de Mellin : étude des lois statistiques de l'imagerie cohérente. Rapport technique 2006D010, Institut Télécom, Télécom ParisTech, Paris, France, 2006.
- [Noc96] L. NOCERA : *Reconstruction du relief en imagerie radar à ouverture de synthèse : stéréoscopie et ombrages*. Thèse de doctorat, Université de Paris 7, Paris, France, 1996.
- [ODZ07a] M. ORTNER, X. DESCOMBES et J. ZERUBIA : A marked point process of rectangles and segments for automatic analysis of digital elevation models. *IEEE Transactions on Pattern Analysis and machine intelligence*, 2007.
- [ODZ07b] M. ORTNER, X. DESCOMBES et J. ZERUBIA : Building outline extraction from digital elevation models using marked point processes. *International Journal of Computer Vision*, 72(2):107–132, 2007.
- [OHB09] OHB-System AG, Bremen, Germany. *SAR-Lupe : The innovative program for satellite-based radar reconnaissance*, 2009.
- [Orb06] Orbital Sciences Corporation, Dulles, VA, USA. *Fact sheet Orbview-3 : Delivering high-resolution images of the Earth*, 2006.
- [Ori03] H. ORIOT : Statistical snakes for building extraction from stereoscopic aerial images. *In ISPRS Archives*, volume 34, pages 65–70, 2003.
- [PAD01] D. PETIT, F. ADRAGNA et J. D. DUROU : The filtering of layover area in high resolution IFSAR for the building extraction. *In SAR Image Analysis, Modeling and Techniques III*, 2001.
- [Pap98] N. PAPANODITIS : *Reconstruction tridimensionnelle de paysages urbains en imagerie satellitale haute résolution*. Thèse de doctorat, Université de Nice Sophia-Antipolis, Nice, France, 1998.
- [PB01] M. PESARESI et J. A. BENEDIKTSSON : A new approach for the morphological segmentation of high-resolution satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 39(2):309–320, 2001.
- [PDR09] G. PONS, L. DENISE et P. RÉFRÉGIER : Hierarchical feature-based classification approach for fast and user-interactive SAR image interpretation. *IEEE Geoscience and Remote Sensing Letters*, 6(1):117–121, 2009.
- [PIS08] V. POULAIN, J. INGLADA et M. SPIGAI : High resolution remote sensing image analysis with exogeneous data : A generic framework. *In IEEE International Geoscience and Remote Sensing Symposium, 2008. IGARSS'08. Proceedings*, volume 2, pages 1025–1028, 2008.
- [PIS⁺09] V. POULAIN, J. INGLADA, M. SPIGAI, J. Y. TOURNERET et P. MARTON : Fusion of high resolution optical and SAR images with vector data bases for change detection. *In IEEE International Geoscience and Remote Sensing Symposium, 2009. IGARSS'09. Proceedings*, volume 4, pages 956–959, 2009.
-

-
- [QD04] M. QUARTULLI et M. DATCU : Stochastic geometrical modeling for built-up area understanding from a single SAR intensity image with meter resolution. *IEEE Transactions on Geoscience and Remote Sensing*, 42(9):1996–2003, 2004.
- [Rap10] RapidEye AG, Brandenburg an der Havel, Germany. *RapidEye standard image product specifications*, 2010. Révision 3.0.
- [Réf04] P. RÉFRÉGIER : *Noise theory and application to physics : From fluctuations to information*. Springer, New York, USA, 2004.
- [RJ98] J. A. RICHARDS et X. JIA : *Remote sensing digital image analysis : An introduction*. Springer, Germany, 1998.
- [SD03] A. K. SHACKELFORD et C. H. DAVIS : Fully automated road network extraction from high-resolution satellite multispectral imagery. In *IEEE International Geoscience and Remote Sensing Symposium, 2003. IGARSS'03. Proceedings*, volume 1, 2003.
- [SDW04] A. K. SHACKELFORD, C. H. DAVIS et X. WANG : Automated 2-D building footprint extraction from high-resolution satellite multispectral imagery. In *IEEE International Geoscience and Remote Sensing Symposium, 2004. IGARSS'04. Proceedings*, volume 3, 2004.
- [Sim02] E. SIMONETTO : *Extraction 3D de structures industrielles sur des images RAMSES haute résolution par radargrammétrie*. Thèse de doctorat, Université de Rennes 1, Rennes, France, 2002.
- [Soe06] U. SOERGEL : Radargrammetric extraction of building features from high resolution multi-aspect SAR data. In *IEEE International Geoscience and Remote Sensing Symposium, 2006. IGARSS'06. Proceedings*, pages 3635–3638, 2006.
- [SOG05] E. SIMONETTO, H. ORIOT et R. GARELLO : Rectangular building extraction from stereoscopic airborne radar images. *IEEE Transactions on Geoscience and Remote Sensing*, 43(10):2386–2395, 2005.
- [SOGLC03] E. SIMONETTO, H. ORIOT, R. GARELLO et J. LE CAILLEC : Radargrammetric processing for 3-D building extraction from high-resolution airborne SAR data. In *IEEE International Geoscience and Remote Sensing Symposium, 2003. IGARSS'03. Proceedings*, volume 3, pages 2002–2004, 2003.
- [Spo03] Spot Image, France. *Spot satellite technical data*, 2003.
- [Spo09a] Spot Image, France. *Formosat 2 : Capacité de revisite quotidienne et haute résolution*, 2009.
- [Spo09b] Spot Image, France. *Kompsat 2 : La solution 1 mètre*, 2009.
- [Spo09c] Spot Image, France. *Pléiades 1 et 2 : La constellation des produits très haute résolution*, 2009.
- [Spo09d] Spot Image, France. *Spot 6 and 7 : High resolution imagery and large area coverage*, 2009.
- [SR10] S. SURI et P. REINARTZ : Mutual-information-based registration of TerraSAR-X and Ikonos imagery in urban areas. *IEEE Transactions on Geoscience and Remote Sensing*, 48(2):939–949, 2010.
-

-
- [SS98] V. K. SHETTIGARA et G. M. SUMMERLING : Height determination of extended objects using shadows in SPOT images. *Photogrammetric Engineering and Remote Sensing*, 64(1):35–44, 1998.
- [SST03] U. STILLA, U. SOERGEL et U. THOENNESSEN : Potential and limits of InSAR data for building reconstruction in built-up areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(1-2):113–123, 2003.
- [ST10] H. SPORTOUCHE et F. TUPIN : A processing chain for simple 3D reconstruction of buildings in urban scenes from high resolution optical and SAR images. In *8th European Conference on Synthetic Aperture Radar, 2010. EUSAR 2010. Proceedings*, 2010.
- [STC07] A. SHABOU, F. TUPIN et F. CHAABANE : Similarity measures between SAR and optic data. In *IEEE International Geoscience and Remote Sensing Symposium, 2007. IGARSS'07. Proceedings*, 2007.
- [STDa] H. SPORTOUCHE, F. TUPIN et L. DENISE : A symmetric scheme for 3D building reconstruction from a couple of HR optical and SAR data. In *Urban Remote Sensing Joint Event, 2011. URBAN'11. Proceedings*. Draft paper accepté.
- [STDb] H. SPORTOUCHE, F. TUPIN et L. DENISE : Extraction and 3D reconstruction of isolated buildings in urban scenes from high-resolution optical and SAR spaceborne images. *IEEE Transactions on Geoscience and Remote Sensing*. Accepted pour publication.
- [STD09a] H. SPORTOUCHE, F. TUPIN et L. DENISE : Building detection by fusion of optical and SAR features in metric resolution data. In *IEEE International Geoscience and Remote Sensing Symposium, 2009. IGARSS'09. Proceedings*, volume 4, pages 769–772, 2009.
- [STD09b] H. SPORTOUCHE, F. TUPIN et L. DENISE : Building extraction and 3D reconstruction in urban areas from high-resolution optical and SAR imagery. In *Urban Remote Sensing Joint Event, 2009. URBAN'09. Proceedings*, 2009.
- [STD10] H. SPORTOUCHE, F. TUPIN et L. DENISE : Building detection and height retrieval in urban areas in the framework of high resolution optical and SAR data fusion. In *IEEE International Geoscience and Remote Sensing Symposium, 2010. IGARSS'10. Proceedings*, 2010.
- [STGT07] U. SOERGEL, A. THIELE, H. GROSS et U. THOENNESSEN : Extraction of bridge features from high-resolution InSAR data and optical images. In *Urban Remote Sensing Joint Event, 2007. URBAN'07. Proceedings*, 2007.
- [STS03] U. SOERGEL, U. THOENNESSEN et U. STILLA : Iterative building reconstruction from multi-aspect InSAR data. In *International Archives of Photogrammetry and Remote Sensing*, volume 34, pages 186–192, 2003.
- [SW03] H. SCHUSTER et U. WEIDNER : A new approach towards quantitative quality evaluation of 3D building models. In *ISPRS Archives*, 2003.
- [Tai04] F. TAILLANDIER : *Reconstruction du bâti en milieu urbain : une approche multi-vues*. Thèse de doctorat, École Polytechnique, Palaiseau, France, 2004.
- [Tai05] F. TAILLANDIER : Automatic building reconstruction from cadastral maps and aerial images. In *IAPRS*, volume 36, pages 105–110, 2005.
-

- [TCS⁺07a] A. THIELE, E. CADARIO, K. SCHULZ, U. THOENNESSEN et U. SOERTEL : Building recognition from multi-aspect high-resolution InSAR data in urban areas. *IEEE Transactions on Geoscience and Remote Sensing*, 45(11):3583–3593, 2007.
- [TCS⁺07b] A. THIELE, E. CADARIO, K. SCHULZ, U. THOENNESSEN et U. SOERTEL : Feature extraction of gable-roofed buildings from multi-aspect high-resolution InSAR data. In *IEEE International Conference on Geoscience and Remote Sensing Symposium, 2007. IGARSS'07. Proceedings*, pages 262–265, 2007.
- [TD07] F. TUPIN et L. DENIS : Recalage et extraction d'informations 3D en optique / radar dans un cadre interférométrique. Rapport technique R-S06/SI/OT 04-010, Institut Télécom, Télécom ParisTech, Paris, France, 2007. Révision 1.
- [THD02] F. TUPIN, B. HOUSHMAND et M. DATCU : Road detection in dense urban areas using SAR imagery and the usefulness of multiple views. *IEEE Transactions on Geoscience and Remote Sensing*, 40(11):2405–2414, 2002.
- [Tis04] C. TISON : *Interférométrie RSO à haute résolution en milieu urbain : application au calcul de MNS urbain*. Thèse de doctorat, École Nationale Supérieure des Télécommunications, Paris, France, 2004.
- [TLB88] R. TOUZI, A. LOPÈS et P. BOUSQUET : A statistical and geometrical edge detector for SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 26(6):764–773, 1988.
- [TMM⁺98] F. TUPIN, H. MAÎTRE, J. F. MANGIN, J. M. NICOLAS et E. PECHERSKY : Detection of linear features in SAR images : application to road network extraction. *IEEE Transactions on Geoscience and Remote Sensing*, 36(2):434–453, 1998.
- [TN02] F. TUPIN et J. M. NICOLAS : Matching criteria for radargrammetry. In *IEEE International Geoscience and Remote Sensing Symposium, 2002. IGARSS'02. Proceedings*, volume 5, pages 2608–2610, 2002.
- [TP94] I. TANNOUS et B. PIKEROEN : Parametric modeling of spaceborne SAR image geometry. Application : SEASAT/SPOT image registration. *Photogrammetric engineering and remote sensing*, 60(6):755–766, 1994.
- [TPCI09] J. Y. TOURNERET, V. POULAIN, M. CHABERT et J. INGLADA : Similarity measures between vector data bases and optical images for change detection. In *IEEE International Geoscience and Remote Sensing Symposium, 2009. IGARSS'09. Proceedings*, volume 2, pages 992–995, 2009.
- [TR03] F. TUPIN et M. ROUX : Detection of building outlines based on the fusion of SAR and optical features. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(1-2):71–82, 2003.
- [TR04] F. TUPIN et M. ROUX : 3D information extraction by structural matching of SAR and optical features. In *ISPRS Archives*, volume 35, 2004.
- [TR05] F. TUPIN et M. ROUX : Markov random field on region adjacency graph for the fusion of SAR and optical data in radargrammetric applications. *IEEE Transactions on Geoscience and Remote Sensing*, 43(8):1920–1928, 2005.
- [TTM04] C. TISON, F. TUPIN et H. MAÎTRE : Retrieval of building shapes from shadows in high resolution SAR interferometric images. In *IEEE International*
-

-
- Geoscience and Remote Sensing Symposium, 2004. IGARSS'04. Proceedings*, volume 3, 2004.
- [TTM07] C. TISON, F. TUPIN et H. MAÎTRE : A fusion scheme for joint retrieval of urban height map and classification from high-resolution interferometric SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 45(2):496–505, 2007.
- [Tup97] F. TUPIN : *Reconnaissance des formes et analyse de scènes en imagerie RSO*. Thèse de doctorat, École Nationale Supérieure des Télécommunications, Paris, France, 1997.
- [Tup03] F. TUPIN : Extraction of 3D information using overlay detection on SAR images. In *IEEE GRSS/ISPRS Joint Workshop on Data Fusion And Remote Sensing over Urban Areas, 2003. URBAN'03. Proceedings*, pages 72–76, 2003.
- [Tup04] F. TUPIN : Merging of SAR and optical features for 3D reconstruction in a radargrammetric framework. In *IEEE International Geoscience and Remote Sensing Symposium, 2004. IGARSS'04. Proceedings*, volume 1, 2004.
- [VC02] S. VINSON et L. D. COHEN : Extraction des bâtiments complexes à partir d'images aériennes et de MNE. In *Actes de la Conférence Reconnaissance des Formes et Intelligence Artificielle 2002. RFIA'02.*, pages 125–134, 2002.
- [Wal07] S. WALKER : Airborne digital imagers. In *ASPRS*, 2007.
- [WE10] J. D. WEYDAHL et K. ELDHUSET : Extracting building heights from TSX stereo SAR acquisitions. In *8th European Conference on Synthetic Aperture Radar, 2010. EUSAR 2010. Proceedings*, 2010.
- [WL06] Y. WANG et H. LIU : Semi automatic extraction of building information and variation detection from high resolution remote sensing images. In *SPIE'06. Proceedings*, volume 6419, 2006.
- [WS08] J. D. WEGNER et U. SOERGEL : Bridge height estimation from combined high-resolution optical and SAR imagery. In *ISPRS Archives*, volume 37, pages 1071–1076, 2008.
- [WST09] J. D. WEGNER, U. SOERGEL et A. THIELE : Building extraction in urban scenes from high resolution InSAR data and optical imagery. In *Urban Remote Sensing Joint Event, 2009. URBAN'09. Proceedings*, 2009.
- [WZGZ04] L. WU, Y. ZHANG, Y. GAO et Y. ZHANG : Tree crown detection and delineation in high resolution RS image : a texture approach discussion. In *IEEE International Geoscience and Remote Sensing Symposium, 2004. IGARSS'04. Proceedings*, volume 6, pages 3841–3844, 2004.
- [WZS04] Y. WEI, Z. ZHAO et J. SONG : Urban building extraction from high-resolution satellite panchromatic image using clustering and edge detection. In *IEEE International Geoscience and Remote Sensing Symposium, 2004. IGARSS'04. Proceedings*, volume 3, pages 2008–2010, 2004.
- [WZS10] J. D. WEGNER, J. R. ZIEHN et U. SOERGEL : Building detection and height estimation from high-resolution InSAR and optical data. In *IEEE International Geoscience and Remote Sensing Symposium, 2010. IGARSS'10. Proceedings*, 2010.
-

- [XJ07] F. XU et Y. Q. JIN : Automatic reconstruction of building objects from multi-aspect meter-resolution SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 45(7):2336–2353, 2007.
- [XP98] C. XU et J. L. PRINCE : Snakes, shapes, and gradient vector flow. *IEEE Transactions on Image Processing*, 7(3):359–369, 1998.
-