



HAL
open science

Approches radiométriques et géométriques pour l'extraction de bâtiments en imagerie satellitaire optique et RSO haute résolution

Edouard Barthelet

► To cite this version:

Edouard Barthelet. Approches radiométriques et géométriques pour l'extraction de bâtiments en imagerie satellitaire optique et RSO haute résolution. Traitement du signal et de l'image [eess.SP]. Télécom Bretagne, Université de Rennes 1, 2013. Français. NNT: . tel-00951191

HAL Id: tel-00951191

<https://theses.hal.science/tel-00951191>

Submitted on 24 Feb 2014

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

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

Sous le sceau de l'Université européenne de Bretagne

Télécom Bretagne

En habilitation conjointe avec l'Université de Rennes 1

Ecole Doctorale – MATISSE

APPROCHES RADIOMÉTRIQUES ET GÉOMÉTRIQUES POUR L'EXTRACTION DE BÂTIMENTS EN IMAGERIE SATELLITAIRE OPTIQUE ET RSO HAUTE RÉOLUTION

Thèse de Doctorat

Mention : Très Honorable

Présentée par **Édouard Barthelet**

Département : Image et Traitement de l'Information

Laboratoire : Lab-STICC - Pôle : CID

Directeur de thèse : Grégoire Mercier

Soutenue le 5 juin 2013

Jury :

M. Nicolas Paparoditis - Directeur de Recherche, IGN, Saint-Mandé (Président)
M. Jordi Inglada - Ingénieur, CNES, Toulouse (Rapporteur)
Mme Florence Tupin - Professeur, Telecom ParisTech, Paris (Rapporteur)
M. Grégoire Mercier - Professeur, Telecom Bretagne, Brest (Directeur de thèse)
M. Léonard Denise - Ingénieur, Thales Communications & Security, Vélizy (Examineur)
M. Philippe Réfrégier - Professeur, École Centrale Marseille, Marseille (Examineur)
M. Sébastien Reynaud - Ingénieur, ONERA, Palaiseau (Examineur)

Résumé

Les travaux réalisés dans le cadre de cette thèse sont relatifs à l'extraction tridimensionnelle de bâtiments à partir d'images satellitaires optiques et RSO haute et très haute résolution.

Une étude des effets géométriques liés à la projection des bâtiments dans les images optiques et RSO est tout d'abord menée. Dans le cadre de cette étude, la signature géométrique d'un bâtiment dans une image est définie. Elle correspond à l'ensemble des structures géométriques de l'image caractéristiques de la présence du bâtiment et peut être décrite, de manière équivalente, par un ensemble de régions ou de segments dans l'image. À l'issue de cette étude, des méthodes sont développées pour simuler la signature géométrique des bâtiments dans les images.

Deux méthodes, qui s'appuient sur les méthodes de simulation précédemment établies, sont ensuite proposées pour l'extraction des bâtiments en monoscopie optique et RSO. Ces méthodes, dont les performances sont évaluées sur données simulées et sur données réelles, répondent aux exigences suivantes :

- prise en compte de la complexité géométrique et radiométrique inhérente aux images de télédétection haute et très haute résolution en milieu urbain ;
- applicabilité des approches développées aussi bien en imagerie optique que RSO, afin de faciliter leur mise en œuvre dans un contexte multicapteur.

La première méthode, qui repose sur une approche radiométrique, réalise une segmentation locale de l'image contrainte par un modèle de bâtiment. Cette méthode s'appuie sur deux critères qui quantifient respectivement l'homogénéité radiométrique des régions constitutives de la signature géométrique d'un bâtiment et l'adéquation des segments de cette signature avec les bords présents dans l'image. Une stratégie consistant à maximiser successivement ces deux critères est alors mise en œuvre afin de déterminer le bâtiment dont la signature géométrique présente une adéquation optimale avec l'image.

La seconde méthode, qui repose sur une approche géométrique, réalise une inversion de modèle à partir d'images binaires résultant de l'extraction des primitives caractéristiques de la présence des bâtiments dans les images optiques et RSO. La densité de probabilité de présence des pixels détectés dans ces images binaires est tout d'abord calculée pour une prise de vue et un modèle de bâtiment donnés. La détection des bâtiments et l'estimation des paramètres du modèle, adopté pour décrire ces bâtiments, sont ensuite réalisées au moyen de tests et d'estimateurs statistiques qui s'appuient sur le calcul de la vraisemblance des pixels effectivement détectés dans les images binaires.

Abstract

This thesis deals with the extraction of 3D buildings from optical and SAR high resolution spaceborne remote sensing images.

The geometric effects, resulting from the projection of 3D buildings in optical and SAR images, are first analyzed. Within this study, the geometric signature of a building in an image is defined. It refers to all geometric features of the image resulting from the presence of the building and can equivalently be described by a set of regions or by a set of segments in the image. From this analysis, methods for the simulation of building geometric signatures in both optical and SAR images are then designed.

Afterwards, two methods, based on the previous building geometric signature simulation methods, are proposed for the extraction of 3D buildings in optical and SAR monoscopic imagery. These methods, which performances are assessed on both simulated and real data, are compliant with the two following requirements:

- the geometric and radiometric complexity of high and very high resolution remote sensing images in urban environment has to be taken into account;
- the proposed approaches can be implemented in optical as well as in SAR imagery, so that they can easily be used in a multi sensor context.

The first method relies on a radiometric approach which performs a local segmentation of the image, while being constrained by a building model. Two criteria, which quantify the match between an image and the geometric signature of a building in this image, are defined. The first one measures the radiometric homogeneity of building geometric signature regions whereas the second one quantifies the match between building geometric signature segments and image edges. A strategy, based on the consecutive maximization of these criteria is then implemented in order to get the building, which geometric signature matches at best with the image.

The second method relies on a geometric approach which performs a model inversion from binary images, resulting from the extraction of building specific primitives in optical and SAR images. The probability density function of location of pixels detected in these binary images is computed *a priori* for a given image acquisition geometry and for a given building model. The detection of buildings as well as the estimation of building model parameters are then performed by means of statistical tests and estimators, which involve the computation of binary image detected pixel likelihood.

Remerciements

Alors que j'écris (non sans soulagement !) les dernières lignes de ce manuscrit, je repense avec gratitude à l'ensemble des personnes, aux qualités humaines et scientifiques remarquables, que j'ai eu la chance de rencontrer au cours de ces trois années de doctorat. Je tiens ici à les remercier pour le concours qu'elles ont apporté à cette thèse.

J'adresse tout d'abord mes plus sincères remerciements à l'ensemble des membres de mon jury et en premier lieu à M. Nicolas Paparoditis, directeur du laboratoire MATIS de l'IGN, qui l'a présidé.

Je remercie également Mme Florence Tupin, Professeur à Telecom ParisTech, et M. Jordi Inglada, Ingénieur au CNES, qui ont accepté de rapporter ces travaux de thèse. Je les remercie pour l'intérêt qu'ils ont porté à ces travaux ainsi que pour leur lecture critique du manuscrit et les remarques constructives qu'ils ont formulées.

Je remercie tout particulièrement M. Philippe Réfrégier, Professeur à l'École Centrale Marseille et examinateur au sein de ce jury, qui m'a enseigné avec passion et rigueur le traitement du signal et des images lors de ma formation à l'École Centrale Marseille.

J'exprime ma profonde gratitude à M. Grégoire Mercier, Professeur à Telecom Bretagne qui a dirigé ces travaux, ainsi qu'à M. Léonard Denise, Ingénieur chez Thales Communications & Security qui a assuré l'encadrement industriel de cette thèse. Je les remercie chaleureusement pour la qualité de leur enseignement, pour leur écoute bienveillante ainsi que pour l'enthousiasme qu'ils ont manifesté vis-à-vis de mes travaux à l'occasion des nombreux échanges techniques que nous avons pu avoir. Je leur suis également reconnaissant pour les encouragements qu'ils m'ont témoignés ainsi que pour la confiance qu'ils m'ont accordée dans la conduite de mes travaux.

Mes sincères remerciements vont également à M. Sébastien Reynaud, Ingénieur d'études à l'ONERA, qui a coencadré mes travaux et en a assuré le suivi quotidien avec un enthousiasme inébranlable et une passion débordante. Je le remercie tout particulièrement pour le soutien scientifique et moral qu'il a su m'apporter, pour les nombreuses idées qui ont émergé de nos échanges et enfin pour sa participation active au cheminement de ce manuscrit.

Je suis également reconnaissant envers mes anciens collègues de Thales Communications & Security pour les excellents souvenirs que je conserverai de ces trois années passées avec eux.

Je remercie tout particulièrement Christian pour m'avoir, en quelques semaines seulement, ouvert les yeux sur la précarité de ma condition, et pour tous les efforts qu'il a par la suite déployés afin de m'en extraire : idées business toutes plus foireuses les unes que les autres, placements juteux sur le Forex sans oublier les multiples idées de reconversion allant de la plomberie jusqu'au notariat. Je le remercie surtout d'avoir fait de ces trois années un moment inoubliable!

Je remercie bien évidemment Laurent, qui a su, par ses conseils avisés, me guider à travers les méandres de la thèse. Je le remercie pour les bouts de canard que nous avons partagé à l'occasion de nos innombrables trajets en train, en bus ou encore en Spérymobile! Je le remercie enfin pour

ses éclaircissements sur cette science obscure qu'est la géographie et pour la relecture minutieuse et critique qu'il a faite du présent manuscrit.

Je n'oublie pas non plus M. R, insatiable mangeur de biscottes et amateur de vieilles croutes, qui a pleinement contribué à l'ambiance mémorable et toujours très classe de ces trois années. Je ne lui tiendrai pas rigueur de la jalousie maladive dont il a fait preuve après mon expédition canadienne sur l'île aux ours. J'aurais, en revanche, apprécié un peu plus de reconnaissance pour l'assurance multirisques que je lui ai souscrite auprès de Medicare...

Je tiens également à remercier très sincèrement Hélène, sans qui je n'aurais probablement pas postulé pour cette thèse et dont les brillants travaux ont été sources d'inspiration. Je la remercie pour sa sympathie et ses conseils éclairés qui m'ont permis d'affronter la difficile première année de thèse.

Je suis également très reconnaissant envers Yannick et François pour l'intérêt qu'ils ont continuellement porté à mes travaux, pour leurs conseils algorithmiques et informatiques toujours très pertinents, sans oublier les interminables séances de debug avec ddd et valgrind pour lesquelles ils se sont toujours montrés très motivés. Je remercie plus particulièrement Yannick pour ses leçons d'humilité et félicite François pour la résistance remarquable dont il a su faire preuve face aux deux affreux.

Pour finir, je remercie Amandine, Sébastien, Damien, Christophe, Roger et tous les membres du département Image Intelligence de Thales Communications & Security pour leur sympathie et leur soutien au cours de ces trois années. Une mention toute particulière à Véro pour sa gentillesse ainsi que pour l'organisation toujours impeccable de mes missions à l'occasion des nombreuses conférences auxquelles j'ai participé.

Je remercie bien évidemment mes collègues thésards de Telecom Bretagne, et tout particulièrement Sarab, Pauline et Luca, pour l'accueil chaleureux qu'ils m'ont réservé à chacune de mes expéditions en terre bretonne et pour les moments que nous avons partagés sur le campus, dans les nombreuses crêperies de Brest ou encore sur l'île d'Ouessant.

Enfin, je remercie mes proches, famille et amis, pour l'infinie compréhension dont ils ont su faire preuve à mon égard ainsi que pour leur soutien indéfectible durant ces trois années.

Table des matières

Résumé	i
Abstract	iii
Remerciements	v
Table des matières	ix
Liste des figures	xv
Liste des tableaux	xvii
Liste des acronymes	xix
Introduction	1
1 Introduction à l'imagerie de télédétection optique et RSO	5
1.1 Imagerie optique	6
1.1.1 Généralités	6
1.1.2 Formation des images	6
1.1.3 Caractéristiques des images	9
1.2 Imagerie RADAR à synthèse d'ouverture	13
1.2.1 Généralités	13
1.2.2 Formation des images	13
1.2.3 Caractéristiques des images	19
1.3 Conclusion	21
2 Signature géométrique des bâtiments en imagerie de télédétection optique et RSO	23
2.1 Introduction	24
2.2 Modélisation des bâtiments	25
2.3 Rendu tridimensionnel des bâtiments en imagerie optique	26
2.3.1 Projections entre l'espace terrain et l'espace image	26
2.3.2 Effets géométriques liés à la projection des bâtiments	29
2.3.3 Simulation de la signature géométrique d'un bâtiment	34

2.3.4	Exemples de signatures géométriques simulées	37
2.4	Rendu tridimensionnel des bâtiments en imagerie RSO	39
2.4.1	Projections entre l'espace terrain et l'espace image	39
2.4.2	Effets géométriques liés à la projection des bâtiments	41
2.4.3	Simulation de la signature géométrique d'un bâtiment	45
2.4.4	Exemples de signatures géométriques simulées	50
2.5	Conclusion	51
3	État de l'art sur l'extraction de bâtiments en imagerie de télédétection optique et RSO	53
3.1	Introduction	54
3.1.1	Problématique de l'extraction de bâtiments	54
3.1.2	Classification des méthodes d'extraction de bâtiments	55
3.2	Extraction des bâtiments en imagerie optique	56
3.2.1	Extraction de primitives	56
3.2.2	Extraction des toits des bâtiments	57
3.2.3	Extraction tridimensionnelle des bâtiments	59
3.3	Extraction des bâtiments en imagerie RSO	61
3.3.1	Extraction de primitives	61
3.3.2	Extraction des emprises au sol des bâtiments	64
3.3.3	Extraction tridimensionnelle des bâtiments	65
3.4	Extraction de bâtiments à partir d'images optiques et RSO	69
3.4.1	Méthodes exploitant la complémentarité d'information entre images optiques et RSO	70
3.4.2	Méthodes exploitant la redondance d'information entre images optiques et RSO	70
3.5	Conclusion	72
4	Approches radiométriques pour l'extraction de bâtiments	75
4.1	Introduction	76
4.2	Adéquation radiométrique des régions de la signature géométrique d'un bâtiment avec une image	77
4.2.1	Travaux antérieurs	77
4.2.2	Expression du critère pour une distribution statistique quelconque	78
4.2.3	Mise en œuvre du critère en imagerie optique	79
4.2.4	Mise en œuvre du critère en imagerie RSO	83
4.3	Adéquation radiométrique des segments de la signature géométrique d'un bâtiment avec une image	88
4.3.1	Mise en œuvre du critère en imagerie optique	88
4.3.2	Mise en œuvre du critère en imagerie RSO	91
4.4	Estimation des paramètres d'un bâtiment par mesure de l'adéquation radiométrique de sa signature géométrique avec une image	95
4.4.1	Estimation des paramètres en imagerie optique	96
4.4.2	Estimation des paramètres en imagerie RSO	101

4.5	Conclusion	103
5	Approches géométriques pour l'extraction de bâtiments	107
5.1	Introduction	108
5.2	Extraction de primitives	110
5.2.1	Imagerie optique	110
5.2.2	Imagerie RSO	112
5.3	Détection grossière des bâtiments	112
5.3.1	Calcul des densités de probabilité de présence des pixels détectés et non détectés dans les images binaires de primitives en présence d'un bâtiment	113
5.3.2	Test de significativité pour la détection des bâtiments	122
5.3.3	Détection grossière à partir des primitives considérées conjointement	129
5.4	Estimation des paramètres d'un bâtiment	135
5.4.1	Positionnement du problème	136
5.4.2	Estimation par maximisation de vraisemblance	137
5.4.3	Stratégie d'estimation	137
5.5	Détection fine des bâtiments	152
5.5.1	Calcul de la vraisemblance généralisée des pixels des images binaires de primitives	153
5.5.2	Stratégie de détection des bâtiments à partir de l'image de vraisemblance généralisée	154
5.5.3	Résultats en imagerie optique	154
5.5.4	Résultats en imagerie RSO	155
5.6	Conclusion	158
	Conclusion	163
	Liste des travaux	167
	Bibliographie	174

Liste des figures

1.1	Géométrie d'un capteur optique <i>wiskbroom</i>	7
1.2	Géométrie d'un capteur optique <i>pushbroom</i>	8
1.3	Effets de distorsion observés sur les images acquises par les capteurs <i>pushbroom</i>	8
1.4	Géométrie d'un capteur optique <i>frame</i>	9
1.5	Relation entre l'IFOV et la taille d'un pixel projeté au sol.	10
1.6	Définition IEEE des bandes spectrales utilisées par les systèmes RADAR.	13
1.7	Géométrie d'un capteur RADAR à synthèse d'ouverture.	14
1.8	Données brutes et image obtenue à l'issue des traitements en distance et en azimut	14
1.9	Modes d'acquisition des capteurs RADAR à synthèse d'ouverture.	19
1.10	Géométrie RADAR et géométrie terrain.	20
2.1	Définition des paramètres du modèle parallélépipédique de bâtiment.	25
2.2	Passage du repère terrain \mathcal{R}_T au repère capteur \mathcal{R}_C par application successive de trois rotations.	26
2.3	Principe du lancer de rayon itératif sur un modèle numérique de terrain.	29
2.4	Définition des repères \mathcal{R}_T et \mathcal{R}_C et représentation des angles ψ , θ et ϕ permettant d'orienter le repère \mathcal{R}_C relativement au repère \mathcal{R}_T	30
2.5	Rendu tridimensionnel d'un bâtiment en imagerie optique.	31
2.6	Effet de déplacement sur une structure verticale lors de sa projection de l'espace terrain dans l'espace image.	32
2.7	Définition de la longueur de l'ombre au sol et de la longueur de l'ombre projetée dans l'espace image.	33
2.8	Simulation de la signature géométrique d'un bâtiment parallélépipédique en imagerie optique.	35
2.9	Simulation de la signature géométrique de bâtiments dans une image QuickBird.	38
2.10	Projection de l'espace image vers l'espace terrain.	40
2.11	Effets de contraction-dilatation en imagerie RSO.	41
2.12	Effets de repliement en imagerie RSO.	42
2.13	Effets d'ombrage en imagerie RSO.	44
2.14	Réflexions multiples sur des structures diédriques et triédriques.	45
2.15	Variation de la largeur apparente d'un bâtiment parallélépipédique vu par un capteur RSO.	46

2.16	Simulation de la signature géométrique d'un bâtiment parallélépipédique en imagerie RSO.	48
2.17	Simulation de la signature géométrique de bâtiments dans une image TerraSAR-X.	50
4.1	Images optiques réelle et simulée à partir desquelles sont calculées les coupes du critère C_I^r	82
4.2	Coupes du critère C_I^r calculées à partir des images optiques de la figure 4.1.	84
4.3	Somme cohérente des amplitudes complexes réfléchies par six réflecteurs élémentaires.	85
4.4	Images RSO réelle et simulée à partir desquelles sont calculées les coupes du critère C_I^r	88
4.5	Coupes du critère C_I^r calculées à partir des images RSO de la figure 4.4.	89
4.6	Définition des régions impliquées dans le calcul du critère d'adéquation des segments de la signature géométrique avec l'image.	90
4.7	Coupes du critère C_I^s calculées à partir des images optiques de la figure 4.1.	92
4.8	Coupes du critère C_I^s calculées à partir des images RSO de la figure 4.4.	94
4.9	Signatures géométriques du bâtiment G dans l'image RSO simulée de la figure 4.4 (b) pour différentes hauteurs.	95
4.10	Évolution de la moyenne quadratique des erreurs finales, observées à l'issue d'une maximisation unidimensionnelle des critères C_I^r et C_I^s en imagerie optique, en fonction de l'erreur initiale.	97
4.11	Résultats de la stratégie d'estimation sur données simulées en imagerie optique : signatures géométriques des neufs bâtiments de la scène d'étude simulées à partir des valeurs estimées et des valeurs vraies des paramètres de ces bâtiments.	99
4.12	Résultats de la stratégie d'estimation sur données réelles en imagerie optique : signatures géométriques des neufs bâtiments de la scène d'étude simulées à partir des valeurs estimées et des valeurs vraies des paramètres de ces bâtiments.	100
4.13	Évolution de la moyenne quadratique des erreurs finales, observées à l'issue d'une maximisation unidimensionnelle des critères C_I^r et C_I^s en imagerie RSO, en fonction de l'erreur initiale.	102
4.14	Résultats de la stratégie d'estimation sur données simulées en imagerie RSO : signatures géométriques des neufs bâtiments de la scène d'étude simulées à partir des valeurs estimées et des valeurs vraies des paramètres de ces bâtiments.	104
4.15	Résultats de la stratégie d'estimation sur données réelles en imagerie RSO : signatures géométriques des neufs bâtiments de la scène d'étude simulées à partir des valeurs estimées et des valeurs vraies des paramètres de ces bâtiments.	105
5.1	Schéma synoptique de la méthode d'extraction proposée.	109
5.2	Image QuickBird sur la région de Marseille, France.	110
5.3	Extraction de primitives en imagerie optique.	111
5.4	Image TerraSAR-X sur la région de Marseille, France.	112
5.5	Extraction de primitives en imagerie RSO.	113
5.6	Densités de probabilité de présence des pixels détectés dans les images binaires en imagerie optique.	115
5.7	Densités de probabilité de présence des pixels détectés dans les images binaires en imagerie RSO.	116
5.8	Densités de probabilité de présence des pixels non détectés dans les images binaires en imagerie optique.	116

5.9	Densités de probabilité de présence des pixels non détectés dans les images binaires en imagerie RSO.	116
5.10	Calcul de la densité de probabilité de présence des pixels détectés dans l'image binaire résultant de l'extraction des contours de toit en imagerie optique par somme de Riemann et par méthode de Monte-Carlo.	120
5.11	Densités de probabilité de présence des pixels détectés et non détectés dans les images binaires de primitives en imagerie optique.	121
5.12	Densités de probabilité de présence des pixels détectés et non détectés dans les images binaires de primitives en imagerie RSO.	121
5.13	Implémentation du calcul des densités de probabilité de présence des pixels détectés et non détectés dans les images binaires de primitives.	122
5.14	Image binaire B_I^0 simulée et carte de log-vraisemblance ℓ^0 associée.	124
5.15	Image binaire B_I^0 réelle et carte de log-vraisemblance ℓ^0 associée.	124
5.16	Image binaire B_I^1 simulée et carte de log-vraisemblance ℓ^1 associée.	124
5.17	Image binaire B_I^1 réelle et carte de log-vraisemblance ℓ^1 associée.	124
5.18	Cartes binaires de détection D_I^0 établies à partir d'images binaires de primitives B_I^0 simulée et réelle.	125
5.19	Cartes binaires de détection D_I^1 établies à partir d'images binaires de primitives B_I^1 simulée et réelle.	125
5.20	Cartes binaires de détection D_I résultant du produit des cartes binaires de détection D_I^0 et D_I^1	125
5.21	Image binaire B_I^0 simulée et carte de log-vraisemblance ℓ^0 associée.	126
5.22	Image binaire B_I^0 réelle et carte de log-vraisemblance ℓ^0 associée.	126
5.23	Image binaire B_I^1 simulée et carte de log-vraisemblance ℓ^1 associée.	127
5.24	Image binaire B_I^1 réelle et carte de log-vraisemblance ℓ^1 associée.	127
5.25	Cartes binaires de détection D_I^0 établies à partir d'images binaires de primitives B_I^0 simulée et réelle.	128
5.26	Cartes binaires de détection D_I^1 établies à partir d'images binaires de primitives B_I^1 simulée et réelle.	128
5.27	Cartes binaires de détection D_I résultant du produit des cartes binaires de détection D_I^0 et D_I^1	128
5.28	Cartes de log-vraisemblance et cartes binaires de détection établies en imagerie optique en considérant conjointement les images binaires de primitives.	135
5.29	Cartes de log-vraisemblance et cartes binaires de détection établies en imagerie optique en considérant conjointement les images binaires de primitives.	135
5.30	Cartes de détection et cartes binaires de détection établies en imagerie RSO en considérant conjointement les images binaires de primitives.	136
5.31	Cartes de détection et cartes binaires de détection établies en imagerie RSO en considérant conjointement les images binaires de primitives.	136
5.32	Images binaires simulées du bâtiment G de la scène d'étude.	138
5.33	Images binaires réelles du bâtiment G de la scène d'étude.	138
5.34	Densités de probabilité de présence conditionnelles pour l'estimation du paramètre w_1 en imagerie optique.	139
5.35	Densités de probabilité de présence conditionnelles pour l'estimation du paramètre w_2 en imagerie optique.	139

5.36	Courbes de log-vraisemblance calculées à partir des images binaires B_I^0 simulée et réelle des figures 5.32 et 5.33 pour l'estimation des paramètres w_1 et w_2	140
5.37	Image binaire B_I^0 simulée pour un bâtiment dont l'estimation des paramètres w_1 et w_2 est sujette à un phénomène d'ambiguïté.	141
5.38	Courbes de log-vraisemblance calculées à partir de l'image binaire B_I^0 simulée de la figure 5.37 pour l'estimation des paramètres w_1 et w_2	141
5.39	Ambiguïté entre les paramètres w_1 et w_2 pour des bâtiments dont l'orientation α est proche de 0 degré ou 90 degrés.	141
5.40	Densités de probabilité de présence conditionnelles pour l'estimation du paramètre w_1 sous l'hypothèse \mathcal{H}_α^1) en imagerie optique.	142
5.41	Densités de probabilité de présence conditionnelles pour l'estimation du paramètre w_1 sous l'hypothèse \mathcal{H}_α^2) en imagerie optique.	142
5.42	Densités de probabilité de présence conditionnelles pour l'estimation du paramètre w_2 sous l'hypothèse \mathcal{H}_α^1) en imagerie optique.	142
5.43	Densités de probabilité de présence conditionnelles pour l'estimation du paramètre w_2 sous l'hypothèse \mathcal{H}_α^2) en imagerie optique.	142
5.44	Courbes de log-vraisemblance calculées à partir de l'image binaire B_I^0 simulée de la figure 5.37 pour l'estimation des paramètres w_1 et w_2	143
5.45	Densités de probabilité de présence conditionnelles pour l'estimation du paramètre α en imagerie optique.	144
5.46	Courbes de log-vraisemblance calculées à partir des images binaires simulées et réelles des figures 5.32 et 5.33 pour l'estimation des paramètres α et h	144
5.47	Densités de probabilité de présence conditionnelles des pixels détectés pour l'estimation du paramètre h en imagerie optique.	145
5.48	Images binaires simulées du bâtiment G de la scène d'étude.	145
5.49	Images binaires réelles du bâtiment G de la scène d'étude.	145
5.50	Discontinuité de la forme de l'écho double d'un bâtiment lorsque son orientation est parallèle à la trace au sol du porteur.	146
5.51	Densités de probabilité de présence conditionnelles pour l'estimation du paramètre w_1 sous l'hypothèse \mathcal{H}_α^1) en imagerie RSO.	147
5.52	Densités de probabilité de présence conditionnelles pour l'estimation du paramètre w_1 sous l'hypothèse \mathcal{H}_α^2) en imagerie RSO.	147
5.53	Densités de probabilité de présence conditionnelles pour l'estimation du paramètre w_2 sous l'hypothèse \mathcal{H}_α^1) en imagerie RSO.	147
5.54	Densités de probabilité de présence conditionnelles pour l'estimation du paramètre w_2 sous l'hypothèse \mathcal{H}_α^2) en imagerie RSO.	147
5.55	Courbes de log-vraisemblance calculées à partir des images binaires B_I^0 simulée et réelle des figures 5.48 et 5.49 pour l'estimation des paramètres w_1 et w_2	148
5.56	Densités de probabilité de présence conditionnelles pour l'estimation du paramètre α en imagerie RSO.	149
5.57	Courbes de log-vraisemblance pour l'estimation des paramètres α et h en imagerie RSO.	149
5.58	Densités de probabilité de présence conditionnelles des pixels détectés pour l'estimation du paramètre h en imagerie RSO.	150
5.59	Densités de probabilité de présence conditionnelles des pixels non détectés pour l'estimation du paramètre h en imagerie RSO.	150
5.60	Courbes de log-vraisemblance calculées à partir des pixels détectés et non détectés dans une image binaire B_I^1 pour l'estimation du paramètre h en imagerie RSO.	150

5.61 Régions impliquées dans le calcul de la log-vraisemblance pour l'estimation de la hauteur d'un bâtiment en imagerie RSO.	151
5.62 Cartes de log-vraisemblance généralisée établies à partir d'images binaires B_I^0 et B_I^1 simulées et réelles.	155
5.63 Résultats de la méthode d'extraction sur données simulées en imagerie optique : simulation des signatures géométriques des bâtiments extraits dans les images binaires de primitives.	156
5.64 Résultats de la méthode d'extraction sur données simulées en imagerie optique : simulation des signatures géométriques des bâtiments extraits dans les images binaires de primitives.	157
5.65 Cartes de log-vraisemblance généralisée établies à partir d'images binaires B_I^0 et B_I^1 simulées et réelles.	158
5.66 Résultats de la méthode d'extraction sur données simulées en imagerie RSO : simulation des signatures géométriques des bâtiments extraits dans les images binaires de primitives.	159
5.67 Résultats de la méthode d'extraction sur données simulées en imagerie RSO : simulation des signatures géométriques des bâtiments extraits dans les images binaires de primitives.	160

Liste des tableaux

2.1	Paramètres des bâtiments de la scène étudiée pour un modèle parallélépipédique. . . .	38
2.2	Incertitude sur la mesure des paramètres d'un bâtiment par saisie manuelle des sommets visibles de ce bâtiment dans l'image QuickBird de la figure 2.9.	38
4.1	Paramètres des lois normales mises en œuvre pour la simulation de l'image optique de la figure 4.1 (b).	82
4.2	Paramètres des lois de Rayleigh mises en œuvre pour la simulation de l'image RSO de la figure 4.4 (b).	87
4.3	Résultats de la stratégie d'estimation sur données simulées en imagerie optique : valeurs estimées et valeurs vraies des paramètres des neuf bâtiments de la scène d'étude. . . .	99
4.4	Résultats de la stratégie d'estimation sur données réelles en imagerie optique : valeurs estimées et valeurs vraies des paramètres des neuf bâtiments de la scène d'étude. . . .	100
4.5	Résultats de la stratégie d'estimation sur données simulées en imagerie RSO : valeurs estimées et valeurs vraies des paramètres des neuf bâtiments de la scène d'étude. . . .	104
4.6	Résultats de la stratégie d'estimation sur données réelles en imagerie RSO : valeurs estimées et valeurs vraies des paramètres des neuf bâtiments de la scène d'étude. . . .	105
5.1	Résultats de la méthode d'extraction sur données simulées en imagerie optique : valeurs estimées et valeurs vraies des paramètres des neuf bâtiments de la scène d'étude. . . .	156
5.2	Résultats de la méthode d'extraction sur données réelles en imagerie optique : valeurs estimées et valeurs vraies des paramètres des neuf bâtiments de la scène d'étude. . . .	157
5.3	Résultats de la méthode d'extraction sur données simulées en imagerie RSO : valeurs estimées et valeurs vraies des paramètres des neuf bâtiments de la scène d'étude. . . .	159
5.4	Résultats de la méthode d'extraction sur données réelles en imagerie RSO : valeurs estimées et valeurs vraies des paramètres des neuf bâtiments de la scène d'étude. . . .	160

Liste des acronymes

AVIRIS	<i>Airborne Visible InfraRed Imaging Spectrometer</i>
CCD	<i>Charge-Couple Devices</i>
CMOS	<i>Complementary Metal Oxide Semi-conductor</i>
ECEF	<i>Earth-Centered Earth-Fixed</i>
FOV	<i>Field Of View</i>
GPS	<i>Global Positioning System</i>
GSD	<i>Ground Sample Distance</i>
HRV	Haute Résolution Visible
HSV	<i>Hue Saturation Value</i>
IFOV	<i>Instantaneous Field Of View</i>
LIDAR	<i>Light Detection And Ranging</i>
MNS	Modèle Numérique de Surface
MNT	Modèle Numérique de Terrain
MSS	<i>Multi Spectral Scanner</i>
NASA	<i>National Aeronautics and Space Administration</i>
NDVI	<i>Normalized Difference Vegetation Index</i>
RADAR	<i>RAdio Detection And Ranging</i>
RPC	<i>Rational Polynomial Coefficients</i>
RSB	Rapport Signal sur Bruit
RSO	RADAR à Synthèse d'Ouverture
SER	Surface Equivalente RADAR
SIG	Système d'Information Géographique
SLC	<i>Single Look Complex</i>
SPOT	Système Probatoire d'Observation de la Terre
SVM	<i>Support Vector Machine</i>
TM	<i>Thematic Mapper</i>
TSV	Teinte Saturation Valeur
UTM	<i>Universal Transverse Mercator</i>
WGS84	<i>World Geodetic System 1984</i>

Introduction

Analyse de scènes urbaines en imagerie satellitaire haute résolution

Les satellites imageurs lancés au cours de la dernière décennie, comme QuickBird et Pléiades-HR en imagerie optique ou encore TerraSAR-X et COSMO-SkyMed en imagerie RSO¹, fournissent des images dont la résolution spatiale est inférieure au mètre. Ces images, dont le contenu informationnel a été considérablement accru en comparaison des images basse, moyenne et haute résolution², offrent de nouvelles possibilités d'exploitation mais font également émerger de nouvelles problématiques.

Alors que l'exploitation des images basse et moyenne résolution est principalement limitée à la détermination de la couverture et de l'occupation des sols, la disponibilité d'images haute et très haute résolution permet d'envisager de nouvelles applications, telles que l'analyse de scènes urbaines, cadre général au sein duquel s'inscrivent les travaux présentés.

L'analyse de scènes urbaines à partir d'images de télédétection consiste à extraire les éléments caractéristiques de l'environnement urbain, tels que les réseaux routiers et ferroviaires, les bâtiments ou encore les véhicules. Cette analyse constitue une problématique difficile en raison de la complexité intrinsèque de ces scènes et de la variabilité de leur rendu au sein des images de télédétection.

La complexité intrinsèque d'une scène urbaine est liée, d'une part, à la variabilité de forme et de dimension des différents éléments constitutifs de cette scène et, d'autre part, à la densité et à la disposition relative de ces éléments. Ainsi, la modélisation des différents éléments de la scène, étape préalable à l'extraction de ces éléments, peut s'avérer délicate et constitue une première difficulté.

La variabilité du rendu des scènes urbaines dans les images de télédétection s'explique par les conditions changeantes dans lesquelles les images sont acquises ainsi que par la nature distincte des capteurs dont elles sont issues. De cette variabilité, il résulte qu'un même objet (et notamment les objets du sur-sol) présente des signatures radiométriques et géométriques différentes dans les images de télédétection, lorsque ces images n'ont pas été acquises dans les mêmes conditions et/ou par le même capteur.

Dans ce contexte, les méthodes développées pour l'exploitation des images de télédétection basse et moyenne résolution, qui s'appuient essentiellement sur la radiométrie des images, ne sont plus adaptées aux problématiques de l'analyse de scènes urbaines. En imagerie satellitaire très haute résolution, la modélisation du contenu de la scène, la prise en compte de la géométrie de prise de vue des images ainsi que des effets géométriques liés à la projection des objets du sur-sol dans ces images sont indispensables pour résoudre les problèmes liés à l'analyse de scènes urbaines.

¹RADAR à Synthèse d'Ouverture

²Il n'existe pas de taxinomie officielle relative à la résolution spatiale des images de télédétection, cependant il est communément admis que les images BR (Basse Résolution) ont une résolution supérieure à 100 mètres, les images MR (Moyenne Résolution) ont une résolution comprise entre 10 et 100 mètres, les images HR (Haute Résolution) ont une résolution comprise entre 1 et 10 mètres tandis que les images THR (Très Haute Résolution) présentent une résolution inférieure au mètre.

La nécessité de développer de nouvelles méthodes adaptées aux spécificités des images haute et très haute résolution, donne lieu à de nombreux travaux. Ceux-ci se focalisent très majoritairement sur les bâtiments, éléments de l'environnement urbain dont l'extraction est la plus complexe mais aussi la plus intéressante (possibilité de création de villes virtuelles en 3D, etc.).

Parmi ces travaux, certaines méthodes réalisent l'extraction de bâtiments à partir de données stéréoscopiques en imagerie optique et de données radargrammétriques ou interférométriques en imagerie RSO. Ces jeux de données multiscopiques permettent de surmonter, en partie, les problèmes inhérents à l'analyse de scènes urbaines en facilitant l'extraction d'informations d'élévation sur la scène. Malheureusement, la disponibilité de tels jeux de données n'est pas toujours garantie, de sorte que l'analyse de scènes urbaines à partir d'une unique image de télédétection très haute résolution reste un enjeu majeur de la recherche en imagerie de télédétection.

Objectifs et contributions

Les travaux réalisés dans le cadre de cette thèse portent sur le développement de deux méthodes d'extraction tridimensionnelle de bâtiments à partir d'une unique image satellitaire, optique ou RSO. Ces méthodes répondent aux deux exigences suivantes :

- prise en compte de la complexité géométrique et radiométrique inhérente aux images de télédétection haute résolution en milieu urbain ;
- applicabilité des méthodes proposées aussi bien en imagerie optique que RSO.

Afin de répondre à la première exigence, imposée par la nature des images, les méthodes reposent sur des algorithmes permettant de simuler la signature géométrique d'un bâtiment dans une image optique ou RSO. La signature géométrique d'un bâtiment dans une image, qui résulte de la projection de ce bâtiment dans l'image, correspond à un ensemble de régions, ou de manière équivalente à un ensemble de segments, au moyen desquels est représenté le bâtiment dans l'image. Ainsi, les méthodes décrites dans le présent manuscrit consistent à quantifier l'adéquation entre la signature géométrique simulée d'un bâtiment et l'image au sein de laquelle il a été projeté.

La seconde exigence est imposée par les problématiques opérationnelles auxquelles sont confrontés certains utilisateurs amenés à travailler aussi bien avec des images optiques que RSO.

En raison du nombre croissant de capteurs satellitaires haute résolution et très haute résolution, les utilisateurs disposent fréquemment de plusieurs images optiques et/ou RSO d'une même scène urbaine. Dans ce contexte multicapteur, la fusion du contenu informationnel des différentes images et la détection des changements entre ces images constituent les deux problématiques majeures auxquelles sont confrontés les utilisateurs. Aux deux problématiques précédentes, s'ajoute également celle du recalage des images. Cette opération consiste à mettre en correspondance les images entre elles et demeure une étape préalable indispensable à la mise en œuvre de méthodes de fusion d'information ou de détection de changements.

Afin de prendre en compte ces trois problématiques, directement reliées à l'analyse de scènes urbaines, les deux méthodes d'extraction proposées ont été conçues de manière à pouvoir être mises en œuvre dans un contexte multicapteur, ce qui constitue l'une des principales contributions de ces travaux.

Organisation du document

Le présent document est organisé en cinq chapitres articulés autour de trois parties distinctes.

La première partie (chapitres 1 et 2) reprend les concepts photogrammétriques et radargrammétriques permettant d'interpréter puis de simuler la signature géométrique des bâtiments dans les images optiques et RSO.

La seconde partie (chapitre 3) comprend un état de l'art des méthodes développées pour l'extraction des bâtiments en imagerie satellitaire et aéroportée.

Enfin, dans la dernière partie (chapitres 4 et 5), les deux méthodes d'extraction développées dans le cadre de ce travail doctoral sont présentées.

Le principe de fonctionnement des capteurs optiques et RSO ainsi que les caractéristiques des images acquises au moyen de ces capteurs font l'objet du chapitre 1.

Les fonctions de localisation, au moyen desquelles sont modélisées les géométries de prise de vue des images optiques et RSO, sont ensuite présentées au chapitre 2. Ces fonctions permettent d'expliquer les effets géométriques affectant les objets du sur-sol lorsqu'ils sont projetés de l'espace terrain dans l'espace image. Après avoir décrit ces effets géométriques, nous proposons des méthodes pour simuler la signature géométrique des bâtiments dans les images optiques et RSO, bâtiments qui sont supposés parallélépipédiques et qui sont décrits au moyen d'un modèle paramétrique. Ces méthodes de simulation sont par la suite utilisées aux chapitres 4 et 5, où elles sont mises en œuvre pour l'extraction des bâtiments.

Un état de l'art des différents travaux réalisés pour l'extraction des bâtiments en imagerie de télédétection optique et RSO haute résolution est ensuite établi au chapitre 3. Plus particulièrement, nous décrivons les méthodes développées en imagerie optique puis celles développées en imagerie RSO avant de présenter les méthodes qui s'intéressent à l'extraction de bâtiments dans un contexte de fusion optique-RSO. Nous constatons l'existence d'une grande diversité de méthodes, qui sont, pour la plupart, spécifiques à un type de capteur (optique ou RSO) ainsi qu'à la nature monoscopique ou multiscopique des données à partir desquelles est réalisée l'extraction.

Nous établissons alors, aux chapitres 4 et 5, deux nouvelles méthodes pour l'extraction de bâtiments à partir d'une unique image satellitaire haute résolution, méthodes qui peuvent être mises en œuvre indifféremment en imagerie optique ou RSO. Ces méthodes reposent respectivement sur des approches radiométrique et géométrique.

La première méthode, décrite au chapitre 4, met en œuvre deux critères qui quantifient l'homogénéité radiométrique des régions constitutives de la signature géométrique d'un bâtiment et l'adéquation des segments de cette signature avec les bords présents dans l'image. Après avoir donné les expressions analytiques de ces critères, expressions qui sont adaptées à la nature du bruit présent dans les images optiques et RSO, nous réalisons une étude qualitative de leur comportement sur données simulées et réelles. Nous étudions ensuite les propriétés de convergence de ces critères et en déduisons une stratégie d'optimisation, stratégie dont les résultats sont présentés sur données simulées et réelles. Si cette première méthode conduit à des résultats probants, la complexité calculatoire qu'elle implique rend difficile sa mise en œuvre à des fins de détection et restreint son utilisation à l'estimation des paramètres de bâtiments préalablement détectés.

Pour pallier à cette limitation, une seconde méthode d'extraction est développée au chapitre 5. Contrairement à la première méthode, pour laquelle l'adéquation de la signature géométrique d'un bâtiment est directement mesurée dans l'image, la seconde méthode réalise cette mesure à partir d'images binaires issues de l'extraction des primitives caractéristiques de la présence des bâtiments dans les images optiques ou RSO. Plus précisément, cette seconde approche repose sur le calcul, pour une géométrie de prise de vue donnée, de la densité de probabilité de présence des pixels détectés dans ces images binaires et sur le calcul d'une vraisemblance géométrique pour quantifier l'adéquation de la signature géométrique d'un bâtiment avec la position des pixels détectés dans les images binaires. La méthode proposée se décompose en quatre étapes qui sont successivement décrites et illustrées sur données simulées et réelles.

Introduction à l'imagerie de télédétection optique et RSO

Sommaire

1.1	Imagerie optique	6
1.1.1	Généralités	6
1.1.2	Formation des images	6
1.1.3	Caractéristiques des images	9
1.2	Imagerie RADAR à synthèse d'ouverture	13
1.2.1	Généralités	13
1.2.2	Formation des images	13
1.2.3	Caractéristiques des images	19
1.3	Conclusion	21

Dans ce chapitre, les notions de base de l'imagerie de télédétection optique et RSO sont rappelées. Nous abordons le fonctionnement des capteurs optiques et RSO, mis en œuvre sur plateformes aéroportées et satellitaires, avant de présenter les caractéristiques des images issues de ces capteurs. Ce chapitre se limite aux notions utiles à la compréhension du présent document. Le lecteur peut se reporter à [RJ98, MBM01, Mai01, KE01, MMB04, Hei04] pour une description plus complète des phénomènes physiques et mathématiques impliqués en imagerie de télédétection optique et RSO.

1.1 Imagerie optique

1.1.1 Généralités

Les capteurs optiques ont pour fonction la mesure de l'énergie rayonnée par une surface dans le domaine optique du spectre électromagnétique, à savoir pour des longueurs d'onde allant de l'ultraviolet ($0.3 \mu\text{m}$) à l'infrarouge ($1 \mu\text{m}$). Ces capteurs sont constitués de détecteurs à semi-conducteurs CCD¹ ou CMOS² qui reçoivent une énergie électromagnétique $E(\omega, \lambda, t)$ intégrée sur un angle solide $\Delta\omega$, pendant une durée Δt et dans une bande spectrale de largeur $\Delta\lambda$. Finalement, la mesure M réalisée par un détecteur s'exprime comme suit :

$$M = G \int_{\Delta\omega} \int_{\Delta\lambda} \int_{\Delta t} E(\omega, \lambda, t) R(\lambda) d\omega d\lambda dt, \quad (1.1)$$

où G et $R(\lambda)$ correspondent respectivement au facteur de gain et à la réponse spectrale du détecteur résultant des caractéristiques électroniques et optiques du capteur.

1.1.2 Formation des images

La focalisation du rayonnement électromagnétique sur les détecteurs est réalisée au moyen d'une optique et d'un système opto-mécanique plus ou moins complexe assurant le balayage de la scène imagée. La méthode de balayage adoptée dépend de la disposition spatiale des détecteurs dans le plan focal de l'objectif. Ainsi, nous distinguons trois types de géométrie de prise de vue pour les capteurs optiques.

1.1.2.1 Les capteurs à géométrie ponctuelle (capteurs *wiskbroom*)

Les capteurs *wiskbroom*, que l'on retrouve par exemple sur les satellites LandSat (capteurs MSS³ et TM⁴), contiennent un seul détecteur, placé dans le plan focal de l'objectif. Comme illustré à la figure 1.1, les pixels constitutifs d'une même ligne de l'image sont acquis au moyen d'un balayage mécanique dans la direction perpendiculaire à la trajectoire du porteur. Ce balayage mécanique peut être réalisé par un miroir oscillant ou bien par une rotation du porteur sur lui même. L'acquisition d'une nouvelle ligne est assurée par le mouvement du porteur ce qui nécessite une synchronisation entre la vitesse du porteur et la vitesse du balayage mécanique. Le balayage d'une ligne, et donc sa durée d'acquisition, doit correspondre au temps nécessaire au porteur pour avancer d'un pixel le long de sa trajectoire.

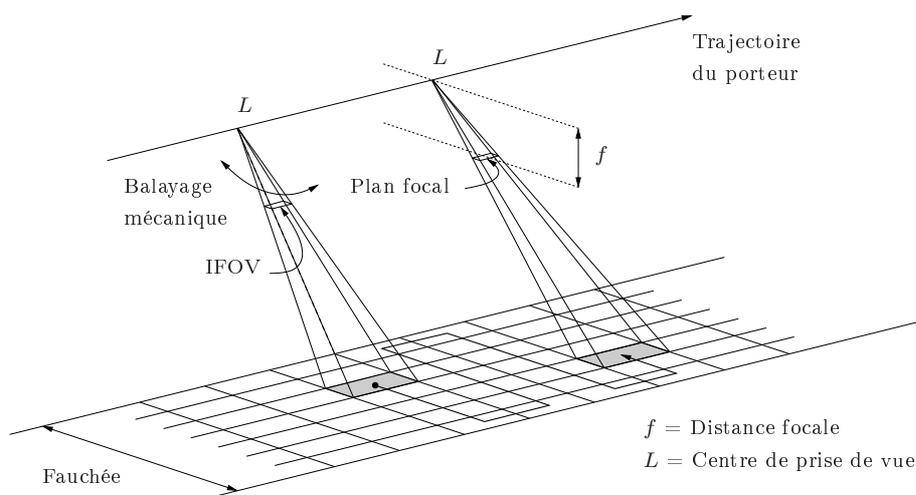
Étant donné qu'ils réalisent l'acquisition des images pixel à pixel, les capteurs *wiskbroom* mettent en œuvre des optiques plus petites et plus simples que celles employées sur les autres types de capteur. Ceci constitue un avantage majeur pour l'acquisition d'images dans les longueurs d'onde non visibles qui nécessite l'utilisation de systèmes optiques plus élaborés. Ces capteurs nécessitent, en revanche, des systèmes mécaniques plus complexes afin de réaliser le balayage nécessaire à l'acquisition des pixels d'une même ligne.

¹ Charge-Couple Devices

² Complementary Metal Oxide Semi-conductor

³ Multi Spectral Scanner

⁴ Thematic Mapper

Figure 1.1 – Géométrie d'un capteur optique *whiskbroom*.

De plus, l'acquisition de l'ensemble des pixels au moyen d'un seul détecteur se traduit par des temps d'intégration plus courts que pour les autres types de capteur et par une sensibilité accrue aux variations d'attitude de la plateforme. En effet, les variations de roulis, de tangage et de lacet du porteur au cours de l'acquisition engendrent des décalages entre les pixels de l'image.

1.1.2.2 Les capteurs à géométrie linéique (capteurs *pushbroom*)

Dans le cas des capteurs *pushbroom*, une barrette de détecteurs, généralement orientée perpendiculairement à la trajectoire du porteur, est placée dans le plan focal de l'objectif. Comme indiqué à la figure 1.2, les pixels constitutifs d'une même ligne de l'image sont acquis simultanément par l'ensemble des détecteurs de la barrette tandis que l'acquisition des différentes lignes est assurée par le mouvement du porteur et donc de la barrette.

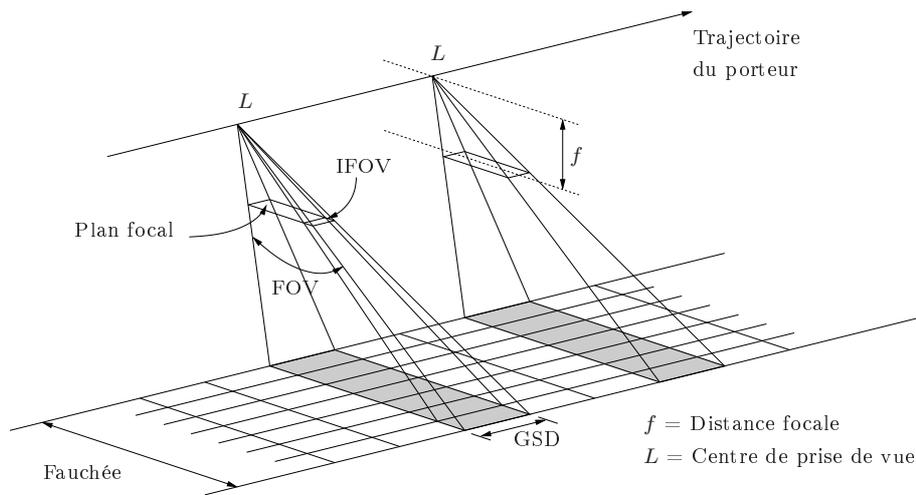
Cette configuration, mise en œuvre sur la plupart des capteurs spatiaux, comme les capteurs HRV⁵ des satellites SPOT⁶ ou les capteurs du satellite Pléiades-HR, et sur certains capteurs aéroportés, comme le capteur ADS40 de Leica, est la plus répandue. En effet, la fabrication des barrettes de détecteurs, utilisées au sein des capteurs *pushbroom*, est plus simple que celle des matrices de détecteurs mise en œuvre par les capteurs *frame*, que nous présentons dans le paragraphe suivant. De plus, les capteurs *pushbroom* ne nécessitent généralement aucun balayage mécanique, contrairement aux capteurs *whiskbroom*.

Comme mentionné précédemment, les lignes d'une image issue d'un capteur *pushbroom* sont acquises successivement si bien que la position du centre de prise de vue L et l'attitude du capteur varient d'une ligne à l'autre en raison du déplacement du porteur. Les variations d'attitude, qui sont négligeables sur plateforme satellitaire mais qui peuvent être importantes sur plateforme aéroportée, engendrent notamment une distorsion de l'image dans la direction perpendiculaire à la trajectoire de la plateforme, comme illustré à la figure 1.3. Une mesure précise de l'attitude et de la position de la plateforme au moyen de systèmes de navigation inertielle, de systèmes GPS⁷ (pour les capteurs aéroportés) ou de senseurs stellaires (pour les capteurs spatiaux) permet de positionner finement les lignes de l'image les unes par rapport aux autres.

⁵Haute Résolution Visible

⁶Système Probatoire d'Observation de la Terre

⁷Global Positioning System

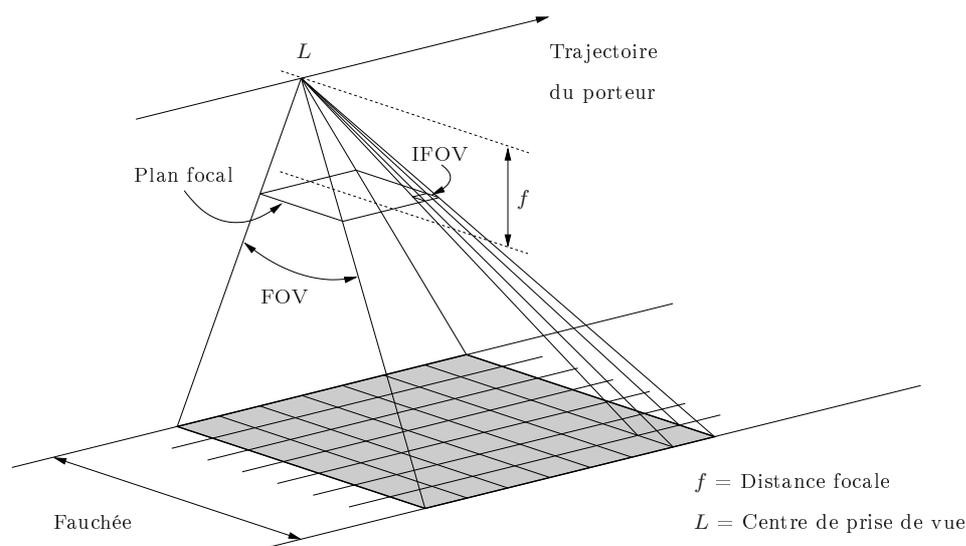
Figure 1.2 – Géométrie d'un capteur optique *pushbroom*.Figure 1.3 – Effets de distorsion observés sur les images acquises par les capteurs *pushbroom* : image originale (a) et image rectifiée (b) (images extraites de [BAE07], ©Jena-Optronik GmbH).

1.1.2.3 Les capteurs à géométrie matricielle (capteurs *frame*)

Les capteurs *frame* sont constitués d'un ensemble de détecteurs disposés matriciellement dans le plan focal de l'objectif. La géométrie de ce type de capteur, illustrée à la figure 1.4, peut être qualifiée d'homogène puisqu'elle implique une acquisition de toute l'image avec le même centre de prise de vue L .

Si les capteurs *frame* sont le plus souvent utilisés sur plateforme aéroportée, ce qui est le cas, par exemple, du capteur UltraCam-Xp de Vexcel Imaging et du capteur CAMv2 de l'IGN, ce type de capteur peut également être mis en œuvre sur plateforme spatiale, comme cela est fait avec le satellite KOSMOS.

Pour finir, notons qu'il existe des capteurs à géométrie matricielle pour lesquels la disposition des détecteurs n'est pas plane mais cylindrique ou sphérique. Les capteurs à géométrie cylindrique, comme le capteur Oméra 40, ont une ouverture de 180 degrés dans la direction perpendiculaire à l'axe du cylindre, ce qui permet d'acquérir des images d'horizon à horizon. Les capteurs à géométrie sphérique, également appelés *fisheyes* circulaires, ont une ouverture de 180 degrés dans toutes les directions.

Figure 1.4 – Géométrie d'un capteur optique *frame*.

Si les capteurs à géométrie cylindrique et sphérique ont l'avantage de présenter des ouvertures plus importantes que les capteurs à géométrie plane, ils introduisent en revanche des effets de distorsion bien plus marqués que les capteurs à géométrie plane.

1.1.3 Caractéristiques des images

Indépendamment du type de capteur avec lequel elles ont été acquises, les images optiques sont caractérisées par un certain nombre de grandeurs physiques qui sont décrites dans cette section. Nous abordons notamment les notions de résolution spatiale, de résolution radiométrique et de sensibilité spectrale, dont dépendent les possibilités d'exploitation de ces images.

1.1.3.1 Résolution spatiale

La résolution spatiale d'une image est une grandeur bidimensionnelle⁸ qui permet de quantifier la capacité d'un capteur à restituer fidèlement les détails de la scène visualisée, détails qui correspondent aux hautes fréquences présentes dans cette scène.

En raison des phénomènes de diffraction et des contraintes de fabrication des capteurs, l'image d'une source lumineuse ponctuelle par un système optique n'est pas un point (cas d'un capteur parfait) mais une tâche⁹ dont les dimensions permettent de définir la résolution spatiale. Si l'on considère deux sources lumineuses ponctuelles de même intensité que l'on rapproche progressivement, il existe une distance entre ces deux sources à partir de laquelle les tâches correspondant à ces sources se recouvrent. La distance en dessous de laquelle il n'est plus possible d'observer l'existence des deux sources correspond alors à la résolution spatiale du capteur.

Dans le cas des capteurs numériques, l'échantillonnage spatial réalisé lors de l'acquisition des données doit vérifier le théorème de Shannon afin de limiter les phénomènes de repliement de spectre (phénomène d'*aliasing*) et ainsi éviter une dégradation de la résolution spatiale des images. Dans la plupart des cas, les limitations physiques des capteurs conduisent naturellement à un

⁸La valeur de la résolution n'est généralement pas identique selon les lignes et les colonnes d'une même image.

⁹Pour des optiques dont les aberrations géométriques ont été corrigées, la tâche observée est liée uniquement aux phénomènes de diffraction et est appelée tâche d'Airy.

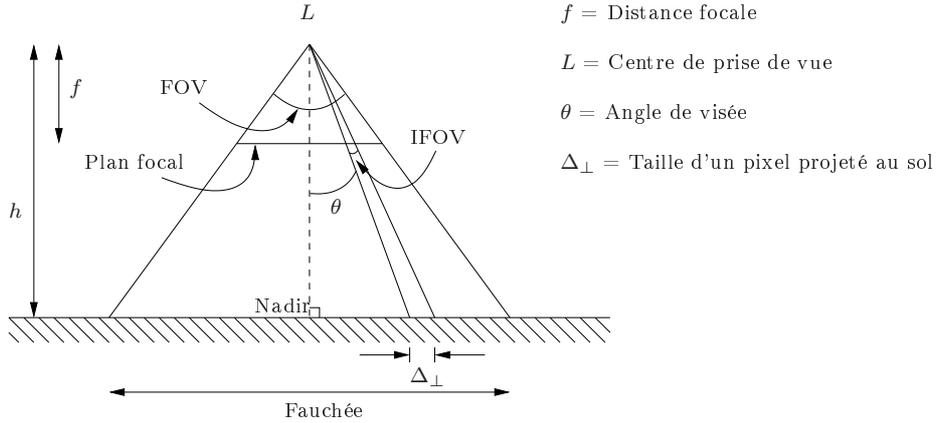


Figure 1.5 – Relation entre l'IFOV et la taille d'un pixel projeté au sol (Δ_{\perp}).

filtrage des hautes fréquences qui est généralement suffisant pour éviter ces phénomènes de repliement.

Comme indiqué à la figure 1.5, l'échantillonnage spatial des données dans la direction perpendiculaire à la trajectoire du porteur (direction *across-track*) est réglé par l'altitude du porteur h , par l'angle de visée θ et par l'IFOV¹⁰, qui correspond à l'angle solide sur lequel est intégrée l'énergie électromagnétique reçue par un détecteur (cf. équation (1.1)).

Dans le cas où l'IFOV est constante, la taille d'un pixel projeté au sol selon la direction *across-track*, couramment désignée par l'acronyme anglo-saxon GSD¹¹ et notée Δ_{\perp} , varie en fonction de l'angle de visée. Cette variation engendre des phénomènes de distorsion géométrique dans la direction *across-track*, phénomènes qui sont d'autant plus marqués que l'altitude du porteur est basse et que la FOV¹² (et donc la fauchée) est importante. Sous hypothèse de Terre plate, l'expression de la taille d'un pixel projeté au sol selon la direction *across-track* en fonction de l'angle de visée est donnée par les équations suivantes :

$$\Delta_{\perp}(\theta) = \frac{\Delta_{\perp}(\theta = 0)}{\cos^2 \theta}, \quad (1.2)$$

$$\Delta_{\perp}(\theta = 0) = 2h \tan\left(\frac{\text{IFOV}}{2}\right) \approx h \text{ IFOV}. \quad (1.3)$$

Il existe des capteurs pour lesquels l'IFOV varie d'un pixel à l'autre de manière à garantir une taille de pixel au sol Δ_{\perp} identique pour tous les pixels de l'image. Sous hypothèse de Terre plate, la variation de l'IFOV avec l'angle de visée est alors régie par les équations suivantes :

$$\text{IFOV}(\theta) = \text{IFOV}(\theta = 0) \cos^2 \theta, \quad (1.4)$$

$$\text{IFOV}(\theta = 0) = 2 \arctan\left(\frac{\Delta_{\perp}}{2h}\right). \quad (1.5)$$

Dans le cas des capteurs *frame*, l'échantillonnage spatial dans la direction parallèle à la trajectoire du porteur (direction *along-track*) est régi par les mêmes équations que dans la direction *across-track*. En revanche, pour les capteurs *wiskbroom* et *pushbroom*, l'échantillonnage spatial dans la direction *along-track* est lié au déplacement du porteur pendant la durée d'intégration des détecteurs (cf. équation (1.1)). L'expression de la taille d'un pixel projeté au sol selon la direction *across-track* (Δ_{\parallel}) est alors donnée par l'équation suivante :

$$\Delta_{\parallel} = V_L \Delta t, \quad (1.6)$$

où V_L et Δt correspondent respectivement à la vitesse du centre de prise de vue L et au temps d'intégration des détecteurs.

¹⁰ *Instantaneous Field Of View*

¹¹ *Ground Sample Distance*

¹² *Field Of View*

Il est à noter que, contrairement à la taille d'un pixel projeté au sol selon la direction *across-track*, la taille d'un pixel projeté au sol selon la direction *along-track* est la même pour tous les pixels de l'image, de sorte qu'il n'y a pas d'effets de distorsion dans la direction *along-track* pour les images issues des capteurs *wiskbroom* et *pushbroom*.

1.1.3.2 Résolution radiométrique

La résolution radiométrique d'une image est liée à la capacité d'un capteur à distinguer deux zones étendues de réflectances voisines. La résolution radiométrique, exprimée en pas de quantification ou en unité physique, est essentiellement limitée par le bruit d'acquisition dont trois composantes distinctes peuvent être identifiées : le bruit de lecture, le bruit thermique et le bruit de photon (ou bruit quantique), ce dernier étant prédominant à flux de photon élevé.

De même que pour la résolution spatiale, l'échantillonnage radiométrique doit satisfaire le théorème de Shannon afin de préserver la résolution radiométrique. La radiométrie des images est usuellement échantillonnée sur 11 ou 12 bits par bande spectrale, ce qui offre respectivement 2048 et 4096 valeurs possibles pour quantifier la dynamique des différentes bandes spectrales. Pour des raisons informatiques, la radiométrie des images est généralement codée sur 16 bits par pixel et par bande spectrale.

1.1.3.3 Sensibilité spectrale

La notion de sensibilité spectrale d'une image est liée au nombre et à la largeur des bandes spectrales constitutives de cette image. Parmi les images produites par les capteurs optiques aéroportés et spatiaux, nous distinguons :

- les images panchromatiques, images monobandes qui résultent d'une intégration de l'énergie électromagnétique reçue sur une large bande spectrale ;
- les images multispectrales qui correspondent à une acquisition sur plusieurs bandes spectrales de faible largeur ;
- les images hyperspectrales, qui correspondent à l'acquisition de quelques dizaines à quelques centaines de bandes spectrales très fines et contiguës.

Dans le cas des images multispectrales et hyperspectrales, la lumière reçue par le capteur est décomposée en plusieurs bandes spectrales. Cette décomposition, réalisée au moyen de filtres colorimétriques ou de prismes, se traduit par une diminution du flux lumineux reçu par les détecteurs et donc par une diminution du RSB¹³. Cette diminution du RSB peut alors être compensée par une augmentation du temps d'intégration du capteur. En raison du déplacement relatif du capteur par rapport à la scène imagée, l'augmentation du temps d'intégration implique, en revanche, une augmentation du pas d'échantillonnage spatial et donc une diminution de la résolution spatiale. Ainsi, à qualité d'image comparable, l'amélioration de la sensibilité spectrale des images se fait au détriment de leur résolution spatiale et réciproquement.

À titre d'exemple, les images panchromatiques fournies par le capteur spatial QuickBird [Dig06] intègrent les longueurs d'onde du spectre électromagnétique comprises entre 450 et 900 nm et présentent un GSD de 61 cm au nadir. Les images multispectrales fournies par ce même capteur sont constituées de quatre bandes spectrales (bleu : 450-520 nm, vert : 520-600 nm, rouge : 630-690 nm et proche infrarouge : 760-900 nm) qui présentent chacune un GSD de 2,44 m au nadir. Enfin, les images fournies par le capteur spatial hyperspectral Hyperion [TRW01] possèdent 220 bandes spectrales contiguës (d'une largeur de 10 nm) qui couvrent les longueurs d'onde du spectre électromagnétique allant de 400 nm à 2500 nm et qui présentent chacune un GSD de 30 m.

Selon l'exploitation qui est faite des images de télédétection, la résolution spatiale est préférée à la sensibilité spectrale ou inversement. Dans le cadre de l'exploitation des images de télédétection pour la surveillance de l'environnement, les images présentant une bonne sensibilité spectrale, mais une faible

¹³Rapport Signal sur Bruit

résolution spatiale, sont souvent préférées puisqu'elles conduisent à de meilleurs résultats de classification et donc à une meilleure estimation de l'occupation des sols. À l'inverse, les images présentant une bonne résolution spatiale, mais une faible sensibilité spectrale, sont privilégiées pour l'analyse de scènes urbaines, étant donné que les objets caractéristiques de ces scènes ont des dimensions métriques ou décimétriques.

1.1.3.4 Niveaux de traitement

Indépendamment des caractéristiques propres aux différents types de capteur, il convient de rappeler que certains traitements, visant à améliorer la géométrie et la radiométrie des images, sont effectués par les organismes producteurs de données. Parmi les traitements usuellement appliqués, nous distinguons notamment :

- les traitements qui recouvrent les corrections appliquées aux métadonnées de positionnement de la plateforme lors de l'acquisition (correction des données de datation et d'orbitographie, lissage de l'attitude du capteur) ;
- les traitements radiométriques qui incluent la suppression des pixels morts, le calibrage des réponses radiométriques relatives des différents capteurs et le calibrage de la radiométrie globale à partir de signaux radiométriques connus ;
- les traitements géométriques parmi lesquels nous distinguons les traitements bas-niveau et les traitements haut-niveau. Les traitements bas-niveau corrigent les images de l'ensemble des distorsions introduites par les composantes optiques et mécaniques du capteur tandis que les traitements haut-niveau consistent à rééchantillonner l'image dans une géométrie terrain (dans un système de coordonnées géographiques ou cartographiques) corrigée des effets du relief et éventuellement des effets liés au éléments du sur-sol.

Chaque producteur de données définit des niveaux de traitement qui lui sont propres de sorte qu'il n'existe pas de typologie de traitement standard. À titre d'exemple, il existe six niveaux de traitement distincts pour les images fournies par les satellites SPOT [EAD02] :

- le niveau 0, qui correspond à la donnée image brute ;
- le niveau 1A, pour lequel seules les corrections radiométriques ont été appliquées ;
- le niveau 1B, pour lequel les traitements géométriques bas-niveau ont également été appliqués ;
- le niveau 2A, pour lequel l'image a été rééchantillonnée dans un référentiel cartographique, sans prise en compte du relief ;
- le niveau 2B, identique au niveau 2A mais pour lequel des points d'appui ont été utilisés ;
- le niveau 3, qui correspond à une image orthorectifiée avec prise en compte des effets du relief par utilisation d'un MNT¹⁴.

S'il n'existe pas de typologie standard, la tendance actuelle est toutefois à la simplification du nombre de niveaux de traitement, de sorte que l'on ne distingue généralement plus que deux niveaux, comme dans le cas des images issues du satellite Pléiades-HR [EAD12] :

- le niveau « capteur parfait », niveau de traitement le plus proche de l'image telle qu'acquise par le capteur. Ce niveau, qui garantit une restitution parfaite des conditions d'acquisition, est destiné aux utilisateurs familiarisés avec les techniques de traitement d'image satellite qui souhaitent appliquer leurs propres méthodes de production (orthorectification ou modélisation 3D par exemple) ;
- le niveau « ortho », pour lequel l'ensemble des traitements décrits précédemment ont été appliqués. Ce niveau de traitement est essentiellement destiné aux utilisateurs finaux pour une intégration au sein de SIG¹⁵.

Sur les images orthorectifiées fournies par les producteurs de données, les objets du sur-sol présentent généralement des effets de distorsion liés à l'utilisation de données d'élévation dont la précision planimétrique et altimétrique est insuffisante. Les objets du sur-sol sont d'autant plus affectés par ces effets de distorsion que leur hauteur est importante et que la prise de vue est oblique. Ainsi, dans le

¹⁴Modèle Numérique de Terrain

¹⁵Système d'Information Géographique

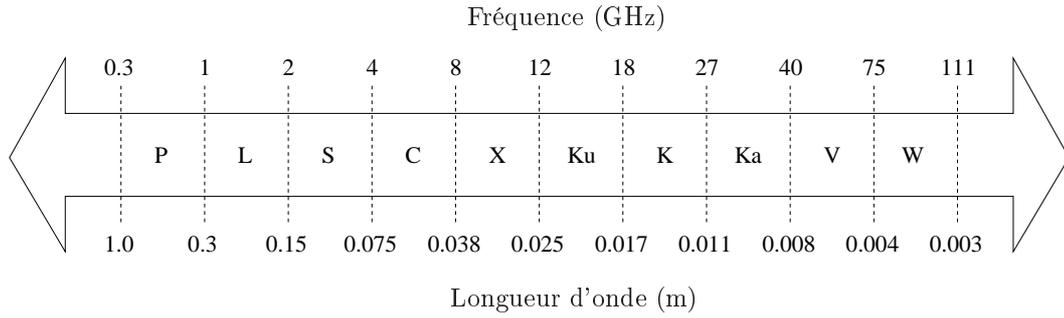


Figure 1.6 – Définition IEEE des bandes spectrales utilisées par les systèmes RADAR.

cadre de l'extraction de bâtiments, il est souvent préférable de travailler sur des images pour lesquelles la géométrie d'acquisition n'a pas été modifiée.

1.2 Imagerie RADAR à synthèse d'ouverture

1.2.1 Généralités

Contrairement aux capteurs optiques qui bénéficient de l'illumination de la scène par une source lumineuse externe au système, les capteurs RSO sont des capteurs actifs qui illuminent la scène en émettant des ondes hyperfréquences. À la phase d'émission succède une phase de réception au cours de laquelle les capteurs RSO mesurent l'énergie de l'onde électromagnétique réfléchie par les réflecteurs présents dans la scène. Les ondes hyperfréquences émises par les systèmes RADAR¹⁶ couvrent des bandes spectrales dont la définition est spécifiée à la figure 1.6.

De même que les capteurs optiques, les capteurs RSO mesurent l'énergie de l'onde électromagnétique réfléchie par la scène visualisée, de sorte que l'équation (1.1) est également valable dans le cas des capteurs RSO. La maîtrise du processus d'émission de l'onde électromagnétique, propre aux capteurs actifs, permet également de mesurer la distance entre le capteur RSO et les réflecteurs présents dans la scène ainsi que l'effet Doppler résultant du mouvement relatif du capteur par rapport aux réflecteurs.

La distance entre le capteur et un réflecteur est simplement déduite du temps mis par l'onde électromagnétique pour effectuer, à la vitesse de la lumière, l'aller et retour entre le capteur et le réflecteur.

En ce qui concerne l'effet Doppler résultant du mouvement relatif du capteur RSO par rapport à un réflecteur, il se traduit par un décalage fréquentiel entre l'onde émise par le capteur et l'onde réfléchie par le réflecteur. Dans le cas où une onde sinusoïdale de fréquence f_e est émise par le capteur, la fréquence f_r de l'onde sinusoïdale réfléchie par le réflecteur est donnée par la relation suivante :

$$f_r = f_e \frac{2V_S}{c} \cos \theta, \quad (1.7)$$

où V_S désigne la vitesse relative de l'antenne par rapport au réflecteur, c est la vitesse de la lumière et θ est l'angle, parfois dénommé angle Doppler, permettant de repérer le réflecteur relativement au porteur, comme indiqué à la figure 1.7.

1.2.2 Formation des images

Pour chaque ligne de l'image, le capteur RSO émet une impulsion pendant une durée T_e avant de passer en mode réception pendant une durée T_r , durée au cours de laquelle il reçoit les échos réfléchis par les réflecteurs qui ont été atteints par le faisceau émis. Les pixels d'une même ligne de l'image, encore appelés « cases distance », intègrent l'énergie des ondes électromagnétiques réfléchies par les

¹⁶ *RADio Detection And Ranging*

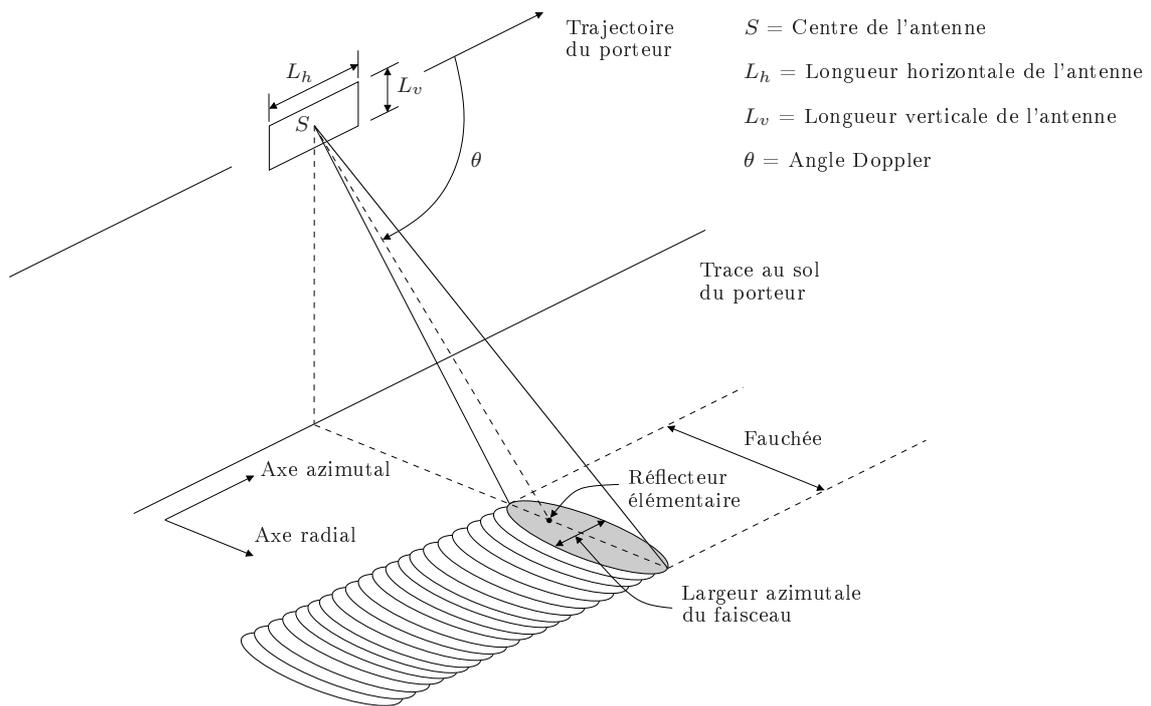


Figure 1.7 – Géométrie d'un capteur RADAR à synthèse d'ouverture.

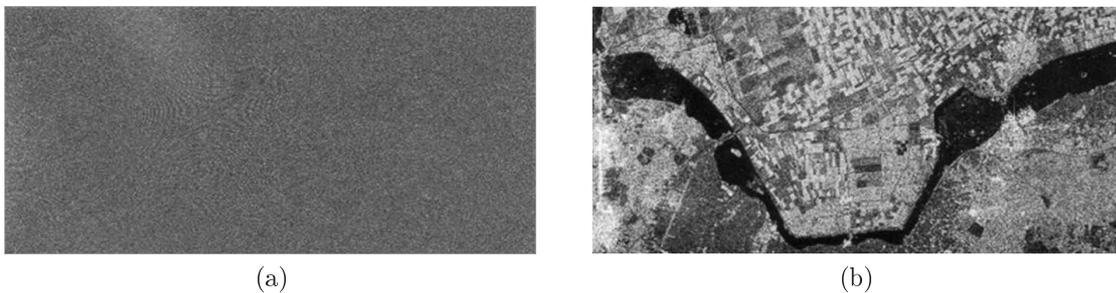


Figure 1.8 – Données brutes issues d'un capteur RSO (a) et image obtenue à l'issue des traitements en distance et en azimut (b) (images extraites de [Den09]).

réflecteurs selon leur distance au capteur. Schématiquement, les pixels de début de ligne correspondent aux premiers échos reçus, c'est-à-dire aux réflecteurs les plus proches du capteurs, tandis que les pixels de fin de ligne correspondent aux réflecteurs les plus éloignés du capteurs. La répétition de ce cycle d'émission-réception, conjuguée au mouvement du porteur, permet d'acquérir l'ensemble des lignes de l'image.

Les données acquises par les capteurs RSO sont des données complexes dont l'amplitude représente l'énergie électromagnétique reçue et l'argument correspond au déphasage entre l'onde émise et l'onde reçue par le capteur. Ces données, qualifiées de données brutes (*raw data*), ne peuvent pas être interprétées directement et nécessitent l'application de traitements par les processeurs RSO pour obtenir des images dont l'exploitation visuelle devient possible, comme illustré à la figure 1.8.

Les deux premières sous-sections sont consacrées à la présentation des traitements en distance et en azimut tandis que les différents modes d'acquisition des capteurs RSO sont décrits dans la dernière sous-section.

1.2.2.1 Traitement en distance et compression d'impulsion

La forme des impulsions émises par un capteur RSO étant connue, le traitement en distance consiste à appliquer au signal reçu un filtrage adapté au signal émis par ce capteur de manière à maximiser le RSB du signal reçu.

La compression d'impulsion consiste à émettre une impulsion modulée linéairement en fréquence¹⁷ plutôt qu'une simple impulsion sinusoïdale, ce qui conduit, comme nous allons le voir, à une amélioration de la résolution radiale des images. La description du principe de la compression d'impulsion que nous présentons ici s'appuie sur les résultats établis par Maître *et al.* [Mai01].

Nous supposons dans ce paragraphe que le signal $s_e(t)$ émis par le capteur RSO est modulé linéairement en fréquence de sorte que sa fréquence instantanée $f_e(t)$ s'écrit comme suit :

$$\forall t \in \left[-\frac{T_e}{2}; \frac{T_e}{2}\right], f_e(t) = f_c + \frac{K_t}{2} t, \quad (1.8)$$

où T_e correspond à la durée de l'impulsion, f_c désigne la fréquence centrale de l'onde émise et $K_t/2$ est le coefficient de variation linéaire de la fréquence de l'onde émise.

L'expression du signal $s_e(t)$ émis par le capteur RSO est alors donnée par la relation suivante :

$$\forall t \in \left[-\frac{T_e}{2}; \frac{T_e}{2}\right], s_e(t) = s_0 \exp[2j\pi(f_e(t) t)], \quad (1.9)$$

$$s_e(t) = s_0 \exp\left[2j\pi\left(f_c t + \frac{K_t t^2}{2}\right)\right], \quad (1.10)$$

où s_0 correspond à l'amplitude du signal émis.

Le traitement en distance consiste à appliquer au signal $s_r(t)$ reçu par le capteur un filtrage adapté au signal émis $s_e(t)$, ce qui conduit à un signal $s_a(t)$ dont l'expression est donnée par l'équation suivante :

$$s_a(t) = \int_{-\infty}^{\infty} s_e^*(t' - t) s_r(t') dt', \quad (1.11)$$

où $s_e^*(t)$ correspond au signal conjugué de $s_e(t)$.

Pour mettre en évidence le gain en résolution radiale obtenu à l'issue de la compression d'impulsion, nous étudions le cas particulier où le signal reçu résulte de la réflexion du signal émis sur un unique réflecteur situé à une distance D du capteur RSO. Le signal $s_r(t)$ reçu par le capteur s'écrit alors :

$$\forall t \in \left[t_D - \frac{T_e}{2}; t_D + \frac{T_e}{2}\right], s_r(t) = \sigma^0 s_0 \exp\left[2j\pi\left(f_c(t - t_D) + \frac{K_t(t - t_D)^2}{2}\right)\right], \quad (1.12)$$

où $t_D = 2D/c$ correspond au temps mis par l'onde pour effectuer l'aller-retour du capteur jusqu'au réflecteur et σ^0 est la surface équivalente RADAR du réflecteur.

Étant données les expressions (1.9) et (1.12) des signaux $s_e(t)$ et $s_r(t)$, la variable d'intégration t' de l'équation (1.11) doit vérifier les deux conditions suivantes :

$$t' \in \left[t - \frac{T_e}{2}; t + \frac{T_e}{2}\right], \quad (1.13)$$

$$t' \in \left[t_D - \frac{T_e}{2}; t_D + \frac{T_e}{2}\right]. \quad (1.14)$$

Si l'on considère le cas où $t > t_D$ (le cas où $t < t_D$ s'écrit de manière similaire), les bornes d'intégration peuvent alors être restreintes à l'intervalle $[t - T_e/2; t_D + T_e/2]$ et l'équation (1.11) se

¹⁷Ce type d'impulsion est couramment désigné par le terme anglo-saxon « chirp ».

réécrit comme suit :

$$s_a(t) = \sigma^0 s_0^2 \exp[2j\pi f_c (t - t_D)] \int_{t - \frac{T_e}{2}}^{t_D + \frac{T_e}{2}} \exp\left[-2j\pi K_t \frac{(t' - t)^2}{2}\right] \exp\left[2j\pi K_t \frac{(t' - t_D)^2}{2}\right] dt'. \quad (1.15)$$

Le calcul de cette intégrale conduit alors à l'expression générale du signal $s_a(t)$:

$$\forall t \in [t_D - T_e; t_D + T_e], s_a(t) = \sigma^0 s_0^2 \exp[2j\pi f_c (t - t_D)] 2T_s(t) \frac{\sin(U(t)T_s(t))}{U(t)T_s(t)}, \quad (1.16)$$

où :

$$U(t) = 2\pi K_t (t - t_D), \quad (1.17)$$

$$T_s(t) = \frac{T_e}{2} - \frac{|t - t_D|}{2}. \quad (1.18)$$

Avec les valeurs classiques de K_t et de T_e , les variations de $T_s(t)$ avec t sont négligeables par rapport à celles de $U(t)$ dans un voisinage de t_D , de sorte que seules les variations de $U(t)$ peuvent être prises en compte au sein du sinus cardinal, $T_s(t)$ étant alors considéré comme constant et égal à $T_e/2$. Finalement, dans un voisinage de t_D , le signal $s_a(t)$ en sortie du filtre adapté s'écrit :

$$s_a(t) \approx \sigma^0 s_0^2 \exp[2j\pi f_c (t - t_D)] T_e \frac{\sin(\pi K_t T_e (t - t_D))}{\pi K_t T_e (t - t_D)}. \quad (1.19)$$

En imagerie optique, la notion de résolution spatiale est relative à la séparabilité de deux sources lumineuses. De manière analogue, la résolution radiale d'une image RSO, notée δ_t ou δ_r selon que l'on travaille en temps ou en distance, est relative à la séparabilité de deux réflecteurs le long de l'axe radial. Si l'on définit cette résolution radiale comme la largeur à -3 dB du lobe principal du sinus cardinal¹⁸ de l'expression (1.19), son expression est donnée par la relation suivante :

$$\delta_t = \frac{0,88}{K_t T_e} = \frac{0,88}{B}, \quad (1.20)$$

où $B = K_t T_e$ correspond à la largeur de bande de l'impulsion émise. Finalement la résolution radiale en distance s'écrit :

$$\delta_r = \delta_t \frac{c}{2} = \frac{0,88}{K_t T_e} \frac{c}{2}. \quad (1.21)$$

La résolution radiale obtenue avec une impulsion modulée en fréquence doit être comparée à celle obtenue avec une simple impulsion sinusoïdale :

$$\delta_r = 0,58 T_e \frac{c}{2}. \quad (1.22)$$

Ainsi, l'émission d'une impulsion modulée en fréquence conduit à une amélioration de la résolution radiale d'un facteur égal à BT_e par rapport à une impulsion sinusoïdale. Il est à noter que la résolution radiale obtenue avec une impulsion modulée en fréquence est inversement proportionnelle à la largeur de bande B de cette impulsion et qu'elle ne dépend pas de la fréquence centrale de l'impulsion.

1.2.2.2 Traitement en azimut et synthèse d'ouverture

Dans cette sous-section, nous présentons le traitement en azimut des images RSO, encore appelé synthèse d'ouverture, qui permet, de manière similaire au traitement en distance, d'accroître la résolution azimutale de ces images. La description que nous faisons ici de la synthèse d'ouverture s'appuie également sur les résultats établis par Maître *et al.* [Maï01].

¹⁸La largeur du sinus cardinal à -3 dB vaut $0,88\pi$.

En raison des phénomènes de diffraction, le faisceau émis par une antenne RADAR présente une ouverture azimutale $\Delta\theta$ à -3 dB qui vaut :

$$\Delta\theta = \frac{\lambda}{L_h}, \quad (1.23)$$

où λ correspond à la longueur d'onde de l'onde émise et L_h désigne la longueur de l'antenne le long de l'axe azimutal, comme indiqué à la figure 1.7. Un même réflecteur est ainsi illuminé pour différentes positions consécutives du capteur RSO, c'est-à-dire qu'il intervient dans plusieurs cycles d'émission-réception du capteur et donc dans plusieurs lignes de l'image.

Dans cette sous-section, nous souhaitons étudier l'évolution du signal $s_a(t)$ obtenu en sortie du filtre adapté en fonction de la position x du capteur RSO le long de l'axe azimutal, de sorte que ce signal est désormais noté $s_a(t, x)$.

L'expression de la distance X parcourue par le capteur entre le moment où le réflecteur commence à être illuminé par l'onde RADAR et le moment où il cesse d'être illuminé par cette onde est donnée par l'équation suivante :

$$X = D_0\Delta\theta = D_0\frac{\lambda}{L_h}, \quad (1.24)$$

où D_0 correspond à la distance du réflecteur à l'axe azimutal.

Si l'on note x la coordonnée du centre S de l'antenne RADAR le long de l'axe azimutal, l'origine étant placée au niveau du projeté orthogonal du réflecteur sur cet axe, la distance D entre le capteur et le réflecteur s'écrit :

$$D(x) = \sqrt{D_0^2 + x^2} \approx D_0 + \frac{x^2}{2D_0}. \quad (1.25)$$

La variation de D pendant l'illumination d'un réflecteur engendre une variation de t_D qui est relié à D par la relation $t_D(x) = 2D(x)/c$.

La dépendance de $s_a(t, x)$, dont l'expression est donnée à l'équation (1.19), par rapport à la variable t_D intervient au niveau du terme exponentiel et du sinus cardinal. Le terme exponentiel présente une forte sensibilité aux variations de D puisqu'une variation de ce terme d'une quantité égale à la longueur d'onde λ de l'onde RADAR engendre une variation de la phase de 2π . En revanche, le sinus cardinal est peu sensible aux variations de D de sorte que sa valeur peut être considérée comme quasiment constante et égale à sa valeur maximale atteinte en $t = t_D$ et valant 1. Le terme prépondérant $\tilde{s}_a(t, x)$ du signal $s_a(t, x)$ observé en sortie du filtre adapté s'écrit donc :

$$\tilde{s}_a(t, x) = \exp[2j\pi f_c(t - t_D(x))] = \exp\left[2j\pi f_c\left(t - \frac{2D(x)}{c}\right)\right]. \quad (1.26)$$

En remplaçant D par son approximation au premier ordre, donnée à l'équation (1.25), le terme prépondérant $\tilde{s}_a(t, x)$ s'écrit alors :

$$\tilde{s}_a(t, x) = \exp\left[2j\pi f_c\left(t - \frac{2D_0}{c}\right)\right] \exp\left[2j\pi f_c \frac{x^2}{D_0 c}\right]. \quad (1.27)$$

Dans cette expression, le terme $\exp[2j\pi f_c x^2/(D_0 c)]$ peut être vu comme un *chirp* de fréquence spatiale centrale nulle et de coefficient de variation $K_x = 2f_c/(D_0 c) = 2/(D_0 \lambda)$.

De manière analogue au traitement en distance, le traitement en azimut, encore appelé synthèse d'ouverture, consiste à appliquer un filtrage adapté le long de l'axe azimutal, c'est-à-dire à t fixé. Alors que le traitement en distance est appliqué sur les lignes de l'image, le long desquelles la position x du capteur est fixée, le traitement en azimut est appliqué sur les colonnes de l'images, le long desquelles le temps t est fixé. Si l'on définit la résolution azimutale δ_a de manière analogue à la résolution radiale,

c'est-à-dire comme la largeur à -3 dB du lobe principal du sinus cardinal obtenu en sortie du filtre adapté, son expression se déduit de l'équation (1.20) par analogie de K_t avec K_x et de T_e avec X :

$$\delta_a = \frac{0,88}{K_x X} = 0,88 \frac{L_h}{2}. \quad (1.28)$$

Cette expression est à comparer à la résolution classique d'une antenne, résolution limitée par les phénomènes de diffraction :

$$\delta_a = D_0 \frac{\lambda}{L_h}. \quad (1.29)$$

Contrairement à la résolution classique atteinte par une antenne, qui dépend de la distance à cette antenne et qui est inversement proportionnelle à la longueur de l'antenne, la résolution obtenue par synthèse d'ouverture dépend uniquement de la longueur azimutale de l'antenne et est d'autant meilleure que l'antenne est petite, ce qui constitue un résultat *a priori* surprenant.

Si l'on reprend l'équation 1.7, qui traduit la variation de fréquence observée lors de la réflexion d'une onde sinusoïdale sur un réflecteur élémentaire, variation liée au mouvement relatif entre le capteur et le réflecteur (effet Doppler), une différentiation de cette équation conduit à la relation suivante :

$$\Delta f_r = \frac{2V_S}{c} [\Delta f_e \cos \theta + f_e \sin \theta |\Delta \theta|]. \quad (1.30)$$

De manière analogue à la résolution radiale (cf. équation (1.20)), la résolution azimutale est d'autant meilleure que la largeur de bande Δf_r de la réponse spectrale du réflecteur est importante. Dans le cas d'un signal modulé linéairement en fréquence (*chirp*), les termes Δf_e et f_e de l'équation (1.30) correspondent respectivement à la largeur de bande et à la fréquence centrale de ce signal. Avec les valeurs usuelles de Δf_e et f_e , qui sont respectivement de l'ordre du Mégahertz et du Gigahertz, la largeur de bande Δf_r est maximale, pour une ouverture azimutale du faisceau $\Delta \theta$ donnée, lorsque θ vaut 90 degrés, ce qui explique que les RADAR imageurs soient des capteurs à visée latérale. L'équation (1.30) se réécrit alors comme suit :

$$\Delta f_r = \frac{2V_S}{c} f_e |\Delta \theta|. \quad (1.31)$$

Finalement, nous déduisons de cette dernière équation que la largeur de bande est d'autant plus grande que l'ouverture azimutale $\Delta \theta$ est grande et donc que la longueur azimutale de l'antenne est faible, expliquant ainsi le résultat *a priori* surprenant établi précédemment.

1.2.2.3 Les différents modes d'acquisition

Comme illustré à la figure 1.9, les capteurs RSO sont caractérisés par trois modes d'acquisition distincts qui permettent de jouer sur la taille et la résolution des images acquises.

Le mode d'acquisition le plus simple est le mode *StripMap* pour lequel la fauchée des images est limitée par l'ouverture en élévation du faisceau émis et la taille azimutale est uniquement limitée par la capacité de stockage du porteur. Les images acquises selon ce mode couvrent usuellement une surface de plusieurs dizaines de kilomètres et présentent une résolution allant de quelques mètres à quelques dizaines de mètres.

Le mode *SoptLight* permet d'améliorer la résolution azimutale des images au moyen d'un dépointage mécanique ou électronique de l'antenne en azimut. Ce dépointage permet d'augmenter la durée d'illumination des réflecteurs et donc d'accroître la largeur de bande Doppler ce qui se traduit par une amélioration de la résolution azimutale par rapport au mode *StripMap*. La fauchée des images est limitée, comme pour le mode *StripMap*, par l'ouverture en élévation du faisceau émis tandis que la taille azimutale des images est limitée par l'amplitude du dépointage de l'antenne.

Le mode *ScanSAR* est un mode d'acquisition permettant d'augmenter la fauchée des images au prix d'une dégradation de leur résolution azimutale. La fauchée globale est découpée en sous-fauchées acquises séquentiellement par un dépointage électronique de l'antenne en élévation. Ce dépointage se traduit par une diminution de la durée d'illumination des réflecteurs et donc par une diminution de la

résolution azimutale. En pratique, la résolution radiale est volontairement dégradée par rapport aux autres modes d'acquisition, de manière à obtenir des images de résolutions comparables en azimut et en distance. Les images acquises selon ce mode, reconstituées par mosaïquage des différentes sous-fauchées, couvrent usuellement plusieurs centaines de kilomètres.

À titre d'exemple, les images acquises par le capteur spatial TerraSAR-X [Inf09] présentent des résolutions radiales et azimutales de 1.2 m et 3.3 m en mode *StripMap*, de 1.2 m et 1.7 m en mode *SpotLight* et de 1.2 m et 1.1 m en mode *High Resolution SpotLight*¹⁹, tandis qu'en mode *ScanSAR*, les résolutions radiales et azimutales sont proches de 18.5 m.

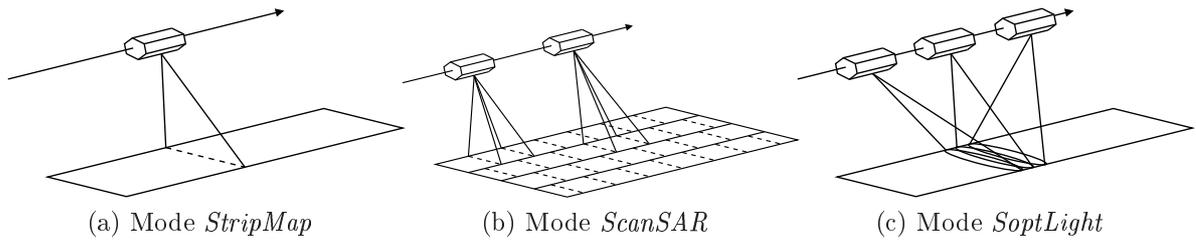


Figure 1.9 – Modes d'acquisition des capteurs RADAR à synthèse d'ouverture.

En imagerie RSO comme en imagerie optique, il est préférable de disposer d'images ayant une bonne résolution spatiale pour l'analyse de scènes urbaines. Ainsi, dans le cadre de l'extraction de bâtiments à partir d'images RSO, les images acquises en mode *SpotLight* seront préférées aux images acquises dans les autres modes.

1.2.3 Caractéristiques des images

Dans cette section nous revenons rapidement sur la notion de résolution spatiale, qui a déjà été largement abordée dans la section précédente, avant de décrire les différents niveaux de traitement associés aux images RSO.

1.2.3.1 Résolution spatiale

Comme mentionné dans la section précédente, les données brutes issues d'un capteur RSO n'ont subi aucun traitement de sorte qu'elles correspondent aux données acquises par un capteur dont la résolution spatiale, en distance comme en azimut, est de plusieurs kilomètres. La mise en œuvre des traitements en distance et en azimut permet alors d'obtenir des images dont la résolution radiale comme azimutale est améliorée d'un facteur allant de quelques centaines à un millier.

Comme pour les images optiques, l'échantillonnage spatial des données doit satisfaire certaines contraintes de manière à préserver la résolution spatiale de ces dernières. Dans la direction radiale, le pas d'échantillonnage Δ_r est relié à la fréquence F_e avec laquelle le signal reçu par l'antenne est échantillonné. Cette fréquence doit permettre une reconstruction du signal modulé en fréquence (*chirp*) en limitant les phénomènes de recouvrement de spectre. Le critère de Shannon impose donc que cette fréquence d'échantillonnage soit supérieure à la largeur de bande $K_t T_e$ du signal reçu. Le pas d'échantillonnage radial Δ_r se déduit de la fréquence d'échantillonnage F_e par la relation suivante :

$$\Delta_r = \frac{c}{2F_e}. \quad (1.32)$$

Dans la direction azimutale, le pas d'échantillonnage Δ_a est réglé par la fréquence de répétition des impulsions F_{RI} qui, pour respecter le théorème de Shannon, doit être supérieure à la bande Doppler

¹⁹En mode *High Resolution SpotLight*, la vitesse du dépointage de l'antenne en azimut est plus élevée qu'en mode *SpotLight* ce qui se traduit par une durée d'illumination des réflecteurs plus importante et donc par une amélioration de la résolution azimutale.

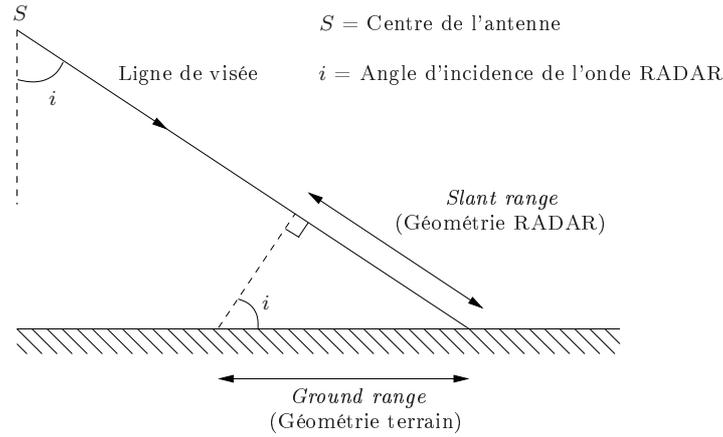


Figure 1.10 – Géométrie RADAR et géométrie terrain.

Δf_r . Le pas d'échantillonnage est relié à la fréquence de répétition des impulsions par la relation suivante :

$$\Delta_a = \frac{V_S}{F_{RI}}, \quad (1.33)$$

où V_S correspond à la vitesse du porteur.

Jusqu'à présent, les résolutions ainsi que les pas d'échantillonnage ont été exprimés en géométrie RADAR, à savoir en distance radiale ou distance oblique (*slant range*) par rapport à la position du capteur. La projection au sol (supposé plat) de ces grandeurs, comme illustré à la figure 1.10, permet d'obtenir leurs expressions en géométrie terrain, ce qui est souvent plus pratique :

$$\Delta_{r,sol} = \frac{c}{2F_e \sin i}, \quad (1.34)$$

$$\delta_{r,sol} = \frac{c}{2K_t T_e \sin i}, \quad (1.35)$$

où i désigne l'angle d'incidence de l'onde émise par le capteur RSO. Les grandeurs relatives à la dimension azimutale ne sont pas affectées par la projection au sol de sorte que leurs expressions en géométrie RADAR et en géométrie terrain sont identiques.

1.2.3.2 Niveaux de traitement

Contrairement aux images optiques pour lesquelles les niveaux de traitement sont relativement variables en fonction des capteurs et des organismes producteurs de données, il existe, pour les images RSO, une typologie standard relative aux niveaux de traitement de sorte que l'on distingue les données brutes, les données monovues complexes et les données multivues.

Comme mentionné à la section 1.2.2, les données brutes (*raw data*), directement issues du capteur RSO, ne sont pas directement exploitables puisqu'elles correspondent à des images acquises par un capteur dont la résolution spatiale, en distance comme en azimut, est de plusieurs kilomètres. Ces données sont fournies pour les utilisateurs souhaitant appliquer eux mêmes les traitements en distance et en azimut.

Les images monovues complexes sont obtenues à l'issue des traitements en distance et en azimut. Ces images, dont l'amplitude est représentative de la scène imagée et permet une exploitation visuelle, possède une géométrie qui retranscrit fidèlement la géométrie de prise de vue du capteur RSO.

Enfin, les images multivues sont des images d'amplitude pour lesquelles une diminution du bruit de speckle, caractéristique des images RSO, est obtenue par sommation incohérente de plusieurs sous-bandes lors des traitements en distance et en azimut, d'où une dégradation de la résolution spatiale de ces images. Ces images sont, le plus souvent, rééchantillonnées en géométrie terrain²⁰ et corrigées des

²⁰La projection est effectuée sans prise en compte du relief, c'est-à-dire sans utilisation de MNT.

effets de distorsions liées à la variation de l'angle d'incidence le long de l'axe radial, variation négligeable dans le cas des capteurs spatiaux mais possiblement importante pour les capteurs aéroportés.

Dans un contexte d'extraction de bâtiments à partir d'images **RSO**, les images monovues complexes sont préférées aux autres types d'images puisqu'elles présentent la meilleure résolution spatiale et que la géométrie de prise de vue n'a pas été modifiée, évitant ainsi les effets de distorsion dont peuvent être affectés les éléments du relief et du sur-sol lors d'un rééchantillonnage en géométrie terrain.

1.3 Conclusion

Nous avons succinctement présenté, dans ce chapitre, les capteurs de télédétection optiques et **RSO**, dont nous avons précisé le fonctionnement et la géométrie d'acquisition, avant d'introduire les principales caractéristiques des images issues de ces capteurs.

Plusieurs notions, relatives à l'imagerie de télédétection et sur lesquelles s'appuient les travaux réalisés au cours de cette thèse, n'ont volontairement pas été abordées et font l'objet de développements détaillés dans les chapitres où elles interviennent.

Plus particulièrement, nous introduisons, au chapitre 2, les fonctions de localisation au moyen desquelles sont modélisées les géométries de prise de vue des images optiques et **RSO**. Ces fonctions de localisation permettent notamment d'expliquer les effets géométriques affectant les éléments du sur-sol, dont les bâtiments, lors de leur projection de l'espace terrain vers l'espace image.

De plus, nous décrivons, au chapitre 4, les phénomènes physiques qui sont à l'origine des bruits observés dans les images optiques et **RSO**. La description de ces phénomènes justifie les lois statistiques qui interviennent dans la méthode d'extraction de bâtiments proposée dans ce même chapitre et au moyen desquelles est modélisé le bruit présent dans les images.

Signature géométrique des bâtiments en imagerie de télédétection optique et RSO

Sommaire

2.1	Introduction	24
2.2	Modélisation des bâtiments	25
2.3	Rendu tridimensionnel des bâtiments en imagerie optique	26
2.3.1	Projections entre l'espace terrain et l'espace image	26
2.3.2	Effets géométriques liés à la projection des bâtiments	29
2.3.3	Simulation de la signature géométrique d'un bâtiment	34
2.3.4	Exemples de signatures géométriques simulées	37
2.4	Rendu tridimensionnel des bâtiments en imagerie RSO	39
2.4.1	Projections entre l'espace terrain et l'espace image	39
2.4.2	Effets géométriques liés à la projection des bâtiments	41
2.4.3	Simulation de la signature géométrique d'un bâtiment	45
2.4.4	Exemples de signatures géométriques simulées	50
2.5	Conclusion	51

Nous nous intéressons, dans ce chapitre, à la signature géométrique des bâtiments dans les images optiques et RSO haute résolution. La signature géométrique d'un bâtiment dans une image correspond à l'ensemble des structures géométriques de l'image caractéristiques de la présence du bâtiment. Elle résulte de la projection de ce bâtiment, défini comme un élément de l'espace terrain (espace tridimensionnel), dans l'espace image (espace bidimensionnel).

Nous introduisons, dans un premier temps, les fonctions de localisation au moyen desquelles sont modélisées les géométries de prise de vue des capteurs optiques et RSO. Nous présentons ensuite les concepts de photogrammétrie et de radargrammétrie permettant d'interpréter les signatures géométriques des bâtiments dans les images issues de ces capteurs. Nous proposons, dans un dernier temps, des méthodes permettant de simuler, pour un modèle de bâtiment et pour une géométrie de prise de vue donnés, la signature géométrique d'un bâtiment dans une image optique ou RSO.

2.1 Introduction

En imagerie de télédétection, les objets de la scène visualisée sont caractérisés, dans l'image, par un ensemble de structures géométriques (structures ponctuelles, linéiques ou surfaciques) qui présentent des attributs radiométriques (radiométrie moyenne, texture) également caractéristiques de ces objets. La combinaison de ces éléments géométriques et radiométriques définissent la signature de l'objet dans cette image.

La composante radiométrique de la signature d'un objet dans une image résulte de l'interaction rayonnement-matière entre l'onde électromagnétique provenant de la source d'illumination et l'objet mais aussi de l'interaction de cette même onde avec le milieu de propagation, à savoir l'atmosphère. La signature radiométrique d'un objet dans une image, qui correspond à la mesure du rayonnement résultant de cette interaction rayonnement-matière, dépend des caractéristiques spectrales du capteur avec lequel l'image a été acquise. Plus précisément, elle dépend du nombre, de la nature et de la largeur des bandes spectrales des images issues de ce capteur.

En ce qui concerne la composante géométrique de la signature d'un objet dans une image, elle résulte de la projection de cet objet, défini comme un élément de l'espace terrain, dans l'espace image. Cette composante dépend de la géométrie de l'objet tridimensionnel imagé ainsi que de la géométrie du capteur qui conditionne la prise de vue selon laquelle l'objet est projeté de l'espace terrain dans l'espace image.

L'objectif de ce chapitre est d'établir des méthodes permettant de simuler la signature géométrique d'un bâtiment dans les images optiques et RSO. De manière générale, la simulation d'images à partir de scènes tridimensionnelles se décompose en deux étapes :

1. La modélisation, dans l'espace terrain, des objets tridimensionnels constitutifs de la scène. Parmi les différentes méthodes de modélisation des objets, nous distinguons, comme cela est décrit dans [LL00], la modélisation par fils de fer (ensemble de points et d'arêtes reliant certains de ces points entre eux), la modélisation par facettes polygonales planes, la modélisation par surfaces gauches (comme les quadriques) ou encore la modélisation au moyen de solides (l'espace est alors divisé en cubes élémentaires, appelés voxels) ;
2. Le rendu tridimensionnel de la scène qui résulte de la projection dans l'espace image des objets définis dans l'espace terrain lors de l'étape de modélisation. La projection est réalisée au moyen de fonctions mathématiques modélisant la géométrie de prise de vue de l'image. La géométrie de prise de vue dépend du type de capteur ainsi que de la géométrie de ce capteur.

Conformément aux deux étapes décrites précédemment, nous présentons, dans ce chapitre, le modèle avec lequel les bâtiments sont décrits dans l'espace terrain, avant d'aborder dans les deux parties suivantes le rendu tridimensionnel des bâtiments dans les images de télédétection optiques et RSO.

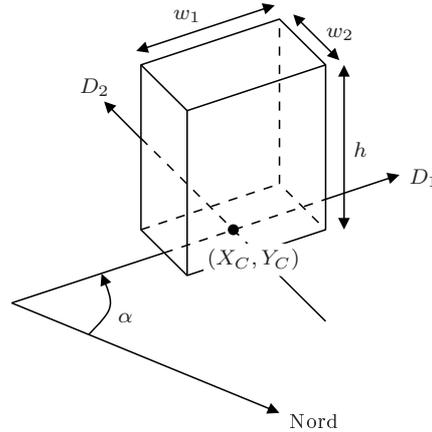


Figure 2.1 – Définition des paramètres du modèle parallélépipédique de bâtiment.

2.2 Modélisation des bâtiments

Dans le présent manuscrit, les bâtiments sont modélisés, dans l'espace terrain, au moyen de parallélépipèdes rectangles, de sorte que seuls les bâtiments présentant une emprise au sol rectangulaire et un toit horizontal sont considérés. Dans le cadre de cette modélisation, un bâtiment est entièrement défini, comme illustré à la figure 2.1, par un vecteur \vec{b} constitué de six paramètres :

$$\vec{b} = (X_C, Y_C, \alpha, w_1, w_2, h)^T. \quad (2.1)$$

X_C et Y_C correspondent aux coordonnées planimétriques du centre de l'emprise au sol du bâtiment dans un repère terrain \mathcal{R}_T lié à la Terre. Ces coordonnées peuvent être exprimées :

- directement au sein d'un système de coordonnées géodésiques, le plus connu étant le système WGS84¹ associé au système de positionnement GPS ;
- ou bien au sein d'un système de coordonnées résultant d'une projection cartographique de la surface de la Terre, comme le système de coordonnées résultant de la projection conforme UTM².

Nous supposons, dans le présent document, disposer d'un MNT, de sorte que la coordonnée altimétrique du centre de l'emprise au sol du bâtiment est donnée par la valeur de ce MNT en (X_C, Y_C) .

α correspond à l'angle entre la direction du Nord géographique et la direction principale du bâtiment. Parmi les deux directions associées à l'emprise au sol rectangulaire du bâtiment, la direction principale (D_1) est celle pour laquelle l'angle formé avec la direction du Nord géographique appartient à l'intervalle $[0; \pi/2[$. La direction secondaire (D_2) est alors définie comme la direction orthogonale à la direction principale.

Enfin, w_1 correspond à la largeur du bâtiment le long de la direction principale D_1 , w_2 correspond à la largeur du bâtiment le long de la direction secondaire D_2 et h correspond à la hauteur du bâtiment.

Par la suite, nous notons \mathcal{E}_b l'ensemble auquel appartiennent les vecteurs \vec{b} décrivant les bâtiments. Cet ensemble peut s'écrire comme le produit cartésien des ensembles auxquels appartiennent les différents paramètres du bâtiment :

$$\mathcal{E}_b = \mathcal{E}_{X_C} \times \mathcal{E}_{Y_C} \times \mathcal{E}_\alpha \times \mathcal{E}_{w_1} \times \mathcal{E}_{w_2} \times \mathcal{E}_h, \quad (2.2)$$

où $\mathcal{E}_{X_C} \times \mathcal{E}_{Y_C}$ est un ensemble bidimensionnel inclus dans \mathbb{R}^2 auquel appartiennent les coordonnées (X_C, Y_C) et $\mathcal{E}_\alpha = [\alpha^{min}; \alpha^{max}]$, $\mathcal{E}_{w_1} = [w_1^{min}; w_1^{max}]$, $\mathcal{E}_{w_2} = [w_2^{min}; w_2^{max}]$, $\mathcal{E}_h = [h^{min}; h^{max}]$ correspondent aux intervalles auxquels appartiennent les paramètres α , w_1 , w_2 et h .

¹ World Geodetic System 1984

² Universal Transverse Mercator

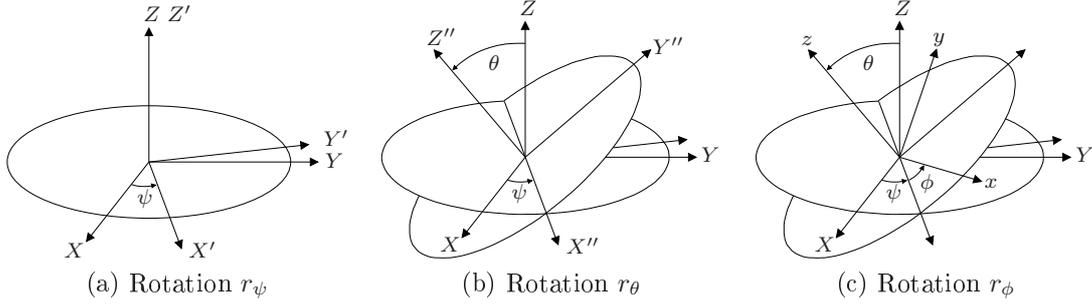


Figure 2.2 – Passage du repère terrain \mathcal{R}_T , défini par le système d'axes (X, Y, Z) au repère capteur \mathcal{R}_C , défini par le système d'axes (x, y, z) , par application successive des rotations r_ψ (a), r_θ (b), r_ϕ (c).

2.3 Rendu tridimensionnel des bâtiments en imagerie optique

Dans cette partie, nous présentons les fonctions mathématiques qui modélisent les projections entre l'espace terrain et l'espace image, avant d'aborder les effets géométriques qui interviennent lors de ces projections. Nous introduisons ensuite une méthode qui réalise la projection dans l'espace image des bâtiments définis, dans l'espace terrain, à l'aide du modèle présenté à la partie 2.2. Nous présentons finalement plusieurs exemples de signatures géométriques de bâtiments simulées au moyen de la méthode précédemment décrite.

2.3.1 Projections entre l'espace terrain et l'espace image

2.3.1.1 Projection de l'espace terrain vers l'espace image

La projection de l'espace terrain vers l'espace image, associé à une image I , est modélisée au moyen d'une fonction notée \mathcal{F}_I^i et appelée fonction de projection ou encore fonction de localisation inverse (en référence à la fonction de localisation directe introduite à la sous-section suivante). Nous donnons ici l'expression de cette fonction pour les capteurs *frame* (cf. section 1.1.2), pour lesquels la prise de vue associée est qualifiée de conique.

Nous notons respectivement (X_L, Y_L, Z_L) et (x_L, y_L, z_L) les coordonnées du centre de prise de vue L dans un repère orthonormé \mathcal{R}_T lié à la Terre et dans un repère orthonormé \mathcal{R}_C lié au capteur. Nous notons également (X_P, Y_P, Z_P) les coordonnées, dans le repère \mathcal{R}_T , d'un point P de l'espace terrain et (x_p, y_p, z_p) les coordonnées, dans le repère \mathcal{R}_C , du point p de l'espace image correspondant au point P .

Les coordonnées (x_p, y_p, z_p) et (X_P, Y_P, Z_P) sont alors reliées par l'équation suivante :

$$\begin{bmatrix} x_p - x_L \\ y_p - y_L \\ z_p - z_L \end{bmatrix} = k\mathbf{M} \begin{bmatrix} X_P - X_L \\ Y_P - Y_L \\ Z_P - Z_L \end{bmatrix}, \quad (2.3)$$

où k est un facteur d'échelle et \mathbf{M} est la matrice de passage du repère terrain \mathcal{R}_T vers le repère capteur \mathcal{R}_C .

Comme l'illustre la figure 2.2, le passage du repère \mathcal{R}_T , défini par le système d'axes (X, Y, Z) , au repère \mathcal{R}_C , défini par le système d'axes (x, y, z) , se fait usuellement par application successive de trois rotations r_ψ , r_θ et r_ϕ , qui désignent respectivement :

- la rotation d'angle ψ autour de l'axe Z du repère \mathcal{R}_T ;
- la rotation d'angle θ autour de l'axe X' , image de l'axe X du repère \mathcal{R}_T par la rotation r_ψ ;
- la rotation d'angle ϕ autour de l'axe Z'' , image de l'axe Z du repère \mathcal{R}_T obtenue par application successive des rotations r_ψ et r_θ .

Finalement, la matrice \mathbf{M} résultant de la composition de ces trois rotations s'exprime comme suit :

$$\mathbf{M} = \begin{bmatrix} \cos \phi \cos \psi - \cos \theta \sin \psi \sin \phi & \cos \phi \sin \psi + \cos \theta \cos \psi \sin \phi & \sin \phi \sin \theta \\ -\sin \phi \cos \psi - \cos \theta \sin \psi \cos \phi & -\sin \phi \sin \psi + \cos \theta \cos \psi \cos \phi & \cos \phi \sin \theta \\ \sin \theta \sin \psi & -\sin \theta \cos \psi & \cos \theta \end{bmatrix}. \quad (2.4)$$

En pratique, le repère lié au capteur \mathcal{R}_C est choisi de sorte que ses deux premiers axes sont contenus dans le plan focal de l'objectif et coïncident avec les axes du repère image \mathcal{R}_I , c'est-à-dire avec les lignes et les colonnes de l'image I . En choisissant l'intersection de la ligne de visée avec le plan focal comme origine du repère \mathcal{R}_C et en orientant le troisième axe dans le sens de la direction de visée, la fonction de localisation inverse \mathcal{F}_I^i dérive de l'équation (2.3) et s'exprime comme suit, après élimination du facteur d'échelle k :

$$\mathcal{F}_I^i \left\{ \begin{array}{l} \mathcal{E}_T \longrightarrow \\ \begin{bmatrix} X_P \\ Y_P \\ Z_P \end{bmatrix} \longrightarrow \end{array} \right. \begin{array}{l} \mathcal{E}_I \\ \begin{bmatrix} x_p \\ y_p \end{bmatrix} \end{array} = \begin{bmatrix} -f \frac{m_{11}(X_P - X_L) + m_{12}(Y_P - Y_L) + m_{13}(Z_P - Z_L)}{m_{31}(X_P - X_L) + m_{32}(Y_P - Y_L) + m_{33}(Z_P - Z_L)} \\ -f \frac{m_{21}(X_P - X_L) + m_{22}(Y_P - Y_L) + m_{23}(Z_P - Z_L)}{m_{31}(X_P - X_L) + m_{32}(Y_P - Y_L) + m_{33}(Z_P - Z_L)} \end{bmatrix}, \quad (2.5)$$

où f désigne la distance focale de l'objectif, m_{ij} correspond au coefficient de la matrice \mathbf{M} situé à la ligne i et à la colonne j et \mathcal{E}_T , \mathcal{E}_I désignent respectivement l'espace terrain et l'espace image.

Dans le cas d'un capteur *pushbroom*, la fonction de localisation inverse fait intervenir les géométries de prise de vue propres à chacune des lignes de l'image. Une barrette pouvant être assimilée à une matrice de détecteurs ne comportant qu'une seule ligne, la géométrie de prise de vue d'une ligne de l'image est régie par la même équation que pour un capteur *frame*. Finalement la fonction de localisation d'un capteur *pushbroom* résulte de l'association des fonctions de localisation correspondant à chacune des lignes de l'image.

Dans le cas des capteurs *wiskbroom*, chaque pixel de l'image possède sa propre géométrie de prise de vue, régie par l'équation (2.5), ce qui complique l'établissement des fonctions de localisation relatives à ces capteurs.

La mise en œuvre pratique de la fonction de localisation inverse, dont l'expression est donnée à l'équation (2.5), nécessite la connaissance d'un ensemble de paramètres physiques relatifs à la géométrie du capteur (distance focale f) ainsi qu'à la position et à l'orientation de ce dernier dans un référentiel lié à la Terre. Ces paramètres sont disponibles dans les métadonnées de l'image lorsque cette dernière est fournie avec un modèle physique.

Pour des raisons de simplicité d'utilisation, une image peut être fournie avec un modèle de remplacement qui approche le modèle physique de l'image et se substitue à ce dernier. Parmi les différents modèles de remplacement, nous distinguons notamment :

- les modèles grille, pour lesquels les valeurs de la fonction de localisation inverse \mathcal{F}_I^i sont fournies pour un ensemble de points de l'espace terrain. La valeur de la fonction de localisation inverse en un point quelconque de l'espace terrain est alors obtenue par interpolation des valeurs de cette fonction calculées sur la grille ;
- les modèles polynomiaux, pour lesquels la fonction de localisation inverse est approchée par un polynôme (modèle simple) ou bien par une fraction rationnelle (modèle RPC^3). Dans le cas des modèles polynomiaux, la seule donnée des coefficients du ou des polynômes suffit alors à définir la fonction de localisation inverse.

³ *Rational Polynomial Coefficients*

2.3.1.2 Projection de l'espace image vers l'espace terrain

Dans cette sous-section, nous introduisons la fonction de localisation directe, notée \mathcal{F}_I^d , qui permet de déterminer les coordonnées dans l'espace terrain \mathcal{E}_T d'un point de l'espace image \mathcal{E}_I .

La matrice de passage \mathbf{M} introduite à la sous-section précédente est une matrice orthogonale de sorte que l'inverse de cette matrice est égale à sa transposée \mathbf{M}^T . L'équation (2.3) peut ainsi être facilement inversée et conduit à l'expression des coordonnées (X_P, Y_P, Z_P) d'un point P de l'espace terrain en fonction des coordonnées (x_p, y_p, z_p) du point p correspondant dans l'espace image :

$$\begin{bmatrix} X_P - X_L \\ Y_P - Y_L \\ Z_P - Z_L \end{bmatrix} = \frac{1}{k} \mathbf{M}^T \begin{bmatrix} x_p - x_L \\ y_p - y_L \\ z_p - z_L \end{bmatrix}. \quad (2.6)$$

L'alignement des axes du repère \mathcal{R}_C sur ceux du repère image \mathcal{R}_I et l'élimination du facteur d'échelle k conduit alors au système d'équations suivant :

$$\begin{cases} X_P - X_L = (Z_P - Z_L) \frac{m_{11}x_p + m_{21}y_p + m_{31}(-f)}{m_{13}x_p + m_{23}y_p + m_{33}(-f)}, \\ Y_P - Y_L = (Z_P - Z_L) \frac{m_{12}x_p + m_{22}y_p + m_{32}(-f)}{m_{13}x_p + m_{23}y_p + m_{33}(-f)}. \end{cases} \quad (2.7)$$

Ainsi, la projection de l'espace image vers l'espace terrain conduit à un système à deux équations et trois inconnues. La solution de ce système est l'ensemble des points situés sur la droite de l'espace terrain qui passe par le centre de prise de vue L et par le point p du plan focal.

En exprimant les coordonnées du point P dans le repère ECEF⁴ et en considérant une modélisation ellipsoïdale de la surface terrestre, la connaissance *a priori* de l'altitude h du point de l'espace terrain fournit une contrainte de la forme :

$$\left(\frac{X_P}{A_T + h} \right)^2 + \left(\frac{Y_P}{B_T + h} \right)^2 + \left(\frac{Z_P}{C_T + h} \right)^2 = 1, \quad (2.8)$$

où A_T , B_T et C_T correspondent aux demi-axes de l'ellipsoïde modélisant la surface terrestre. L'intersection de la droite définie à l'équation (2.7) avec l'ellipsoïde augmenté, défini par l'équation (2.8), fournit au plus deux solutions analytiques. La seule physiquement acceptable est celle pour laquelle la distance au centre de prise de vue L est la plus faible.

L'altitude du point de l'espace terrain correspondant au point de l'espace image n'étant pas connue *a priori*, la contrainte précédente ne peut pas être utilisée directement. En pratique, la position dans l'espace terrain du point correspondant au point de l'espace image est déterminée au moyen d'un algorithme de lancer de rayon sur un MNT ou un MNS⁵, l'équation du rayon étant donnée par l'équation (2.7).

Parmi les différentes méthodes de lancer de rayon, nous distinguons notamment la méthode itérative consistant à déterminer l'intersection de ce rayon avec des ellipsoïdes qui approchent localement le MNT ou le MNS. Comme l'illustre la figure 2.3, cette méthode consiste à déterminer l'intersection du rayon et de l'ellipsoïde associé à une hauteur initiale h_0 fournissant ainsi une première solution (X_0, Y_0, Z_0) . Le procédé est alors répété à l'altitude h_1 , altitude fournie par le MNT pour le point (X_0, Y_0, Z_0) , conduisant ainsi à un point (X_1, Y_1, Z_1) et ainsi de suite. L'algorithme est arrêté lorsque la différence de hauteur entre deux itérations successives est inférieure à un seuil prédéfini.

Lorsqu'une image est fournie avec un modèle de remplacement, le rayon associé au pixel (x_p, y_p) de l'image ne peut être déterminé, de sorte que la méthode de lancer de rayon décrite précédemment ne peut être mise en œuvre. La localisation directe est alors calculée au moyen d'une inversion numérique de la fonction de localisation inverse \mathcal{F}_I^i .

⁴ Earth-Centered Earth-Fixed

⁵ Modèle Numérique de Surface

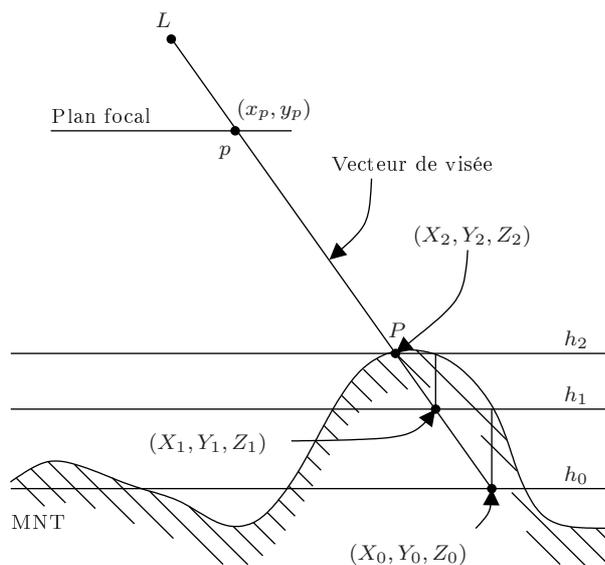


Figure 2.3 – Principe du lancer de rayon itératif sur un modèle numérique de terrain, présenté dans le cas d’une convergence en trois itérations.

2.3.2 Effets géométriques liés à la projection des bâtiments

Dans cette section, nous présentons les concepts photogrammétriques qui expliquent la signature géométrique résultant de la projection d’un bâtiment de l’espace terrain dans l’espace image. Nous abordons successivement les effets de perspective, les effets de déplacement et les effets d’ombrage auxquels sont sujets les objets du sur-sol, et plus particulièrement les bâtiments. Ces différentes notions sont présentées, sans perte de généralité, dans le cas particulier d’une prise de vue conique.

Avant d’aborder ces différentes notions, nous précisons les repères \mathcal{R}_T et \mathcal{R}_C qui sont adoptés dans cette section pour exprimer les coordonnées des points de l’espace terrain et de l’espace image.

Le repère terrain \mathcal{R}_T qui a été choisi est le repère géographique local. L’origine de ce repère est définie par le nadir N , point qui correspond, comme l’illustre la figure 2.4, à la projection orthogonale du centre de prise de vue L sur l’ellipsoïde. Les axes X et Y de ce repère sont contenus dans le plan tangent à l’ellipsoïde au point N , ces axes étant orientés respectivement vers l’Est et le Nord. L’axe Z est porté par la droite orthogonale à l’ellipsoïde qui relie le nadir N au centre de prise de vue L et est orienté de N vers L .

La définition du repère capteur \mathcal{R}_C est identique à celle adoptée à la section précédente. Son origine est définie par l’intersection de la ligne de visée avec le plan focal, comme le montre la figure 2.4. Les axes x et y sont contenus dans le plan focal de l’objectif et confondus avec ceux du repère image \mathcal{R}_I , tandis que l’axe z est porté par la ligne de visée et est orienté dans la direction de visée.

Étant données ces définitions, les angles ψ , θ et ϕ , introduits à la section précédente et qui permettent d’orienter le repère \mathcal{R}_C relativement au repère \mathcal{R}_T , sont représentés à la figure 2.4. La direction principale, qui est définie par la droite passant par le point n (projection du point N dans l’espace image) et le point principal (point d’intersection entre l’axe optique et le plan image), est également représentée à la figure 2.4.

2.3.2.1 Effets de perspective

La fonction de localisation directe, introduite à la section précédente, correspond à une projection perspective de l’espace terrain dans l’espace image, projection pour laquelle les longueurs, les angles, les rapports de longueurs et le parallélisme ne sont pas conservés. Ainsi, la projection d’un contour de

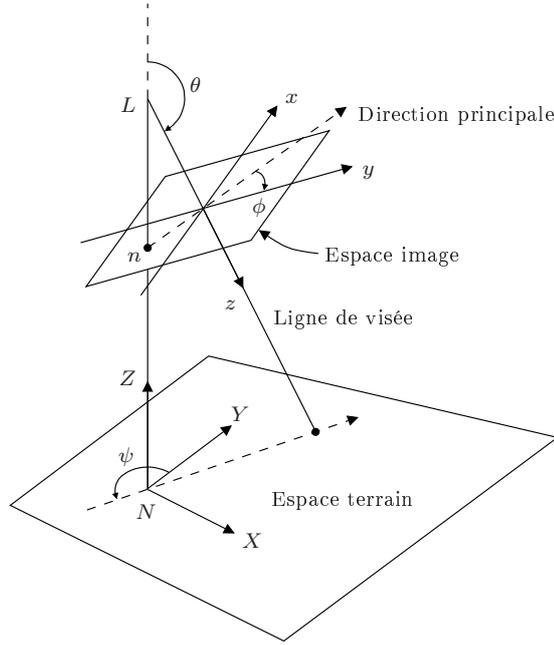


Figure 2.4 – Définition des repères \mathcal{R}_T et \mathcal{R}_C et représentation des angles ψ , θ et ϕ permettant d'orienter le repère \mathcal{R}_C relativement au repère \mathcal{R}_T .

toit rectangulaire dans l'espace terrain est un contour quadrilatéral quelconque dans l'espace image, comme illustré à la figure 2.5 (a).

Des droites qui sont parallèles dans l'espace terrain ne le sont plus dans l'espace image et se croisent en un point unique, appelé point de fuite, qui est caractéristique de la direction des droites parallèles dans l'espace terrain. Ainsi, il existe, dans une image, autant de points de fuite que de jeux de droites parallèles dans l'espace terrain.

Comme illustré à la figure 2.5 (a), les points de fuite (a et b) associés aux droites parallèles qui sont horizontales dans l'espace terrain sont alignés, dans l'espace image, sur une droite définissant l'horizon. De plus, le point de fuite associé aux droites verticales dans l'espace terrain est confondu, comme illustré à la figure 2.4, avec le point n , qui correspond à la projection du nadir dans l'espace image.

Dans le cadre d'une approximation de la projection perspective par une projection orthogonale, le parallélisme des droites est conservé, de sorte que la projection d'un contour de toit rectangulaire dans l'espace terrain est désormais un parallélogramme dans l'espace image, comme illustré à la figure 2.5 (b). Les angles ne sont, en revanche, toujours pas conservés et il peut être montré qu'un angle droit dans l'espace terrain est projeté dans l'espace image en un angle β dont l'expression est donnée par la formule suivante :

$$\beta = \arctan\left(\frac{\mu}{\nu}\right), \quad (2.9)$$

avec :

$$\mu = 1 + \tan^2(\theta) \cos^2(\phi - \alpha), \quad (2.10)$$

$$\nu = \tan^2(\theta) \cos(\phi - \alpha) \sin(\phi - \alpha), \quad (2.11)$$

où α désigne l'angle que fait l'un des segments qui porte l'angle β avec l'axe x de l'image, comme illustré à la figure 2.5 (b).

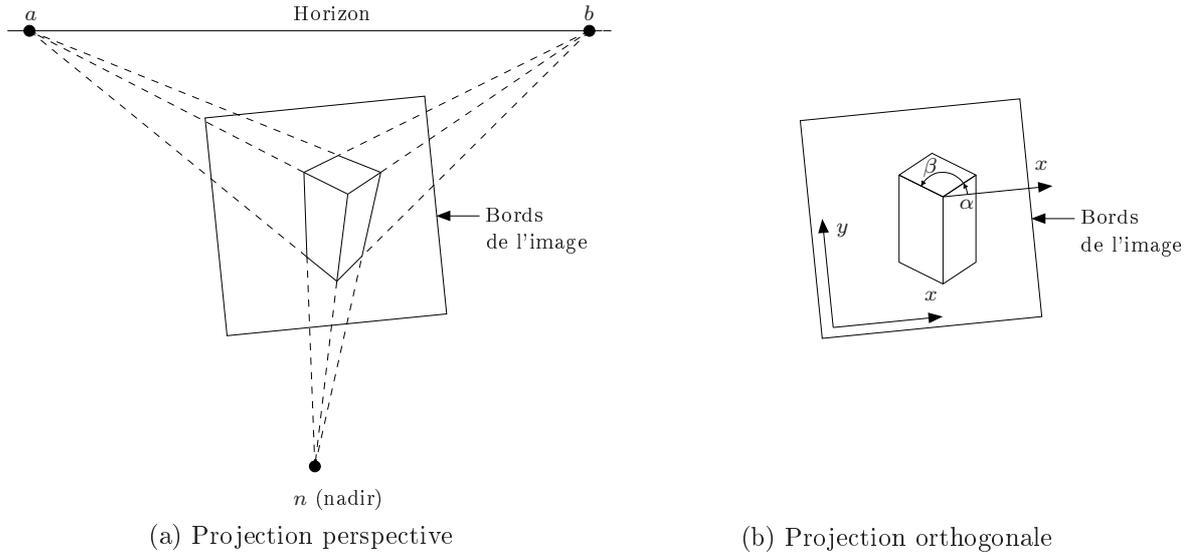


Figure 2.5 – Rendu tridimensionnel d'un bâtiment en imagerie optique dans le cas d'une projection perspective (a) et d'une projection orthogonale (b) (images extraites de [MBM01]).

2.3.2.2 Effets de déplacement

Comme nous l'avons vu à la sous-section précédente, une structure verticale dans l'espace terrain est projetée dans l'espace image selon une droite qui passe par le point n , projection dans l'espace image du nadir N . Dans cette sous-section, nous souhaitons établir la relation entre la hauteur (dans l'espace terrain) d'une structure verticale et la longueur du déplacement (dans l'espace image) entre le haut et le bas de cette structure.

Pour ce faire, nous considérons une structure verticale de hauteur h , définie dans l'espace terrain par les points S et T , qui correspondent respectivement aux extrémités basse et haute de cette structure.

Comme illustré à la figure 2.6 (a), les points s et t , qui résultent de la projection des points S et T dans l'espace image, sont portés par une droite qui passe par le point n . L'angle formé par cette droite avec la direction principale est noté γ .

Comme indiqué à la figure 2.6 (b), nous notons q le projeté orthogonal du centre de prise de vue L sur la droite (st) et Q le point de l'espace terrain associé à ce point.

Il peut être montré que les angles δ , ϵ et κ , formés par la droite (LN) avec les droites (LQ) , (LS) et (LT) , s'expriment comme suit :

$$\delta = \frac{\pi}{2} - \arccos(\sin \theta \cos \gamma), \quad (2.12)$$

$$\epsilon = \arctan\left(\frac{D}{H}\right), \quad (2.13)$$

$$\kappa = \arctan\left(\frac{D}{H-h}\right), \quad (2.14)$$

où H correspond à l'altitude du centre de prise de vue L et D est la distance de la structure verticale au nadir.

La distance d_{st} entre les points s et t peut alors s'exprimer en fonction de la distance focale f et des angles δ , ϵ , κ et θ :

$$d_{st} = -f \frac{\cos \delta}{\cos \theta} (\tan(\kappa - \delta) - \tan(\epsilon - \delta)). \quad (2.15)$$

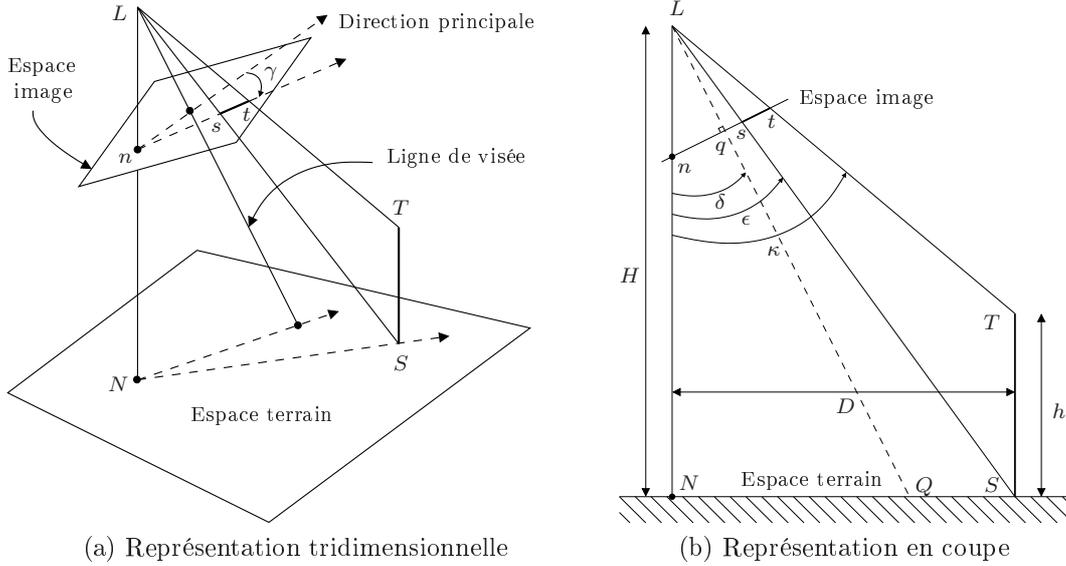


Figure 2.6 – Effet de déplacement sur une structure verticale lors de sa projection de l'espace terrain dans l'espace image.

D'après l'équation précédente, établie dans le cas rigoureux d'une projection perspective, la longueur du déplacement d_{st} dans l'espace image entre le haut et le bas de la structure verticale dépend de la hauteur h de cette structure et de sa distance D au nadir N , par l'intermédiaire des angles ϵ et κ . Une structure verticale localisée au niveau du nadir ne subit aucun effet de déplacement tandis qu'une structure verticale située au bord d'une image apparaît en s'inclinant selon la direction du nadir en s'éloignant de ce point.

Ces observations ressortent plus nettement dans le cas particulier d'une prise de vue verticale ($\theta = \pi$ et $\delta = 0$), où l'équation (2.15) se réécrit comme suit :

$$d_{st} = f \left(\frac{D}{H-h} - \frac{D}{H} \right) = d_{nt} \frac{h}{H}, \quad (2.16)$$

où d_{nt} correspond à la distance, dans l'espace image, entre les points n et t .

Dans le cadre d'une approximation de la prise de vue par une projection orthogonale, l'effet de déplacement est homogène sur toute l'image. La longueur du déplacement dans l'espace image entre le haut et le bas de la structure verticale ne dépend plus de la position de cette structure dans l'image. Pour une prise de vue verticale ($\theta = \pi$), ce déplacement est donc considéré comme nul pour l'ensemble des objets présents dans l'image. En revanche, pour une prise de vue oblique ($\theta \neq \pi$), le déplacement d_{st} est donné par la formule suivante :

$$d_{st} = hk \sin \theta, \quad (2.17)$$

où k est le facteur d'échelle entre l'espace terrain et l'espace image, introduit à la section précédente.

Ainsi, il est possible de mesurer la hauteur d'un bâtiment à partir d'une unique image en exploitant les effets de déplacement, hormis lorsque le bâtiment se trouve au nadir puisque qu'il n'est, dans ce cas, sujet à aucun effet de déplacement.

2.3.2.3 Effets d'ombrage

L'ombre projetée par un bâtiment résulte de l'absence d'illumination directe d'une partie de la scène par le soleil en raison de la présence de ce bâtiment. La projection, de l'espace terrain vers l'espace image, de l'ombre projetée au sol par un bâtiment est également régie par la fonction de

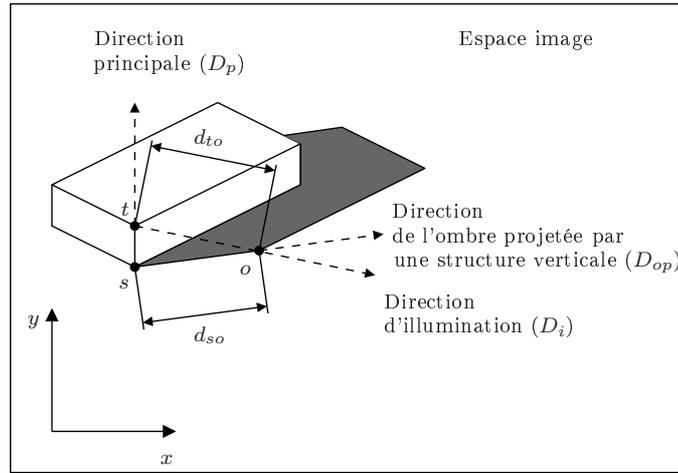


Figure 2.7 – Définition de la longueur de l'ombre au sol (d_{so}) et de la longueur de l'ombre projetée (d_{to}) dans l'espace image (image extraite de [LN98]).

localisation définie à la section 2.3.1. Ainsi, la projection des ombres conduit aux mêmes effets de perspective et de déplacement que ceux observés lors de la projection des bâtiments. En supposant que le sol est localement plat dans un voisinage des bâtiments considérés, hypothèse qui est faite dans la suite de ce manuscrit, l'ombre projetée par un bâtiment correspond à une surface plane d'altitude constante. Ainsi, les ombres ne sont pas affectées par les effets de déplacement lors de la projection de l'espace terrain vers l'espace image.

La hauteur h d'un bâtiment peut être reliée à la longueur de l'ombre au sol d_{so} , qui correspond, comme indiqué à la figure 2.7, à la distance dans l'espace image entre un point de l'emprise au sol (point s) et le point homologue du contour de l'ombre projetée (point o) :

$$d_{so} = hk \tan i, \quad (2.18)$$

où k est le facteur d'échelle entre l'espace terrain et l'espace image, introduit à la section précédente, et i correspond à l'angle d'incidence des rayons solaires.

La hauteur h d'un bâtiment peut également être reliée à la longueur de l'ombre projetée d_{to} , qui correspond, comme indiqué à la figure 2.7, à la distance dans l'espace image entre un point du contour du toit (point t) et le point homologue du contour de l'ombre projetée (point o). Dans le cadre d'une approximation de la projection perspective par une projection orthogonale, la longueur de l'ombre projetée d_{to} est reliée à la hauteur h par l'expression suivante :

$$d_{to} = \begin{cases} hk \tan i & \text{si } \theta = \pi \\ \frac{hk \sin(\theta - i)}{\cos i} & \text{si } \begin{cases} \theta \neq \pi \\ \omega_{D_{op}} = \omega_{D_i} = \omega_{D_p} + \pi \end{cases} \\ -\frac{hk \sin(i + \theta)}{\cos i} & \text{si } \begin{cases} \theta \neq \pi \\ \omega_{D_{op}} = \omega_{D_i} = \omega_{D_p} \end{cases} \\ \frac{hk \sin \theta |\sin(\omega_{D_{op}} - \omega_{D_p})|}{|\sin(\omega_{D_{op}} - \omega_{D_i})|} & \text{sinon} \end{cases}, \quad (2.19)$$

où ω_{D_p} , ω_{D_i} et $\omega_{D_{op}}$ correspondent respectivement aux angles formés par l'axe x de l'image avec la direction principale, la direction d'illumination et la direction de l'ombre projetée par une structure verticale.

Dans le cas d'une prise de vue verticale ($\theta = \pi$), les effets de déplacement sur les arêtes verticales sont nuls (pour une projection orthogonale), de sorte que les points s et t de la figure 2.7 sont confondus. Les distances d_{to} et d_{so} sont alors égales et correspondent à la longueur de l'ombre au sol : $d_{to} = d_{so} = hk \tan i$, d'où l'expression de l'équation (2.19).

Dans le cas d'une prise de vue oblique ($\theta \neq \pi$), l'expression générale de la hauteur h est obtenue par application de la loi des sinus dans le triangle formé par les points t , s et o . Dans le cas particulier où ces points sont alignés, la longueur d_{to} s'écrit comme la somme ou la différence (selon l'ordre dans lequel les points sont alignés) des longueurs d_{st} et d_{so} , ce qui conduit aux expressions de l'équation (2.19).

2.3.3 Simulation de la signature géométrique d'un bâtiment

Dans cette section, nous décrivons la méthode mise en œuvre pour simuler la signature géométrique d'un bâtiment dans une image. L'objectif est de modéliser au mieux les effets relatifs à la projection des bâtiments qui ont été décrits à la section 2.3.2.

Nous supposons disposer d'un bâtiment \vec{b} défini dans l'espace terrain \mathcal{E}_T au moyen des paramètres du modèle parallélépipédique introduit à la partie 2.2.

Nous notons $\mathcal{S}_I(\vec{b})$ la signature géométrique du bâtiment \vec{b} dans une image I , dont la géométrie de prise de vue est supposée connue. De manière générale, la signature d'un bâtiment dans une image peut être décrite indifféremment comme un ensemble de N_r régions R_j (contenant éventuellement des trous) ou de N_s segments S_k définis dans l'espace image \mathcal{E}_I :

$$\mathcal{S}_I(\vec{b}) = \{R_j, j = 1 \dots N_r\}, \quad (2.20)$$

$$\mathcal{S}_I(\vec{b}) = \{S_k, k = 1 \dots N_s\}. \quad (2.21)$$

Dans le cas particulier d'une modélisation parallélépipédique des bâtiments, la signature géométrique de ces bâtiments s'écrit :

$$\mathcal{S}_I(\vec{b}) = \{R_T, R_O\} \cup \{R_F^i, i = 1 \dots N_f\}, \quad (2.22)$$

où N_f correspond au nombre de façades visibles dans l'image I et R_T , R_O , R_F^i désignent respectivement les régions de l'image I associées au toit, à l'ombre et à la $i^{\text{ème}}$ façade visible du bâtiment.

La méthode que nous proposons pour simuler la signature géométrique $\mathcal{S}_I(\vec{b})$ résultant de la projection du bâtiment \vec{b} dans l'image I se décompose, comme illustré à la figure 2.8 :

1. Projection de l'emprise au sol du bâtiment de l'espace terrain \mathcal{E}_T vers l'espace image \mathcal{E}_I ;
2. Détermination de la région R_T correspondant au toit du bâtiment ;
3. Détermination des régions R_F^i correspondant aux façades visibles du bâtiment ;
4. Détermination de la région R_O correspondant à l'ombre projetée par le bâtiment.

2.3.3.1 Projection de l'emprise au sol du bâtiment de l'espace terrain vers l'espace image

La première étape de la méthode consiste à projeter dans l'espace image \mathcal{E}_I les points définissant l'emprise au sol du bâtiment dans l'espace terrain \mathcal{E}_T , comme illustré à la figure 2.8 (a). Ces points, dont les coordonnées dans l'espace terrain sont calculées à partir des coordonnées (X_C, Y_C) du centre de l'emprise au sol et des paramètres α , w_1 et w_2 , sont projetés dans l'espace image au moyen de la fonction de localisation \mathcal{F}_I^i .

D'après l'équation (2.5), la projection d'un point de l'espace terrain vers l'espace image au moyen de la fonction de localisation nécessite la connaissance de l'altitude de ce point, dont nous supposons ici qu'elle est fournie par un MNT. En pratique, il est courant que le MNT fournisse des altitudes différentes pour les différents points de l'emprise au sol ce qui engendre des effets de déplacements différents pour ces points lors de la projection de l'espace terrain vers l'espace image. Il en résulte

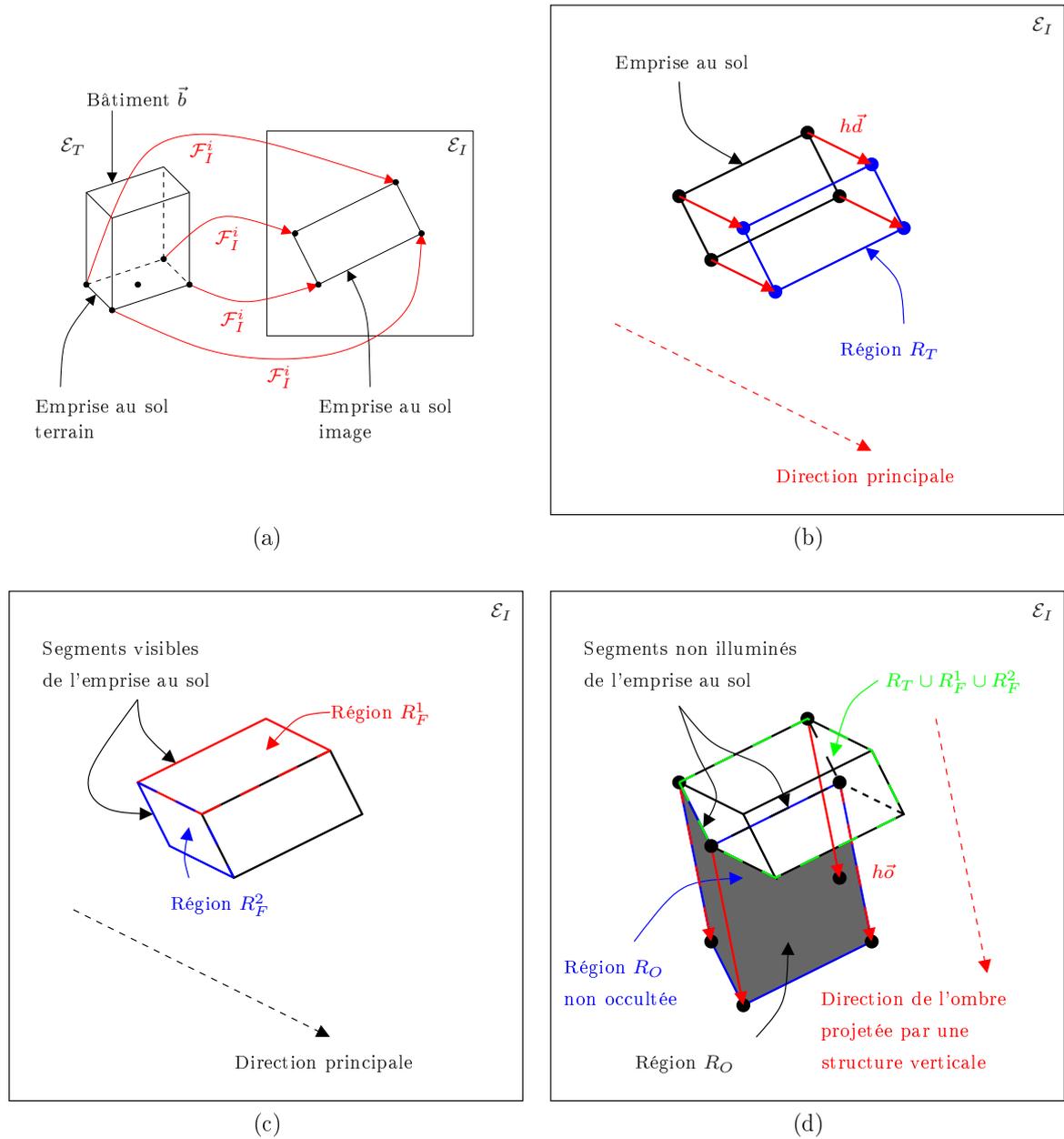


Figure 2.8 – Simulation de la signature géométrique d'un bâtiment parallélépipédique en imagerie optique : projection de l'emprise au sol de l'espace terrain vers l'espace image (a), détermination de la région R_T (b), détermination des régions R_F^i (c) et détermination de la région ombre R_O (d).

finaleme nt une distorsion de l'emprise au sol dans l'image qui est liée uniquement à l'imperfection du MNT utilisé. Afin de remédier à ce phénomène de distorsion, les points de l'emprise au sol sont projetés dans l'espace image en utilisant une altitude identique pour tous les points, à savoir l'altitude du centre de l'emprise au sol.

Notons que l'emprise au sol ainsi définie dans l'espace image, à l'issue de cette première étape, présente une forme quadrilatérale quelconque qui intègre les effets de perspective mis en évidence à la sous-section 2.3.2.1.

2.3.3.2 Détermination de la région R_T correspondant au toit du bâtiment

La deuxième étape de la méthode consiste à définir la région R_T de l'espace image correspondant au toit du bâtiment à partir de l'emprise au sol définie dans l'espace image à l'étape précédente. Nous considérons, lors de cette seconde étape, que la projection perspective peut être approchée par une projection orthogonale, de sorte que les structures verticales sont toutes parallèles à la direction principale. En pratique, la direction principale est déterminée en considérant deux points de l'espace image a et b , de coordonnées respectives (x_a, y_a) et (x_b, y_b) . Ces points résultent de la projection de deux points terrain A et B , dont les coordonnées planimétriques sont identiques ($X_B = X_A$ et $Y_B = Y_A$) et dont la différence d'altitude est arbitrairement fixée à une valeur H ($Z_B = Z_A + H$) :

$$(x_a, y_a) = \mathcal{F}_I^i(X_A, Y_A, Z_A), \quad (2.23)$$

$$(x_b, y_b) = \mathcal{F}_I^i(X_B, Y_B, Z_B). \quad (2.24)$$

L'effet de déplacement observé lors de la projection d'une structure verticale dans l'image, dont nous supposons qu'il est linéairement lié à la hauteur de cette structure, est alors quantifié au moyen du vecteur \vec{d} correspondant au déplacement associé à une structure verticale d'une hauteur de 1 mètre :

$$\vec{d} = \frac{\vec{ab}}{H}. \quad (2.25)$$

Comme le montre la figure 2.8 (b), les points de l'espace image définissant le contour du toit, et donc la région R_T , sont finalement déterminés par translation, dans l'espace image, des points de l'emprise au sol d'un vecteur $h\vec{d}$, où h correspond à la hauteur du bâtiment.

2.3.3.3 Détermination des régions R_F^i correspondant aux façades visibles du bâtiment

La troisième étape de la méthode consiste à déterminer les façades visibles du bâtiment. Pour ce faire, nous déterminons les segments de l'emprise au sol qui sont visibles en utilisant la direction principale de l'image. À chacun de ces segments visibles est associé un des segments du contour du toit, définissant ainsi une région de l'image correspondant à une façade visible, comme l'illustre la figure 2.8 (c).

2.3.3.4 Détermination de la région R_O correspondant à l'ombre projetée du bâtiment

La région R_O correspondant à l'ombre projetée par le bâtiment est finalement déterminée dans la dernière étape de cette méthode.

La direction de l'ombre projetée par une structure verticale dans l'espace image est tout d'abord calculée à partir de la direction d'illumination contenue dans les métadonnées de l'image. Nous notons \vec{o} le vecteur reliant la base d'une structure verticale de 1 mètre de hauteur à l'extrémité de l'ombre projetée par cette structure (la norme de ce vecteur est donnée par l'équation (2.18)). L'ombre projetée par chacune des arêtes verticales du bâtiment est alors déterminée en translatant les points de l'emprise au sol d'un vecteur $h\vec{o}$.

En considérant les segments de l'emprise au sol non illuminés par le soleil, nous déterminons ensuite les régions correspondant à l'ombre projetée par les façades non illuminées par le soleil. L'union de ces

régions conduit alors à la région R_O , identifiée par la couleur bleue dans la figure 2.8 (d) et pour laquelle une occlusion éventuelle par une partie du bâtiment n'est pas prise en compte. La partie de l'ombre occultée par le bâtiment, région résultant de l'intersection entre les régions R_O et $R_T \cup R_F^1 \cup R_F^2$, est finalement retranchée à la région R_O , ce qui conduit à la région grisée de la figure 2.8 (d).

2.3.4 Exemples de signatures géométriques simulées

Dans cette section, nous mettons en œuvre la méthode proposée à la section précédente afin de simuler les signatures géométriques de plusieurs bâtiments dans une image optique.

L'image dont nous disposons, présentée à la figure 2.9, a été acquise par le capteur spatial Quick-Bird le 22 mars 2002 sur la région de Marseille (France). Cette image, dont la géolocalisation est réalisée au moyen d'un modèle de remplacement RPC, est une image panchromatique dont la résolution est comprise entre 0.9 et 1 mètre. L'angle formé par la ligne de visée avec la direction du nadir est de 19.2 degrés et l'angle d'incidence des rayons solaires est proche de 45 degrés.

La scène étudiée comporte neuf bâtiments de forme parallélépipédique que nous numérotions de A à I . En utilisant le modèle de bâtiment introduit à la partie 2.2, chacun de ces bâtiments peut être décrit par un ensemble de six paramètres. Les valeurs de ces paramètres, présentées au tableau 2.1, ont été déterminées par saisie manuelle, dans l'image QuickBird, des sommets visibles du parallélépipède.

La précision avec laquelle les paramètres des bâtiments ont été déterminés peut être reliée à la précision de l'opération de pointé. Nous notons a , b et c , d les points de l'espace image correspondant respectivement aux extrémités d'une arête horizontale et d'une arête verticale d'un bâtiment parallélépipédique. Les incertitudes Δd_{ab} et Δd_{cd} associées à la mesure des longueurs d_{ab} et d_{cd} entre ces couples de points est donnée par la relation : $\Delta d_{ab} = \Delta d_{cd} = \sqrt{\Delta p_l^2 + \Delta p_c^2}$, où Δp_l et Δp_c désignent les erreurs de pointé selon les lignes et les colonnes de l'image. Les expressions des incertitudes ΔX_C , ΔY_C , Δw_1 , Δw_2 , $\Delta \alpha$ et Δh associées à la mesure des paramètres d'un bâtiment par saisie manuelle, dans l'image, des sommets visibles du parallélépipède sont alors données par les relations suivantes :

$$\Delta X_C = \Delta Y_C = \frac{\Delta d_{ab}}{\sqrt{2}k}, \quad (2.26)$$

$$\Delta w_1 = \Delta w_2 = \frac{\Delta d_{ab}}{k}, \quad (2.27)$$

$$\Delta \alpha = \frac{\Delta d_{ab}}{d_{ab}}, \quad (2.28)$$

$$\Delta h = \frac{\Delta d_{cd}}{k \sin \theta}, \quad (2.29)$$

où k et θ désignent respectivement le facteur d'échelle entre l'espace terrain et l'espace image et l'angle formé par la direction du nadir avec la ligne de visée.

En supposant que les erreurs de pointé liées à la saisie d'un point dans l'image sont identiques selon les lignes et les colonnes, nous avons finalement : $\Delta d_{ab} = \Delta d_{cd} = \sqrt{2}\Delta p$, avec $\Delta p = \Delta p_l = \Delta p_c$. Les incertitudes liées à la mesure des paramètres d'un bâtiment par saisie manuelle de ses sommets visibles dans l'image QuickBird de la figure 2.9 sont données au tableau 2.2 pour une erreur de pointé de un pixel selon les lignes et les colonnes ($\Delta p = 1$).

La figure 2.9 représente les signatures géométriques résultant de la projection des bâtiments, définis au tableau 2.1, dans l'image QuickBird. Nous constatons la bonne adéquation entre les signatures géométriques simulées et les bâtiments présents dans l'image, validant ainsi la méthode présentée à la section précédente. Notons toutefois que la méthode proposée suppose que les bâtiments sont isolés, de sorte qu'elle ne prend pas en compte les occlusions éventuelles entre les différents bâtiments. Nous pouvons notamment constater à la figure 2.9 que les ombres projetées par les bâtiments C et I sont en partie occultées par les bâtiments B et H , phénomène qui n'est pas modélisé par la méthode proposée.

Label	X_C (m)	Y_C (m)	α (degrés)	w_1 (m)	w_2 (m)	h (m)
A	702995.0	4795064.0	83.0	20.0	20.0	22.0
B	703025.0	4795026.0	83.0	20.0	20.0	15.0
C	703042.0	4794999.0	83.0	20.0	20.0	15.0
D	703053.0	4795074.0	83.0	20.0	20.0	22.0
E	703088.5	4795050.0	83.0	20.0	20.0	22.0
F	703120.0	4795076.0	83.0	20.0	20.0	22.0
G	703159.0	4795086.0	83.0	20.0	20.0	22.0
H	703236.0	4795092.0	83.0	20.0	20.0	15.0
I	703264.0	4795072.0	83.0	20.0	20.0	15.0

Tableau 2.1 – Paramètres des bâtiments de la scène étudiée pour un modèle parallélépipédique. Les coordonnées (X_C, Y_C) du centre de l'emprise au sol des bâtiments sont données dans le repère UTM (Zone 31 Nord).

ΔX_C (m)	ΔY_C (m)	$\Delta \alpha$ (degrés)	Δw_1 (m)	Δw_2 (m)	Δh (m)
0.6	0.6	2.7	0.9	0.9	2.7

Tableau 2.2 – Incertitude sur la mesure des paramètres d'un bâtiment par saisie manuelle des sommets visibles de ce bâtiment dans l'image QuickBird de la figure 2.9.



Figure 2.9 – Simulation de la signature géométrique de bâtiments dans une image QuickBird (©DigitalGlobe).

2.4 Rendu tridimensionnel des bâtiments en imagerie RSO

Comme en imagerie optique, nous introduisons, dans cette partie, les fonctions de localisation qui modélisent les projections entre l'espace terrain et l'espace image dans le cas des capteurs RSO. Nous décrivons ensuite les effets géométriques liés à la projection de l'espace terrain vers l'espace image avant de proposer une méthode pour simuler la signature géométrique des bâtiments dans les images RSO. Nous illustrons enfin la méthode développée en présentant des signatures géométriques de bâtiments obtenues au moyen de cette méthode.

2.4.1 Projections entre l'espace terrain et l'espace image

2.4.1.1 Projection de l'espace image vers l'espace terrain

Dans cette sous-section, nous introduisons la fonction de localisation directe \mathcal{F}_I^d qui permet de déterminer les coordonnées dans l'espace terrain \mathcal{E}_T d'un point dont les coordonnées sont connues dans l'espace image \mathcal{E}_I .

Nous notons (x_p, y_p) les coordonnées d'un point p de l'espace image, dans le repère \mathcal{R}_I lié à l'image, et (X_P, Y_P, Z_P) les coordonnées du point P de l'espace terrain correspondant, dans un repère \mathcal{R}_T lié à la Terre.

Nous supposons, par la suite, que le premier axe du repère \mathcal{R}_I est l'axe radial et que le second est l'axe azimutal, conformément à la géométrie des capteurs RSO décrite à la figure 1.7.

À partir des coordonnées (x_p, y_p) d'un point p de l'espace image, il est possible de déterminer la date t à laquelle a été acquis l'écho associé à ce point ainsi que la distance R séparant le capteur et le point P de l'espace terrain correspondant :

$$t = \frac{1}{F_{RI}} y_p + t_0, \quad (2.30)$$

$$R = \frac{c}{2F_e} x_p + R_0, \quad (2.31)$$

où c désigne la vitesse de la lumière, t_0 est la date d'acquisition de la première ligne de l'image et R_0 est la distance entre le capteur et le point de l'espace terrain qui apparaît dans la première case distance d'une ligne. Enfin, F_{RI} et F_e correspondent respectivement à la fréquence de répétition des impulsions et à la fréquence d'échantillonnage.

Une interpolation des éphémérides contenues dans les métadonnées de l'image permet de déterminer la position $S(t)$ du capteur dans le repère \mathcal{R}_T à la date t . Connaissant la position du capteur et la distance le séparant du point terrain cherché P , nous obtenons une première équation (équation de la sphère distance) qui contraint la position de ce point terrain sur une sphère centrée en $S(t)$ et de rayon R :

$$(X_P - X_S(t))^2 + (Y_P - Y_S(t))^2 + (Z_P - Z_S(t))^2 = R^2, \quad (2.32)$$

où $(X_S(t), Y_S(t), Z_S(t))$ désignent les coordonnées du point $S(t)$ dans le repère \mathcal{R}_T .

L'angle Doppler, noté θ conformément aux notations introduites à la figure 1.7, correspond à l'angle formé, à la date t , par le vecteur vitesse de la plateforme $\vec{V}_S(t)$ et le vecteur $\vec{S(t)P}$ reliant le capteur au point terrain cherché. La connaissance de cet angle Doppler fournit une seconde équation (équation du cône Doppler) qui contraint la position du point terrain P sur un cône de sommet $S(t)$ et de génératrice $\vec{V}_S(t)$:

$$\vec{V}_S \cdot \vec{S\dot{P}} = |\vec{V}_S| |\vec{S\dot{P}}| \cos \theta. \quad (2.33)$$

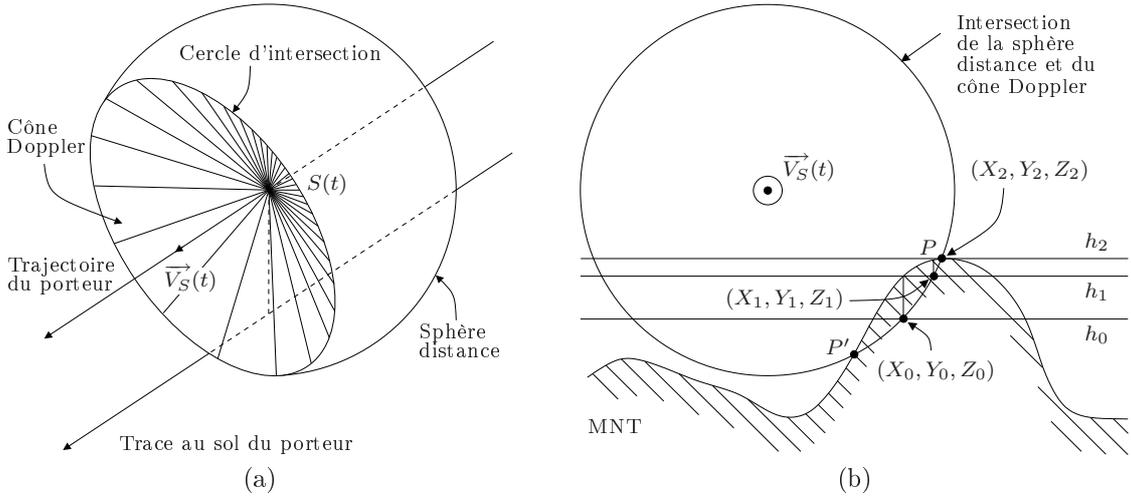


Figure 2.10 – Projection de l'espace image vers l'espace terrain : cercle d'intersection de la sphère distance et du cône Doppler (a) (image extraite de [MBM01]) et intersection de ce cercle avec un modèle numérique de terrain (b).

Le développement de cette équation conduit finalement à l'équation suivante :

$$\begin{aligned} \dot{X}_S(t)(X_P - X_S(t)) + \dot{Y}_S(t)(Y_P - Y_S(t)) + \dot{Z}_S(t)(Z_P - Z_S(t)) = \\ \cos \theta \left(\dot{X}_S(t)^2 + \dot{Y}_S(t)^2 + \dot{Z}_S(t)^2 \right)^{\frac{1}{2}} \left((X_P - X_S(t))^2 + (Y_P - Y_S(t))^2 + (Z_P - Z_S(t))^2 \right)^{\frac{1}{2}}, \quad (2.34) \end{aligned}$$

où $(\dot{X}_S(t), \dot{Y}_S(t), \dot{Z}_S(t))$ désignent les coordonnées du vecteur vitesse de la plateforme $\vec{V}_S(t)$ dans le repère \mathcal{R}_T .

En pratique, les images RSO fournies par les producteurs de données sont à « Doppler nul », c'est-à-dire que la datation associée à ces images correspond à celle d'une image qui aurait été acquise avec un angle Doppler de 90 degrés. Dans ce cas, le cône Doppler est un plan orthogonal au vecteur vitesse de la plateforme et le second membre de l'équation (2.34) est nul.

L'ensemble des points de l'espace terrain satisfaisant le système d'équations formé par les équations (2.32) et (2.34) est le cercle résultant de l'intersection de la sphère distance avec le cône Doppler, comme indiqué à la figure 2.10 (a). En supposant connue l'altitude h_P du point terrain cherché P , une modélisation ellipsoïdale de la surface terrestre (équation (2.8)) fournit une troisième équation (équation de surface). L'intersection du cercle sus mentionné avec l'ellipsoïde augmenté conduit finalement à deux solutions dans l'espace terrain. Ces solutions sont associées respectivement à une visée gauche et une visée droite du capteur RSO. Le système formé par l'équation de la sphère distance, l'équation du cône Doppler et l'équation de surface est un système de trois équations non linéaires à trois inconnues (X_P , Y_P et Z_P). Ce système peut être résolu analytiquement, cependant les résolutions numériques, telles que la méthode de Raphson ou la méthode de dichotomie, sont souvent préférées à la résolution analytique en raison de leur simplicité de mise en œuvre.

En pratique, l'altitude h_P du point terrain n'est pas connue *a priori*. L'intersection du cercle résultant de l'intersection sphère-cône avec un MNT ou un MNS peut être déterminée, comme illustré à la figure 2.10 (b), au moyen d'une méthode itérative similaire à celle présentée à la section 2.3.1.

Pour conclure, notons que l'intersection du MNT avec le cercle résultant de l'intersection sphère-cône peut conduire à plusieurs points distincts de l'espace terrain. Cette configuration, représentée à la figure 2.10 (b) (points P et P'), correspond à un effet de repliement et ne permet pas de distinguer les deux points dont l'énergie rétrodiffusée vers le capteur RSO intervient au sein de la même case distance dans l'image.

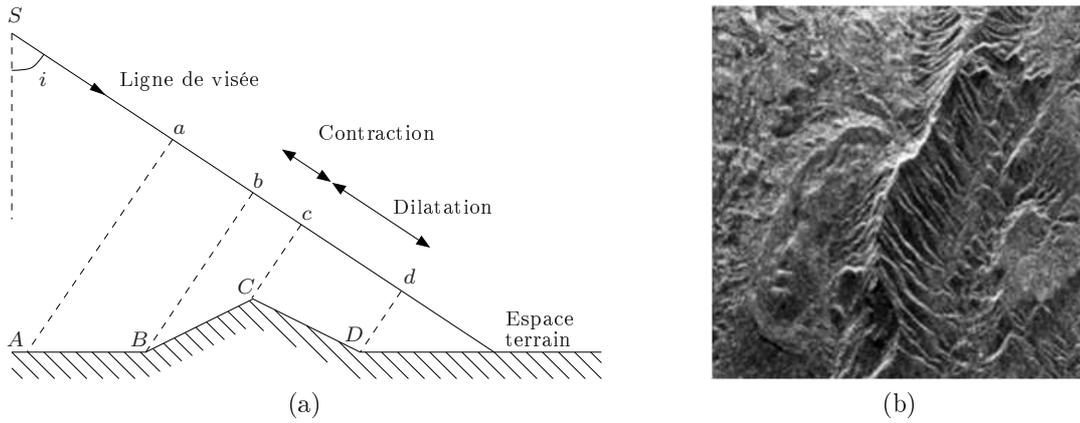


Figure 2.11 – Effets de contraction-dilatation en imagerie RSO : schéma illustratif (a) et exemple de contraction-dilatation lié à la présence d'une crête (b) (image extraite de [Den09]).

2.4.1.2 Projection de l'espace terrain vers l'espace image

La projection de l'espace terrain vers l'espace image est modélisée par la fonction de localisation inverse \mathcal{F}_I^i . Cette fonction permet de déterminer les coordonnées (x_p, y_p) du point p de l'espace image correspondant à un point P de l'espace terrain de coordonnées (X_P, Y_P, Z_P) connues.

La projection de l'espace terrain vers l'espace image est réalisée au moyen des équations de la sphère distance et du cône Doppler (équations (2.32) et (2.34)). La date t est déterminée numériquement à partir de l'équation du cône Doppler. La distance R , associée au point P , se déduit alors directement de l'équation de la sphère distance. Les coordonnées (x_p, y_p) du point p de l'espace image sont finalement déterminées par inversion des relations (2.30) et (2.31).

2.4.2 Effets géométriques liés à la projection des bâtiments

Dans cette section, nous décrivons les concepts radargrammétriques qui expliquent la signature géométrique résultant de la projection d'un bâtiment de l'espace terrain dans l'espace image. Nous abordons successivement les effets de contraction-dilatation, les effets de repliement et les effets d'ombrage auxquels sont sujets les éléments du relief ainsi que les objets du sur-sol, et plus particulièrement les bâtiments.

2.4.2.1 Effets de contraction-dilatation

Les effets de contraction (effets de *foreshortening*), qui se traduisent par un raccourcissement des distances en géométrie RADAR, interviennent lorsque l'onde est rétrodiffusée par une surface orientée dans la direction du capteur et dont la pente α est inférieure à l'angle d'incidence i de l'onde. Lorsque la pente de la surface est supérieure à l'angle d'incidence de l'onde, des effets de repliement (effets de *layover*), que nous décrivons à la sous-section suivante, apparaissent.

Les effets de dilatation, qui se traduisent par un allongement des distances en géométrie RADAR, interviennent lorsque l'onde est rétrodiffusée par une surface orientée dans la direction opposée à celle du capteur et dont la pente α est inférieure à $(\pi/2 - i)$. Lorsque la pente est supérieure à $(\pi/2 - i)$, des effets d'ombrage, décrits dans les paragraphes suivants, interviennent.

Nous considérons trois segments $[AB]$, $[BC]$ et $[CD]$ de longueur égale dans l'espace terrain ($d_{AB} = d_{BC} = d_{CD} = d$) mais présentant des pentes différentes, comme indiqué à la figure 2.11 (a). Le segment $[AB]$ est horizontal tandis que les deux autres segments forment un angle α avec l'horizontale, le segment $[BC]$ étant orienté dans la direction du capteur et le segment $[CD]$ étant orienté dans la direction opposée à celle du capteur. La longueur des segments $[ab]$, $[bc]$ et $[cd]$ résultant de la projection

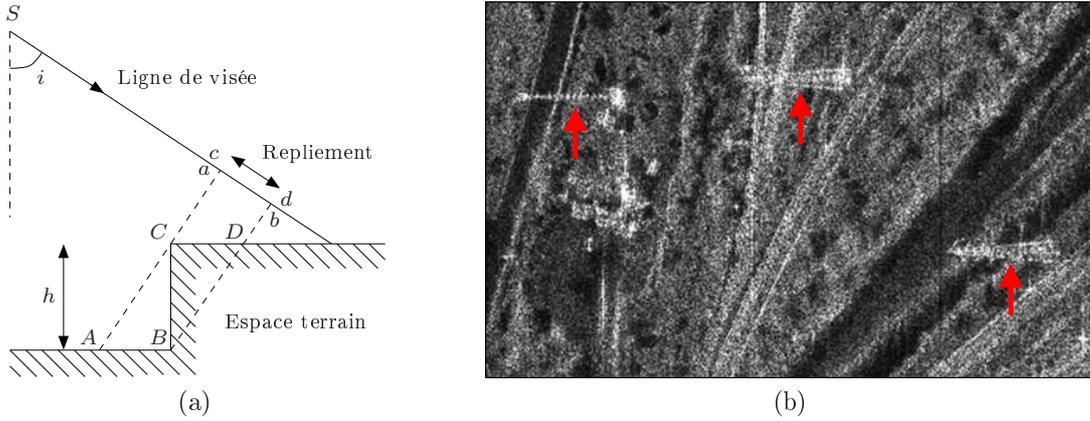


Figure 2.12 – Effets de repliement en imagerie RSO : schéma illustratif (a) et exemple de repliements liés à la présence de pylônes électriques (b) (image extraite de [Den09]).

des segments $[AB]$, $[BC]$ et $[CD]$ en géométrie RADAR est donnée par les formules suivantes :

$$d_{ab} = d \sin i, \quad (2.35)$$

$$d_{bc} = d \sin(i - \alpha), \quad (2.36)$$

$$d_{cd} = d \sin(i + \alpha). \quad (2.37)$$

Nous constatons que le segment $[BC]$ est sujet, lors de sa projection en géométrie RADAR, à un effet de contraction ($d_{bc} < d_{ab}$) tandis que le segment $[CD]$ est sujet à un phénomène de dilatation ($d_{cd} > d_{ab}$).

Lorsque la pente du terrain est orientée dans la direction du capteur, l'énergie électromagnétique réfléchiée est concentrée dans moins de cases distance que pour un terrain plat de sorte que la radiométrie de ces cases distance est plus importante que pour un terrain plat. Le phénomène inverse est observé dans le cas d'une pente orientée dans la direction opposée à celle du capteur. La figure 2.11 (b), qui représente une crête, illustre ce phénomène de contraction-dilatation.

2.4.2.2 Effets de repliement

Comme mentionné à la sous-section précédente, les effets de repliement (effets de *layover*) se caractérisent par la rétrodiffusion simultanée d'ondes électromagnétiques en provenance de plusieurs surfaces disjointes. Ces effets interviennent lorsque la pente du terrain est supérieure à l'angle d'incidence i de l'onde RADAR.

Lorsque l'effet de repliement résulte de la présence d'une structure verticale, la hauteur de cette structure peut être reliée à la longueur du repliement observé en géométrie RADAR. À la figure 2.12 (a), la présence d'une structure verticale de hauteur h dans l'espace terrain engendre un effet de repliement correspondant à la rétrodiffusion simultanée des segments $[AB]$, $[BC]$ et $[CD]$. Les segments $[ab]$, $[bc]$ et $[cd]$ résultant de leur projection en géométrie RADAR sont donc confondus. La longueur d du repliement ($d = d_{ab} = d_{bc} = d_{cd}$) est reliée à la hauteur h par la relation suivante :

$$d = h \cos i. \quad (2.38)$$

D'un point de vue radiométrique, la rétrodiffusion simultanée de plusieurs surfaces de l'espace terrain se traduit par une augmentation de l'énergie électromagnétique reçue par le capteur et donc par une radiométrie plus importante, comme l'illustre la figure 2.12 (b) qui présente plusieurs zones de repliement liées à la présence de pylônes électriques dans la scène imagée.

2.4.2.3 Effets d'ombrage

Les effets d'ombrage apparaissent lorsqu'aucune énergie n'est rétrodiffusée vers le capteur RSO. Ce phénomène se produit lorsqu'une partie de la scène imagée est occultée par un objet présent dans la scène. Les effets d'ombrage, dont une illustration est proposée à la figure 2.13 (d), sont davantage prononcés avec les capteurs aéroportés qu'avec les capteurs satellitaires pour lesquels les angles d'incidence sont en général plus faibles.

En imagerie RSO, il est possible, comme en imagerie optique, de relier la longueur d'une ombre à la hauteur de l'objet ayant engendré cette ombre. La longueur de l'ombre engendrée par un objet dépend des valeurs relatives de la hauteur h et de la largeur apparente l_a de cet objet, la largeur apparente d'un objet étant définie comme la largeur de cet objet le long de la direction *across-track*. Comme illustré à la figure 2.13, nous distinguons les objets occultants, les objets partiellement occultants et les objets non occultants.

Un objet est considéré comme occultant (figure 2.13 (a)), pour une prise de vue donnée, lorsque la hauteur h de cet objet est inférieure à une hauteur limite h_l dont l'expression fait intervenir la largeur apparente l_a de cet objet ainsi que l'angle d'incidence i de l'onde RADAR :

$$h_l = l_a \tan i. \quad (2.39)$$

Dans le cas d'un objet occultant, la longueur d_{fe} de l'ombre RADAR engendré par l'objet est reliée à la hauteur h de cet objet par la relation suivante :

$$d_{fe} = \frac{h}{\cos i}. \quad (2.40)$$

Un objet est considéré comme partiellement occultant (figure 2.13 (b)), pour une prise de vue donnée, lorsque la hauteur h de cet objet est supérieure à la hauteur limite h_l . La longueur d_{gc} de l'ombre engendrée par cet objet est alors déterminée en considérant les longueurs d_{gf} , d_{fe} et d_{ec} :

$$d_{gc} = d_{gf} + d_{fe} - d_{ec}, \quad (2.41)$$

avec :

$$d_{gf} = \frac{h}{\cos i}, \quad (2.42)$$

$$d_{fe} = l_a \sin i, \quad (2.43)$$

$$d_{ec} = h \cos i. \quad (2.44)$$

En remplaçant d_{gf} , d_{fe} et d_{ec} par leurs expressions respectives, l'équation (2.41) se réécrit finalement comme suit :

$$d_{gc} = h \sin i \tan i + l_a \sin i. \quad (2.45)$$

Contrairement au cas des objets occultants, la longueur de l'ombre engendrée par les objets partiellement occultants fait intervenir la largeur apparente de ces objets. Il est donc nécessaire de connaître la largeur apparente d'un objet si l'on souhaite estimer sa hauteur à partir de la longueur de l'ombre RADAR associée.

Enfin, un objet de largeur apparente nulle ($l_a = 0$) sera considéré comme un objet non occultant. La longueur d_{db} de l'ombre engendrée par cet objet est déterminée en considérant les longueurs d_{dc} et d_{cb} :

$$d_{db} = d_{dc} - d_{cb}, \quad (2.46)$$

avec :

$$d_{dc} = \frac{h}{\cos i}, \quad (2.47)$$

$$d_{cb} = h \cos i. \quad (2.48)$$

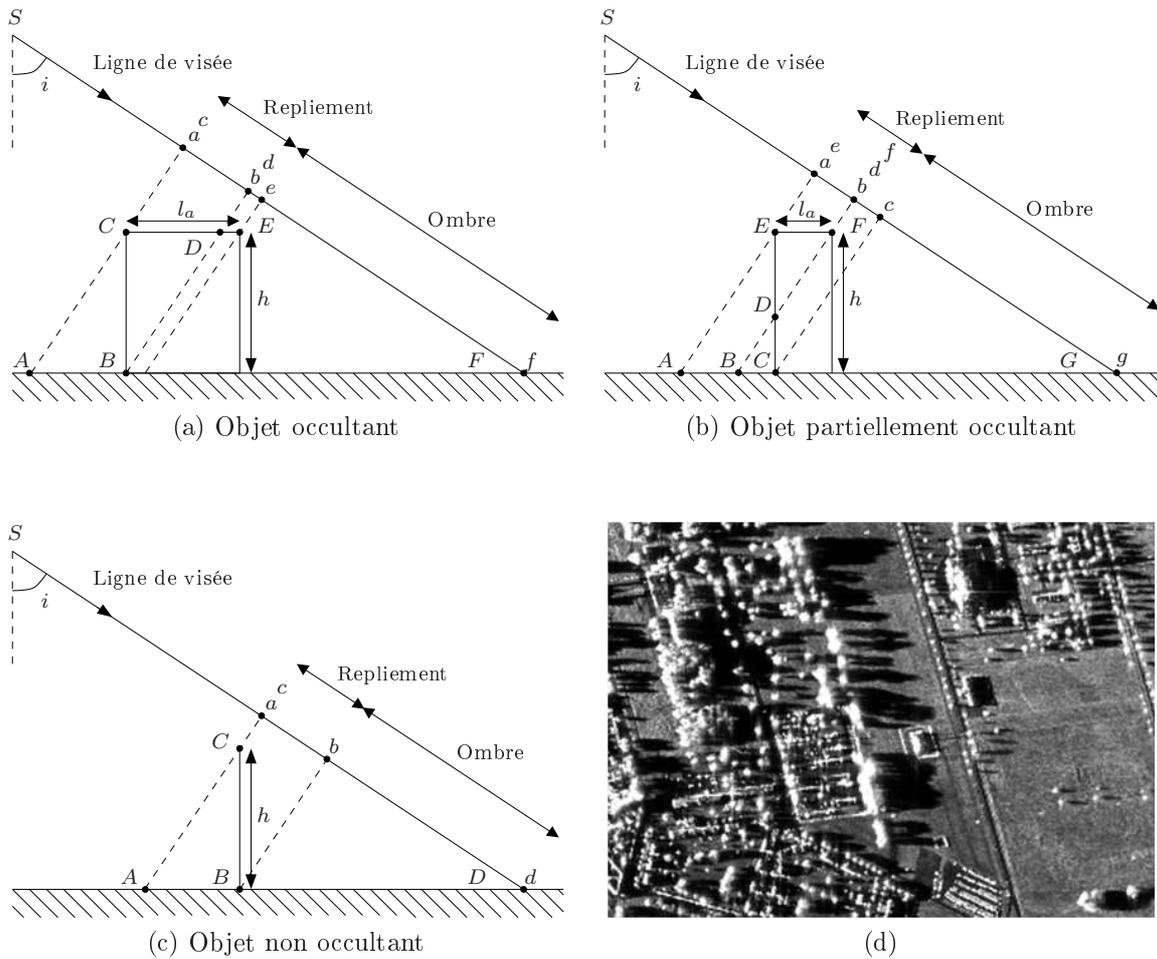


Figure 2.13 – Effets d’ombrage en imagerie RSO : schémas illustratifs (a,b,c) et exemple d’ombres RADAR liées à la présence de végétation (d) (image extraite de [Den09]).

L’équation (2.46) se réécrit alors comme suit :

$$d_{db} = h \sin i \tan i. \quad (2.49)$$

Étant donné que la largeur apparente d’un objet vu par un capteur RSO peut varier lors du déplacement de ce capteur, il est possible que cet objet présente un comportement hybride. Un même objet peut ainsi être vu alternativement comme un objet occultant, un objet partiellement occultant ou un objet non occultant pour différentes positions azimutales du capteur RSO. Le comportement hybride d’un objet est compatible avec les équations (2.40), (2.45) et (2.49), à partir desquelles nous pouvons constater qu’il y a continuité entre le modèle d’objet occultant et le modèle d’objet partiellement occultant en $l_a = 0$, ainsi qu’entre le modèle d’objet partiellement occultant et le modèle d’objet non occultant en $h = h_l$.

2.4.2.4 Effets liés aux réflexions multiples sur les structures diédriques et triédriques

La présence dans l’espace terrain de structures diédriques, formées par l’intersection de deux plans orthogonaux, se traduit dans l’espace image par la présence de structures linéiques présentant une forte radiométrie. Comme illustré à la figure 2.14 (a), l’onde RADAR réfléchi par une structure diédrique à l’issue de deux réflexions spéculaires (une réflexion sur chacun des plans constitutifs de la structure diédrique) est une onde plane dont la direction est opposée à celle de l’onde incidente. Ce résultat est obtenu en montrant que la différence de marche entre les rayons tracés à la figure 2.14 (a) est

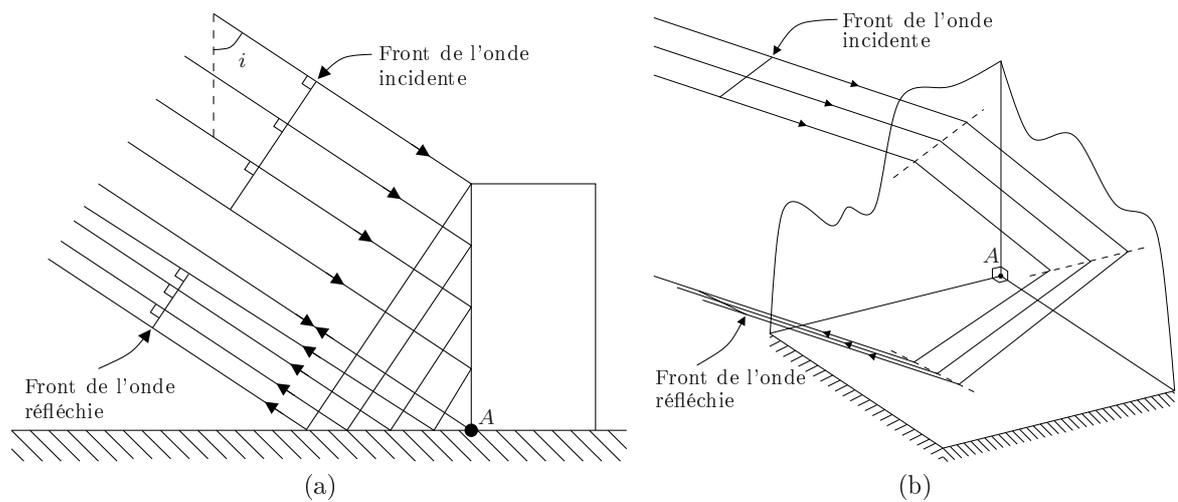


Figure 2.14 – Réflexions multiples sur une structure dièdrique (a) et sur une structure trièdrique(b).

nulle quel que soit l'angle d'incidence i de l'onde **RADAR**. Ainsi, les ondes réfléchies par la structure dièdrique sont en phase et arrivent en même temps au niveau du capteur **RSO**. Tout se passe comme si une seule onde était réfléchi au niveau du point A , l'énergie de cette onde résultant de la remise en phase des ondes réfléchies par l'ensemble de la structure.

Nous déduisons de ces observations que les structures dièdriques formées par les façades d'un bâtiment avec le sol supposé plat engendrent, dans l'espace image, des lignes brillantes au niveau des segments visibles de l'emprise au sol de ce bâtiment. En pratique, l'hypothèse d'orthogonalité entre les façades et le sol n'est pas parfaitement vérifiée, de sorte que les lignes brillantes présentent une largeur non nulle.

Comme illustré à la figure 2.14 (b), le phénomène observé au niveau des structures trièdriques est identique à celui observé au niveau des structures dièdriques à la différence que l'onde plane réfléchi résulte de trois réflexions spéculaires (une réflexion sur chacun des plans constitutifs de la structure trièdrique) et que les manifestations dans l'espace image de ces structures sont des points brillants et non plus des lignes brillantes.

Les structures trièdriques sont, le plus souvent, formées par le sol avec deux façades d'un même bâtiment. Cependant, avec le modèle parallélépipédique de bâtiment adopté à la partie 2.2, de telles structures ne sont pas observées. Ainsi, seules sont considérées, par la suite, les lignes brillantes résultant de la présence de structures dièdriques.

2.4.3 Simulation de la signature géométrique d'un bâtiment

En s'appuyant sur les concepts introduits à la section 2.4.2, nous proposons dans cette section une méthode pour simuler la signature géométrique en imagerie **RSO** d'un bâtiment décrit au moyen du modèle parallélépipédique de la section 2.2.

Comme mentionné à la section précédente, la largeur apparente et la hauteur d'un bâtiment peuvent varier pour différentes positions azimutales d'un capteur **RSO**, de sorte que ce bâtiment peut être vu alternativement comme un objet occultant, partiellement occultant ou non occultant.

Dans le cas d'un bâtiment parallélépipédique, la hauteur h du bâtiment est constante et seule varie la largeur apparente l_a du bâtiment. Comme illustré à la figure 2.15, ce bâtiment est tout d'abord vu par le capteur **RSO** comme un objet non occultant puisque sa largeur apparente est nulle. La largeur apparente augmente alors linéairement avec le déplacement du capteur le long de l'axe azimutal jusqu'à atteindre une valeur maximale l_a^{max} . Tant que la largeur apparente est inférieure à $h / \tan i$, le bâtiment

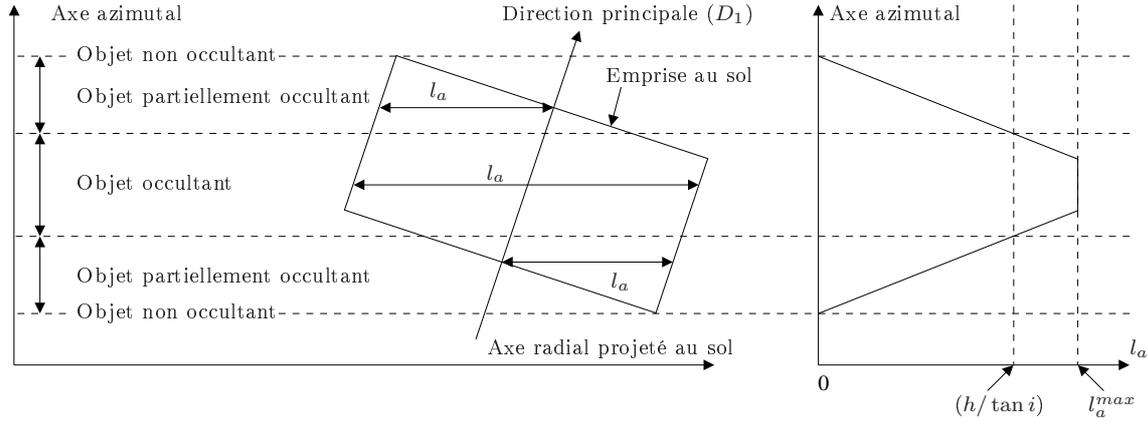


Figure 2.15 – Variation de la largeur apparente d'un bâtiment parallélépipédique vu par un capteur RSO.

est vu par le capteur RSO comme un objet partiellement occultant avant d'être finalement vu comme un objet occultant dès que la largeur apparente est supérieure à $h/\tan i$. La largeur apparente diminue ensuite linéairement de sorte que l'objet est de nouveau vu comme un objet partiellement occultant avant d'être finalement vu comme un objet non occultant lorsque sa largeur apparente s'annule.

Dans le cas où la largeur apparente maximale l_a^{max} est inférieure à la valeur limite $h/\tan i$, le modèle d'objet occultant n'intervient pas dans la signature géométrique du bâtiment. C'est le cas des gratte-ciel qui sont plus hauts que larges, contrairement aux bâtiments industriels qui sont plus larges que hauts et pour lesquels le modèle d'objet occultant intervient.

Dans le cas particulier où l'axe azimutal est parallèle à la direction principale (D_1) ou à la direction secondaire (D_2) du bâtiment, la largeur apparente de ce bâtiment est constante de sorte que la signature géométrique de ce bâtiment est décrite à l'aide d'un seul modèle (occultant ou partiellement occultant).

Lorsqu'un bâtiment est vu comme un objet occultant pour une position azimutale donnée du capteur RSO, sa signature géométrique selon l'axe radial se décompose, comme illustré à la figure 2.13 (a), en trois segments distincts, segments pour lesquels différentes surfaces de l'espace terrain interviennent dans la rétrodiffusion de l'onde RADAR vers le capteur. Le segment $[ab]$ de l'espace image correspond à la rétrodiffusion simultanée des segments $[AB]$, $[BC]$ et $[CD]$ de l'espace terrain, segments associés respectivement au sol, à la façade et à une partie du toit du bâtiment. Le segment $[de]$ correspond à la rétrodiffusion du segment $[DE]$ qui constitue la partie restante du toit, tandis que le segment $[fe]$ correspond à l'ombre RADAR et n'est donc associé à aucune surface de l'espace terrain. La longueur de ces différents segments est donnée par les équations suivantes :

$$d_{ab} = h \cos i, \quad (2.50)$$

$$d_{de} = l_a \sin i - h \cos i, \quad (2.51)$$

$$d_{fe} = \frac{h}{\cos i}. \quad (2.52)$$

Lorsqu'un bâtiment est vu comme un objet partiellement occultant pour une position donnée du capteur RSO, sa signature géométrique selon l'axe radial se décompose, comme illustré à la figure 2.13 (b), en trois segments distincts. Le segment $[ab]$ de l'espace image correspond à la rétrodiffusion simultanée des segments $[AB]$, $[DE]$ et $[EF]$ de l'espace terrain, segments associés respectivement au sol, à une partie de la façade et au toit du bâtiment. Le segment $[cd]$ correspond à la rétrodiffusion du segment $[CD]$ associé à la partie restante de la façade et le segment $[gc]$ correspond à l'ombre

RADAR. La longueur de ces différents segments est donnée par les équations suivantes :

$$d_{ab} = l_a \sin i, \quad (2.53)$$

$$d_{cd} = h \cos i - l_a \sin i, \quad (2.54)$$

$$d_{gc} = h \sin i \tan i + l_a \sin i. \quad (2.55)$$

Enfin, lorsqu'un bâtiment est vu comme un objet non occultant pour une position donnée du capteur RSO, sa signature géométrique selon l'axe radial se décompose, comme illustré à la figure 2.13 (c), en deux segments distincts. Le premier segment, le segment $[ab]$, résulte de la rétrodiffusion simultanée des segments $[AB]$ et $[BC]$, qui correspondent respectivement au sol et à la façade du bâtiment tandis que le second segment, le segment $[db]$, correspond à l'ombre RADAR. La longueur de ces deux segments est donnée par les expressions suivantes :

$$d_{ab} = h \cos i, \quad (2.56)$$

$$d_{db} = h \sin i \tan i. \quad (2.57)$$

Finalement, la signature géométrique d'un bâtiment parallélépipédique dans une image RSO est obtenue à partir des signatures géométriques établies pour les différentes positions azimutales du capteur. En reprenant les mêmes notations que celles introduites à la section 2.3.3, la signature géométrique $\mathcal{S}_I(\vec{b})$ d'un bâtiment parallélépipédique \vec{b} dans une image RSO peut alors s'écrire comme un ensemble de régions, éventuellement vides, définies dans l'espace image \mathcal{E}_I :

$$\mathcal{S}_I(\vec{b}) = \{R_{TFS}, R_T, R_{ED}, R_O\} \cup \{R_{FS}^i, i = 1 \dots N_{fs}\}, \quad (2.58)$$

où :

- la région R_{TFS} correspond au phénomène de repliement résultant de la rétrodiffusion simultanée du sol, du toit et des façades du bâtiment ;
- la région R_T correspond à la rétrodiffusion du toit du bâtiment ;
- la région R_{ED} correspond à la double réflexion de l'onde RADAR sur les structures diédriques formées par le sol et les façades du bâtiment ;
- la région R_O correspond à l'ombre RADAR engendrée par la présence du bâtiment ;
- les régions R_{FS}^i correspondent au phénomène de repliement résultant de la rétrodiffusion simultanée du sol et des façades du bâtiment.

La méthode qui est proposée pour simuler la signature géométrique d'un bâtiment connu \vec{b} dans une image I se décompose en quatre étapes principales :

1. Projection de l'emprise au sol du bâtiment de l'espace terrain \mathcal{E}_T vers l'espace image \mathcal{E}_I ;
2. Détermination des régions R_{TFS} , R_T et R_{FS}^i correspondant à la rétrodiffusion du toit, des façades et du sol ;
3. Détermination de la région R_O correspondant à l'ombre RADAR ;
4. Détermination de la région R_{ED} correspondant à la double réflexion sur le dièdre formé par les façades et le sol.

2.4.3.1 Projection de l'emprise au sol du bâtiment de l'espace terrain vers l'espace image

La première étape de la méthode consiste à projeter dans l'espace image \mathcal{E}_I les points définissant l'emprise au sol du bâtiment dans l'espace terrain \mathcal{E}_T , comme illustré à la figure 2.16 (a). Cette étape est rigoureusement identique à celle présentée à la section 2.3.3 pour la simulation de la signature géométrique d'un bâtiment en imagerie optique.

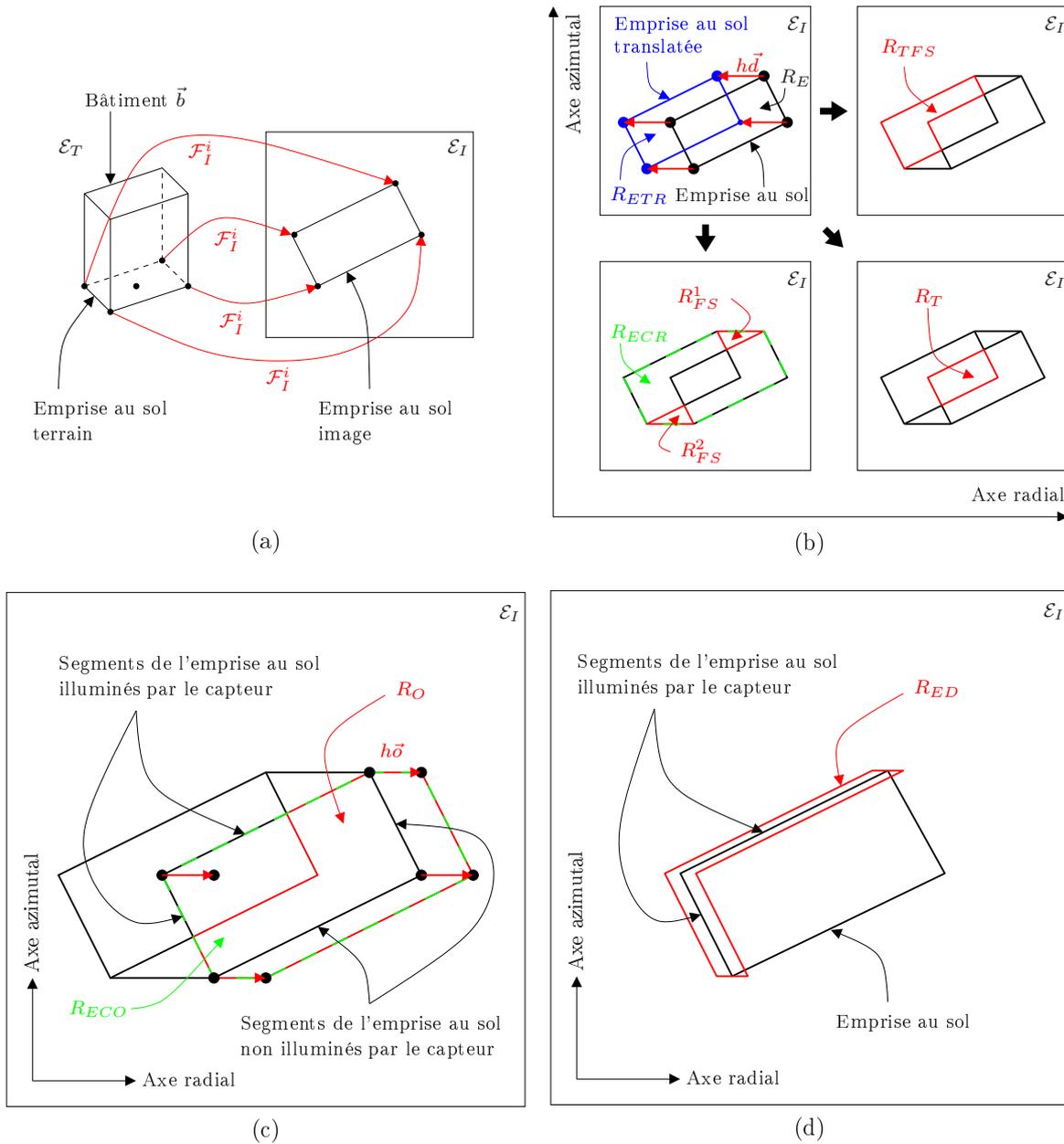


Figure 2.16 – Simulation de la signature géométrique d'un bâtiment parallélépipédique en imagerie RSO : projection de l'emprise au sol de l'espace terrain vers l'espace image (a), détermination des régions R_{TFS} , R_T et R_{FS}^i (b), détermination de la région R_O (c) et détermination de la région écho double R_{ED} (d).

2.4.3.2 Détermination des régions R_{TFS} , R_T et R_{FS}^i correspondant à la rétrodiffusion du toit, des façades et du sol

La deuxième étape de la méthode consiste à définir les régions de l'image associées aux différents effets de repliement.

Conformément à l'équation (2.38), l'emprise au sol définie à l'étape précédente est translatée, dans l'espace image \mathcal{E}_I , d'un vecteur $h\vec{d}$ où \vec{d} est le vecteur de norme $\cos i/\Delta_r$ orienté selon la direction opposée à celle de l'axe radial, Δ_r désignant le pas d'échantillonnage radial en géométrie RADAR. Le contour ainsi obtenu, qui correspond en fait à la projection du contour du toit du bâtiment de l'espace terrain vers l'espace image, permet de définir l'ensemble des régions au sein desquelles interviennent les phénomènes de repliement.

Nous notons R_E et R_{ETR} les régions définies respectivement par le contour de l'emprise au sol et le contour de l'emprise au sol translatée. Comme indiqué à la figure 2.16 (b), la région R_T , qui correspond à la rétrodiffusion du toit, est définie, lorsqu'elle existe, par l'intersection des régions R_E et R_{ETR} :

$$R_T = R_E \cap R_{ETR}. \quad (2.59)$$

Cette région existe lorsque le bâtiment est vu comme un objet occultant pour au moins une position azimutale du capteur RSO, c'est-à-dire lorsque la largeur apparente maximale du bâtiment est supérieure à $h/\tan i$. Dans le cas contraire, cette région est réduite à l'ensemble vide.

Contrairement à la région R_T , la région R_{TFS} , qui correspond aux effets de repliement faisant intervenir le sol, les façades et le toit du bâtiment, est toujours définie. Cette région est obtenue en soustrayant la région R_T , lorsqu'elle existe, à la région R_{ETR} :

$$R_{TFS} = R_{ETR} \setminus R_T. \quad (2.60)$$

En considérant la région R_{ECR} , définie par l'enveloppe convexe de l'ensemble de points formé par les sommets de l'emprise au sol et de l'emprise au sol translatée, les régions R_{FS}^i sont obtenues en soustrayant les régions R_E et R_{ETR} à la région R_{ECR} :

$$\{R_{FS}^i, i = 1 \dots N_{fs}\} = R_{ECR} \setminus \{R_E \cup R_{ETR}\}. \quad (2.61)$$

Lorsque la largeur apparente maximale du bâtiment est strictement supérieure à $h/\tan i$, les régions R_E et R_{ETR} sont disjointes, de sorte que seule une région R_{FS} est définie ($N_{fs} = 1$). Dans le cas contraire, deux régions R_{FS} sont définies ($N_{fs} = 2$).

2.4.3.3 Détermination de la région R_O correspondant à l'ombre RADAR

Au cours de la troisième étape, l'emprise au sol est translatée d'un vecteur $h\vec{o}$ où \vec{o} est le vecteur de norme $\sin i \tan i/\Delta_r$ orienté selon la direction radiale, comme illustré à la figure 2.16 (c). En notant R_{ECO} la région définie par l'enveloppe convexe de l'ensemble des points formé par les sommets de l'emprise au sol et de cette emprise au sol translatée, la région R_O est alors obtenue en soustrayant la région R_T , lorsqu'elle est non vide, à la région R_{ECO} :

$$R_O = R_{ECO} \setminus R_T. \quad (2.62)$$

2.4.3.4 Détermination de la région R_{ED} correspondant à la double réflexion sur le dièdre formé par les façades et le sol

Nous avons vu à la section 2.4.2 que la double réflexion de l'onde RADAR sur une structure diédrique se traduit, dans l'espace image, par des structures linéiques brillantes. En pratique, les plans associés au sol et aux façades du bâtiment ne sont pas parfaitement orthogonaux, de sorte que

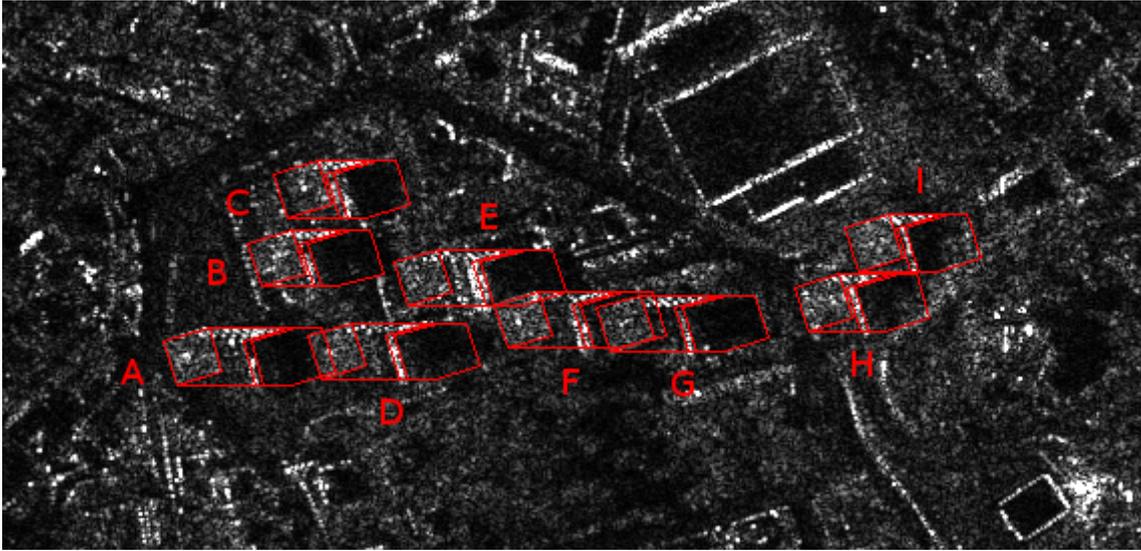


Figure 2.17 – Simulation de la signature géométrique de bâtiments dans une image TerraSAR-X (©Astrium Geo-Information Services).

les structures linéiques résultant de la double réflexion de l'onde RADAR sur ces plans présentent une certaine largeur, largeur qui peut être reliée à l'angle réel formé par ces plans. N'ayant pas de connaissance *a priori* sur l'angle formé par ces plans, nous fixons arbitrairement la largeur de ces structures linéiques à une valeur constante ϵ_{ED} .

Comme indiqué à la figure 2.16 (d), la région R_{ED} est obtenue en considérant les segments de l'emprise au sol illuminés par le capteur RSO. Ces segments sont dilatés symétriquement selon la direction radiale de la largeur constante fixée précédemment.

À l'issue de cette dernière étape, la région R_{ED} est retranchée à l'ensemble des autres régions (R_{TFS} , R_T , R_{FS}^i et R_O) avec lesquelles elle est en intersection.

2.4.4 Exemples de signatures géométriques simulées

Dans cette section, nous mettons en œuvre la méthode proposée à la section précédente afin de simuler les signatures géométriques de plusieurs bâtiments dans une image RSO.

L'image dont nous disposons, présentée à la figure 2.9, a été acquise par le capteur RSO spatial TerraSAR-X le 09 décembre 2007 sur la région de Marseille (France). Cette image, dont la géolocalisation est réalisée via le modèle physique du capteur, a été acquise en mode *SpotLight HR* et présente une résolution de 1.2 mètres selon l'axe radial et une résolution de 1.1 mètres selon l'axe azimutal. Enfin, l'image a été acquise avec un angle d'incidence de 31.8 degrés en milieu de fauchée.

La scène étudiée est la même que celle de la section 2.3.4, de sorte que les paramètres du modèle parallélépipédique permettant de décrire les bâtiments de cette scène sont identiques à ceux présentés au tableau 2.1. Seules les coordonnées (X_C, Y_C) du centre de l'emprise au sol des bâtiments ont été modifiées de manière à corriger les erreurs de recalage entre l'image QuickBird et l'image TerraSAR-X.

Nous constatons à la figure 2.17 que les signatures géométriques simulées au moyen de la méthode décrite à la section 2.4.3 présentent une bonne adéquation avec les bâtiments présents dans l'image. Comme en imagerie optique, la méthode proposée suppose que les bâtiments sont isolés si bien que les phénomènes d'occlusion entre les différents bâtiments constitutifs de la scène ne sont pas pris en compte. Ceci se traduit par des recouvrements entre les signatures géométriques correspondant à des bâtiments distincts. La figure 2.17 montre notamment que les ombres associées au bâtiments *A* et *F* sont en partie occultées par les bâtiments *D* et *G*.

2.5 Conclusion

Dans ce chapitre, nous avons étudié les aspects géométriques relatifs au rendu tridimensionnel des bâtiments en imagerie optique et en imagerie RSO. Nous avons tout d'abord décrit les effets géométriques affectant les bâtiments lors de leur projection de l'espace terrain vers l'espace image. Nous avons ensuite proposé deux méthodes permettant de simuler la signature géométrique des bâtiments parallélépipédiques dans les images optiques et RSO.

Hormis au chapitre 3, relatif à l'état de l'art sur l'extraction tridimensionnelle de bâtiments en imagerie optique et RSO, nous ne considérons, dans la suite du présent manuscrit, que des bâtiments parallélépipédiques pouvant être décrits au moyen du modèle paramétrique introduit à la partie 2.2.

Les méthodes qui ont été proposées, aux sections 2.3.3 et 2.4.3, pour simuler la signature géométrique des bâtiments dans les images optiques et RSO, sont exploitées aux chapitres 4 et 5, où elles permettent de modéliser les effets géométriques liés à la projection des bâtiments dans ces images.

État de l'art sur l'extraction de bâtiments en imagerie de télédétection optique et RSO

Sommaire

3.1	Introduction	54
3.1.1	Problématique de l'extraction de bâtiments	54
3.1.2	Classification des méthodes d'extraction de bâtiments	55
3.2	Extraction des bâtiments en imagerie optique	56
3.2.1	Extraction de primitives	56
3.2.2	Extraction des toits des bâtiments	57
3.2.3	Extraction tridimensionnelle des bâtiments	59
3.3	Extraction des bâtiments en imagerie RSO	61
3.3.1	Extraction de primitives	61
3.3.2	Extraction des emprises au sol des bâtiments	64
3.3.3	Extraction tridimensionnelle des bâtiments	65
3.4	Extraction de bâtiments à partir d'images optiques et RSO	69
3.4.1	Méthodes exploitant la complémentarité d'information entre images optiques et RSO	70
3.4.2	Méthodes exploitant la redondance d'information entre images optiques et RSO	70
3.5	Conclusion	72

Dans ce chapitre, nous présentons un panorama des différentes approches permettant de réaliser une extraction tridimensionnelle de bâtiments à partir d'images de télédétection optiques et RSO, acquises indifféremment par des capteurs aéroportés ou spatiaux. Après une première partie introductive au sein de laquelle nous soulignons les enjeux relatifs à l'extraction tridimensionnelle de bâtiments, nous présentons, dans la seconde partie, les méthodes développées en imagerie optique pour extraire les bâtiments à partir d'une unique image ou de plusieurs images acquises avec des angles de prise de vue différents. Dans la troisième partie, nous décrivons les méthodes qui réalisent une extraction de bâtiments en imagerie RSO à partir d'une unique image ou de plusieurs images en conditions interférométriques ou radargrammétriques. Enfin, dans la dernière partie, les méthodes développées pour l'extraction des bâtiments dans un contexte multicapteur optique-RSO sont présentées. Notons que les approches développées pour l'extraction de bâtiments à partir de données LIDAR¹ ou de données RSO polarimétriques ne sont pas étudiées dans le présent chapitre.

3.1 Introduction

La notion d'extraction de bâtiments recouvre deux aspects : la détection des bâtiments et la reconstruction tridimensionnelle de ces derniers. La détection consiste à trouver les bâtiments présents dans une image tandis que la reconstruction tridimensionnelle correspond à l'estimation de leur forme et de leurs dimensions. Comme le souligne [May99] dans son état de l'art sur l'extraction de bâtiments en imagerie optique aéroportée, ces deux aspects de l'extraction de bâtiments sont fortement liés. Ainsi, seule une reconstruction précise des bâtiments permet d'atteindre des performances de détection satisfaisantes, notamment en termes de fausses alarmes.

3.1.1 Problématique de l'extraction de bâtiments

L'extraction de bâtiments à partir d'images de télédétection optiques et RSO est une tâche difficile en raison de la complexité de la scène et des données à partir desquelles l'extraction est réalisée.

3.1.1.1 Complexité d'une scène

La complexité d'une scène dépend de son contenu, c'est-à-dire de la nature, de la variabilité et de la densité des objets présents dans cette scène. Ainsi, l'extraction de bâtiments en milieu semi urbain est souvent plus aisée qu'en milieu urbain dense pour lequel les phénomènes d'occlusion entre bâtiments peuvent devenir critiques notamment dans le cas de prises de vue obliques.

3.1.1.2 Complexité des données

La complexité des données à partir desquelles est réalisée l'extraction de bâtiments est définie par le nombre, la nature et les caractéristiques des images disponibles.

Nombre d'images. Les performances en extraction de bâtiments sont d'autant meilleures que le nombre d'images disponibles sur la scène est grand et que les angles sous lesquels est visualisée la scène dans les images sont différents. Ainsi, les jeux de données multiscopiques et stéréoscopiques, dans le cas de l'imagerie optique, ou radargrammétriques et interférométriques, dans le cas de l'imagerie RSO, conduisent à des performances en extraction de bâtiments supérieures à celles obtenues en monoscopie. Schématiquement, la redondance d'information présente au sein de ces jeux de données est exploitée pour accroître les performances en détection alors que l'apport d'information altimétrique permet une reconstruction tridimensionnelle plus robuste.

¹ *Light Detection And Ranging*

Nature des images. Selon la nature des capteurs avec lesquels ont été acquises les images, l'extraction des bâtiments demeure plus ou moins complexe. Ainsi, l'extraction de bâtiments en imagerie optique conduit généralement à de meilleurs résultats que ceux obtenus en imagerie RSO. En effet, les images RSO présentent usuellement un RSB plus faible que les images optiques. De plus, la géométrie de prise de vue des images RSO conduit à des effets de distorsion sur les objets du sur-sol plus marqués que dans les images optiques (cf. section 1.2.3).

Caractéristiques des images. Comme mentionné au chapitre 1, la résolution spatiale des images, qui conditionne l'observabilité des bâtiments, demeure une caractéristique essentielle de ces images dans le cadre de l'extraction de bâtiments. La qualité des images, en termes de contraste et de bruit, constitue également une caractéristique dont dépend l'observabilité des bâtiments. Enfin, les performances en extraction dépendent de l'angle sous lequel la scène est visualisée. En effet, les effets géométriques affectant les objets du sur-sol (occlusion, déplacement, repliement), qui apparaissent notamment dans le cas des prises de vue avec incidence rasante, compliquent en général l'extraction des bâtiments bien qu'ils constituent une source d'information sur la hauteur de ces bâtiments.

3.1.2 Classification des méthodes d'extraction de bâtiments

Parmi les critères qui permettent de différencier les méthodes d'extraction de bâtiments, en imagerie optique comme en imagerie RSO, nous distinguons le modèle de bâtiment adopté et la stratégie d'extraction mise en œuvre.

Modèle de bâtiment. La modélisation des bâtiments correspond à la façon dont est introduite l'information *a priori* que l'on a sur la géométrie des bâtiments à extraire. Les modèles de bâtiments peuvent être bidimensionnels lorsque l'on cherche à extraire uniquement les emprises au sol ou les toits des bâtiments, ou bien tridimensionnels si l'on cherche également à déterminer la hauteur des bâtiments. La complexité des modèles de bâtiments adoptés est un critère important puisqu'il conditionne la généralité et la robustesse des méthodes d'extraction. Ainsi, les méthodes basées sur des modèles simples ont l'avantage d'être robustes mais sont restreintes à l'extraction de certains types de bâtiment, tandis que les méthodes basées sur des modèles plus complexes présentent une plus grande généralité au détriment de leur robustesse.

Stratégie d'extraction. La stratégie d'extraction désigne la méthodologie mise en œuvre pour réaliser l'extraction des bâtiments. Les différentes stratégies d'extraction peuvent être classées en deux grandes catégories : les approches ascendantes (*bottom-up*) et les approches descendantes (*top-down*).

Les approches ascendantes, qui sont les plus répandues, reposent sur une extraction de caractéristiques radiométriques et/ou géométriques dans les images. Ces caractéristiques sont alors utilisées pour établir des hypothèses de bâtiments en imposant des contraintes géométriques, généralement faibles, sur la forme de ces hypothèses. Ces hypothèses sont éventuellement régularisées et/ou fusionnées avant d'être finalement validées ou rejetées lors d'une étape consistant à quantifier l'adéquation de ces hypothèses avec l'image et/ou les caractéristiques extraites. Parmi les approches ascendantes, nous distinguons notamment :

- les approches basées sur une extraction de primitives caractéristiques de la présence des bâtiments, primitives à partir desquelles des structures géométriques de plus en plus complexes sont construites par agrégation successive de ces primitives ;
- et les approches basées sur une classification des pixels d'une image ou des régions issues d'une segmentation de l'image à partir de vecteurs de caractéristiques, représentatifs des propriétés radiométriques et/ou géométriques d'un voisinage des pixels ou des régions considérées.

Contrairement aux approches ascendantes, pour lesquelles les contraintes géométriques imposées lors de la génération des hypothèses sont relativement faibles, les approches descendantes imposent des contraintes géométriques beaucoup plus fortes sur la forme des bâtiments extraits. Dans le cadre des approches descendantes, il est souvent supposé que les bâtiments à extraire peuvent être décrits par un ou plusieurs modèles, bidimensionnels ou tridimensionnels, connus *a priori*. Ces approches

consistent alors à détecter les bâtiments présents dans la scène, à déterminer les modèles décrivant le mieux ces bâtiments au sein de l'ensemble des modèles admissibles, et à estimer les paramètres des modèles décrivant les bâtiments détectés. Ces différentes étapes peuvent être réalisées successivement ou conjointement.

3.2 Extraction des bâtiments en imagerie optique

Dans cette partie, nous nous intéressons à l'extraction des bâtiments à partir d'une ou plusieurs images optiques haute ou très haute résolution, acquises au moyen de capteurs satellitaires ou aéroportés. Nous abordons tout d'abord la problématique de l'extraction des primitives caractéristiques de la présence des bâtiments dans les images optiques, primitives sur lesquelles s'appuient de nombreuses méthodes d'extraction. Nous présentons ensuite successivement les méthodes qui réalisent une extraction des contours des toits des bâtiments puis celles qui réalisent une extraction tridimensionnelle de ces derniers.

3.2.1 Extraction de primitives

Le terme « primitive » désigne l'ensemble des éléments ponctuels, linéiques ou surfaciques d'une image qui constituent la signature géométrique d'un bâtiment dans cette image et qui sont caractéristiques de la présence de ce bâtiment.

3.2.1.1 Primitives ponctuelles

En reprenant les mêmes notations que celles introduites à la section 2.3.3, les primitives ponctuelles d'un bâtiment \bar{b} dans une image I correspondent aux intersections entre les différents segments S_k constitutifs de la signature géométrique $\mathcal{S}_I(\bar{b})$ de ce bâtiment dans l'image I (cf. équation (2.21)).

Ces primitives sont le plus souvent détectées à l'issue de la détection de primitives linéiques comme des jonctions en « L », des jonctions en « T » ou des jonctions triédriques entre ces segments comme cela est fait dans [MS94, LN95, Shu96]. Notons que les jonctions triédriques ne sont visibles que dans le cas d'une prise de vue oblique contrairement aux jonctions en « L » et en « T » qui interviennent dans les prises de vue verticales et les prises de vue obliques.

3.2.1.2 Primitives linéiques

Les primitives linéiques d'un bâtiment dans une image, primitives qui sont les plus utilisées pour l'extraction de bâtiments [LN98, KS08], correspondent aux segments S_k constitutifs de la signature géométrique du bâtiment dans l'image.

Ces primitives correspondent aux contours de l'ombre projetée par un bâtiment ou aux arêtes des bâtiments, dans l'hypothèse d'une géométrie de bâtiment rectiligne. Ces arêtes peuvent représenter la jonction entre un mur et le toit d'un bâtiment, entre deux murs d'un bâtiment ou bien entre le mur d'un bâtiment et le sol.

Elles sont le plus souvent extraites au moyen de méthodes fondées sur l'opérateur gradient, comme cela est fait par le détecteur de Nevatia-Babu [NB80] ou le détecteur de Canny [Can86] qui demeure, encore aujourd'hui, l'un des détecteurs de contours les plus utilisés.

3.2.1.3 Primitives surfaciques

Les primitives surfaciques d'un bâtiment dans une image correspondent aux régions R_j constitutives de la signature du bâtiment dans l'image (cf. équation (2.20)). Comme décrit au chapitre précédent, ces primitives surfaciques sont associées aux façades des bâtiments, aux toits des bâtiments ou encore aux ombres projetées par les bâtiments. Ces différents éléments, dont il est souvent fait

l'hypothèse qu'ils présentent une homogénéité en termes de réflectance, se traduisent dans l'image par des régions de radiométrie ou de texture homogène, qui sont le plus souvent extraites par segmentation. L'hypothèse d'homogénéité de réflectance est toutefois souvent mise en défaut, en raison de la présence de fenêtres sur les façades des bâtiments ou de superstructures, telles que les cheminées, sur les toits des bâtiments.

3.2.2 Extraction des toits des bâtiments

Nous décrivons, dans cette section, les méthodes qui réalisent une extraction du contour du toit d'un bâtiment, qui, dans le cas des prises de vue verticales, correspond également à l'emprise au sol de ce bâtiment. La plupart de ces méthodes s'appuient sur les éléments géométriques, morphologiques, contextuels, radiométriques ou encore spectraux, indiquant la présence de bâtiments, pour détecter et reconstruire les contours des toits des bâtiments. Comme mentionné en introduction, nous distinguons les méthodes ascendantes, pour lesquelles les contours de toit sont générées à partir de ces éléments caractéristiques en imposant de faibles contraintes sur la forme des contours, et les méthodes descendantes, qui consistent à vérifier la consistance de ces caractéristiques avec un ensemble de formes admissibles.

3.2.2.1 Les approches ascendantes

L'une des techniques les plus répandues parmi les approches ascendantes est l'agrégation de primitives, introduite par [MN89]. Cette technique repose sur la construction successive de groupes de primitives de plus en plus complexes par organisation perceptive de primitives bas-niveau. Des hypothèses de segments sont tout d'abord générées à partir des contours extraits de l'image par le détecteur de Nevatia-Babu [NB80]. Des hypothèses de couples de segments parallèles sont ensuite générées à partir des hypothèses de segments. Des structures en U et des structures rectangulaires sont finalement générées selon le même procédé de sorte que l'on dispose, à l'issue de cette première étape, d'un ensemble d'hypothèses de contours de toits rectilignes. À cette première étape de génération d'hypothèses, succède une étape de sélection de ces hypothèses par mise en œuvre d'un réseau de satisfaction de contraintes. Il s'agit de quantifier, d'une part, l'adéquation géométrique entre les hypothèses et, d'autre part, l'adéquation des hypothèses avec les contours et les coins extraits de l'image. Dans le cas monoscopique, la dernière étape de cette méthode consiste à combiner les contours rectangulaires précédemment sélectionnés, au moyen de règles de fusion géométriques simples, afin d'obtenir des contours rectilignes et plus uniquement rectangulaires. Dans le cas stéréoscopique, les étapes de génération et sélection d'hypothèses sont réalisées indépendamment sur chacune des deux images constitutives du couple. Les hypothèses sélectionnées sur chacune des deux images sont alors mises en correspondance en exploitant les propriétés de la géométrie épipolaire.

Les méthodes par agrégation de primitives présentent une robustesse accrue par rapport aux approches par segmentation, pour lesquelles les contours sont fermés (par construction) mais ne conduisent pas toujours aux bons objets. Ces approches sont également plus performantes que les méthodes par tracé de contours, qui ne fonctionnent que pour des objets simples et pour lesquelles l'explosion combinatoire (au niveau des intersections de contours) peut vite devenir limitante.

Parmi les approches ascendantes, nous distinguons également les approches fondées sur la classification de pixels ou de régions obtenues à l'issue d'une segmentation de l'image [SC00, MZ05, CBV04, KN04, LP05, IS10, CHA10]. Chaque pixel ou région de l'image est caractérisé par un vecteur d'attributs qui regroupe des attributs géométriques, radiométriques ou encore contextuels. Les pixels ou régions de l'image, représentés par leurs vecteurs d'attributs, sont alors affectés à l'une ou l'autre des classes « bâtiment » ou « non-bâtiment » au moyen d'algorithmes de classification qui nécessitent le plus souvent une phase d'apprentissage sur un ensemble de bâtiments dont les contours ont été extraits manuellement. À l'issue de cette classification, les contours des différentes régions identifiées comme étant des bâtiments correspondent alors aux contours des toits des bâtiments.

Dans [MZ05], chacune des régions, issues d'une segmentation par croissance de régions, est caractérisée par un vecteur regroupant des attributs géométriques (relatif à la taille et la forme de la région), des attributs radiométriques (mode et moyenne de la valeur des pixels de la région dans la composante *hue* de l'espace colorimétrique HSV²) et des attributs contextuels (relatifs à la présence d'ombres et d'autres bâtiments dans un proche voisinage de la région considérée). Un classifieur linéaire est finalement mis en œuvre pour classifier les vecteurs d'attributs, et donc les régions, selon les classes « bâtiment » et « non-bâtiment ».

Dans [CBV04], les pixels de l'image sont représentés par leurs profils morphologiques différentiels, vecteurs d'attributs qui caractérisent les pixels en termes d'intensité radiométrique et de contexte géométrique. Ces vecteurs d'attributs sont alors classifiés au moyen d'un réseau de neurones ou d'un algorithme de classification par logique floue, algorithmes de classification nécessitant une phase d'apprentissage sur un ensemble de bâtiments extraits manuellement.

Enfin, il existe des méthodes, que l'on peut qualifier de mixtes, qui présentent des aspects d'agrégation de primitives et des aspects de classification/segmentation pour l'extraction des contours de toit. À titre d'exemple, Cote *et al.* [CS12] propose une méthode pour l'extraction de contours de toit en monoscopie optique (avec prise de vue verticale) qui repose sur une extraction de coins par détecteur de Harris, coins qui sont ensuite sélectionnés selon des critères de texture pour finalement conduire à la génération d'hypothèses de contours de toit. Les hypothèses de contours sont alors optimisées localement au moyen d'une segmentation par contours actifs implicites (*level-set*) avant d'être validées ou rejetées selon des critères géométriques et radiométriques quantifiant l'adéquation avec la donnée image.

3.2.2.2 Les approches descendantes

Contrairement aux approches ascendantes, pour lesquelles seules des contraintes de rectilinéarité sont parfois imposées, les approches descendantes imposent des contraintes géométriques beaucoup plus fortes sur la forme des contours des toits extraits. Ces contraintes de formes impliquent une plus grande robustesse des méthodes d'extraction qui présentent, en contrepartie, une plus faible généralité.

Dans [STD09], Sportouche *et al.* proposent une méthode d'extraction de contours rectangulaires en monoscopie optique, basée sur l'utilisation des profils morphologiques différentiels et de la transformée de Hough. Un ensemble de régions d'intérêt est tout d'abord défini par recherche de structures rectangulaires au sein des images issues du calcul des profils morphologiques différentiels. Cette recherche est réalisée hiérarchiquement en partant des images correspondant aux objets de grande dimension : lorsqu'un objet présente un recouvrement avec son rectangle englobant supérieur à un seuil donné, son empreinte est extraite puis masquée dans l'ensemble des images correspondant à une taille d'élément structurant plus faible. Une transformée de Hough est ensuite mise en œuvre à partir d'une image binaire issue d'un seuillage de l'image gradient pour extraire les lignes de l'image. Des hypothèses de contours de toit rectangulaires sont finalement générées à partir des lignes extraites et seules sont conservées celles qui maximisent un critère d'adéquation avec les contours de l'image. On retrouve donc ici une approche similaire aux approches par agrégation de primitives, qui impose cependant la forme des contours générés.

Dans un contexte de détection de changements à partir d'un couple d'images optiques multispectrales haute résolution, Benedek *et al.* [BDZ12] proposent une méthode d'extraction de contours de toit rectangulaires basée sur les processus ponctuels marqués. Une première étape de détection grossière est réalisée en recherchant des régions pour lesquelles la radiométrie est homogène et cohérente avec une information *a priori* sur la couleur des toits recherchés. Ces régions doivent également posséder dans leur voisinage une ombre dont la position est cohérente avec la position du soleil lors de l'acquisition. Enfin, l'histogramme des directions des vecteurs gradients, calculé localement, doit présenter deux modes séparés de 90 degrés. À l'issue de cette première étape, un

² Hue Saturation Value

schéma itératif de génération-validation d'hypothèses est mis en œuvre dans le cadre des processus ponctuels marqués afin de minimiser une énergie globale. Cette énergie quantifie la compatibilité des hypothèses entre elles ainsi que l'adéquation des hypothèses avec l'image, les critères d'adéquation étant les mêmes que ceux utilisés pour la détection grossière.

Les approches développées par [SPY06, KP09, ST10] sont relativement similaires aux approches ascendantes par classification à la différence qu'elles imposent des *a priori* sur la forme des contours par l'intermédiaire de bibliothèques contenant un ensemble de formes admissibles.

La méthode proposée par [ST10], consiste à définir des régions d'intérêt par l'exploitation conjointe d'un MNS normalisé³ et des résultats d'une classification supervisée (SVM⁴) mise en œuvre à partir d'une image multispectrale. Pour chacune des régions d'intérêt, les contours des toits sont ensuite extraits au moyen d'un filtre de Canny et d'une vectorisation par transformée de Hough pour laquelle les contours sont supposés rectangulaires, circulaires ou ellipsoïdaux.

Dans [KP09], Karantzas *et al.* proposent une méthode d'extraction de contours de toits en monoscopie optique au moyen d'une segmentation par contours actifs implicites intégrant un ensemble d'*a priori* sur la forme de ces contours. Les auteurs soulignent la nécessité d'introduire ces *a priori* de forme sans lesquels l'inhomogénéité radiométrique des bâtiments conduit à des résultats insatisfaisants. La méthode proposée est basée sur une mesure de similarité de contours par homographie projective ainsi que sur une approche variationnelle permettant de considérer l'ensemble des formes *a priori* lors la segmentation.

3.2.3 Extraction tridimensionnelle des bâtiments

Dans cette section, nous abordons la problématique de l'extraction tridimensionnelle de bâtiments, qui implique l'extraction des emprises au sol (ou des contours de toits) et l'estimation de la hauteur de ces bâtiments. Ces deux tâches, qui peuvent être traitées séquentiellement, sont le plus souvent réalisées conjointement de manière à améliorer les performances en détection, aussi bien en termes de probabilité de fausses alarmes que de probabilité de détection. Nous abordons successivement les méthodes d'extraction ascendantes, pour lesquelles les bâtiments 3D sont reconstruits à partir de caractéristiques géométriques, radiométriques et contextuelles avec des contraintes géométriques faibles, et les méthodes d'extraction descendantes pour lesquelles un ou plusieurs modèles *a priori* sont imposés lors de la reconstruction des bâtiments 3D.

3.2.3.1 Les approches ascendantes

Les approches par agrégation de primitives, initialement développées pour l'extraction des contours de toits, ont fait l'objet de plusieurs travaux [LHN94, LN95, NLH97, LN98, FKL⁺98, KHN00, NN01] visant à étendre ces approches à l'extraction tridimensionnelle de bâtiments à partir d'une ou plusieurs images.

Dans [LN98], Lin *et al.* proposent une méthode de reconstruction tridimensionnelle en monoscopie basée sur l'exploitation conjointe de primitives 2D et 3D. Des hypothèses de contours de toits sont tout d'abord générées, comme cela est fait dans [MN89], par agrégation de primitives linéiques, issues d'une extraction des contours de l'image. Les hypothèses générées, en forme de parallélogramme (afin de modéliser en partie les effets géométriques liés à la projection d'un bâtiment dans une image, comme décrit à la section 2.3.2), sont ensuite sélectionnées selon des critères d'adéquation avec les primitives ponctuelles (jonctions en « T », jonctions en « L » et jonctions trihédriques) et linéiques extraites de l'image. Cette sélection locale d'hypothèses est suivie d'une sélection globale permettant de supprimer les hypothèses dupliquées et celles incluses dans d'autres hypothèses de contours. À l'issue de cette étape de sélection d'hypothèses, dont l'objectif est d'en restreindre le nombre par des considérations

³À la différence d'un modèle numérique de terrain, un modèle numérique de surface prend en compte les objets du sur-sol. Un modèle numérique de surface normalisé est obtenu par soustraction d'un modèle numérique de terrain à un modèle numérique de surface de manière à récupérer la hauteur des objets du sur-sol.

⁴Support Vector Machine

géométriques simples, les hypothèses sont finalement validées ou rejetées par vérification de la présence de façades et d'ombres projetées conformément aux positions, supposées connues, du capteur et du soleil. La vérification de la présence de façades et d'ombres projetées, réalisée pour différentes hauteurs candidates (dont les valeurs varient, par pas de 1 mètre, entre une hauteur minimale et une hauteur maximale fixées par l'utilisateur), repose sur la recherche de primitives linéiques rectilignes et de primitives ponctuelles caractéristiques de la présence de ces éléments dans l'image. Les hypothèses validées par la présence d'ombres et de façades (hypothèses pour lesquelles une information de hauteur est désormais disponible) sont finalement combinées de manière à construire des bâtiments plus complexes composés de sous structures parallélépipédiques et éventuellement de superstructures.

La méthode développée par Izadi *et al.* [IS12] pour l'extraction tridimensionnelle de bâtiments en monoscopie optique est relativement proche de celle développée par [LN98] à la différence que les contours de toit sont polygonaux et sont générés en recherchant, par programmation dynamique, des boucles fermées au sein d'un graphe dont les arêtes correspondent aux contours rectilignes extraits de l'image et les nœuds correspondent aux intersections observées et potentielles entre ces contours rectilignes.

D'autres approches [MS94, Shu96], relativement similaires aux approches par agrégation de primitives, exploitent les propriétés de la projection perspective (et notamment l'existence de points de fuite pour les structures parallèles) de manière à identifier les contours de l'image correspondant à des structures verticales ou horizontales. L'information relative à l'orientation verticale ou horizontale des contours permet alors de contraindre la génération des hypothèses de bâtiments.

Lorsque plusieurs images (deux ou plus) sont disponibles, la plupart des méthodes proposées [BM99, JPD00, FMR01, VD01, CJDS03, Tai05] réalisent l'extraction des bâtiments à partir de surfaces tridimensionnelles issues d'un MNS, construit par stéréoscopie ou multiscopie à partir des images disponibles. D'autres approches, comme celle développée par Oriot *et al.* [Ori03], travaillent directement dans l'espace des disparités associé au couple d'images disponibles, sans avoir recourt à la construction d'un MNS intermédiaire.

Dans [FMR01], Fradkin *et al.* proposent une méthode d'extraction tridimensionnelle de bâtiments en milieu urbain dense à partir d'un ensemble d'images multispectrales multivues. La première étape de cette méthode consiste à construire, pour chacune des images disponibles, des surfaces tridimensionnelles, établies au moyen d'une segmentation de l'image et d'une mise en correspondance stéréoscopique des images. Les surfaces tridimensionnelles issues des différentes images disponibles sont alors fusionnées, selon leurs propriétés radiométriques, géométriques et spatiales. Pour chacune des images disponibles, un nuage de points 3D est généré dans l'espace objet à partir de ces surfaces tridimensionnelles. Les façades des bâtiments sont ensuite détectées et extraites lors d'une seconde étape au cours de laquelle les points 3D précédemment générés sont projetés au sol et regroupés au sein d'un accumulateur. Les cellules de l'accumulateur contenant au moins deux points indiquent la présence d'une structure verticale. Les segments correspondant à l'emplacement planimétrique des façades sont alors extraits par mise en œuvre d'une transformée de Hough à partir de l'image binaire obtenue par seuillage de l'accumulateur. Les bâtiments sont finalement reconstruits par combinaison des surfaces 3D extraites dans la première étape et des façades détectées lors de la seconde étape.

L'approche développée par Taillandier [Tai05] réalise une reconstruction tridimensionnelle des bâtiments présentant des toits bi-pentes ou plats (non nécessairement horizontaux) à partir d'un ensemble d'images et de données cadastrales, spécifiant l'emprise au sol des bâtiments. Cette approche consiste tout d'abord à associer à chaque segment constitutif de l'emprise au sol d'un bâtiment deux plans obliques orientés arbitrairement à -45 et 45 degrés par rapport à l'horizontale. Un graphe 3D est ensuite établi, graphe au sein duquel les nœuds correspondent aux différents plans et les arêtes correspondent aux relations admissibles entre ces plans. Des hypothèses de toit admissibles sont alors extraites de ce graphe 3D par recherche des cliques maximales⁵. Un critère de corrélation centré est alors défini de manière à quantifier l'adéquation de ces hypothèses avec un MNS établi par corrélation entre les différentes images disponibles. L'hypothèse pour laquelle le critère de corrélation est maximal

⁵ Au sein d'un graphe, une clique est définie comme un ensemble de nœuds qui sont tous reliés deux à deux par une arête. Une clique est dite maximale si l'ajout d'un nœud quelconque à cette clique met en défaut la propriété précédente.

est finalement retenue et les paramètres associés à cette hypothèse (hauteur des façades associées à chacun des segments de l'emprise au sol et pente des différents pans du toit) sont estimés par minimisation de la distance (au sens de la norme L1) entre l'hypothèse de toit et le MNS.

Dans [Ori03], Oriot *et al.* ont développé une méthode basée sur la segmentation sous contraintes de l'espace des disparités associées à un couple d'images stéréoscopiques. Une première étape de détection supervisée conduit à la définition d'un ensemble de régions d'intérêt contenant chacune un bâtiment. L'espace des disparités associé à chacune de ces régions est alors segmenté par contours actifs en deux zones distinctes : bâtiment et sol. Le contour polygonal optimal est obtenu par minimisation de coûts de corrélation définis sur les régions toit et sol tout en imposant une adéquation avec les contours de l'image (terme d'attache aux données) et en favorisant les contours rectilignes (termes de régularisation).

3.2.3.2 Les approches descendantes

Parmi les approches descendantes proposées pour l'extraction tridimensionnelle de bâtiments, nous distinguons la méthode développée par Karantzas *et al.* [KP10], qui repose sur une segmentation de l'image contrainte par des modèles 3D de bâtiments, et les méthodes qui s'appuient sur les processus ponctuels marqués [LDZPD06, ODZ07, TBBD10].

Dans [KP10], Karantzas *et al.* proposent une méthode d'extraction tridimensionnelle de bâtiments à partir d'une image panchromatique ou multispectrale et d'un MNS. Cette méthode, qui dérive de l'approche proposée dans [KP09] pour l'extraction de contours de toit, est basée sur une segmentation conjointe, par contours actifs implicites, de l'image et du MNS (considéré comme une image) afin de distinguer les régions « bâtiments » des régions « fond ». L'énergie globale minimisée au cours de la segmentation fait intervenir trois termes distincts dont le premier quantifie l'homogénéité radiométrique des régions de la partition courante au sein de l'image et du MNS. Le deuxième terme quantifie l'adéquation de la forme des régions « bâtiment » avec un ensemble de formes admissibles, la mise en correspondance étant réalisée par homographie projective. Enfin, le dernier terme quantifie l'adéquation de la surface tridimensionnelle associée à chaque région « bâtiment » avec un ensemble de modèles de toits admissibles. Une minimisation par descente de gradient conduit alors à une reconstruction tridimensionnelle des bâtiments.

Des approches fondées sur les processus ponctuels marqués [LDZPD06, ODZ07, TBBD10] ont également été développées pour réaliser l'extraction tridimensionnelle de bâtiments à partir de MNS et/ou d'images haute résolution. Dans [ODZ07], Ortner *et al.* présentent une méthode d'extraction d'emprises au sol rectangulaires à partir d'un MNS, pouvant être généré par mise en correspondance stéréoscopique de deux images. La méthode repose sur la minimisation par recuit simulé de l'énergie globale associée à un ensemble d'emprises au sol, cette énergie prenant en compte l'adéquation de chaque emprise avec le MNS (en considérant notamment le gradient de ce MNS) et la cohérence géométrique des emprises entre elles. La hauteur des bâtiments détectés est alors déterminée en calculant la valeur moyenne du MNS à l'intérieur des emprises au sol associées à ces bâtiments.

3.3 Extraction des bâtiments en imagerie RSO

Nous nous intéressons, dans cette partie, à l'extraction des bâtiments à partir d'une ou plusieurs images RSO haute ou très haute résolution, acquises par des capteurs satellitaires ou aéroportés. Nous abordons, dans un premier temps, la problématique de l'extraction des primitives caractéristiques de la présence des bâtiments dans les images RSO et sur lesquelles s'appuient la plupart des méthodes d'extraction. Nous présentons, dans un second temps, un ensemble de méthodes qui ont été développées pour l'extraction des emprises au sol. Dans un dernier temps, nous décrivons plusieurs techniques qui réalisent l'extraction tridimensionnelle des bâtiments à partir d'une unique image ou à partir de plusieurs images multiaspects ou en conditions interférométriques.

3.3.1 Extraction de primitives

Comme en imagerie optique, les bâtiments se manifestent au sein des images RSO par un ensemble de primitives ponctuelles, linéiques et surfaciques qui constituent la signature géométrique des bâtiments dans ces images. Dans cette section, nous décrivons ces différentes primitives et présentons les méthodes qui ont été établies pour les extraire.

3.3.1.1 Primitives ponctuelles

Les primitives ponctuelles sont des points brillants qui traduisent une forte rétrodiffusion de l'onde RADAR vers le capteur en raison de la présence de structures triédriques (cf. section 2.4.2) ou d'objets de faibles dimensions avec des propriétés diélectriques particulières. Les points brillants d'une image RSO peuvent résulter de la présence de bâtiments mais aussi d'autres objets de l'environnement urbain, de sorte que ces primitives ponctuelles sont relativement peu caractéristiques de la présence des bâtiments. Plusieurs détecteurs [LTN90, PAD00, MSS02] ont été développés pour détecter ces points brillants qui demeurent cependant peu utilisés pour la détection des bâtiments.

3.3.1.2 Primitives linéiques

Les primitives linéiques, caractéristiques de la présence de bâtiments dans une image, sont des lignes brillantes issues de la réflexion de l'onde RADAR sur des structures présentant une géométrie ou des propriétés diélectriques particulières. Nous distinguons :

- les lignes brillantes qui résultent d'une forte rétrodiffusion de l'onde RADAR sur des structures rectilignes présentant des propriétés diélectriques particulières, comme les gouttières métalliques ou encore les bandes d'arêtier métalliques se trouvant à la jonction entre deux pans d'un toit ;
- les lignes brillantes qui résultent de deux réflexions spéculaires de l'onde RADAR (écho double) sur les plans constitutifs des structures diédriques formées par les façades du bâtiment avec le sol. Comme indiqué à la section 2.4.2, la géométrie des structures diédriques implique une remise en phase de l'onde reçue par une façade orientée vers le capteur quel que soit l'angle d'incidence de l'onde RADAR. Ainsi, l'énergie électromagnétique réfléchi par cette façade est intégrée au sein de la même case distance, ce qui se traduit dans l'image résultante par une forte intensité radiométrique ;
- les lignes brillantes qui résultent de trois réflexions (deux réflexions spéculaires et une réflexion lambertienne) de l'onde RADAR (écho triple) sur les plans constitutifs des structures diédriques introduites précédemment. Ces lignes brillantes, qui correspondent à des réflexions façade-sol-façade ou sol-façade-sol, n'engendrent pas, comme c'est le cas pour les échos doubles, une remise en phase de l'onde reçue par la façade, de sorte que l'énergie réfléchi est répartie sur plusieurs cases distance. De plus, ces échos triples font intervenir une réflexion lambertienne pour laquelle l'énergie de l'onde réfléchi est relativement faible par rapport à l'énergie de l'onde incidente ce qui se traduit par une faible intensité radiométrique de ces lignes brillantes dans l'image.

Parmi les méthodes d'extraction de lignes brillantes en imagerie RSO, nous pouvons citer les travaux de Touzi *et al.* [TLB88], de Tupin *et al.* [TMM⁺98] ou, plus récemment, ceux de Brett *et al.* [BG11].

Sous hypothèse d'une modélisation du bruit de *speckle* par une loi Gamma, Touzi *et al.* [TLB88] établissent la loi statistique du rapport des radiométries moyennes calculées dans deux régions voisines d'une image RSO en intensité. La présence d'un bord est décidée lorsque la valeur de ce rapport, calculé pour différentes orientations relatives des deux régions, est supérieur à un seuil déterminé en fonction de la probabilité de fausse alarme du détecteur de bords.

Tupin *et al.* [TMM⁺98] ont développé un détecteur de lignes brillantes basé sur la fusion par somme symétrique associative de deux détecteurs de bords. Le premier détecteur mis en œuvre est celui de Touzi *et al.* [TLB88] tandis que le second détecteur repose sur la corrélation croisée entre les

échantillons présents dans le voisinage du pixel testé avec des échantillons virtuels qui seraient observés en présence d'un bord. La fusion de ces deux détecteurs permet d'exploiter la précision de localisation du second détecteur et la robustesse du premier détecteur aux inhomogénéités radiométriques dans le voisinage des bords.

Enfin, Brett *et al.* [BG11] ont développé un détecteur de lignes brillantes pour les images RSO très haute résolution. Ce détecteur repose sur la détection de points correspondant à des *extrema* locaux de la radiométrie de l'image calculés dans un espace échelle. La mise en œuvre d'un classifieur bayésien permet alors de décider de l'appartenance des points détectés à une ligne brillante selon la valeur de leur radiométrie et de la courbure radiométrique locale.

3.3.1.3 Primitives surfaciques

La présence de bâtiments dans une image RSO se traduit par la présence de primitives surfaciques, régions radiométriquement homogènes qui constituent la signature géométrique de ces bâtiments dans l'image. Parmi ces primitives surfaciques, nous distinguons, conformément aux observations de la section 2.4.2 :

- les régions ombres pour lesquelles aucune onde n'est rétrodiffusée vers le capteur RSO, de sorte que ces régions présentent une faible radiométrie liée uniquement au bruit thermique du capteur ;
- et les régions associées à un phénomène de repliement traduisant une rétrodiffusion simultanée du sol et de plusieurs éléments constitutifs d'un bâtiment (façade et/ou toit), de sorte que l'énergie reçue par le capteur et donc la radiométrie de ces régions est relativement importante.

L'extraction de ces primitives surfaciques peut être réalisée en monoscopie comme en interférométrie par diverses méthodes identifiées ci-après.

En monoscopie RSO, l'extraction des différentes primitives surfaciques peut être réalisée, comme cela est fait par Bennett *et al.* [BB03] pour l'extraction des ombres RADAR, au moyen d'une segmentation par contours actifs statistiques. Cette méthode permet de séparer une région polygonale d'intérêt d'une région de fond par minimisation de la longueur du contours polygonal et par maximisation de la vraisemblance statistique des pixels contenus dans les deux régions (pour une modélisation statistique du bruit de *speckle* donnée). Contrairement à la méthode précédente qui conduit à une partition de l'image en deux régions, la mise en œuvre d'une grille active statistique [GBR03], dont l'évolution est régie par la minimisation de la complexité stochastique de l'image⁶, conduit à une partition de l'image en plusieurs régions.

D'autres méthodes sont basées sur des considérations géométriques et radiométriques relativement simples, comme celle développée par Bolter [Bol00] qui réalise l'extraction des ombres en appliquant successivement un filtre passe-bas et un seuillage radiométrique sur l'image d'amplitude. L'image binaire ainsi obtenue est filtrée en supprimant les régions de petite taille et les contours des régions sont finalement régularisés par mise en œuvre d'opérateurs morphologiques.

Dans le cas où l'on dispose de données interférométriques, plusieurs méthodes [Tis04, CON05], exploitant conjointement les images d'amplitude et l'image de cohérence⁷, ont été développées pour l'extraction des primitives surfaciques.

Tison *et al.* [Tis04] proposent une méthode basée sur une classification markovienne conjointe des images d'amplitude et de l'image de cohérence. Les zones d'ombres correspondent alors à la classe

⁶La notion de complexité stochastique est précisée à la partie 4.2.

⁷En notant I_1 et I_2 les variables aléatoires décrivant respectivement les valeurs (complexes) des pixels des deux images constitutives d'un couple interférométrique, la cohérence γ entre ces deux images est définie comme suit : $\gamma = \frac{|E[I_1 I_2^*]|}{\sqrt{E[I_1 I_1^*] E[I_2 I_2^*]}}$, où $E[X]$ correspond à l'espérance de la variable aléatoire X et X^* désigne la valeur complexe conjuguée de X . La cohérence est une mesure de corrélation qui permet de quantifier la qualité d'un interférogramme, de caractériser l'interaction onde-matière des différentes surfaces ou encore de détecter les changements qui impliquent une modification de phase.

présentant la plus faible radiométrie au sein des images d'amplitude et la plus faible cohérence tandis que les zones de toit et de façade correspondent aux classes présentant une forte cohérence.

Pour l'extraction des ombres, Cellier *et al.* [CON05] réalisent tout d'abord un filtrage anti-speckle (filtre *Mean Shift*) des images d'amplitude et de l'image de cohérence puis construisent une pyramide d'images binaires pour chacune de ces images par seuillage radiométrique régulier. À chacune des images binaires de la pyramide associée à l'image de cohérence sont associées deux images binaires issues respectivement des pyramides associées aux deux images d'amplitude. Parmi les images binaires résultant du seuillage d'une image d'amplitude, celle qui est associée à une image binaire donnée de la pyramide résultant du seuillage de l'image de cohérence est celle pour laquelle le pourcentage de pixels labellisés dans les deux images comme appartenant à une ombre est maximal. Pour chacune des images binaires de la pyramide résultant du seuillage de l'image de cohérence, l'intersection de cette image avec ses deux images binaires associées est finalement réalisée. La pyramide résultante contient alors des images binaires qui correspondent à une extraction des ombres pour différentes valeurs des probabilités de détection et de fausse alarme.

3.3.2 Extraction des emprises au sol des bâtiments

Nous présentons dans cette section, un ensemble de méthodes permettant d'extraire les emprises au sol des bâtiments à partir de données monoscopiques, radargrammétriques ou interférométriques. Ces méthodes, dont la plupart se restreignent à des emprises au sol de forme rectangulaire, reposent sur une extraction des primitives linéiques et surfaciques au moyen des méthodes introduites à la section précédente.

En monoscopie RSO, Michaelsen *et al.* [MSSD08] et Ferro *et al.* [FBB09, FBB10] ont développé des méthodes d'extraction relativement proches des approches ascendantes proposées en imagerie optique pour l'extraction des contours de toit.

Dans [MSSD08], Michaelsen *et al.* proposent une méthode pour l'extraction d'emprises au sol polygonales qui est basée sur une agrégation de primitives ponctuelles et linéiques par organisation perceptive. La méthode proposée se décompose en quatre étapes principales. La première étape consiste à appliquer des prétraitements visant à faciliter l'extraction de primitives ponctuelles et linéiques lors de la seconde étape. Ainsi, les points brillants sont extraits de l'image résultant d'une sous résolution d'un facteur cinq suivi d'une ouverture morphologique appliqués à l'image originale. Les lignes brillantes sont quant à elles extraites au moyen d'un filtre gradient moyen [För94] à partir de l'image résultant de l'application d'une fermeture morphologique sur l'image initiale, traitement visant à renforcer les lignes fines et fermer d'éventuelles discontinuités au sein de ces lignes. La troisième étape de cette méthode repose sur une agrégation des primitives précédemment extraites de manière à former des structures plus complexes à partir desquelles un ensemble d'emprises au sol sont finalement générées lors de la dernière étape.

Dans [FBB09, FBB10], Ferro *et al.* proposent une méthode pour la détection des bâtiments et la reconstruction des emprises au sol associées, dont la forme est supposée être rectangulaire. Un filtre Gamma MAP [LNTL93] est tout d'abord appliqué à l'image afin de réduire le bruit de *speckle*. Des primitives linéiques et surfaciques sont ensuite extraites à partir de l'image filtrée : les lignes brillantes sont extraites à l'aide de l'opérateur de Steger [Ste98], les régions brillantes sont extraites par mise en œuvre du détecteur de Tupin *et al.* [TMM⁺98] et les régions d'ombre sont extraites au moyen du détecteur de Cellier *et al.* [CON05]. Les primitives extraites sont alors vectorisées et caractérisées par un ensemble d'attributs radiométriques et géométriques. À partir de ces attributs, des fonctions d'appartenance établies empiriquement permettent d'associer (avec un certain degré d'appartenance) chacune des primitives brillantes extraites (ligne ou région) à l'une des six classes sémantiques pouvant correspondre à cette primitive brillante : façade, écho double, structure linéique du toit, fort écho de la région toit, toit plat et toit bi-pente. Des hypothèses d'emprises au sol sont ensuite générées à partir des structures extraites et de structures composites, établies par composition de structures élémentaires présentant des caractéristiques similaires. L'adéquation de chacune des hypothèses avec l'image est finalement quantifiée par un score dont le calcul est basé sur le degré

d'appartenance et la position relative des différentes structures constitutives de l'hypothèse. Seules sont alors conservées les hypothèses présentant un score supérieur à un seuil défini par l'utilisateur.

En radargrammétrie, Simonetto *et al.* [SOGC03, SOG05] proposent une méthode pour l'extraction des emprises au sol correspondant à des bâtiments parallélépipédiques de grandes dimensions. La première étape de cette méthode consiste à extraire les lignes brillantes présentes dans chacune des images constitutives du couple. Les images binaires résultant de l'extraction de ces lignes brillantes sont obtenues en fusionnant les images binaires issues du détecteur de lignes brillantes de Tupin *et al.* [TMM⁺98] et d'une classification des images en deux classes (pixels brillants et pixels sombres) par mise en œuvre d'une chaîne de Markov. La seconde étape de cette méthode est relative à la détection des échos doubles (en forme de « L ») au sein des images binaires associées aux lignes brillantes. Dans l'hypothèse où tous les bâtiments sont orientés suivant la même direction, cette direction est estimée au moyen d'une transformée de Hough globale appliquée sur toute l'image binaire. Des transformées de Hough locales sont ensuite mises en œuvre afin d'extraire les structures orientées selon cette direction globale et perpendiculairement à cette direction, structures à partir desquelles les intersections en forme de « X », de « L » ou de « T » sont détectées. Lors d'une troisième étape, les intersections détectées indépendamment à partir de chacune des images binaires sont mises en correspondance par programmation dynamique [BD62]. Seules sont alors conservées les intersections pour lesquelles les segments constitutifs sont à la même hauteur (déterminée par stéréoscopie), permettant ainsi d'éliminer les intersections entre lignes brillantes correspondant à des structures au sol et des structures en hauteur. Chacun des segments constitutifs d'une intersection est alors modélisé par une région rectangulaire dont la position, l'orientation et les dimensions sont estimées par maximisation d'un critère basé sur l'appartenance des pixels détectés dans les deux images binaires à cette région et sur l'appartenance des pixels non détectés à une région fond. Les emprises au sol, supposées rectangulaires, sont alors générées à partir de ces segments dont la longueur a été précédemment estimée.

Enfin, Thiele *et al.* [TCS⁺07b] ont développé une méthode pour l'extraction des emprises au sol de bâtiment à partir de deux jeux de données interférométriques multiaspects, pour lesquels les trajectoires des porteurs sont orthogonales. Après une première étape de recalage par maximisation de cohérence, les quatre images d'amplitude disponibles (deux images par jeu de données) sont fusionnées au moyen des opérateurs maximum, minimum et moyenne conduisant ainsi à trois images d'amplitude. Les lignes brillantes sont alors extraites de l'image issue de la fusion par l'opérateur maximum au moyen de l'opérateur de Steger [Ste98], tandis que les contours des toits et des ombres sont extraits des images résultant de la fusion par les opérateurs moyenne et minimum par mise en œuvre de l'opérateur de Canny [Can86] et d'une méthode de segmentation [TCS⁺06]. Des hypothèses de contours (en forme de quadrilatère) sont ensuite générées à partir des primitives extraites en imposant un ensemble de contraintes géométriques ainsi qu'une cohérence des hypothèses générées avec les géométries de prise de vue des deux jeux de données (au moins un des segments orientés face aux capteurs doit correspondre à une ligne brillante et au moins un des segments non illuminés par les capteurs doit correspondre à un des contours de l'ombre). Finalement, les hypothèses de contours sont validées ou rejetées selon deux critères. Le premier critère est relatif à la taille des bâtiments de sorte que les contours dont l'aire est inférieure à un seuil sont rejetés. Le second critère est relatif à la variabilité du MNS, calculé à partir des couples interférométriques, au sein d'un même contour de sorte que les contours pour lesquels la variance de la hauteur est trop élevée sont rejetés. En cas de recouvrement entre deux emprises au sol, celle qui présente la meilleur adéquation avec les images, au sens des deux critères précédents, est conservée.

3.3.3 Extraction tridimensionnelle des bâtiments

Dans cette section, nous abordons l'extraction tridimensionnelle des bâtiments qui recouvre essentiellement la problématique de l'estimation de la hauteur des bâtiments. Nous abordons successivement les méthodes qui réalisent l'estimation de la hauteur des bâtiments dont l'emprise au sol est supposée connue et les méthodes qui permettent de réaliser une extraction tridimensionnelle complète des

bâtiments. Comme à la section précédente, nous présentons des méthodes pour lesquelles l'extraction est réalisée à partir de données monoscopiques, radargrammétriques ou interférométriques.

3.3.3.1 Estimation de la hauteur des bâtiments

Plusieurs méthodes ont été développées pour l'estimation de la hauteur d'un bâtiment dont l'emprise au sol est connue. Hormis la méthode de Franceschetti *et al.* [FIR02, FGI+06, FGI+07, GIR10], qui repose sur la modélisation du champs électromagnétique réfléchi par un bâtiment, les méthodes que nous présentons ici sont basées sur la mesure de l'adéquation entre les données (monoscopiques ou interférométriques) à partir desquelles sont réalisées l'estimation et des données simulées pour différentes hypothèses de hauteur.

Parmi les méthodes développées en monoscopie RSO pour l'estimation de la hauteur d'un bâtiment d'emprise au sol connue, nous pouvons citer les travaux de Franceschetti *et al.* [FIR02], de Brunner *et al.* [BLBG10] ou encore ceux de Sportouche *et al.* [STD11].

Dans [FIR02], Franceschetti *et al.* exploitent conjointement les théories de l'optique géométrique et de l'optique ondulatoire afin d'établir l'expression du champs électromagnétique réfléchi par un bâtiment de forme parallélépipédique illuminé par un capteur RSO. Le champs réfléchi résulte de réflexions simples sur le sol (modélisé par une surface rugueuse), le toit et les façades du bâtiment (modélisés par des surfaces lisses) ainsi que de réflexions multiples entre le sol et les façades du bâtiment qui forment une structure diédrique. Alors que les réflexions sur les structures du bâtiment sont considérées comme spéculaires, celles sur le sol sont considérées comme lambertiennes ou spéculaires selon la rugosité de ce dernier. Une inversion de l'expression du champs électromagnétique conduit finalement, comme cela est fait dans [FGI+06], à une expression de la hauteur d'un bâtiment en fonction :

- de la SER⁸ associée à l'écho double d'un bâtiment ;
- des dimensions planimétriques du bâtiment ;
- des propriétés diélectriques du sol et des façades du bâtiment ;
- des paramètres du processus stochastique permettant de modéliser la rugosité du sol ;
- de l'orientation relative du bâtiment et du capteur ;
- de l'angle d'incidence de l'onde RADAR.

Cette approche analytique est intéressante puisqu'elle permet de quantifier l'influence des différents paramètres physiques impliqués dans la formation de l'écho double des bâtiments. Elle demeure, en revanche, difficilement exploitable pour l'estimation de la hauteur d'un bâtiment aux vues du nombre et de la nature des paramètres dont la connaissance *a priori* est nécessaire pour pouvoir mettre en œuvre cette estimation.

La méthode proposée par Brunner *et al.* [BLBG10] repose sur la simulation d'images RSO pour différentes hypothèses de hauteur et sur une mesure de l'adéquation entre ces images simulées et l'image réelle. Les images RSO simulées, pour lesquelles la géométrie de prise de vue est identique à celle de l'image réelle, sont obtenues au moyen d'un simulateur basé sur une méthode de lancé de rayons. Ce simulateur prend en compte les réflexions multiples mais néglige en revanche le bruit de *speckle* résultant de la sommation cohérente des ondes réfléchies par les réflecteurs élémentaires. L'adéquation entre une image simulée et l'image réelle est mesurée en calculant l'information mutuelle entre les variables aléatoires représentant les niveaux de gris des pixels des deux images. L'image simulée pour laquelle l'information mutuelle avec l'image réelle est la plus élevée conduit alors à la valeur estimée de la hauteur du bâtiment.

Dans [STD11], Sportouche *et al.* proposent une méthode basée sur le calcul d'une vraisemblance radiométrique généralisée pour quantifier l'adéquation entre l'image RSO et les signatures géométriques résultant de la projection des bâtiments dans cette image pour différentes hypothèses de hauteur. Comme cela est fait à la section 2.4.3, les bâtiments, dont la forme est supposée être parallélépipédique et dont l'emprise au sol est connue, sont projetés dans l'image RSO au moyen de trois modèles de rétrodiffusion : le modèle occultant, le modèle partiellement occultant et le modèle

⁸Surface Equivalente RADAR

non occultant. La signature géométrique résultant de la projection d'un bâtiment dans l'image RSO est alors constituée de plusieurs régions qui présentent, lorsque l'hypothèse de hauteur est la bonne, une radiométrie homogène dont la variabilité est liée uniquement au bruit de *speckle*. L'adéquation entre l'image RSO et la signature résultant de la projection d'un bâtiment dans cette image est mesurée en calculant la vraisemblance généralisée des niveaux de gris des pixels constitutifs des différentes régions de la signature géométrique. En supposant que les niveaux de gris des pixels d'une même région sont des réalisations indépendantes d'une variable aléatoire suivant une loi gamma (pour une image en intensité), la log-vraisemblance des pixels de cette région s'écrit comme la somme des log-vraisemblances de ces pixels. La log-vraisemblance de l'ensemble des pixels de la signature correspond alors à la somme des log-vraisemblances calculées pour les différentes régions constitutives de la signature géométrique. Finalement, la hauteur estimée est la hauteur pour laquelle la signature géométrique associée présente la vraisemblance la plus élevée.

Enfin, Thiele *et al.* [THC10] ont développé une approche permettant d'estimer la hauteur d'un bâtiment parallélépipédique d'emprise au sol connue à partir d'un couple d'images RSO en conditions interférométriques. Cette approche repose sur la simulation, pour différentes hypothèses de hauteur, de l'interférogramme obtenu à partir des deux images. La méthode mise en œuvre pour la simulation de l'interférogramme, méthode présentée dans [TCS⁺07a], consiste à sommer les phases interférométriques résultant de la rétrodiffusion simultanée du sol et de un ou plusieurs éléments constitutifs du bâtiment. Finalement, l'adéquation entre l'interférogramme simulé et l'interférogramme calculé à partir des images réelles est mesurée au moyen d'une corrélation. L'hypothèse de hauteur pour laquelle la corrélation entre les deux interférogrammes est maximale correspond alors à la valeur estimée de la hauteur du bâtiment.

3.3.3.2 Extraction tridimensionnelle

Les méthodes présentées à la sous-section précédente permettent d'estimer la hauteur des bâtiments dans le cas où l'emprise au sol de ces derniers est supposée connue. Les approches présentées dans cette sous-section réalisent conjointement l'extraction de l'emprise au sol des bâtiments ainsi que l'estimation de leur hauteur à partir d'une ou plusieurs images RSO. Dans cette sous-section, nous abordons successivement les méthodes réalisant une extraction tridimensionnelle des bâtiments à partir d'une unique image RSO, celles qui réalisent cette extraction à partir d'un ensemble d'images multiaspects et enfin celles basées sur des données interférométriques.

Parmi les méthodes développées pour l'extraction tridimensionnelle de bâtiments à partir d'une unique image RSO, nous distinguons les approches développées par Quartulli *et al.* [QD04] et par Le Moigne *et al.* [MTN07].

Dans [QD04], Quartulli *et al.* proposent une méthode basée sur une représentation hiérarchique de la scène imagée et sur la mise en œuvre d'un algorithme d'optimisation stochastique permettant de déterminer la représentation qui explique au mieux l'image RSO. La scène est représentée par un ensemble de bâtiments, bâtiments qui sont décomposés en un ensemble de facettes, qui sont elles mêmes décomposées en un ensemble de pixels. À chaque représentation de la scène est associée une énergie prenant en compte l'interaction entre les différents bâtiments constitutifs de cette représentation et l'adéquation des facettes constitutives des bâtiments avec les échos doubles et les zones de repliement présents dans l'image. L'énergie d'interaction entre les différents bâtiments d'une représentation est basée sur le calcul de l'aire d'intersection entre les emprises au sol de ces bâtiments, de sorte que les représentations présentant des chevauchements entre les emprises au sol des bâtiments sont pénalisées. En ce qui concerne l'adéquation de la représentation avec l'image RSO, elle est mesurée en considérant l'ensemble des régions de l'image qui, aux vues de la disposition relative des facettes et du capteur RSO, sont sensées présenter une forte rétrodiffusion. La mesure d'adéquation fait alors intervenir des rapports de vraisemblance visant à déterminer l'appartenance de chacun des pixels constitutifs de ces régions à une zone de forte rétrodiffusion. La mise en œuvre d'un algorithme d'optimisation par recuit simulé permet finalement de déterminer la représentation présentant l'énergie la plus faible qui explique au mieux l'image observée.

Dans [MTN07], Le Moigne *et al.* ont développé une méthode qui repose sur la segmentation, par contours actifs statistiques, de la zone de repliement liée à la présence d'un bâtiment à partir des images d'intensité et de phase issues de l'image RSO. L'initialisation du contour de la zone de repliement est réalisée en considérant la position du capteur RSO et des ombres, extraites au moyen d'une classification markovienne de l'image selon la méthode de Tison *et al.* [Tis04]. L'évolution du contour est ensuite régie par un critère faisant intervenir conjointement les images d'intensité et de phase. Ce critère, basé sur les travaux de Chesnaud *et al.* [CRB99], repose sur le calcul de la log-vraisemblance des niveaux de gris des pixels des images d'intensité et de phase (dont il est fait l'hypothèse qu'ils suivent respectivement une loi gamma et une loi normale) et permet de quantifier l'adéquation du contour avec l'image. Le contour résultant est alors régularisé en réduisant le nombre de nœuds au moyen d'un critère de longueur minimale de description (*minimum description length*) [GBR03]. La hauteur du bâtiment est finalement estimée grossièrement en calculant la différence de phase entre les pixels correspondant à une réflexion par le toit du bâtiment et les pixels correspondant à une réflexion par le sol.

En ce qui concerne l'extraction tridimensionnelle des bâtiments à partir d'un ensemble d'images multiaspects, les méthodes développées par Bolter [Bol00] et Bennett *et al.* [BB03] sont relativement proches. Ces approches réalisent une reconstruction tridimensionnelle des bâtiments, dont la forme est supposée parallélépipédique, à partir des primitives surfaciques caractéristiques des bâtiments (ombres et/ou zones de repliement) qui sont extraites des images au moyen des méthodes introduites à la section 3.3.1. Pour chacune des lignes des images disponibles, les longueurs associées aux zones d'ombre et de repliement le long de cette ligne permettent de déterminer, pour un modèle de toit donné, la hauteur et la largeur apparente du bâtiment le long de cette ligne. Les résultats obtenus indépendamment à partir des lignes des différentes images disponibles peuvent alors être fusionnés au sein d'un accumulateur bidimensionnel en géométrie terrain ce qui conduit à une reconstruction tridimensionnelle du bâtiment.

Parmi les approches développées pour l'extraction tridimensionnelle des bâtiments à partir de données interférométriques, nous distinguons :

- les méthodes, comme celles développées par Gamba *et al.* [GHS00], qui réalisent l'extraction des bâtiments à partir d'un MNS issu des données interférométriques ;
- les méthodes, comme celles proposées par Tison *et al.* [TTM04] et Cellier *et al.* [CON06], qui exploitent conjointement les images d'amplitude et l'interférogramme pour réaliser cette extraction.

Les travaux de Gamba *et al.* [GHS00] reposent sur le partitionnement du MNS issu des données interférométriques en régions planes. Les régions présentant une altitude supérieure à celle du sol correspondent alors aux bâtiments de la scène, bâtiments pour lesquels l'emprise au sol et la hauteur correspondent respectivement au contour de la région et à l'altitude moyenne au sein de cette région. La première étape de la méthode proposée repose sur une segmentation des lignes constitutives du MNS (considéré comme une image) en segments. Cette segmentation est réalisée en divisant itérativement les segments en deux parties jusqu'à ce qu'aucun point du profil altimétrique correspondant à la ligne du MNS ne se trouve à une distance supérieure à un seuil donné. La seconde étape de la méthode consiste à réaliser une segmentation du MNS à partir des segments précédemment extraits par une méthode de croissance de région permettant d'agréger les segments dont la distance à une région, dans l'espace tridimensionnel, est inférieure à un seuil donné. Les graines permettant d'initialiser l'algorithme sont constituées par trois segments adjacents appartenant à trois lignes distinctes et présentant une longueur minimale fixée par un seuil. Finalement, la dernière étape consiste à approcher chaque région par le plan horizontal le plus proche.

Dans [TTM04], Tison *et al.* proposent une méthode dont la première étape consiste à extraire les zones d'ombre à partir de l'image d'amplitude au moyen de la méthode décrite à la section 3.3.1. La position et la longueur de la zone d'ombre le long d'une ligne de l'image fournissent respectivement la position de la façade la plus éloignée du capteur et une première estimation de la hauteur du bâtiment. La seconde étape de la méthode proposée repose sur une partition de chaque ligne de l'image en prenant conjointement en compte l'image d'amplitude et l'image de phase. Sur chacune des lignes, les positions

extrémales de l'ombre étant connues, il ne reste plus qu'à déterminer les deux positions permettant de délimiter la région de recouvrement et la région correspondant à une rétrodiffusion du toit seul. La partition des lignes est réalisée en minimisant une énergie exprimée comme la somme pondérée de deux termes :

- le premier terme assure la cohérence entre la hauteur du bâtiment mesurée à partir de la longueur de l'ombre et celle déduite, pour une partition candidate, de la différence de phase entre la région correspondant à une rétrodiffusion du toit et celle correspondant à une rétrodiffusion du sol ;
- le second terme quantifie l'adéquation des niveaux de gris des pixels constitutifs de la ligne de l'image d'amplitude avec la partition candidate au moyen d'un calcul de vraisemblance statistique. La vraisemblance des niveaux de gris des pixels d'une ligne est calculée en supposant que ces niveaux de gris suivent une loi de Fisher dont les paramètres sont estimées par apprentissage pour différentes classes : sol, végétation sombre, végétation brillante, toit sombre, toit de radiométrie moyenne et toit brillant.

Finalement, les partitions réalisées indépendamment sur chacune des lignes de l'image conduisent à des régions dont les contours sont bruités de sorte que la dernière étape de cette méthode consiste à régulariser les contours obtenus.

Pour conclure avec les méthodes d'extraction tridimensionnelle à partir de données interférométriques, nous présentons les travaux de Cellier *et al.* [CON06] qui reposent sur un système de gestion d'hypothèses générées à partir de primitives linéiques et surfaciques extraites conjointement à partir des images d'amplitude et de phase. Les lignes brillantes (correspondant à des échos simples ou doubles rétrodiffusés par les structures du toit, de la façade ou par les structures diédriques) et les ombres sont extraites au moyen des méthodes développées respectivement par Simonetto *et al.* [SOG05] et Cellier *et al.* [CON05], méthodes présentées aux sections 3.3.2 et 3.3.1. Les façades du bâtiment sont alors reconstruites à partir des segments résultant de la vectorisation des primitives en considérant l'ensemble des régions délimitées, dans la direction du capteur, par un segment résultant de la présence d'une ligne brillante et, dans la direction opposée au capteur, par un segment correspondant au bord d'une ombre. Parmi ces régions, seules sont conservées celles pour lesquelles la hauteur du bâtiment calculée à partir de la longueur de cette région selon l'axe des distances est cohérente avec la hauteur déduite de la valeur moyenne de la phase interférométrique sur cette région. L'agencement spatial dans l'espace terrain entre les éléments tridimensionnels extraits (façades et segments) est ensuite représenté au sein d'un graphe où les segments correspondant aux extrémités des façades sont représentés par des nœuds uniques tandis que les autres segments sont représentés par plusieurs nœuds correspondant aux différentes interprétations sémantiques pouvant être associées à ces segments (extrémité de l'ombre proche ou éloignée du capteur, extrémité du toit proche ou éloignée du capteur, structure diédrique). La construction initiale de ces graphes repose sur les structures tridimensionnelles présentant une homogénéité de leur agencement selon l'axe des distances pour différentes positions azimutales. Les graphes sont ensuite simplifiés, par suppression de nœuds ou d'arêtes, selon des règles simples qui intègrent l'information *a priori* sur l'agencement des primitives caractéristiques des bâtiments dans les images RSO. Les arêtes du graphe correspondant à un toit sont ensuite modifiées, par une introduction éventuellement de nouveaux nœuds, en considérant les informations altimétriques fournies par l'interférogramme. Pour finir, les incohérences topologiques au sein du graphe (comme les croisements d'arêtes) sont supprimées conduisant ainsi aux bâtiments extraits.

3.4 Extraction de bâtiments à partir d'images optiques et RSO

Dans cette dernière partie, nous nous intéressons à la problématique de l'extraction de bâtiments à partir d'un ensemble d'images optiques et RSO. Parmi les méthodes développées dans ce contexte multicapteur, il convient de distinguer :

- les méthodes qui exploitent la complémentarité des informations contenues au sein des images optiques et RSO. Ces méthodes réalisent une extraction partielle des bâtiments en imagerie optique puis complètent cette extraction en imagerie RSO ou inversement ;

- les méthodes qui exploitent la redondance des informations contenues au sein des images optiques et RSO. Ces méthodes s'appuient conjointement sur les informations contenues dans ces images pour réaliser l'extraction des bâtiments.

3.4.1 Méthodes exploitant la complémentarité d'information entre images optiques et RSO

Parmi les méthodes qui exploitent la complémentarité entre les images optiques et RSO, nous pouvons citer les travaux de Sportouche *et al.* [STD11] qui permettent de réaliser une extraction tridimensionnelle de bâtiments à partir d'une image optique et d'une image RSO. Les emprises au sol des bâtiments sont extraites à partir de l'image optique et les hauteurs des bâtiments sont estimées à partir de l'image RSO.

L'approche proposée repose sur une extraction des contours de toit des bâtiments en monoscopie optique, réalisée conformément à la méthode décrite à la section 3.2.2. La prise de vue de l'image optique étant quasi verticale, les emprises au sol des bâtiments sont supposées être superposées aux contours des toits extraits. Les emprises au sol détectées sont alors projetées de l'espace image, associé à l'image optique, vers l'espace terrain au moyen de la fonction de localisation directe de cette image (cf. section 2.3.1) et d'un MNT. Ces emprises sont ensuite projetées de l'espace terrain vers l'espace image, associé à l'image RSO, au moyen de la fonction de localisation inverse (cf. section 2.4.1) et de ce même MNT. En raison des imperfections du MNT, des incertitudes sur les paramètres physiques nécessaires à la mise en œuvre des fonctions de localisation et de l'hypothèse d'une prise de vue verticale en imagerie optique qui n'est pas toujours vérifiée, des erreurs de localisation sont observées lors de la projection des emprises au sol de l'image optique dans l'image RSO. Comme nous l'avons vu à la section 2.4.2, les segments de l'emprise au sol d'un bâtiment, qui correspondent aux arêtes de structures diédriques, se traduisent dans une image RSO par une forte radiométrie lorsqu'ils sont illuminés par le capteur RSO. Sportouche *et al.* mettent ainsi en œuvre une méthode de recalage fin, basée sur une maximisation de la radiométrie moyenne le long des segments de l'emprise au sol illuminés par le capteur RSO, afin de remédier aux erreurs de localisation. La hauteur des bâtiments est finalement estimée à partir de l'image RSO par maximisation de vraisemblance, comme décrit à la section 3.3.3.

3.4.2 Méthodes exploitant la redondance d'information entre images optiques et RSO

Dans cette section, nous présentons plusieurs méthodes exploitant la redondance d'information en imagerie optique et RSO pour l'extraction de bâtiments ou d'emprises au sol de bâtiments. Nous décrivons tout d'abord deux méthodes [TR04, PIS⁺11] qui réalisent l'extraction à partir d'une image optique et d'une image RSO avant de présenter deux autres méthodes [TR05, WHTS11] qui s'appuient sur une image optique et un couple d'images RSO multivues ou interférométriques.

Dans [TR04], Tupin *et al.* explorent la possibilité d'extraire des informations tridimensionnelles par mise en correspondance de primitives extraites dans une image optique et dans une image RSO. Les lignes brillantes et les points brillants sont tout d'abord extraits de l'image RSO au moyen des détecteurs de Tupin [TMM⁺98] et de Lopes [LTN90]. Ces primitives sont alors projetées dans l'image optique au moyen des fonctions de localisation des images pour un ensemble de hauteurs possibles. Les primitives extraites dans l'image RSO sont supposées correspondre à des discontinuités radiométriques dans l'image optique. Ainsi, pour chaque primitive, seule est retenue la hauteur qui maximise un score d'adéquation de cette primitive avec les images d'amplitude et d'orientation du gradient calculé dans l'image optique. Étant donné que la plupart des primitives extraites de l'image RSO correspondent aux arêtes de structures diédriques localisées au sol, la hauteur du sol, supposé plat, est estimée en calculant le mode de l'histogramme des hauteurs associées aux différentes primitives. Les primitives dont la hauteur est proche de la hauteur du sol peuvent être utilisées pour établir un MNT (au moyen d'une triangulation de Delaunay par exemple) mais ne contiennent pas d'information sur la hauteur

des objets du sur-sol et ne sont donc pas considérées par la suite. La hauteur des objets du sur-sol est finalement estimée en moyennant la hauteur de l'ensemble des primitives restantes situées à l'intérieur des emprises au sol des bâtiments, emprises supposées connues *a priori*.

Poulain *et al.* [PIS⁺11] proposent une méthode pour la création ou la mise à jour d'une base de données cartographiques, contenant les emprises au sol des bâtiments, à partir d'une image optique et d'une image RSO. La première étape de cette méthode consiste à générer des hypothèses d'emprises au sol de bâtiments en vectorisant les contours des régions issues d'une segmentation multiéchelle de l'image optique par l'algorithme *Mean-Shift* [CM99]. Dans le cas de la mise à jour d'une base de données, les emprises au sol, présentes dans la base de données existante, sont également ajoutées à l'ensemble des hypothèses issues de la segmentation de l'image optique. Lors de la seconde étape, des primitives géométriques et radiométriques, caractéristiques de la présence des bâtiments, sont extraites des images optiques et RSO, primitives à partir desquelles sont calculés des scores permettant de quantifier l'adéquation de chacune des hypothèses d'emprises au sol avec les images optique et RSO. Un de ces scores consiste, par exemple, à calculer le pourcentage de pixels de végétation contenus à l'intérieur d'une hypothèse d'emprise au sol, les pixels de végétation étant identifiés dans l'image optique au moyen de l'indice NDVI⁹. Un autre de ces scores repose sur le rapport des radiométries moyennes de l'image RSO calculées dans un voisinage des segments illuminés par le capteur (zone de repliement supposée) et dans un voisinage des segments non illuminés par le capteur (zone d'ombre supposée). Dans une troisième étape, les scores obtenus pour chacune des hypothèses d'emprises au sol à partir des différentes primitives extraites sont fusionnés dans le cadre de la théorie de l'évidence de Dempster-Shafer [Dem67, Sha76], de manière à obtenir un unique score quantifiant l'adéquation d'une hypothèse avec l'ensemble des primitives. La décision consistant à accepter ou à rejeter une hypothèse est finalement prise en comparant ce score à un seuil dont dépend les performances du détecteur de bâtiments.

Dans [WHTS11], Wegner *et al.* présentent une méthode pour la détection de bâtiments et la reconstruction des emprises au sol associées à partir d'une image optique multispectrale orthorectifiée et d'un couple d'image RSO en conditions interférométriques. Cette méthode repose sur une classification en deux classes (« bâtiment » et « non-bâtiment ») de vecteurs regroupant des caractéristiques radiométriques et géométriques extraites à partir de l'image optique et du couple d'images RSO. La classification est réalisée au moyen de champs aléatoires conditionnels [LMP01] nécessitant une phase d'apprentissage. Le vecteur de caractéristiques est constitué des valeurs moyennes des composantes rouge, vert, teinte et saturation¹⁰ de l'image optique, calculées dans un voisinage de chaque pixel de cette image, ainsi que de la variance de la composante teinte calculée dans ce même voisinage. Les descripteurs de Haralick [HSD73] et d'autres caractéristiques basées sur l'histogramme des orientations du vecteur gradient, calculés dans ce voisinage, sont également ajoutés à ce vecteur de caractéristiques. En ce qui concerne les caractéristiques extraites à partir du couple d'images RSO, les lignes brillantes sont tout d'abord extraites au moyen du détecteur de Tupin *et al.* [TMM⁺98]. Après une étape de régularisation visant à supprimer les lignes de faible longueur et à rattacher les lignes disjointes orientées selon la même direction, seules sont conservées les lignes pour lesquelles la valeur moyenne de la phase interférométrique calculée le long de cette ligne correspond à la hauteur du terrain. Ceci permet de conserver uniquement les lignes brillantes liées à la présence d'une structure diédrique. À chaque pixel de l'image optique est finalement associée une valeur traduisant la distance du pixel à la ligne brillante la plus proche extraite des images RSO et projetée en géométrie terrain en utilisant la hauteur déduite de la phase interférométrique. Wegner *et al.* montrent notamment que l'ajout de cette valeur au vecteur de caractéristiques du pixel, vecteur à partir duquel est réalisée la classification, conduit à une amélioration des performances en détection de bâtiments.

Enfin, Tupin *et al.* [TR05] ont développé une méthode pour la reconstruction tridimensionnelle de bâtiments à toit plat en milieu semi-urbain à partir d'une image optique et d'un couple d'images RSO multivues. La première étape de l'approche proposée repose sur une extraction des primitives linéiques et ponctuelles dans les deux images RSO, comme cela est fait dans [TR04]. Les primitives linéiques extraites sont alors mises en correspondance par corrélation, après que les images aient

⁹ *Normalized Difference Vegetation Index*

¹⁰ Composantes de l'espace colorimétrique TSV¹¹, plus couramment désigné par l'acronyme anglo-saxon HSV.

été rééchantillonnées en géométrie épipolaire. En supposant que la plupart des primitives linéiques correspondent à des structures au sol, la hauteur du sol, supposé plat, est estimée en calculant le mode de l'histogramme des hauteurs issues de la mesure des disparités entre les primitives mises en correspondance. Seules sont alors conservées les paires de primitives pour lesquelles la hauteur associée est supérieure à la hauteur du sol. Lors d'une seconde étape, ces primitives sont projetées d'une des deux images RSO vers l'image optique au moyen des fonctions de localisation. La troisième étape de la méthode proposée consiste à sur-segmenter l'image optique au moyen de l'algorithme de Suk *et al.* [SC83] et à construire un graphe à partir des régions issues de cette sur-segmentation. À chacune des ces régions est associé un nœud au sein du graphe, ces nœuds étant reliés en fonction des relations d'adjacence entre les différentes régions. À chaque nœud du graphe est également associé l'ensemble des couples de primitives RSO dont la projection dans l'image optique appartient à la région représentée par ce nœud. Dans la dernière étape de cette méthode, la hauteur associée à chacune des régions est finalement déterminée au moyen de champs de Markov mis en œuvre sur le graphe précédemment construit. Ceci permet de prendre en compte l'information de hauteur fournie par le couple d'images RSO ainsi que les relations d'adjacence entre les régions.

3.5 Conclusion

Un état de l'art relatif à l'extraction des bâtiments en imagerie de télédétection optique et RSO a été établi dans ce chapitre.

Nous avons tout d'abord identifié les primitives ponctuelles, linéiques et surfaciques caractéristiques de la présence des bâtiments dans les images optiques et RSO, et avons brièvement décrit les différentes méthodes au moyen desquelles elles sont extraites.

L'extraction de primitives est une problématique étroitement liée à celle de l'extraction de bâtiments. La plupart des techniques établies pour l'extraction de bâtiments s'appuient en effet sur ces primitives pour générer et/ou valider les hypothèses bidimensionnelles ou tridimensionnelles de bâtiments. À titre d'exemple, nous pouvons citer les méthodes développées par Lin *et al.* [LN98], en imagerie optique, et par Michaelsen *et al.* [MSSD08], en imagerie RSO, qui reposent sur une agrégation de primitives pour l'extraction des bâtiments.

Nous avons ensuite présenté diverses méthodes qui réalisent une extraction bidimensionnelle ou tridimensionnelle de bâtiments à partir d'une ou plusieurs images optiques (partie 3.2), d'une ou plusieurs images RSO (partie 3.3) ou encore à partir d'un ensemble d'images optiques et RSO considérées conjointement (partie 3.4). De cette étude, il ressort trois éléments principaux que nous détaillons ci-après.

Premièrement, il existe une grande diversité de méthodes permettant de réaliser l'extraction des bâtiments en imagerie de télédétection haute et très haute résolution. Comme nous l'avons mentionné en introduction (section 3.1.2), ces méthodes se distinguent par le modèle qu'elles adoptent pour décrire les bâtiments et par la stratégie mise en œuvre pour réaliser l'extraction.

Deuxièmement, il apparaît que les méthodes développées sont, le plus souvent, spécifiques au nombre et à la nature des images à partir desquelles est réalisée l'extraction. Ainsi, ces méthodes diffèrent selon la nature monoscopique ou multiscopique des données mais aussi selon le type de capteur (optique ou RSO) au moyen desquelles ces données ont été acquises.

Enfin, la comparaison quantitative des performances de ces méthodes s'avère délicate en raison de l'absence de jeux de tests référents.

Nous rappelons que l'objectif de ce travail doctoral, qui a été exposé dans l'introduction, consiste à établir des méthodes d'extraction qui répondent aux deux exigences suivantes :

- prise en compte de la complexité géométrique et radiométrique inhérente aux images de télédétection haute résolution en milieu urbain ;

- applicabilité des méthodes proposées aussi bien en imagerie optique que RSO, afin de faciliter leur mise en œuvre dans un contexte multicapteur optique-RSO.

Alors que la plupart des méthodes présentées dans ce chapitre répondent à la première exigence, seules quelques unes [TR04, TR05, STD11, PIS⁺11, WHTS11], qui ont été décrites dans la partie 3.4, répondent à la seconde. L'extraction de bâtiments dans un contexte multicapteur reste donc un problème ouvert auquel nous proposons deux solutions qui sont décrites aux chapitres 4 et 5. Ces deux méthodes, qui s'appuient sur une modélisation parallélépipédique des bâtiments (cf. partie 2.2), permettent de réaliser une extraction tridimensionnelle de ces derniers à partir d'une unique image pouvant indifféremment être une image optique ou une image RSO.

Approches radiométriques pour l'extraction de bâtiments

Sommaire

4.1	Introduction	76
4.2	Adéquation radiométrique des régions de la signature géométrique d'un bâtiment avec une image	77
4.2.1	Travaux antérieurs	77
4.2.2	Expression du critère pour une distribution statistique quelconque	78
4.2.3	Mise en œuvre du critère en imagerie optique	79
4.2.4	Mise en œuvre du critère en imagerie RSO	83
4.3	Adéquation radiométrique des segments de la signature géométrique d'un bâtiment avec une image	88
4.3.1	Mise en œuvre du critère en imagerie optique	88
4.3.2	Mise en œuvre du critère en imagerie RSO	91
4.4	Estimation des paramètres d'un bâtiment par mesure de l'adéquation radiométrique de sa signature géométrique avec une image	95
4.4.1	Estimation des paramètres en imagerie optique	96
4.4.2	Estimation des paramètres en imagerie RSO	101
4.5	Conclusion	103

Nous présentons, dans ce chapitre, une méthode permettant de réaliser l'extraction tridimensionnelle de bâtiments parallélépipédiques en monoscopie optique et RSO. L'approche proposée met en œuvre les méthodes développées au chapitre 2 pour simuler la signature géométrique des bâtiments dans les images optiques et RSO. L'adéquation de ces signatures géométriques avec l'image au sein de laquelle elles ont été simulées est alors quantifiée au moyen de deux critères radiométriques, dont la maximisation permet d'extraire les bâtiments. La signature géométrique d'un bâtiment dans une image peut être vue comme une partition locale de cette image, de sorte que la méthode proposée correspond en fait à une segmentation de l'image, contrainte par la géométrie de prise de vue et le modèle de bâtiment adopté.

4.1 Introduction

Dans ce chapitre, nous décrivons une méthode pour l'extraction tridimensionnelle de bâtiments parallélépipédiques en monoscopie optique et RSO.

Cette méthode repose sur les méthodes développées au chapitre 2 pour simuler la signature géométrique des bâtiments parallélépipédiques dans les images optiques et RSO. Ceci permet de prendre en compte, conformément aux objectifs fixés en introduction, la complexité radiométrique et géométrique qui caractérise les images de télédétection en milieu urbain.

La méthode proposée s'appuie également sur deux critères radiométriques qui quantifient l'adéquation de la signature géométrique d'un bâtiment avec l'image, optique ou RSO, au sein de laquelle elle a été simulée :

- le premier critère, fondé sur une « approche région », quantifie l'adéquation de la radiométrie des différentes régions constitutives de la signature géométrique d'un bâtiment avec un modèle statistique de bruit donné ;
- le second critère, fondé sur une « approche bord », quantifie l'adéquation des différents segments constitutifs de la signature géométrique d'un bâtiment avec les bords présents dans l'image.

L'approche développée consiste alors à déterminer le bâtiment dont la signature géométrique présente une adéquation optimale avec l'image, au sens des critères précédents.

Dans les deux premières parties, nous définissons les deux critères sus mentionnés. Nous établissons tout d'abord les expressions analytiques de ces critères en imagerie optique ainsi qu'en imagerie RSO, en s'appuyant notamment sur une modélisation des phénomènes physiques qui expliquent la présence de bruit dans ces images. Nous analysons ensuite le comportement de ces critères sur un exemple, au sein d'images simulée et réelle, afin d'illustrer la capacité de ces critères à quantifier l'adéquation d'un bâtiment avec une image.

Dans la dernière partie, nous étudions quantitativement le comportement des deux critères et en déduisons une stratégie d'optimisation, qui repose sur une maximisation successive de ces critères. Nous illustrons finalement les performances de cette stratégie sur données simulées et sur données réelles.

Comme cela a été présenté dans [BLRD11], l'approche proposée peut être mise en œuvre pour réaliser une détection-estimation conjointe des bâtiments dans une image optique ou RSO. La stratégie d'optimisation des deux critères précédemment introduits est alors appliquée à partir d'un ensemble de positions réparties régulièrement sur toute l'image. À l'issue de l'optimisation de ces critères, un test, fondé sur les valeurs optimales des critères, est finalement appliqué afin de déterminer si un bâtiment est présent ou non.

Cette approche pour la détection-estimation conjointe des bâtiments fournit des résultats satisfaisants mais engendre une complexité calculatoire trop importante pour pouvoir être mise en œuvre dans un contexte opérationnel. Ainsi, nous supposons, dans ce chapitre, que la position du centre de l'emprise au sol d'un bâtiment à extraire est connue approximativement (avec une précision de 2 mètres), de sorte que nous nous limitons à l'estimation des paramètres de ce bâtiment. La

position du centre de l'emprise au sol peut être déterminée de manière automatique, au moyen de l'une des méthodes de détection présentée au chapitre 3, ou bien de manière supervisée, par saisie manuelle d'un opérateur.

Notons enfin que nous nous restreignons, dans ce chapitre, au cas des images panchromatiques en imagerie optique et au cas des images monovues d'amplitude en imagerie RSO. La méthode d'extraction proposée peut toutefois être facilement étendue à d'autres types d'images, pour peu que l'on puisse modéliser statistiquement le bruit de ces images.

4.2 Adéquation radiométrique des régions de la signature géométrique d'un bâtiment avec une image

Nous présentons, dans cette première partie, un critère permettant de mesurer l'adéquation des régions constitutives de la signature géométrique d'un bâtiment avec l'image au sein de laquelle cette signature a été simulée. Ce critère repose sur le calcul, pour un modèle statistique de bruit donné, de la vraisemblance généralisée des niveaux de gris des pixels contenus dans les différentes régions de la signature géométrique du bâtiment.

Après avoir rappelé les travaux antérieurs sur lesquels s'appuie le critère proposé, nous donnons son expression générale pour un modèle statistique de bruit quelconque. Le critère est ensuite successivement mis en œuvre en imagerie optique puis en imagerie RSO. Nous donnons notamment l'expression de ce critère pour des modèles de bruit spécifiques aux images optiques et RSO puis étudions le comportement de ce dernier sur données simulées ainsi que sur données réelles.

4.2.1 Travaux antérieurs

Le critère présenté dans cette partie repose sur les travaux de Chesnaud *et al.* [CRB99] qui ont introduit la méthode des contours actifs statistiques pour la segmentation d'images. La méthode des contours actifs permet de segmenter un objet en faisant évoluer un contour initial de manière à le faire coïncider avec les contours de l'objet, l'évolution de ce contour étant guidée par la minimisation d'une fonction d'énergie. Avec les contours actifs classiques, introduits par Kass *et al.* [KWT88], l'évolution du contour actif est réalisée en évaluant le gradient d'une fonction de potentiel le long de ce contour. Contrairement à cette « approche bord », les contours actifs statistiques consistent à considérer les deux régions délimitées par le contour actif. Cette « approche région », particulièrement adaptée aux images au sein desquelles l'extraction des contours est compliquée du fait d'un niveau de bruit élevé, repose sur une maximisation de la vraisemblance généralisée des niveaux de gris des pixels constitutifs des deux régions, pour un modèle de bruit donné.

Les contours actifs statistiques, qui ne considèrent que deux régions et ne permettent de segmenter qu'un unique objet, ont par la suite été étendus par Galland *et al.* [GBR03] à la segmentation d'images en un nombre quelconque de régions au moyen d'une grille active statistique. L'approche proposée repose sur la minimisation, par modification itérative de la grille, de la complexité stochastique de l'image, notion introduite par Rissanen [Ris78]. Dans le cas d'une grille polygonale et pour un modèle statistique de bruit donné, la complexité stochastique de l'image s'écrit comme la somme de trois termes :

$$\Delta = \Delta_C + \Delta_P + \Delta_S, \quad (4.1)$$

où Δ_C est le nombre de bits nécessaire pour coder la description de la grille polygonale, Δ_P est le nombre de bits nécessaires pour coder les paramètres des densités de probabilités décrivant la radiométrie de chacune des régions de cette grille et Δ_S est le nombre de bits nécessaires pour coder les niveaux de gris des régions de la grille au moyen des densités de probabilité précédentes. Le terme Δ_S , dont il est montré qu'il correspond à l'opposé de la log-vraisemblance des niveaux de gris de l'image, reflète l'homogénéité radiométrique des régions de la grille tandis que les termes Δ_C et Δ_P reflètent la complexité de la grille, permettant ainsi de limiter le nombre de régions et de nœuds au sein de la grille.

Comme nous l'avons décrit à la section 3.3.3, Le Moigne *et al.* [MTN07] ont développé une méthode, s'appuyant sur les travaux de Chesnaux *et al.* et de Galland *et al.*, pour la reconstruction tridimensionnelle de bâtiments à partir d'une unique image RSO. Cette méthode repose sur une segmentation de la zone de repliement des bâtiments au moyen des contours actifs statistiques en considérant conjointement les images d'amplitude et de phase.

Sportouche *et al.* [STD11] ont également proposé une méthode, qui s'appuie sur ces mêmes travaux, pour estimer la hauteur des bâtiments à partir d'une unique image RSO d'amplitude, l'emprise au sol des bâtiments étant supposée rectangulaire et connue *a priori*. L'approche proposée pour estimer la hauteur d'un bâtiment consiste à simuler les signatures géométriques de ce bâtiment pour différentes hauteurs possibles. La hauteur estimée est alors celle pour laquelle la log-vraisemblance des niveaux de gris des différentes régions de la signature associée est maximale.

Le critère que nous présentons dans cette section est similaire à celui mis en œuvre par Sportouche *et al.* [STD11] et correspond, au signe près, au terme Δ_S de la complexité stochastique, dont l'expression est donnée à l'équation (4.1).

Comme nous l'avons vu au chapitre 2, la signature géométrique d'un bâtiment parallélépipédique dans une image peut être décrite par un ensemble de régions. En imagerie RSO, selon qu'un bâtiment est vu comme un objet occultant ou partiellement occultant, le nombre de régions constitutives de la signature géométrique associée à ce bâtiment peut varier. De même, en imagerie optique, le nombre de façades visibles d'un bâtiment, et donc le nombre de régions constitutives de sa signature géométrique, varie selon la position relative de ce bâtiment avec le capteur. En pratique, pour une prise de vue donnée, le nombre de régions décrivant la signature géométrique d'un bâtiment varie tout au plus de un à deux en imagerie optique comme en imagerie RSO. Ainsi, nous considérons par la suite que la signature géométrique d'un bâtiment dans une image, qui peut être vue comme une grille de partitionnement locale de cette image, présente une complexité constante, c'est-à-dire que les termes Δ_C et Δ_P sont considérés comme constants. Le critère que nous présentons correspond donc, à un signe et un terme constant près, à la complexité stochastique locale de l'image.

À la partie 4.4, nous proposons d'étendre l'utilisation de ce critère à l'extraction tridimensionnelle complète des bâtiments, c'est-à-dire à la détermination de l'ensemble des paramètres du modèle parallélépipédique de bâtiment, en monoscopie optique ainsi qu'en monoscopie RSO.

4.2.2 Expression du critère pour une distribution statistique quelconque

En reprenant les notations introduites à la section 2.3.3, nous notons $\mathcal{S}_I(\vec{b})$ la signature géométrique résultant de la projection du bâtiment \vec{b} dans l'image I , qui peut indifféremment être une image optique ou une image RSO. Cette signature est décrite par un ensemble de N_r régions R_j définies dans l'espace image \mathcal{E}_I :

$$\mathcal{S}_I(\vec{b}) = \{R_j, j = 1 \dots N_r\}. \quad (4.2)$$

Nous supposons que les niveaux de gris des pixels de la région R_j peuvent être décrits par une variable aléatoire X_j , dont la loi p_j est paramétrée par le vecteur de paramètres $\vec{\theta}_j$ défini dans l'espace Θ . En considérant les niveaux de gris des pixels contenus dans la région R_j comme des réalisations indépendantes de la variable aléatoire X_j , la vraisemblance de ces niveaux de gris, notée L_j , s'écrit :

$$L_j(\vec{\theta}_j) = L\left(\{I(k) | k \in R_j\} | \vec{\theta}_j\right) = \prod_{k \in R_j} p_j(I(k)), \quad (4.3)$$

où $I(k)$ désigne le niveau de gris du pixel k de l'image I . Pour des raisons calculatoires et numériques, la log-vraisemblance des niveaux de gris, notée ℓ_j , est usuellement préférée à la vraisemblance L_j de ces niveaux de gris et s'écrit :

$$\ell_j(\vec{\theta}_j) = \ln \left[L_j(\vec{\theta}_j) \right] = \sum_{k \in R_j} \ln p_j(I(k)). \quad (4.4)$$

L'expression de la log-vraisemblance ℓ_j , donnée à l'équation (4.4), fait intervenir le vecteur de paramètres $\vec{\theta}_j$ qui n'est pas connu *a priori*. Une estimation de ce vecteur par maximisation de vraisemblance conduit au vecteur $\vec{\theta}_j^{MV}$:

$$\vec{\theta}_j^{MV} = \underset{\vec{\theta}_j \in \Theta}{\operatorname{argmax}} \ell_j(\vec{\theta}_j), \quad (4.5)$$

vecteur à l'aide duquel est calculée la log-vraisemblance généralisée $\ell_j(\vec{\theta}_j^{MV})$ des niveaux de gris des pixels contenus dans la région R_j . La log-vraisemblance généralisée permet de quantifier l'adéquation des niveaux de gris observés au sein de la région R_j avec la loi p_j et constitue ainsi une mesure de l'homogénéité radiométrique de la région R_j .

Le critère $\mathcal{C}_I^r(\vec{b})$, que nous étudions dans cette partie, s'écrit de la manière suivante :

$$\mathcal{C}_I^r(\vec{b}) = \ell_0(\vec{\theta}_0^{MV}) + \sum_{j=1}^{N_r} \ell_j(\vec{\theta}_j^{MV}), \quad (4.6)$$

où $\ell_j(\vec{\theta}_j^{MV})$ désigne la log-vraisemblance généralisée des pixels contenus dans la région R_j de la signature géométrique $\mathcal{S}_I(\vec{b})$ et $\ell_0(\vec{\theta}_0^{MV})$ correspond à la log-vraisemblance généralisée des pixels contenus dans une région R_0 englobant la signature géométrique.

En supposant qu'un bâtiment \vec{b}_0 est effectivement présent dans une région R_0 de l'image I , une estimation des paramètres de ce bâtiment peut être réalisée en maximisant le critère $\mathcal{C}_I^r(\vec{b})$, conduisant ainsi au bâtiment \vec{b}_0 :

$$\vec{b}_0 = \underset{\vec{b} \in \mathcal{E}_b}{\operatorname{argmax}} \mathcal{C}_I^r(\vec{b}), \quad (4.7)$$

où \mathcal{E}_b désigne, conformément aux notations introduites à la partie 2.2, l'espace au sein duquel est défini le vecteur de paramètres \vec{b} . Il est important de préciser que la région englobante R_0 est fixée lors de la recherche du vecteur de paramètres \vec{b} maximisant le critère $\mathcal{C}_I^r(\vec{b})$, de sorte que le calcul du critère pour différents vecteurs \vec{b} fait intervenir les mêmes pixels de l'image I .

4.2.3 Mise en œuvre du critère en imagerie optique

Dans cette section, nous présentons les phénomènes physiques à l'origine des fluctuations radiométriques observées en imagerie optique au sein d'une région de réflectance uniforme. Nous rappelons le modèle statistique couramment utilisé pour décrire ces fluctuations et donnons l'expression analytique du critère statistique précédemment introduit pour ce modèle de bruit.

4.2.3.1 Modélisation statistique du bruit en imagerie optique

Les images acquises par les capteurs optiques sont altérées par trois bruits distincts : le bruit de photon, le bruit thermique et le bruit intrinsèque, qui engendrent les fluctuations radiométriques observées au sein des régions de réflectance uniforme.

Le bruit de photons, encore appelé bruit quantique ou bruit de grenaille, s'explique par une modélisation corpusculaire de la lumière selon laquelle une onde lumineuse est constituée de particules élémentaires, appelées photons, dont l'énergie E vaut :

$$E = h\nu, \quad (4.8)$$

où h et ν correspondent respectivement à la constante de Planck¹ et à la fréquence de l'onde lumineuse supposée monochromatique.

¹ $h = 6,62.10^{-34} \text{ J.s}$

Dans le cadre de cette modélisation corpusculaire de la lumière, le nombre n de photons arrivant sur un détecteur pendant une durée T est décrit par une variable aléatoire discrète distribuée suivant une loi de Poisson, de paramètre λ_n :

$$p_n(n|\lambda_n) = \exp(-\lambda_n) \frac{\lambda_n^n}{n!}. \quad (4.9)$$

L'espérance et la variance d'une variable aléatoire distribuée selon une loi de Poisson sont égales au paramètre λ de cette distribution ($E[n] = \lambda_n$ et $\text{var}[n] = \lambda_n$), de sorte que λ_n peut être relié à la puissance P de l'onde lumineuse reçue par le détecteur, la durée d'intégration τ de ce détecteur et l'énergie $h\nu$ d'un photon par la relation suivante :

$$\lambda_n = E[n] = \frac{P\tau}{h\nu}. \quad (4.10)$$

Dans un capteur électro-optique, le flux de photons reçus par un photodétecteur est converti en courant électrique i_q dont l'expression est donnée par la relation suivante :

$$i_q = \frac{me}{\tau}, \quad (4.11)$$

où m correspond au nombre de photons détectés par le détecteur et convertis en électrons pendant la durée τ et e est la charge électrique élémentaire. Les fluctuations de m sont directement reliées aux fluctuations de n par l'intermédiaire de l'efficacité quantique du détecteur η , définie comme le rapport du nombre moyen de photons détectés sur le nombre moyen de photons reçus par le détecteur. Les fluctuations de m sont donc décrites, comme celles de n , par une distribution de Poisson de paramètre $\lambda_m = \eta\lambda_n$, distribution qui peut être approximée, à flux de photons élevé, par une distribution normale de valeur moyenne et de variance égales à λ_m . Finalement, les fluctuations du courant électrique i_q , mesuré en sortie du détecteur, sont décrites par une distribution normale de paramètres μ_{i_q} et d'écart type σ_{i_q} :

$$p_{i_q}(i_q|\mu_{i_q}, \sigma_{i_q}) = \frac{1}{\sqrt{2\pi}\sigma_{i_q}} \exp\left[-\frac{(i_q - \mu_{i_q})^2}{2\sigma_{i_q}^2}\right], \quad (4.12)$$

avec :

$$\mu_{i_q} = \frac{e}{\tau} E[m] = \frac{e}{\tau} \lambda_m = \frac{\eta e P}{h\nu}, \quad (4.13)$$

$$\sigma_{i_q}^2 = \left(\frac{e}{\tau}\right)^2 \text{var}[m] = \left(\frac{e}{\tau}\right)^2 \lambda_m = \frac{e}{\tau} \mu_{i_q}. \quad (4.14)$$

Le rapport de la valeur moyenne μ_{i_q} du courant électrique observé en sortie du détecteur et de la puissance lumineuse P reçue par ce détecteur définit alors la sensibilité S de ce détecteur :

$$S = \frac{\mu_{i_q}}{P} = \frac{\eta e}{h\nu}. \quad (4.15)$$

En ce qui concerne le bruit thermique, il résulte de l'agitation thermique des électrons à l'intérieur du détecteur photoélectrique, dont la résistance équivalente est notée R . Les fluctuations du courant i_t résultant de cette agitation thermique sont modélisées par une distribution normale dont la moyenne μ_{i_t} est nulle et dont la variance $\sigma_{i_t}^2$ est proportionnelle à la température absolue T :

$$\mu_{i_t} = 0, \quad (4.16)$$

$$\sigma_{i_t}^2 = \frac{1}{2\tau} \frac{4kT}{R}, \quad (4.17)$$

où k désigne la constante de Boltzmann².

² $k = 1,38.10^{-23} \text{ J/K}$

Enfin, le bruit intrinsèque est lié à l'observation, en l'absence de flux optique, d'un courant électrique i_o en sortie du détecteur, courant qui est qualifié de courant d'obscurité et dont les fluctuations sont également décrites par une distribution normale. La valeur moyenne μ_{i_o} et la variance $\sigma_{i_o}^2$ de ces fluctuations sont des caractéristiques propres à chaque détecteur et sont données par les équations suivantes :

$$\mu_{i_o} \neq 0, \quad (4.18)$$

$$\sigma_{i_o}^2 = \frac{1}{2\tau} S^2 \text{NEP}^2, \quad (4.19)$$

où S correspond à la sensibilité du détecteur, définie à l'équation (4.15), et NEP désigne le *Noise Equivalent Power*, grandeur caractéristique du détecteur exprimée en $W.Hz^{-1/2}$.

Le courant i observé en sortie d'un détecteur photoélectrique résulte de la somme des courants i_q , i_t et i_o , courants dont les fluctuations sont indépendantes et décrites par des distributions normales. Ainsi, les fluctuations du courant i , auxquelles sont directement reliées les fluctuations de la radiométrie des images optiques, sont décrites par une distribution normale dont la valeur moyenne μ_i et la variance σ_i^2 sont données par les relations suivantes :

$$\mu_i = \mu_{i_q} + \mu_{i_t} + \mu_{i_o} = SP + \mu_{i_o}, \quad (4.20)$$

$$\sigma_i^2 = \sigma_{i_q}^2 + \sigma_{i_t}^2 + \sigma_{i_o}^2 = \frac{1}{2\tau} \left(2eSP + \frac{4kT}{R} + S^2 \text{NEP}^2 \right). \quad (4.21)$$

Finalement, les niveaux de gris des pixels d'une région de réflectance uniforme sont décrits par une variable aléatoire I reliée au courant i par un facteur de gain G , de sorte que I suit une loi normale, de paramètres μ et σ :

$$p_I(I|\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left[-\frac{(I - \mu)^2}{2\sigma^2} \right], \quad (4.22)$$

où les paramètres μ et σ peuvent être reliés aux grandeurs physiques caractéristiques du capteur et de l'acquisition par les relations suivantes :

$$\mu = G(SP + \mu_{i_o}), \quad (4.23)$$

$$\sigma^2 = \frac{G^2}{2\tau} \left(2eSP + \frac{4kT}{R} + S^2 \text{NEP}^2 \right). \quad (4.24)$$

4.2.3.2 Expression du critère pour une distribution normale

En utilisant l'expression de la loi normale qui est supposée décrire les niveaux de gris des pixels de la région R_j , la log-vraisemblance de ces niveaux de gris, dont l'expression est donnée à l'équation (4.4), se réécrit :

$$\ell_j(\mu_j, \sigma_j) = \frac{1}{2\sigma_j^2} \sum_{k \in R_j} (I(k) - \mu_j)^2 - \frac{N_{R_j}}{2} \ln(2\pi) - N_{R_j} \ln \sigma_j. \quad (4.25)$$

L'estimation par maximum de vraisemblance des paramètres μ_j et σ_j conduit aux expressions suivantes :

$$\mu_j^{MV} = \frac{1}{N_{R_j}} \sum_{k \in R_j} I(k), \quad (4.26)$$

$$\sigma_j^{MV} = \sqrt{\frac{1}{N_{R_j}} \sum_{k \in R_j} (I(k) - \mu_j^{MV})^2}, \quad (4.27)$$

à partir desquelles est déterminée l'expression de la log-vraisemblance généralisée des niveaux de gris :

$$\ell_j(\mu_j^{MV}, \sigma_j^{MV}) = \frac{1}{2\sigma_j^{MV2}} \sum_{k \in R_j} (I(k) - \mu_j^{MV})^2 - \frac{N_{R_j}}{2} \ln(2\pi) - N_{R_j} \ln \sigma_j^{MV} \quad (4.28)$$

$$\ell_j(\mu_j^{MV}, \sigma_j^{MV}) = \frac{N_{R_j}}{2} (1 - \ln(2\pi)) - N_{R_j} \ln \sigma_j^{MV}. \quad (4.29)$$

Paramètres	Régions				
	R_T	R_F^1	R_F^2	R_O	R_0
μ_j	300	130	150	90	350
σ_j	30	20	20	20	50

Tableau 4.1 – Paramètres des lois normales mises en œuvre pour la simulation de l'image optique de la figure 4.1 (b).

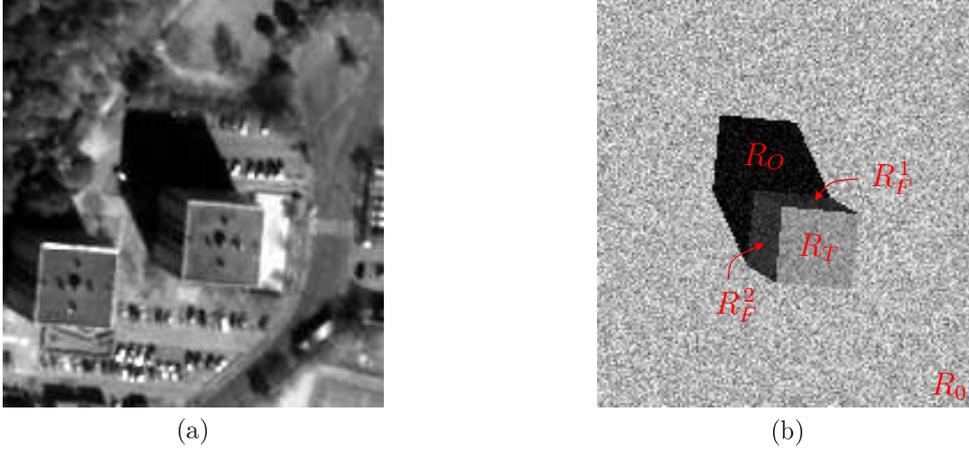


Figure 4.1 – Images optiques réelle (a) et simulée (b) à partir desquelles sont calculées les coupes du critère \mathcal{C}_I^r .

Le critère statistique $\mathcal{C}_I^r(\vec{b})$, caractérisant l'adéquation d'un bâtiment \vec{b} avec l'image optique I au sein de laquelle il a été projeté, s'exprime alors comme la somme des log-vraisemblances généralisées calculées pour chacune des régions de la signature géométrique $\mathcal{S}_I(\vec{b})$ ainsi que pour la région englobante R_0 .

4.2.3.3 Étude du critère sur données simulées et sur données réelles

Nous considérons, dans cette sous-section, les deux images de la figure 4.1, à partir desquelles des coupes du critère \mathcal{C}_I^r , selon les sous-espaces de l'espace des paramètres \mathcal{E}_b , sont calculées de manière à illustrer le comportement de ce critère.

L'image de la figure 4.1 (a) est extraite de l'image QuickBird, présentée à la figure 2.9, et représente le bâtiment G , dont les paramètres, donnés au tableau 2.1, sont ici rappelés : $X_C = 703159.0$ m, $Y_C = 4795086.0$ m, $\alpha = 83.0$ degrés, $w_1 = 20.0$ m, $w_2 = 20.0$ m et $h = 22.0$ m.

L'image de la figure 4.1 (b) représente ce même bâtiment G au sein d'une image simulée, qui a été obtenue en simulant la signature géométrique de ce bâtiment pour une géométrie de prise de vue identique à celle de l'image QuickBird. Les niveaux de gris des pixels contenus dans chacune des régions constitutives de la signature géométrique ainsi que dans la région englobante résultent de tirages aléatoires selon des lois normales, conformément au modèle statistique de bruit adopté précédemment pour décrire la radiométrie des images optiques. Les paramètres de ces lois, estimés à partir de l'image réelle de la figure 4.1 (a), sont donnés au tableau 4.1.

La figure 4.2 présente les coupes du critère \mathcal{C}_I^r selon les espaces \mathcal{E}_{X_C} , \mathcal{E}_{Y_C} , \mathcal{E}_α , \mathcal{E}_{w_1} , \mathcal{E}_{w_2} et \mathcal{E}_h . Ces coupes sont calculées à partir des images réelle et simulée de la figure 4.1 en fixant les paramètres du bâtiment à leurs valeurs vraies à l'exception d'un des paramètres pour lequel nous testons l'ensemble des valeurs admissibles que peut prendre ce paramètre.

Nous constatons que les coupes obtenues à partir de l'image simulée (courbes rouges) présentent toutes un maximum au niveau de la valeur vraie du paramètre associé à la coupe considérée, validant ainsi la capacité du critère \mathcal{C}_I^r à quantifier l'adéquation entre la signature géométrique d'un bâtiment et l'image au sein de laquelle elle a été simulée.

En ce qui concerne les courbes obtenues à partir de l'image réelle (courbes bleues), la position du maximum diffère légèrement de la valeur vraie des paramètres associés. Nous observons ainsi une erreur de 1 à 2 mètres pour les paramètres X_C , Y_C , w_1 , w_2 , h et de 5 degrés pour le paramètre α . Ces erreurs sont comparables aux incertitudes, présentées au tableau 2.2, qui résultent d'une extraction supervisée des bâtiments par une saisie manuelle des sommets visibles de ces bâtiments. Deux raisons peuvent expliquer les écarts observés :

- l'inhomogénéité radiométrique des régions constitutives de la signature géométrique. Cette inhomogénéité, qui résulte de la présence de superstructures sur le toit, de fenêtres sur les façades ou encore de végétation à l'intérieur de l'ombre, met en défaut l'hypothèse selon laquelle la variabilité des niveaux de gris des pixels d'une région est uniquement liée au bruit ;
- l'inadéquation entre le bruit présent dans l'image et le modèle de bruit adopté. Nous avons supposé que le bruit, responsable des variations des niveaux de gris des pixels d'une région de réflectance uniforme, pouvait être modélisé par une loi normale. En pratique, le bruit observé n'est jamais parfaitement gaussien, ce qui peut expliquer, dans une moindre mesure, les écarts observés.

Si les coupes unidimensionnelles du critère \mathcal{C}_I^r présentées ici ne permettent pas de caractériser le comportement de ce critère dans l'espace \mathcal{E}_b , elles illustrent néanmoins sa régularité et laissent à penser qu'il présente un vaste bassin d'attraction, ce qui devrait faciliter l'optimisation de ce critère.

4.2.4 Mise en œuvre du critère en imagerie RSO

Comme en imagerie optique, nous donnons, dans cette section, une brève description des phénomènes physiques permettant d'expliquer le bruit de *speckle* caractéristique des images RSO. Nous rappelons ensuite le modèle statistique couramment utilisé pour décrire ce bruit dans les images RSO monovues complexes. Nous nous appuyons alors sur ce modèle statistique de bruit pour établir l'expression analytique du critère introduit à la section 4.2.2 dans le cas des images RSO. Finalement, le comportement du critère est étudié sur données simulées puis sur données réelles.

4.2.4.1 Modélisation statistique du bruit en imagerie RSO

Dans une image RSO, les pixels d'une région présentant des caractéristiques physiques de rétro-diffusion stationnaires, présentent une grande variabilité radiométrique liée au bruit de *speckle*. La variabilité radiométrique des régions homogènes s'explique par la nature cohérente des ondes RADAR dont la longueur d'onde est du même ordre que la taille caractéristique des réflecteurs élémentaires, ce qui conduit à un phénomène d'interférences entre les ondes rétrodiffusées par ces réflecteurs élémentaires. Ce phénomène d'interférences, qui est parfaitement déterministe au niveau microscopique, devient aléatoire au niveau macroscopique où la disposition des réflecteurs élémentaires est supposée aléatoire au sein d'une cellule de résolution de l'image RSO.

En supposant qu'une cellule de résolution contient un nombre fini N_d de réflecteurs (ou diffuseurs) élémentaires, l'amplitude complexe z , mesurée dans l'image RSO au niveau du pixel correspondant à cette cellule de résolution, résulte de la somme cohérente des amplitudes complexes z_n réfléchies par chacun des ces réflecteurs élémentaires :

$$z = A \exp(j\phi) = \sum_{n=1}^{N_d} z_n = \sum_{n=1}^{N_d} A_n \exp(j\phi_n), \quad (4.30)$$

où A , ϕ désignent respectivement l'amplitude et la phase de l'amplitude complexe z et A_n , ϕ_n correspondent respectivement à l'amplitude et la phase de l'amplitude complexe z_n réfléchi par le $n^{\text{ième}}$ réflecteur élémentaire. Comme illustré à la figure 4.3, l'amplitude z est alors obtenue en

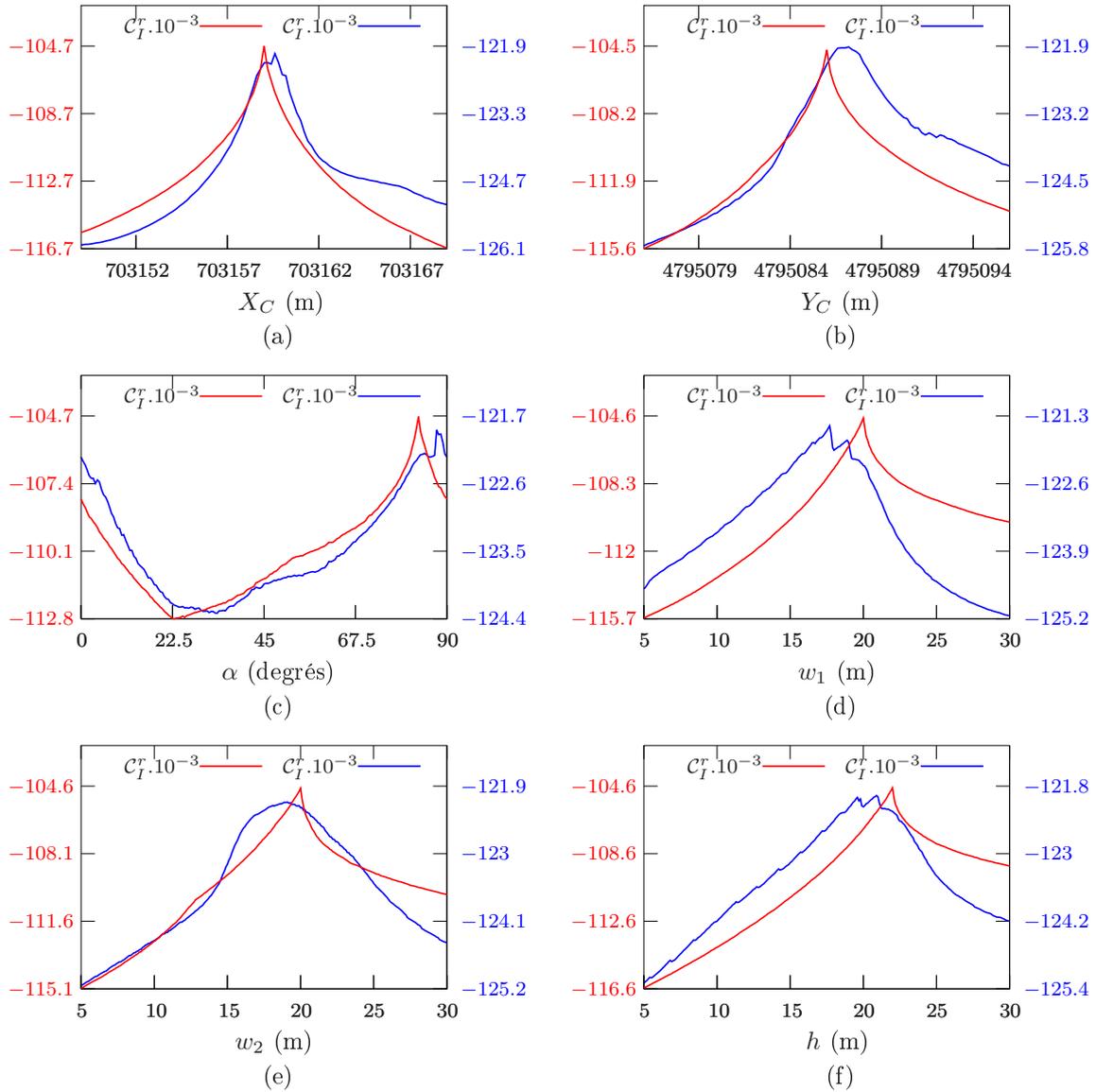


Figure 4.2 – Coupes du critère C_I^r calculées à partir de l'image optique réelle de la figure 4.1 (a) (courbes bleues) et de l'image optique simulée de la figure 4.1 (b) (courbes rouges).

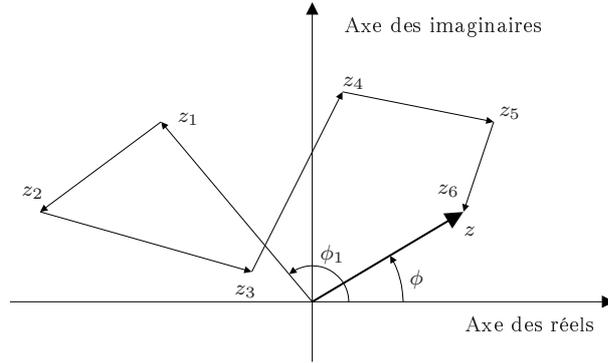


Figure 4.3 – Somme cohérente des amplitudes complexes réfléchies par six réflecteurs élémentaires (image extraite de [Mai01]).

sommant, dans le plan complexe, les N_d vecteurs associés aux réflecteurs élémentaires.

Par la suite, nous nous plaçons dans le cadre du modèle de *speckle* totalement développé, modèle qui repose, comme cela est décrit dans les travaux de Maître *et al.* [Mai01], sur cinq hypothèses :

1. Le nombre N_d de réflecteurs par cellule de rétrodiffusion est grand ;
2. L'amplitude complexe réfléchie par un réflecteur est indépendante de celles réfléchies par les autres réflecteurs, ce qui revient à négliger les interactions entre ces réflecteurs lors de l'interaction onde matière ;
3. Pour un réflecteur donné, l'amplitude et la phase de l'amplitude complexe réfléchie par ce réflecteur sont indépendantes, ce qui revient à supposer une structuration aléatoire du milieu diffusant ;
4. Les amplitudes A_n associées aux réflecteurs élémentaires sont modélisées au moyen de variables aléatoires identiquement distribuées dont les moments d'ordre 1 et 2 sont égaux ;
5. Les phases ϕ_n associées aux réflecteurs élémentaires sont uniformément distribuées dans $[-\pi; \pi]$, ce qui revient à considérer la surface par laquelle est réfléchie l'onde RADAR comme rugueuse par rapport à la longueur d'onde.

Dans le cadre du modèle de *speckle* totalement développé, l'amplitude et la phase associées à un réflecteur élémentaire sont considérées comme des variables aléatoires de sorte que l'amplitude complexe z caractérisant l'onde RADAR rétrodiffusée par une cellule de rétrodiffusion est également une variable aléatoire dont nous souhaitons déterminer la distribution. Pour ce faire, il suffit de déterminer les distributions de z_r et z_i qui désignent respectivement la partie réelle et la partie imaginaire de z et dont les expressions sont déduites de l'équation (4.30) :

$$z_r = \sum_{n=1}^{N_d} A_n \cos \phi_n, \quad (4.31)$$

$$z_i = \sum_{n=1}^{N_d} A_n \sin \phi_n. \quad (4.32)$$

D'après les hypothèses du modèle de *speckle* totalement développé, nous pouvons écrire :

$$E(z_r) = \sum_{n=1}^{N_d} E(A_n)E(\cos \phi_n) = 0, \quad (4.33)$$

$$E(z_i) = \sum_{n=1}^{N_d} E(A_n)E(\sin \phi_n) = 0, \quad (4.34)$$

$$E(z_r^2) = E(z_i^2) = \sum_{n=1}^{N_d} \frac{1}{2} E(A_n^2) = \frac{1}{2} R, \quad (4.35)$$

$$E(z_r z_i) = 0, \quad (4.36)$$

où $E(X)$ désigne l'espérance de la variable aléatoire X et R est défini comme le pouvoir réflecteur de la cellule de résolution, pouvoir réflecteur qui augmente logiquement avec le nombre de réflecteurs élémentaires N_d et avec la SER moyenne $E(A_n^2)$ des réflecteurs élémentaires.

Si le nombre de réflecteurs élémentaires par cellule de résolution est suffisamment grand, les variables aléatoires z_r et z_i , qui s'écrivent comme la somme de N_d variables aléatoires indépendantes et identiquement distribuées, sont distribuées, d'après le théorème de la limite centrale, selon une loi normale :

$$p_{z_r}(z_r|R) = \frac{1}{\sqrt{\pi R}} \exp\left(-\frac{z_r^2}{R}\right), \quad (4.37)$$

$$p_{z_i}(z_i|R) = \frac{1}{\sqrt{\pi R}} \exp\left(-\frac{z_i^2}{R}\right). \quad (4.38)$$

D'après les équations (4.36), (4.37) et (4.38), les variables aléatoires z_r et z_i sont décorrélatées et distribuées selon des lois normales, ce qui implique l'indépendance entre ces variables aléatoires.

En effectuant les changements de variables $A = \sqrt{z_r^2 + z_i^2}$ et $\phi = \arctan(z_r/z_i)$, nous pouvons finalement montrer que l'amplitude A et la phase ϕ de l'amplitude complexe z suivent respectivement une loi de Rayleigh, de paramètre $\sigma = \sqrt{R/2}$, et une loi uniforme sur l'intervalle $[-\pi; \pi]$:

$$p_A(A|\sigma) = \frac{A}{\sigma^2} \exp\left(-\frac{A^2}{2\sigma^2}\right) 1_{\mathbb{R}^+}(A), \quad (4.39)$$

$$p_\phi(\phi) = \frac{1}{2\pi} 1_{[-\pi; \pi]}(\phi), \quad (4.40)$$

où $1_K(\cdot)$ désigne la fonction indicatrice sur l'intervalle K .

L'espérance $E(A)$ et la variance $\text{var}(A)$ de l'amplitude A sont alors données par les relations suivantes :

$$E(A) = \sqrt{\frac{\pi}{2}} \sigma, \quad (4.41)$$

$$\text{var}(A) = \frac{(4 - \pi)}{2} \sigma^2 = \frac{(4 - \pi)}{4} R, \quad (4.42)$$

relations à partir desquelles nous constatons que la variance de A est proportionnelle au pouvoir réflecteur R de la cellule de résolution, ce qui justifie que le bruit de *speckle* soit qualifié de bruit multiplicatif.

4.2.4.2 Expression du critère pour une distribution de Rayleigh

En utilisant l'expression analytique de la distribution de Rayleigh, donnée à l'équation (4.39), la log-vraisemblance des niveaux de gris des pixels contenus dans la région R_j , dont l'expression est donnée à l'équation (4.4), se réécrit comme suit :

$$\ell_j(\sigma_j) = \sum_{k \in R_j} \ln(I(k)) - \frac{1}{2\sigma_j^2} \sum_{k \in R_j} I(k)^2 - 2N_{R_j} \ln(\sigma_j), \quad (4.43)$$

Régions Paramètre	R_{TFS}	R_{FS}^i	R_{ED}	R_O	R_0
σ_j	150	90	400	40	70

Tableau 4.2 – Paramètres des lois de Rayleigh mises en œuvre pour la simulation de l'image RSO de la figure 4.4 (b).

où N_{R_j} désigne le nombre de pixels contenus dans la région R_j .

Il est facile de montrer que l'estimation du paramètre σ_j par l'estimateur du maximum de vraisemblance conduit à l'expression suivante :

$$\sigma_j^{MV} = \sqrt{\frac{1}{2N_{R_j}} \sum_{k \in R_j} I(k)^2}. \quad (4.44)$$

La log-vraisemblance généralisée des niveaux de gris des pixels de la région R_j s'écrit alors :

$$\ell_j(\sigma_j^{MV}) = \sum_{k \in R_j} \ln(I(k)) - N_{R_j} \ln \left[\sum_{k \in R_j} I(k)^2 \right] + N_{R_j} (\ln 2 + \ln N_{R_j} - 1). \quad (4.45)$$

Le critère statistique s'écrit alors, conformément à l'équation (4.6), comme la somme des log-vraisemblances généralisées calculées sur les différentes régions R_j ainsi que sur la région englobante R_0 .

4.2.4.3 Étude du critère sur données simulées et sur données réelles

Dans cette sous-section, nous caractérisons le comportement du critère \mathcal{C}_f^r en imagerie RSO sur les images simulée et réelle de la figure 4.4.

L'image de la figure 4.4 (a), qui est extraite de l'image TerraSAR-X de la figure 2.17, représente le même bâtiment que celui considéré pour l'étude du critère \mathcal{C}_f^r en imagerie optique. Nous rappelons que les paramètres de ce bâtiment valent : $X_C = 703159.0$ m, $Y_C = 4795086.0$ m, $\alpha = 83.0$ degrés, $w_1 = 20.0$ m, $w_2 = 20.0$ m et $h = 22.0$ m.

L'image de la figure 4.4 (b) représente également le bâtiment G et a été obtenue par simulation de la signature géométrique de ce bâtiment en considérant la même géométrie de prise de vue que l'image TerraSAR-X. Conformément au modèle de bruit adopté pour décrire la radiométrie des images RSO monovues d'amplitude, nous effectuons des tirages aléatoires selon des lois de Rayleigh, dont les paramètres sont spécifiés au tableau 4.2, afin de simuler les niveaux de gris des pixels appartenant aux différentes régions de la signature ainsi qu'à la région englobante.

Nous présentons à la figure 4.5 les coupes du critère \mathcal{C}_f^r selon les espaces \mathcal{E}_{X_C} , \mathcal{E}_{Y_C} , \mathcal{E}_α , \mathcal{E}_{w_1} , \mathcal{E}_{w_2} et \mathcal{E}_h , coupes qui ont été calculées à partir des images de la figure 4.4.

Les courbes obtenues à partir de l'image simulée (courbes rouges) présentent toutes un maximum global au niveau de la valeur vraie du paramètre associé à la coupe considérée. En revanche, les courbes obtenues à partir de l'image réelle (courbes bleues) présentent, pour certaines (figures 4.5 (a), 4.5 (e) et 4.5 (f)), un maximum global qui ne correspond pas à la valeur vraie du paramètre associé à la coupe considérée, cette valeur vraie se traduisant alors par un maximum local plus ou moins marqué. Ceci s'explique par la présence de structures proches du bâtiment G , et notamment par la présence du bâtiment F , qui affecte l'aptitude du critère \mathcal{C}_f^r à quantifier l'adéquation de la signature géométrique simulée d'un bâtiment avec l'image.

Par rapport aux courbes observées en imagerie optique, les courbes obtenues ici sont nettement moins régulières (présence de nombreux *maxima* locaux) et présentent, pour certaines, un maximum global dont la position est erronée. L'optimisation de ce critère s'avère donc plus délicate, ce qui se

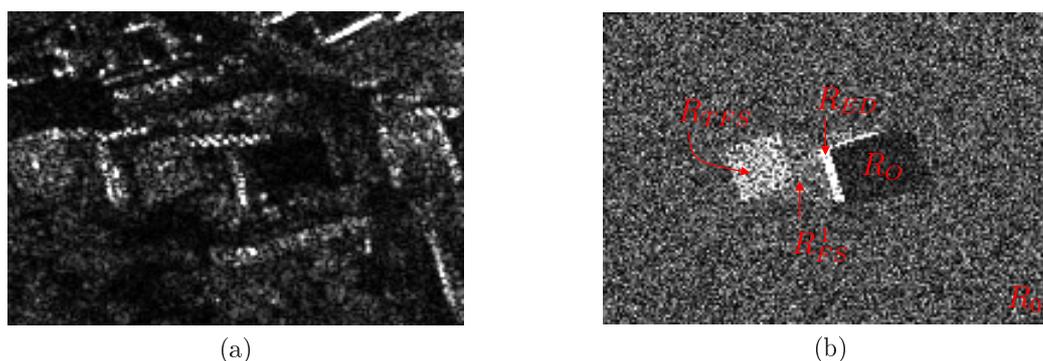


Figure 4.4 – Images RSO réelle (a) et simulée (b) à partir desquelles sont calculées les coupes du critère C_I^r .

traduit par une estimation moins robuste des paramètres des bâtiments, comme nous le constatons à la section 4.4.2.

4.3 Adéquation radiométrique des segments de la signature géométrique d'un bâtiment avec une image

Dans cette partie, nous proposons un second critère visant à quantifier l'adéquation de la signature géométrique d'un bâtiment avec une image optique ou RSO. Contrairement au critère présenté à la partie précédente qui repose sur une « approche région », le critère introduit ici repose sur une « approche bord » quantifiant l'adéquation des segments constitutifs de la signature géométrique avec les bords présents dans l'image. Nous donnons successivement les expressions de ce critère dans le cas de l'imagerie optique puis dans le cas de l'imagerie RSO, expressions qui sont adaptées à la nature additive ou multiplicative du bruit présent dans ces images.

4.3.1 Mise en œuvre du critère en imagerie optique

4.3.1.1 Expression du critère pour un bruit additif

L'état de l'art dressé par Ziou *et al.* [ZT98] sur la détection de contours en imagerie optique conduit à l'identification de deux types de méthode :

- les méthodes qui reposent sur le calcul de la dérivée première de l'image et donc sur le calcul de différences radiométriques locales, calcul réalisé au moyen de l'opérateur gradient appliqué le plus souvent après un prétraitement visant à débruiter l'image. Ces méthodes, parmi lesquelles figure le filtre de Canny [Can86], consistent alors à chercher au sein de l'image les pixels pour lesquels la norme du vecteur gradient est élevée, indiquant ainsi la présence d'un bord ;
- les méthodes qui reposent sur le calcul de la dérivée seconde de l'image, calcul réalisé au moyen de l'opérateur Laplacien également appliqué après un débruitage de l'image. Les bords de l'image sont alors détectés, comme cela est fait par le détecteur de Marr-Hildreth [MH80], en recherchant les pixels pour lesquels cette dérivée seconde s'annule.

En s'inspirant des détecteurs de contours, fondés sur le calcul de la dérivée première de l'image, nous quantifions, dans cette sous-section, l'adéquation d'un segment de la signature géométrique d'un bâtiment avec l'image en calculant la différence des radiométries moyennes calculées dans deux régions situées de part et d'autre de ce segment et s'appuyant sur ce segment.

Nous reprenons ici les notations introduites à la section 2.3.3 et décrivons ainsi la signature géométrique $S_I(\vec{b})$ du bâtiment \vec{b} dans l'image I comme un ensemble de N_s segments S_k définis dans

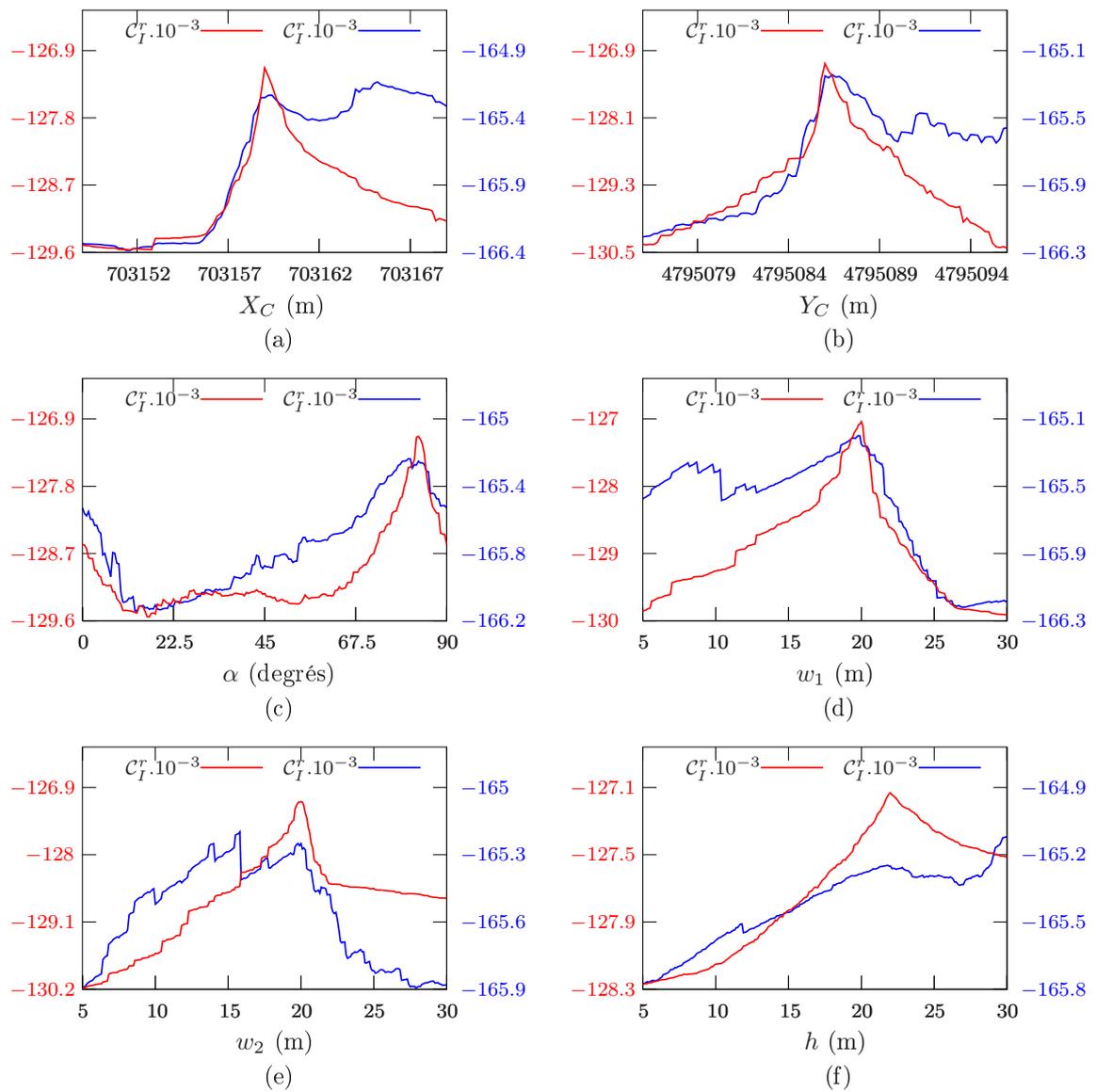


Figure 4.5 – Coupes du critère C_I^r calculées à partir de l'image RSO réelle de la figure 4.4 (a) (courbes bleues) et de l'image RSO simulée de la figure 4.4 (b) (courbes rouges).

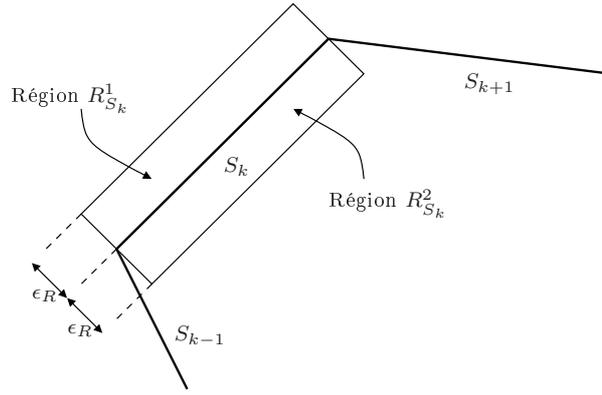


Figure 4.6 – Définitions des régions $R_{S_k}^1$ et $R_{S_k}^2$ situées de part et d'autre du segment S_k .

l'espace image \mathcal{E}_I :

$$\mathcal{S}_I(\vec{b}) = \{S_k, k = 1 \dots N_s\}. \quad (4.46)$$

Comme indiqué à la figure 4.6, nous notons $R_{S_k}^1$ et $R_{S_k}^2$ les deux régions de l'image I situées de part et d'autre du segment S_k et s'appuyant sur ce segment, leur largeur étant définie par le paramètre ϵ_R , dont la valeur est fixée arbitrairement à 3 pixels. L'adéquation du segment S_k avec l'image I est alors mesurée en calculant la différence d_{S_k} des radiométries moyennes calculées dans ces deux régions :

$$d_{S_k} = \left| \frac{1}{N_{R_{S_k}^1}} \sum_{l \in R_{S_k}^1} I(l) - \frac{1}{N_{R_{S_k}^2}} \sum_{l \in R_{S_k}^2} I(l) \right|, \quad (4.47)$$

où $N_{R_{S_k}^1}$ et $N_{R_{S_k}^2}$ désignent le nombre de pixels contenus respectivement dans les régions $R_{S_k}^1$ et $R_{S_k}^2$.

Le critère $\mathcal{C}_I^s(\vec{b})$ s'exprime alors comme la fusion conjonctive, par le biais d'une moyenne géométrique, des différences de radiométrie observées au niveau de chaque segment S_k de la signature :

$$\mathcal{C}_I^s(\vec{b}) = \sqrt[N_s]{\prod_{k=1}^{N_s} d_{S_k}} = \left[\prod_{k=1}^{N_s} \left| \frac{1}{N_{R_{S_k}^1}} \sum_{l \in R_{S_k}^1} I(l) - \frac{1}{N_{R_{S_k}^2}} \sum_{l \in R_{S_k}^2} I(l) \right| \right]^{1/N_s}. \quad (4.48)$$

4.3.1.2 Étude du critère sur données simulées et sur données réelles

Comme pour le critère \mathcal{C}_I^r , nous présentons, dans cette sous-section, des coupes du critère \mathcal{C}_I^s afin de caractériser son comportement en imagerie optique. Ces coupes, présentées à la figure 4.7, ont été établies à partir des images réelle et simulée de la figure 4.1.

Les coupes obtenues à partir de l'image simulée (courbes rouges) montrent la validité de l'approche proposée puisque la position du maximum global de ces courbes correspond bien à la valeur vraie du paramètre associé à chacune des courbes. De même, les coupes établies à partir de l'image réelle (courbes bleues) fournissent une estimation non biaisée des différents paramètres, à l'exception des paramètres Y_C et w_2 pour lesquels une erreur de 2 à 3 mètres est observée. Cette erreur est liée à une légère inadéquation des segments délimitant la région ombre R_O de la signature géométrique simulée du bâtiment G avec les bords présents dans l'image, comme cela peut être observé à la figure 2.9.

Le critère \mathcal{C}_I^r , présenté à la partie précédente, peut être caractérisé de global car il fait intervenir tous les pixels contenus dans la signature géométrique du bâtiment dont on cherche à quantifier l'adéquation avec l'image. À l'inverse, le critère \mathcal{C}_I^s ne fait intervenir que les pixels de l'image contenus dans un voisinage des segments de la signature géométrique du bâtiment, de sorte que ce critère peut être qualifié de local. Ainsi, nous constatons très logiquement que le critère \mathcal{C}_I^s est beaucoup

moins régulier que le critère \mathcal{C}_I^r et que son bassin d'attraction est plus étroit que celui du critère \mathcal{C}_I^r . Sur données réelles, le critère \mathcal{C}_I^s présente en revanche une plus grande sélectivité que le critère \mathcal{C}_I^r et conduit à des estimations non biaisées des paramètres, contrairement au critère \mathcal{C}_I^r , pour lequel l'inhomogénéité radiométrique des différentes régions constitutives du bâtiment engendre des biais sur l'estimation des paramètres.

4.3.2 Mise en œuvre du critère en imagerie RSO

4.3.2.1 Expression du critère pour un bruit multiplicatif

En imagerie RSO, l'expression du critère \mathcal{C}_I^s est analogue à celle établie dans le cas de l'imagerie optique, à la différence que nous calculons le rapport des radiométries quadratiques moyennes, estimées dans les régions $R_{S_k}^1$ et $R_{S_k}^2$, et non plus la différence des radiométries moyennes.

Nous notons X_1 et X_2 les variables aléatoires résultant de la moyenne quadratique des niveaux de gris des régions $R_{S_k}^1$ et $R_{S_k}^2$:

$$X_1 = \frac{1}{N_{R_{S_k}^1}} \sum_{k \in R_{S_k}^1} I(k)^2, \quad (4.49)$$

$$X_2 = \frac{1}{N_{R_{S_k}^2}} \sum_{k \in R_{S_k}^2} I(k)^2. \quad (4.50)$$

Nous avons vu à la section 4.2.4 que la radiométrie des images RSO d'amplitude au sein des régions présentant des caractéristiques physiques de rétrodiffusion stationnaires pouvait être décrite, dans le cadre de l'hypothèse de *speckle* pleinement développé, par une distribution de Rayleigh. Sous cette hypothèse, la variable aléatoire G , définie comme la différence des moyennes quadratiques X_1 et X_2 ($G = X_1 - X_2$), est distribuée, d'après les travaux de Touzi *et al.* [TLB88], selon la loi de probabilité suivante :

$$p_G(G|P_1, P_2) = \begin{cases} \frac{1}{\Gamma(N)^2} \left[\frac{P_1 P_2}{N^2} \right]^{-N} \exp\left(\frac{-GN}{P_1}\right) H(G) & \text{pour } G \geq 0, \\ \frac{1}{\Gamma(N)^2} \left[\frac{P_1 P_2}{N^2} \right]^{-N} \exp\left(\frac{-GN}{P_2}\right) H(-G) & \text{pour } G < 0, \end{cases} \quad (4.51)$$

avec :

$$H(G) = \left[\sum_{i=0}^{N-1} C_{N-1}^i \frac{G^{N-1} \Gamma(i+N)}{(N/P_1 + N/P_2)^{i+N}} \right], \quad (4.52)$$

où $\Gamma(\cdot)$ désigne la fonction Gamma, N correspond au nombre de pixels des régions $R_{S_k}^1$ et $R_{S_k}^2$ ($N = N_{R_{S_k}^1} = N_{R_{S_k}^2}$) et C_l^j désigne le nombre de combinaisons de j éléments parmi l . Enfin, P_1 et P_2 correspondent respectivement aux puissances moyennes rétrodiffusées par les régions $R_{S_k}^1$ et $R_{S_k}^2$:

$$\forall k \in R_{S_k}^1, E(I(k)^2) = P_1, \quad (4.53)$$

$$\forall k \in R_{S_k}^2, E(I(k)^2) = P_2. \quad (4.54)$$

Sous les mêmes hypothèses, Touzi *et al.* établissent également la loi de probabilité de la variable aléatoire R , définie comme le rapport des moyennes quadratiques X_1 et X_2 ($R = X_1/X_2$) :

$$p_R(R|P_1, P_2) = p_R(R|(P_1/P_2)) = \frac{\Gamma(2N)}{\Gamma(N)^2} \left[\frac{(P_1/P_2)^N}{(R + P_1/P_2)^{2N}} \right] R^{N-1} \mathbf{1}_{[0;+\infty[}(R). \quad (4.55)$$

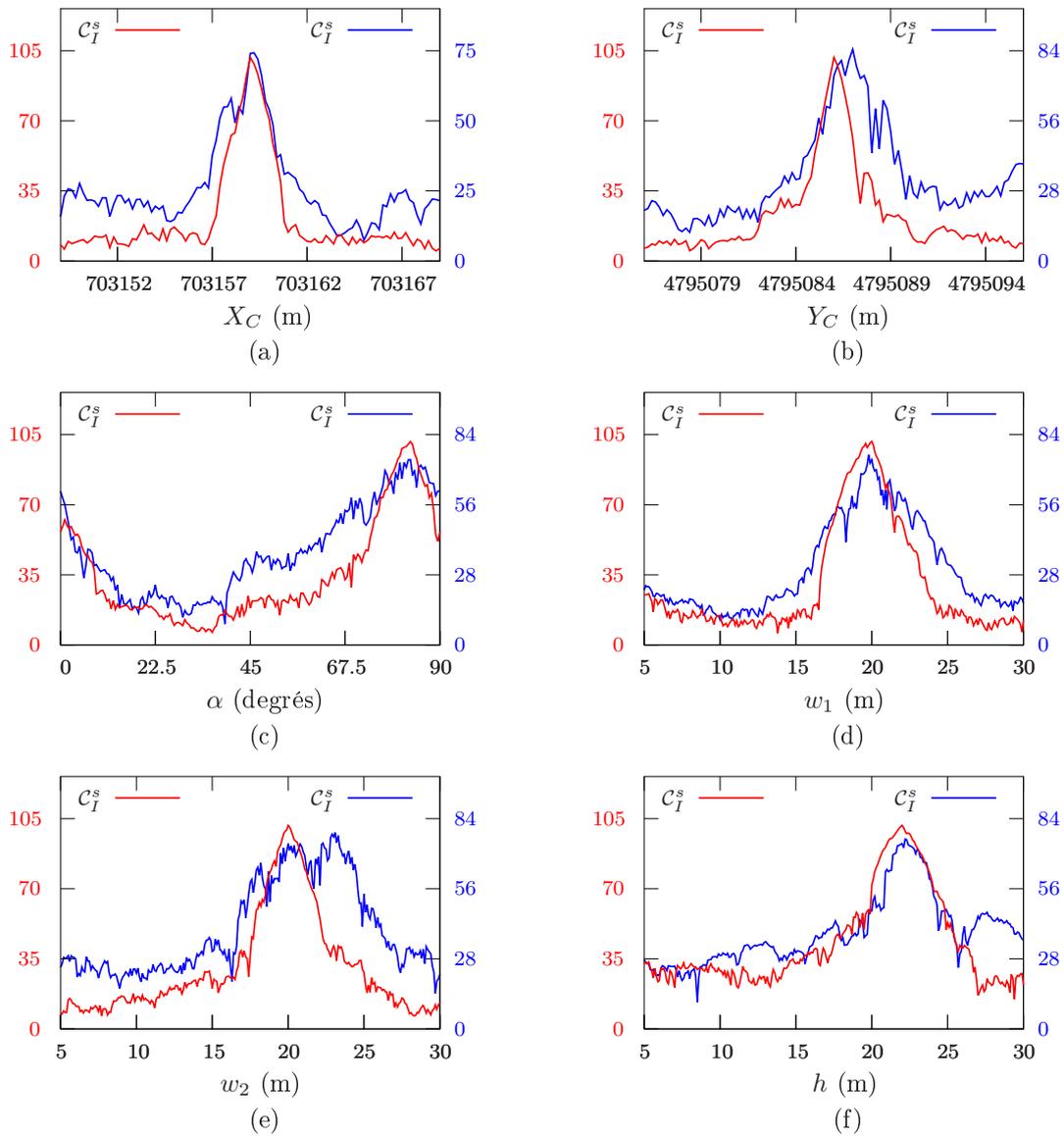


Figure 4.7 – Coupes du critère C_I^s calculées à partir de l'image optique réelle de la figure 4.1 (a) (courbes bleues) et de l'image optique simulée de la figure 4.1 (b) (courbes rouges).

En l'absence de bord, les puissances moyennes P_1 et P_2 rétrodiffusées par les régions $R_{S_k}^1$ et $R_{S_k}^2$ sont égales ($P_1 = P_2 = P$). Les distributions des variables aléatoires G et R , calculées dans ce cas particulier, permettent de calculer, pour un taux de fausse alarme donné, les seuils de détection T_G et T_R des détecteurs de bords fondés respectivement sur le calcul de G et de R . La distribution de G , calculée pour $P_1 = P_2 = P$, présente une dépendance en P , de sorte que le seuil de détection T_G dépend également de P . La détection de bords par l'intermédiaire de la différence des radiométries quadratiques moyennes ne se fait donc pas à taux de fausses alarmes constant. À l'inverse, la distribution de R , calculée pour $P_1 = P_2 = P$, est indépendante de P . Le seuil T_R est donc indépendant de P , de sorte que la détection de bords au moyen du rapport des radiométries quadratiques moyennes se fait à taux de fausses alarmes constant.

De ces observations, Touzi *et al.* concluent que le rapport des radiométries quadratiques moyennes est davantage adapté à la nature multiplicative du bruit de *speckle* que la différence des radiométries quadratiques moyennes pour la détection des bords en imagerie RSO. Ce résultat justifie l'approche que nous mettons ici en œuvre pour quantifier l'adéquation d'un segment S_k de la signature géométrique d'un bâtiment avec une image RSO.

Nous notons r_{S_k} le rapport des radiométries quadratiques moyennes calculées dans les régions $R_{S_k}^1$ et $R_{S_k}^2$, rapport qui quantifie l'adéquation du segment S_k avec l'image et dont la définition garantit que $r_{S_k} \geq 1$:

$$r_{S_k} = \begin{cases} \frac{N_{R_{S_k}^2} \sum_{k \in R_{S_k}^1} I(k)^2}{N_{R_{S_k}^1} \sum_{k \in R_{S_k}^2} I(k)^2} & \text{si } \frac{1}{N_{R_{S_k}^1}} \sum_{k \in R_{S_k}^1} I(k)^2 \geq \frac{1}{N_{R_{S_k}^2}} \sum_{k \in R_{S_k}^2} I(k)^2, \\ \frac{N_{R_{S_k}^1} \sum_{k \in R_{S_k}^2} I(k)^2}{N_{R_{S_k}^2} \sum_{k \in R_{S_k}^1} I(k)^2} & \text{sinon.} \end{cases} \quad (4.56)$$

Le critère $\mathcal{C}_I^s(\vec{b})$, qui résulte, comme en imagerie optique, de la fusion conjonctive des scores obtenus sur chacun des segments S_k , s'exprime alors comme suit :

$$\mathcal{C}_I^s(\vec{b}) = \sqrt[N_s]{\prod_{k=1}^{N_s} r_{S_k}}. \quad (4.57)$$

4.3.2.2 Étude du critère sur données simulées et sur données réelles

Dans cette sous-section, nous présentons à la figure 4.8 les coupes du critère \mathcal{C}_I^s établies à partir des images RSO réelle et simulée de la figure 4.4.

Nous constatons, hormis à la figure 4.8 (f), que chacune des coupes du critère \mathcal{C}_I^s présente, sur données simulées comme sur données réelles, un maximum global situé approximativement au niveau de la valeur vraie du paramètre associé à la coupe considérée, ce qui valide l'approche proposée pour quantifier l'adéquation des segments de la signature géométrique d'un bâtiment avec une image RSO.

Nous pouvons observer à la figure 4.8 (f) que les coupes du critère \mathcal{C}_I^s selon l'espace \mathcal{E}_h présentent, sur données réelles comme sur données simulées, un maximum global pour $h \approx 12$ mètres alors que la hauteur du bâtiment G est de 22 mètres. Afin d'expliquer les raisons pour lesquelles le maximum de cette courbe n'est pas atteint au niveau de la valeur vraie du paramètre h , nous représentons, à la figure 4.9, les signatures géométriques des bâtiments G qui ont été simulées pour différentes hauteurs.

Pour $h = 7$ mètres, une partie du bâtiment est vue par le capteur RSO comme un objet totalement occultant, de sorte que la signature géométrique du bâtiment possède une région R_T qui correspond, conformément aux notations introduites à la section 2.4.3, à la rétrodiffusion du toit du bâtiment.

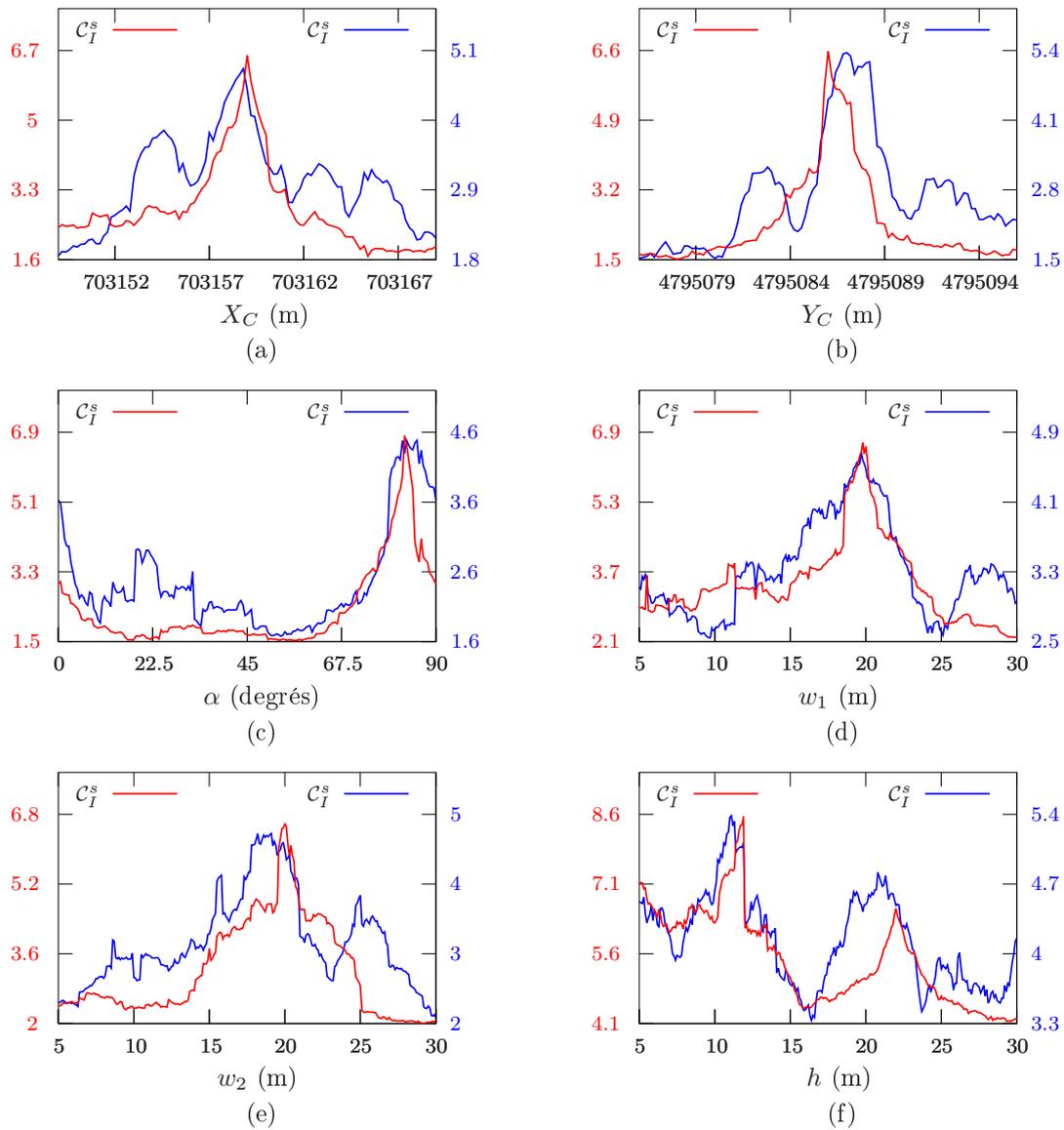


Figure 4.8 – Coupes du critère C_I^s calculées à partir de l'image RSO réelle de la figure 4.4 (a) (courbes bleues) et de l'image RSO simulée de la figure 4.4 (b) (courbes rouges).

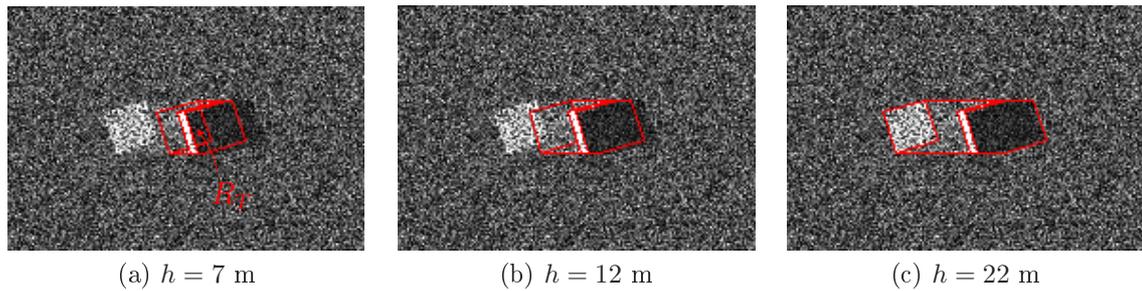


Figure 4.9 – Signatures géométriques du bâtiment G avec une hauteur $h = 7$ m (a), $h = 12$ m (b) et $h = 22$ m (c) dans l'image RSO simulée de la figure 4.4 (b).

Pour $h = 22$ mètres, le bâtiment est vu par le capteur comme un objet non occultant pour l'ensemble des positions azimutales de ce dernier, hormis aux extrémités du bâtiments où il est vu comme un objet partiellement occultant, de sorte que la région R_{TS} n'est pas définie.

Pour $h = 12$ mètres, la région R_T est certes définie mais est quasiment réduite à un segment. Parmi les quatre segments qui définissent cette région, deux sont confondus, en partie, avec l'un des segments délimitant la région R_{ED} associée à l'écho double du bâtiment. Le rapport des radiométries quadratiques moyennes au niveau de ces deux segments est alors très élevé puisqu'ils se trouvent en fait à la frontière entre la région R_{ED} de radiométrie élevée et la région R_O de radiométrie faible, ce qui explique que le critère \mathcal{C}_I^s atteigne son maximum pour cette valeur de h .

Pour conclure, des observations similaires à celles de la sous-section 4.3.2.2 peuvent être faites. Nous constatons notamment que le critère \mathcal{C}_I^s est moins régulier et présente un bassin d'attraction plus étroit que le critère \mathcal{C}_I^r . Il présente, en revanche, une courbure plus importante que ce dernier au niveau de son maximum et conduit ainsi à une estimation plus précise des différents paramètres.

4.4 Estimation des paramètres d'un bâtiment par mesure de l'adéquation radiométrique de sa signature géométrique avec une image

Dans les deux parties précédentes, les critères \mathcal{C}_I^r et \mathcal{C}_I^s ont été définis et leur capacité à quantifier l'adéquation de la signature géométrique d'un bâtiment avec l'image au sein de laquelle elle a été simulée a été validée sur données simulées et sur données réelles.

Dans cette dernière partie, nous supposons qu'un bâtiment a été détecté dans l'image, au moyen d'une méthode automatique ou supervisée, et que la position du centre de son emprise au sol est connue approximativement, c'est-à-dire que les paramètres X_C et Y_C sont connus avec une précision de 1 à 2 mètres. Nous établissons une méthode, fondée sur l'optimisation successive des critères \mathcal{C}_I^r et \mathcal{C}_I^s , pour estimer les paramètres de ce bâtiment. La méthode consiste à déterminer le vecteur de paramètres \vec{b} pour lequel la signature géométrique correspondante $\mathcal{S}_I(\vec{b})$ présente une adéquation maximale avec l'image, au sens des deux critères précédents.

Après avoir mené une étude quantitative visant à déterminer le bassin d'attraction des critères \mathcal{C}_I^r et \mathcal{C}_I^s , nous établissons, en imagerie optique et en imagerie RSO, une stratégie d'optimisation qui repose sur la maximisation successive de ces deux critères. Les résultats de la stratégie proposée sont finalement présentés sur données simulées et sur données réelles.

4.4.1 Estimation des paramètres en imagerie optique

4.4.1.1 Détermination des bassins d'attraction des critères C_I^r et C_I^s sur données simulées

Dans cette sous-section, nous considérons un ensemble de 500 bâtiments parallélépipédiques dont les paramètres sont tirés aléatoirement dans l'espace des paramètres \mathcal{E}_b selon une loi uniforme.

Pour chacun de ces bâtiments, nous simulons, comme cela a été fait à la sous-section 4.2.3.3, l'image de ce bâtiment qui serait obtenue si elle avait été acquise par le capteur QuickBird dans les mêmes conditions de prise de vue que l'image de la figure 2.9. Pour ce faire, la signature géométrique du bâtiment est tout d'abord simulée au moyen de l'algorithme de projection présenté à la section 2.3.3. Les niveaux de gris des pixels contenus dans les différentes régions constitutives de cette signature sont ensuite générés par tirages aléatoires selon des lois normales, dont les paramètres sont précisés au tableau 4.1. Nous obtenons ainsi une image semblable à celle de la figure 4.1 (b).

Pour chacune des images précédemment simulées, les critères C_I^r et C_I^s sont maximisés au moyen d'un algorithme de descente de gradient, le vecteur gradient étant calculé par différence finie. Afin de déterminer les bassins d'attraction de ces critères, nous proposons d'étudier l'influence de l'initialisation de l'algorithme de descente de gradient sur la précision avec laquelle sont estimés les paramètres des bâtiments.

Nous présentons, à la figure 4.10 (a), la moyenne quadratique des erreurs finales observées à l'issue d'une maximisation unidimensionnelle des critères C_I^r et C_I^s dans l'espace \mathcal{E}_{X_C} , l'ensemble des paramètres étant initialisés à leurs valeurs vraies, hormis le paramètre X_C pour lequel l'erreur initiale varie entre 0.5 et 4 mètres. De manière analogue, les courbes des figures 4.10 (b), 4.10 (c), 4.10 (d), 4.10 (e) et 4.10 (f) sont obtenues par maximisation unidimensionnelle des critères C_I^r et C_I^s dans les espaces \mathcal{E}_{Y_C} , \mathcal{E}_α , \mathcal{E}_{w_1} , \mathcal{E}_{w_2} et \mathcal{E}_h .

D'après la figure 4.10, il apparaît très nettement que le critère C_I^r présente un bassin d'attraction plus vaste que le critère C_I^s , ce qui est cohérent avec le comportement de ces critères que nous avons décrit aux sous-sections 4.2.3.3 et 4.3.1.2. À l'inverse du critère C_I^s , qui ne fait intervenir que les pixels situés dans un voisinage des segments de la signature géométrique, le critère C_I^r implique l'ensemble des pixels contenus dans la région englobante R_0 . De plus, le critère C_I^r se caractérise par une grande régularité, contrairement au critère C_I^s qui présente de nombreux *maxima* locaux. Ces deux éléments permettent d'expliquer que le bassin d'attraction du critère C_I^r soit plus important que celui du critère C_I^s .

Plus précisément, nous pouvons remarquer que le critère C_I^r présente un vaste bassin d'attraction quel que soit l'espace considéré, à l'exception de l'espace \mathcal{E}_α au sein duquel le critère présente un bassin d'attraction relativement étroit, puisque l'algorithme de descente de gradient doit être initialisé à une distance inférieure à 20 degrés pour obtenir une moyenne quadratique des erreurs inférieure à 5 degrés. En ce qui concerne le critère C_I^s , nous pouvons considérer que son bassin d'attraction présente une demi largeur d'environ 2 mètres dans les espaces \mathcal{E}_{X_C} et \mathcal{E}_{Y_C} , et d'environ 4 mètres dans les espaces \mathcal{E}_{w_1} , \mathcal{E}_{w_2} et \mathcal{E}_h . Enfin, ce critère présente un bassin d'attraction encore plus étroit que ne l'est celui du critère C_I^r dans l'espace \mathcal{E}_α , puisqu'il faut initialiser l'algorithme de descente de gradient à une distance inférieure à 10 degrés pour obtenir une moyenne quadratique des erreurs finales inférieure à 5 degrés.

L'étude réalisée dans cette sous-section a été menée indépendamment dans les différents espaces \mathcal{E}_{X_C} , \mathcal{E}_{Y_C} , \mathcal{E}_α , \mathcal{E}_{w_1} , \mathcal{E}_{w_2} et \mathcal{E}_h . Elle ne permet donc pas de définir le bassin d'attraction des critères C_I^r et C_I^s dans l'espace 6-dimensionnel \mathcal{E}_b résultant du produit cartésien de ces espaces, mais elle permet néanmoins de déterminer les dimensions maximales de ce dernier.

4.4.1.2 Stratégie d'estimation

Aux vues des propriétés des critères C_I^r et C_I^s en imagerie optique, propriétés qui ont été étudiées aux sous-sections 4.2.3.3, 4.3.1.2 et 4.4.1.1, nous proposons ici d'établir une stratégie visant à esti-

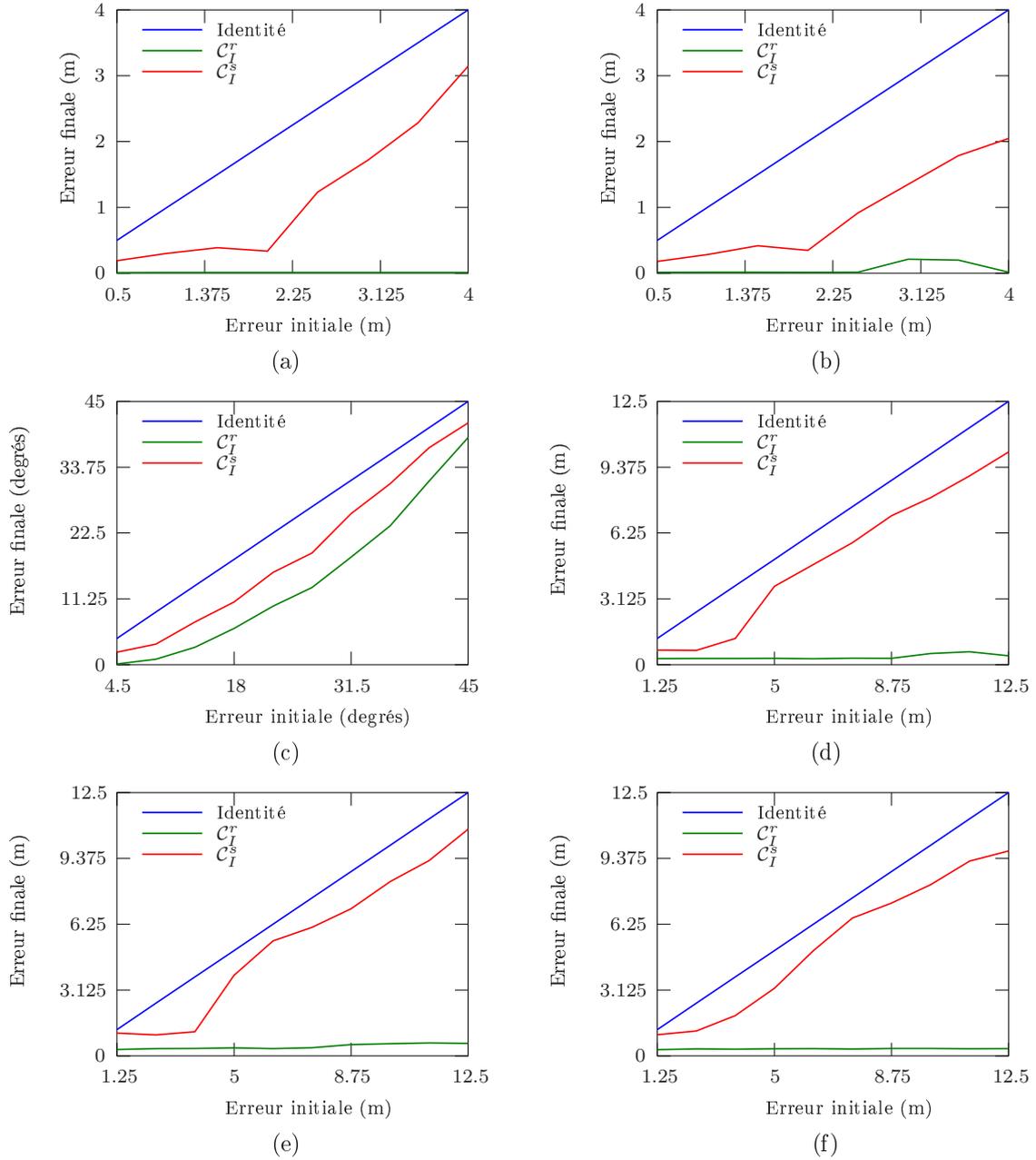


Figure 4.10 – Évolution de la moyenne quadratique des erreurs finales, observées à l'issue d'une maximisation unidimensionnelle des critères C_I^r et C_I^s dans les espaces \mathcal{E}_{X_C} (a), \mathcal{E}_{Y_C} (b), \mathcal{E}_α (c), \mathcal{E}_{w_1} (d), \mathcal{E}_{w_2} (e), \mathcal{E}_h (f), en fonction de l'erreur initiale.

mer les paramètres d'un bâtiment parallélépipédique en maximisant successivement ces deux critères. L'objectif de cette stratégie d'estimation est de tirer le meilleur parti de chacun des critères \mathcal{C}_I^r et \mathcal{C}_I^s afin d'estimer le plus précisément possible les paramètres d'un bâtiment tout en conservant une complexité calculatoire raisonnable. Nous souhaitons notamment exploiter la forte attractivité du critère \mathcal{C}_I^r et la sélectivité (sur données réelles) du critère \mathcal{C}_I^s .

La stratégie proposée consiste tout d'abord à maximiser le critère \mathcal{C}_I^r dans l'espace \mathcal{E}_b par mise en œuvre d'un algorithme de descente de gradient à partir d'un ensemble d'états initiaux répartis sur une grille régulière de l'espace \mathcal{E}_b . Les pas de cette grille régulière sont déterminés d'après les dimensions du bassin d'attraction, établies à la sous-partie précédente, dans chacun des espaces \mathcal{E}_{X_C} , \mathcal{E}_{Y_C} , \mathcal{E}_α , \mathcal{E}_{w_1} , \mathcal{E}_{w_2} et \mathcal{E}_h .

Dans la suite de ce chapitre, nous considérons que les paramètres des bâtiments parallélépipédiques prennent leur valeur au sein des ensembles suivants :

$$\mathcal{E}_{X_C} = [X_C^0 - 2; X_C^0 + 2], \quad (4.58)$$

$$\mathcal{E}_{Y_C} = [Y_C^0 - 2; Y_C^0 + 2], \quad (4.59)$$

$$\mathcal{E}_\alpha = [0; 90], \quad (4.60)$$

$$\mathcal{E}_{w_1} = \mathcal{E}_{w_2} = \mathcal{E}_h = [5; 30], \quad (4.61)$$

où X_C^0 et Y_C^0 correspondent aux coordonnées approximatives du centre de l'emprise au sol du bâtiment, déterminées lors de l'étape de détection préalable à la mise en œuvre de cette stratégie d'estimation.

Étant données les valeurs admissibles des paramètres et les dimensions du bassin d'attraction du critère \mathcal{C}_I^r , l'ensemble \mathcal{E}_b^i des états initiaux, à partir desquels est mis en œuvre l'algorithme de maximisation par descente de gradient, est défini comme le produit cartésien des ensembles $\mathcal{E}_{X_C}^i$, $\mathcal{E}_{Y_C}^i$, \mathcal{E}_α^i , $\mathcal{E}_{w_1}^i$, $\mathcal{E}_{w_2}^i$ et \mathcal{E}_h^i définis comme suit :

$$\mathcal{E}_{X_C}^i = \{X_C^0 - 2; X_C^0; X_C^0 + 2\}, \quad (4.62)$$

$$\mathcal{E}_{Y_C}^i = \{Y_C^0 - 2; Y_C^0; Y_C^0 + 2\}, \quad (4.63)$$

$$\mathcal{E}_\alpha^i = \{0; 5; \dots; 80; 85\}, \quad (4.64)$$

$$\mathcal{E}_{w_1}^i = \mathcal{E}_{w_2}^i = \mathcal{E}_h^i = \{17.5\}. \quad (4.65)$$

Après avoir maximisé le critère \mathcal{C}_I^r à partir des 162 états initiaux précédemment définis, le critère \mathcal{C}_I^s est finalement maximisé par descente de gradient, l'algorithme étant initialisé à l'aide du jeu de paramètres issu de la maximisation du critère \mathcal{C}_I^r .

4.4.1.3 Résultats sur données simulées et sur données réelles

Dans cette sous-section, nous présentons les résultats de la stratégie d'estimation mise en œuvre sur données simulées et sur données réelles.

Les tableaux 4.3 et 4.4 présentent les valeurs estimées des paramètres des bâtiments de la scène d'étude introduite au chapitre 2. L'estimation de ces paramètres a été réalisée à partir de l'image simulée de la figure 4.11 (a) et de l'image réelle de la figure 4.12 (a).

Les résultats obtenus sur données simulées sont tout à fait satisfaisants puisque les erreurs d'estimation maximales pour les paramètres X_C , Y_C , α , w_1 , w_2 et h sont respectivement de 0.2 mètres, 0.4 mètres, 1.1 degrés, 0.4 mètres, 0.8 mètres et 0.3 mètres, hormis pour les bâtiments H et I dont la proximité engendre des erreurs plus importantes.

Les erreurs d'estimation maximales observées sur données réelles sont respectivement de 1.8 mètres, 3.4 mètres, 3.4 degrés, 2.9 mètres, 3.8 mètres et 2.7 mètres pour les paramètres X_C , Y_C , α , w_1 , w_2 et h . Aux vues de la précision avec laquelle a été établie la vérité terrain (cf. tableau 2.2), les résultats obtenus sont relativement bons et illustrent la robustesse de la méthode malgré la mise en défaut de l'hypothèse d'uniformité radiométrique des différentes régions constitutives de la signature géométrique d'un bâtiment. Rappelons que cette hypothèse est mise en défaut en raison de la présence de fenêtres sur les façades, de superstructures sur les toits ou encore en raison des phénomènes d'occlusion des ombres par d'autres bâtiments ou bien par la végétation.

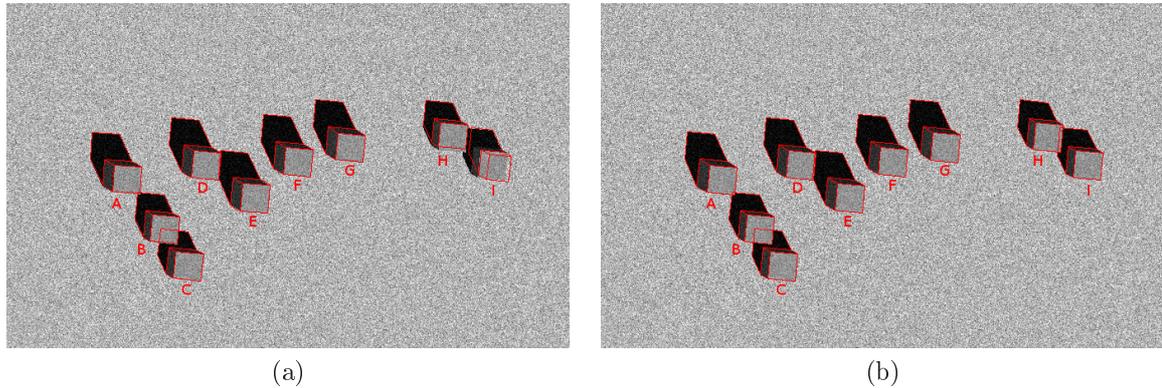


Figure 4.11 – Résultats de la stratégie d'estimation sur données simulées en imagerie optique : signatures géométriques des neuf bâtiments de la scène d'étude simulées à partir des valeurs estimées (a) et des valeurs vraies (b) des paramètres de ces bâtiments.

Label		X_C (m)	Y_C (m)	α (degrés)	w_1 (m)	w_2 (m)	h (m)
A	Vérité terrain	702995.0	4795064.0	83.0	20.0	20.0	22.0
	Estimation	702994.9	4795064.4	83.9	19.8	19.2	21.8
B	Vérité terrain	703025.0	4795026.0	83.0	20.0	20.0	15.0
	Estimation	703024.9	4795026.4	83.2	19.6	19.2	14.8
C	Vérité terrain	703042.0	4794999.0	83.0	20.0	20.0	15.0
	Estimation	703042.2	4794999.1	81.9	20.0	19.6	15.1
D	Vérité terrain	703053.0	4795074.0	83.0	20.0	20.0	22.0
	Estimation	703052.9	4795074.1	83.4	19.6	19.9	21.7
E	Vérité terrain	703088.5	4795050.0	83.0	20.0	20.0	22.0
	Estimation	703088.6	4795050.0	83.8	19.7	20.4	21.7
F	Vérité terrain	703120.0	4795076.0	83.0	20.0	20.0	22.0
	Estimation	703120.0	4795076.2	83.0	20.0	19.6	22.0
G	Vérité terrain	703159.0	4795086.0	83.0	20.0	20.0	22.0
	Estimation	703159.0	4795086.2	83.1	19.9	19.5	22.0
H	Vérité terrain	703236.0	4795092.0	83.0	20.0	20.0	15.0
	Estimation	703236.0	4795093.2	83.1	19.7	17.6	14.7
I	Vérité terrain	703264.0	4795072.0	83.0	20.0	20.0	15.0
	Estimation	703267.5	4795070.6	82.2	15.9	18.5	17.5

Tableau 4.3 – Résultats de la stratégie d'estimation sur données simulées en imagerie optique : valeurs estimées et valeurs vraies des paramètres des neuf bâtiments de la scène d'étude.



Figure 4.12 – Résultats de la stratégie d'estimation sur données réelles en imagerie optique : signatures géométriques des neuf bâtiments de la scène d'étude simulées à partir des valeurs estimées (a) et des valeurs vraies (b) des paramètres de ces bâtiments.

Label		X_C (m)	Y_C (m)	α (degrés)	w_1 (m)	w_2 (m)	h (m)
A	Vérité terrain	702995.0	4795064.0	83.0	20.0	20.0	22.0
	Estimation	702994.4	4795064.5	80.7	18.3	19.8	20.1
B	Vérité terrain	703025.0	4795026.0	83.0	20.0	20.0	15.0
	Estimation	703024.7	4795026.0	85.2	19.9	19.0	12.8
C	Vérité terrain	703042.0	4794999.0	83.0	20.0	20.0	15.0
	Estimation	703042.3	4794998.3	83.5	19.7	19.4	12.3
D	Vérité terrain	703053.0	4795074.0	83.0	20.0	20.0	22.0
	Estimation	703053.6	4795073.8	81.8	20.3	20.8	20.1
E	Vérité terrain	703088.5	4795050.0	83.0	20.0	20.0	22.0
	Estimation	703087.8	4795051.9	82.6	18.6	19.4	21.3
F	Vérité terrain	703120.0	4795076.0	83.0	20.0	20.0	22.0
	Estimation	703119.6	4795075.6	84.3	20.4	19.0	20.3
G	Vérité terrain	703159.0	4795086.0	83.0	20.0	20.0	22.0
	Estimation	703158.9	4795089.4	80.5	19.6	16.2	21.3
H	Vérité terrain	703236.0	4795092.0	83.0	20.0	20.0	15.0
	Estimation	703234.2	4795092.2	86.4	17.1	19.0	14.0
I	Vérité terrain	703264.0	4795072.0	83.0	20.0	20.0	15.0
	Estimation	703265.2	4795070.6	85.0	20.3	20.7	12.8

Tableau 4.4 – Résultats de la stratégie d'estimation sur données réelles en imagerie optique : valeurs estimées et valeurs vraies des paramètres des neuf bâtiments de la scène d'étude.

4.4.2 Estimation des paramètres en imagerie RSO

Comme en imagerie optique, nous présentons, dans cette section, une méthode pour l'estimation des paramètres de bâtiments parallélépipédiques en imagerie RSO. Dans la première sous-section, nous étudions quantitativement le comportement des critères \mathcal{C}_I^r et \mathcal{C}_I^s afin de déterminer les dimensions de leur bassin d'attraction. Dans la seconde sous-section, nous nous appuyons sur ces résultats pour établir une stratégie d'estimation qui repose sur la maximisation successive de ces deux critères. Les performances de cette stratégie d'estimation sont finalement présentées sur données simulées et sur données réelles dans la dernière sous-section.

4.4.2.1 Détermination des bassins d'attraction des critères \mathcal{C}_I^r et \mathcal{C}_I^s sur données simulées

En adoptant la même démarche que celle de la sous-section 4.4.1.1, nous estimons, par maximisation des critères \mathcal{C}_I^r et \mathcal{C}_I^s , les paramètres d'un ensemble de 500 bâtiments à partir d'images RSO simulées. Nous observons alors l'influence de l'initialisation de l'algorithme d'optimisation sur la moyenne quadratique des erreurs d'estimation. Notons que la géométrie de prise de vue utilisée pour la simulation des images RSO est celle de l'image TerraSAR-X de la figure 2.17.

Les courbes de la figure 4.13 montrent clairement que le critère \mathcal{C}_I^r présente, pour les mêmes raisons que celles évoquées en imagerie optique, un bassin d'attraction plus large que ne l'est celui du critère \mathcal{C}_I^s .

Nous constatons également, d'après les figures 4.13 (a) et 4.13 (b), que les critères \mathcal{C}_I^r et \mathcal{C}_I^s présentent un comportement anisotrope au sein de l'espace $\mathcal{E}_{X_C} \times \mathcal{E}_{Y_C}$ puisque les dimensions de leur bassin d'attraction diffèrent dans les espaces \mathcal{E}_{X_C} et \mathcal{E}_{Y_C} . En fait, les dimensions du bassin d'attraction de ces critères dans les espaces \mathcal{E}_{X_C} et \mathcal{E}_{Y_C} dépendent de l'orientation relative du repère image \mathcal{R}_I par rapport au repère terrain \mathcal{R}_T , de sorte que les résultats établis sont spécifiques à la géométrie de prise de vue des images. Par la suite, nous considérons que les demi largeurs des bassins d'attraction des critères \mathcal{C}_I^r et \mathcal{C}_I^s sont identiques dans les espaces \mathcal{E}_{X_C} et \mathcal{E}_{Y_C} et valent 1 mètre.

Comme en imagerie optique, les bassins d'attraction des critères \mathcal{C}_I^r et \mathcal{C}_I^s dans l'espace \mathcal{E}_α sont très étroits puisque l'algorithme de descente de gradient doit être initialisé à des distances respectives de 9 degrés et de 13.5 degrés pour obtenir une moyenne quadratique des erreurs d'estimation inférieure à 5 degrés. La faible attractivité des critères \mathcal{C}_I^r et \mathcal{C}_I^s dans l'espace \mathcal{E}_α se traduit, dans la stratégie d'estimation de la sous-section suivante, par la nécessité d'un parcours quasi-exhaustif de cet espace lors de la maximisation de ces critères.

4.4.2.2 Stratégie d'estimation

Comme en imagerie optique, la stratégie d'estimation mise en œuvre repose sur une maximisation successive des critères \mathcal{C}_I^r et \mathcal{C}_I^s au moyen d'un algorithme de descente de gradient. Cette stratégie permet d'exploiter l'attractivité du critère \mathcal{C}_I^r et la précision d'estimation du critère \mathcal{C}_I^s .

Seule diffère la grille régulière définissant l'ensemble \mathcal{E}_b^i des vecteurs de paramètres à partir desquels est initialisé l'algorithme de descente de gradient pour la maximisation du critère \mathcal{C}_I^r . L'attractivité du critère \mathcal{C}_I^r est plus faible en imagerie RSO qu'en imagerie optique, ce qui s'explique par un niveau de bruit plus élevé en imagerie RSO. La grille régulière mise en œuvre en imagerie RSO est donc plus fine que celle mise en œuvre en imagerie optique. L'ensemble \mathcal{E}_b^i des états initiaux, établi à partir des résultats de la sous-section précédente, est défini par le produit cartésien des ensembles suivants :

$$\mathcal{E}_{X_C}^i = \{X_C^0 - 2; X_C^0; X_C^0 + 2\}, \quad (4.66)$$

$$\mathcal{E}_{Y_C}^i = \{Y_C^0 - 2; Y_C^0; Y_C^0 + 2\}, \quad (4.67)$$

$$\mathcal{E}_\alpha^i = \{0; 5; \dots; 80; 85\}, \quad (4.68)$$

$$\mathcal{E}_{w_1}^i = \mathcal{E}_{w_2}^i = \mathcal{E}_h^i = \{9.17; 17.51; 25.85\}, \quad (4.69)$$

où X_C^0 et Y_C^0 correspondent aux coordonnées approximatives du centre de l'emprise au sol du bâtiment, déterminées lors de l'étape de détection préalable à la mise en œuvre de cette stratégie d'estimation.

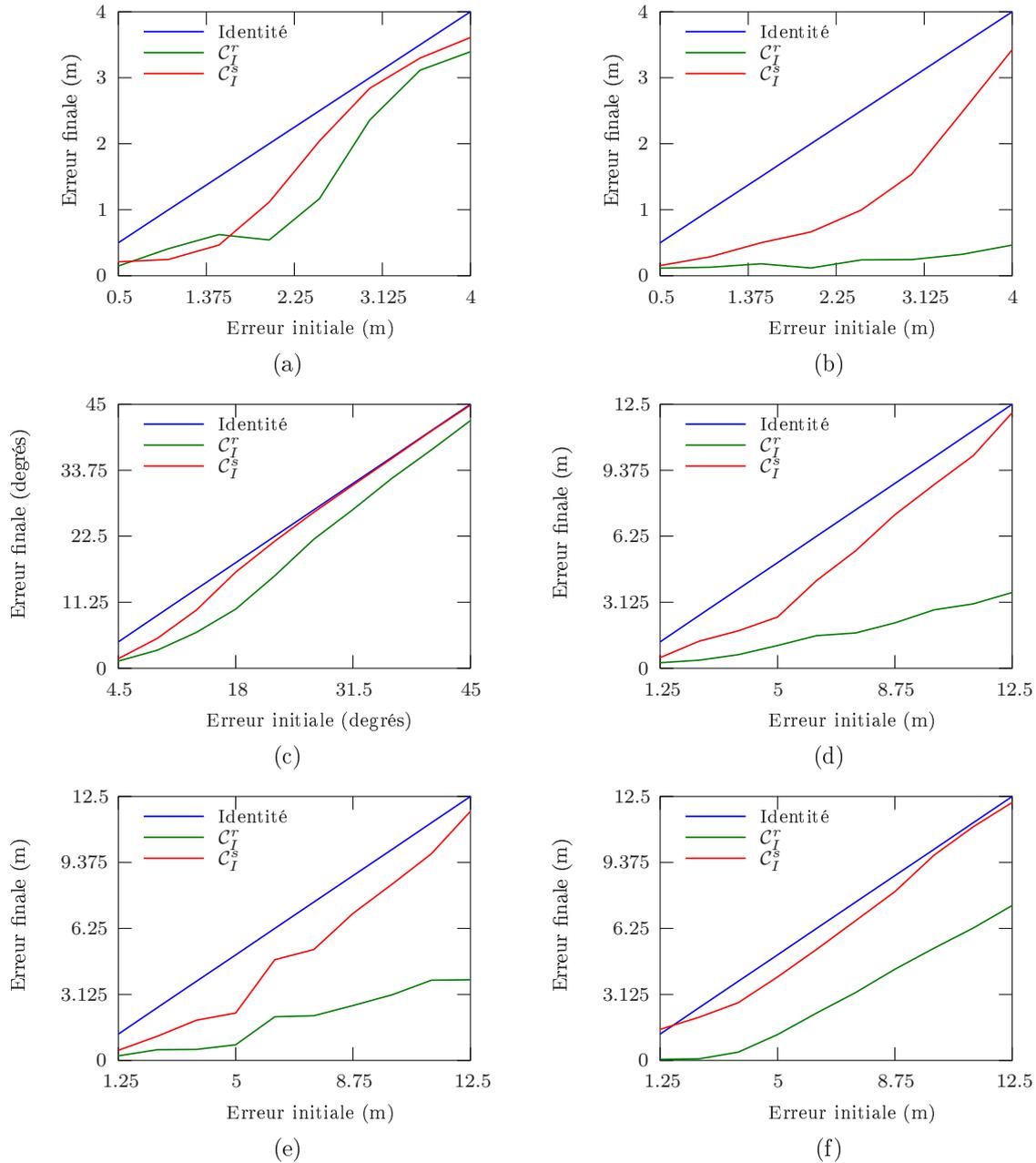


Figure 4.13 – Évolution de la moyenne quadratique des erreurs finales, observées à l'issue d'une maximisation unidimensionnelle des critères C_I^r et C_I^s dans les espaces \mathcal{E}_{X_C} (a), \mathcal{E}_{Y_C} (b), \mathcal{E}_α (c), \mathcal{E}_{w_1} (d), \mathcal{E}_{w_2} (e), \mathcal{E}_h (f), en fonction de l'erreur initiale.

Alors que l'ensemble \mathcal{E}_b^i mis en œuvre pour la maximisation du critère \mathcal{C}_I^r en imagerie optique comporte 162 éléments, celui mis en œuvre en imagerie RSO en comporte 4374, ce qui implique une complexité calculatoire beaucoup plus importante mais nécessaire aux vues de la faible régularité du critère \mathcal{C}_I^r en imagerie RSO.

Notons, pour finir, que cette stratégie d'estimation permet de contourner le problème mis en évidence à la sous-section 4.3.2.2, où nous avons constaté que la position du maximum global de la coupe du critère \mathcal{C}_I^s selon l'espace \mathcal{E}_h était erronée, la position vraie correspondant à un maximum local de ce critère. Dans le cadre de la stratégie proposée, la maximisation du critère \mathcal{C}_I^r , qui précède celle du critère \mathcal{C}_I^s , conduit à un jeu de paramètres proche de la solution. Ainsi, l'algorithme de descente de gradient mis en œuvre pour la maximisation du critère \mathcal{C}_I^s est initialisé dans le bassin d'attraction du maximum local correspondant à la solution désirée.

4.4.2.3 Résultats sur données simulées et sur données réelles

Dans cette dernière sous-section, nous présentons les résultats de la stratégie d'estimation mise en œuvre sur données simulées et sur données réelles.

Les tableaux 4.5 et 4.6 présentent les valeurs estimées des paramètres des bâtiments de la scène d'étude introduite au chapitre 2. L'estimation de ces paramètres a été réalisée à partir de l'image simulée de la figure 4.14 (a) et de l'image réelle de la figure 4.15 (a).

La mise en œuvre de la stratégie d'estimation sur données simulées conduit à des résultats relativement satisfaisants pour les bâtiments B , C , H et I , puisque les erreurs d'estimation maximales observées sur ces quatre bâtiments sont respectivement de 0.5 mètres, 0.5 mètres, 0.7 degrés, 2.2 mètres, 0.9 mètres et 1.3 mètres pour les paramètres X_C , Y_C , α , w_1 , w_2 et h . Pour les autres bâtiments de la scène d'étude, pour lesquels des phénomènes de recouvrement interviennent, les résultats sont mitigés ce qui traduit la robustesse insuffisante de cette approche aux phénomènes de recouvrement.

Pour finir, la qualité des résultats établis à partir des données réelles est peu satisfaisante puisque les erreurs relatives à l'estimation de la position du centre de l'emprise au sol et des dimensions des bâtiments peuvent aller jusqu'à une dizaine de mètres. Il convient toutefois de rappeler la complexité de l'extraction de bâtiments en monoscopie RSO étant donné la nature bruitée des images, la variabilité de la signature géométrique des bâtiments et les recouvrements entre ces bâtiments.

4.5 Conclusion

Dans ce chapitre, une méthode a été établie pour l'extraction tridimensionnelle des bâtiments parallélépipédiques en monoscopie optique et RSO. L'approche proposée, qui correspond à une segmentation de l'image contrainte par un modèle de bâtiment, repose sur la maximisation successive de deux critères radiométriques quantifiant l'adéquation de la signature géométrique d'un bâtiment avec l'image au sein de laquelle elle a été simulée.

La mise en œuvre de cette méthode en imagerie optique conduit à des résultats probants sur données simulées comme sur données réelles. Dans ce dernier cas, l'orientation des bâtiments est estimée avec une erreur inférieure à 5 degrés et les dimensions des bâtiments sont estimées avec des erreurs inférieures à 2 mètres. Ces précisions d'estimation sont satisfaisantes en comparaison de celles obtenues lors d'une extraction manuelle des bâtiments par saisie des sommets visibles de ces bâtiments.

En imagerie RSO, la méthode proposée conduit, en revanche, à des résultats mitigés aussi bien sur données simulées que sur données réelles. Ceci s'explique notamment par la faible robustesse de l'approche proposée aux phénomènes d'occlusion entre bâtiments.

Afin de qualifier objectivement les performances de la méthode développée, il conviendrait de la mettre en œuvre sur un ensemble d'images présentant des géométries de prise de vue variées puis de comparer ces résultats à ceux obtenus au moyen d'autres méthodes proposées dans la littérature.

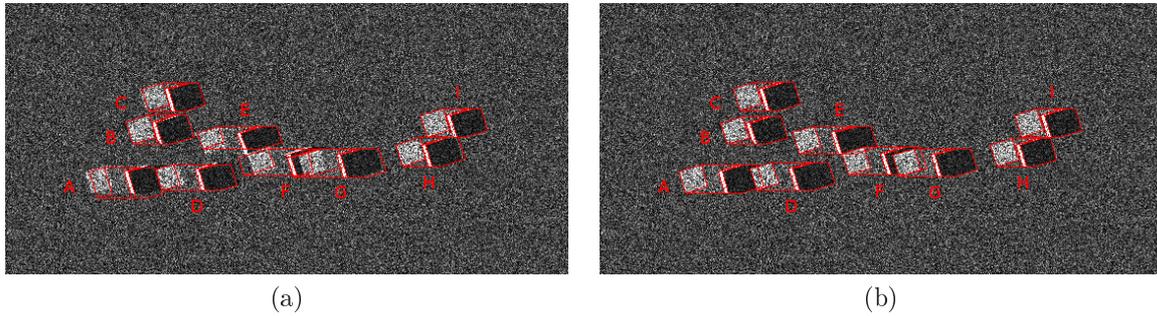


Figure 4.14 – Résultats de la stratégie d'estimation sur données simulées en imagerie RSO : signatures géométriques des neuf bâtiments de la scène d'étude simulées à partir des valeurs estimées (a) et des valeurs vraies (b) des paramètres de ces bâtiments.

Label		X_C (m)	Y_C (m)	α (degrés)	w_1 (m)	w_2 (m)	h (m)
A	Vérité terrain	702995.0	4795064.0	83.0	20.0	20.0	22.0
	Estimation	702994.4	4795066.7	83.2	13.6	24.5	19.8
B	Vérité terrain	703025.0	4795026.0	83.0	20.0	20.0	15.0
	Estimation	703025.5	4795025.5	82.8	21.1	20.9	14.6
C	Vérité terrain	703042.0	4794999.0	83.0	20.0	20.0	15.0
	Estimation	703041.8	4794999.3	83.1	19.4	20.6	13.7
D	Vérité terrain	703053.0	4795074.0	83.0	20.0	20.0	22.0
	Estimation	703051.5	4795074.5	83.8	16.5	19.0	22.2
E	Vérité terrain	703088.5	4795050.0	83.0	20.0	20.0	22.0
	Estimation	703088.1	4795047.7	81.8	19.5	14.9	22.5
F	Vérité terrain	703120.0	4795076.0	83.0	20.0	20.0	22.0
	Estimation	703133.39	4795077.7	83.4	27.4	16.5	27.5
G	Vérité terrain	703159.0	4795086.0	83.0	20.0	20.0	22.0
	Estimation	703158.3	4795086.4	83.7	18.5	20.6	28.4
H	Vérité terrain	703236.0	4795092.0	83.0	20.0	20.0	15.0
	Estimation	703237.5	4795092.0	82.3	22.2	19.9	15.0
I	Vérité terrain	703264.0	4795072.0	83.0	20.0	20.0	15.0
	Estimation	703264.3	4795072.0	82.9	20.3	19.9	15.0

Tableau 4.5 – Résultats de la stratégie d'estimation sur données simulées en imagerie RSO : valeurs estimées et valeurs vraies des paramètres des neuf bâtiments de la scène d'étude.

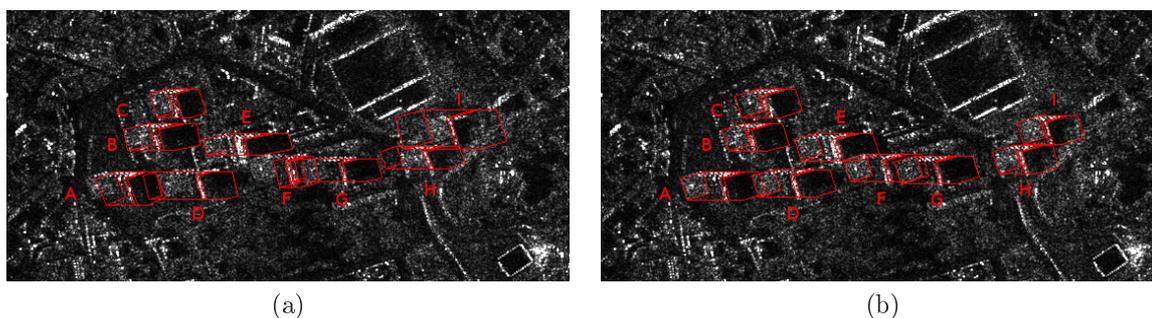


Figure 4.15 – Résultats de la stratégie d'estimation sur données réelles en imagerie RSO : signatures géométriques des neuf bâtiments de la scène d'étude simulées à partir des valeurs estimées (a) et des valeurs vraies (b) des paramètres de ces bâtiments.

Label		X_C (m)	Y_C (m)	α (degrés)	w_1 (m)	w_2 (m)	h (m)
A	Vérité terrain	702995.0	4795064.0	83.0	20.0	20.0	22.0
	Estimation	702993.9	4795065.6	77.8	17.6	24.8	15.9
B	Vérité terrain	703025.0	4795026.0	83.0	20.0	20.0	15.0
	Estimation	703027.9	4795026.2	88.1	23.6	19.4	18.3
C	Vérité terrain	703042.0	4794999.0	83.0	20.0	20.0	15.0
	Estimation	703044.3	4794999.3	82.8	14.4	22.2	12.9
D	Vérité terrain	703053.0	4795074.0	83.0	20.0	20.0	22.0
	Estimation	703049.0	4795075.1	85.7	11.5	22.0	29.8
E	Vérité terrain	703088.5	4795050.0	83.0	20.0	20.0	22.0
	Estimation	703102.9	4795048.4	88.2	26.9	14.4	21.8
F	Vérité terrain	703120.0	4795076.0	83.0	20.0	20.0	22.0
	Estimation	703127.2	4795079.0	88.3	18.1	24.1	7.7
G	Vérité terrain	703159.0	4795086.0	83.0	20.0	20.0	22.0
	Estimation	703156.3	4795085.8	85.3	14.2	17.1	30.0
H	Vérité terrain	703236.0	4795092.0	83.0	20.0	20.0	15.0
	Estimation	703233.8	4795089.6	78.9	15.4	16.0	24.3
I	Vérité terrain	703264.0	4795072.0	83.0	20.0	20.0	15.0
	Estimation	703267.6	4795072.3	84.3	25.6	29.5	28.9

Tableau 4.6 – Résultats de la stratégie d'estimation sur données réelles en imagerie RSO : valeurs estimées et valeurs vraies des paramètres des neuf bâtiments de la scène d'étude.

Contrairement à la plupart des méthodes présentées au chapitre 3, la méthode développée dans ce chapitre s'appuie directement sur la radiométrie des images sans qu'une extraction des primitives caractéristiques de la présence des bâtiments ne soit réalisée. L'extraction de primitives constitue souvent une étape très paramétrée, c'est-à-dire qu'elle nécessite que soient fixés, le plus souvent arbitrairement, un certain nombre de paramètres. En revanche, l'extraction de primitives permet de faciliter la détection des bâtiments dans les images. Ainsi, l'approche proposée présente un nombre restreint de paramètres, ce qui constitue un avantage, mais sa mise en œuvre à des fins de détection s'avère délicate.

Comme nous l'avons précédemment évoqué, la méthode proposée présente une faible robustesse aux phénomènes d'occlusion entre les bâtiments, ce qui se traduit par une dégradation des performances d'extraction dans le cas des prises de vue obliques où ces phénomènes sont très prononcés. La prise en compte de ces phénomènes d'occlusion, lors de la simulation de la signature géométrique des bâtiments, pourrait permettre de solutionner ce problème et constitue une première amélioration à apporter à la méthode proposée.

La seconde limitation de cette méthode est liée au modèle relativement simple adopté pour décrire les bâtiments. Ce modèle restreint en effet l'utilisation de la méthode à l'extraction de bâtiments parallélépipédiques. Les deux critères, sur lesquels repose la méthode proposée pour quantifier l'adéquation de la signature géométrique d'un bâtiment avec une image, peuvent être définis de manière similaire pour des modèles paramétriques de bâtiment plus complexes. Ainsi, une extension de la méthode pour l'extraction de bâtiments en « U », en « L » ou en « T » est tout à fait envisageable.

Enfin, la troisième limitation de la méthode est relative à la complexité calculatoire qu'elle implique, notamment en imagerie RSO, et qui restreint son utilisation à l'estimation des paramètres de bâtiments préalablement détectés. Il conviendrait d'étudier le gain en complexité calculatoire que pourraient apporter d'autres méthodes, et notamment les méthodes stochastiques, pour l'optimisation des deux critères sur lesquels est fondée la méthode.

Approches géométriques pour l'extraction de bâtiments

Sommaire

5.1	Introduction	108
5.2	Extraction de primitives	110
5.2.1	Imagerie optique	110
5.2.2	Imagerie RSO	112
5.3	Détection grossière des bâtiments	112
5.3.1	Calcul des densités de probabilité de présence des pixels détectés et non détectés dans les images binaires de primitives en présence d'un bâtiment	113
5.3.2	Test de significativité pour la détection des bâtiments	122
5.3.3	Détection grossière à partir des primitives considérées conjointement	129
5.4	Estimation des paramètres d'un bâtiment	135
5.4.1	Positionnement du problème	136
5.4.2	Estimation par maximisation de vraisemblance	137
5.4.3	Stratégie d'estimation	137
5.5	Détection fine des bâtiments	152
5.5.1	Calcul de la vraisemblance généralisée des pixels des images binaires de primitives	153
5.5.2	Stratégie de détection des bâtiments à partir de l'image de vraisemblance généralisée	154
5.5.3	Résultats en imagerie optique	154
5.5.4	Résultats en imagerie RSO	155
5.6	Conclusion	158

Dans ce chapitre, nous proposons une seconde méthode pour l'extraction tridimensionnelle des bâtiments parallélépipédiques en monoscopie optique et RSO. Contrairement à la méthode établie au chapitre précédent, qui réalise l'extraction des bâtiments directement à partir des images, les bâtiments sont extraits à partir d'images binaires, résultant de l'extraction des primitives caractéristiques de la présence des bâtiments dans ces images. L'approche proposée s'appuie sur les méthodes développées au chapitre 2 pour simuler les primitives observées dans ces images binaires, puis établir la densité de probabilité de présence des pixels détectés dans ces images binaires. Cette densité de probabilité permet alors de calculer une vraisemblance géométrique au moyen de laquelle est réalisée une détection-estimation conjointe des bâtiments.

5.1 Introduction

Nous présentons, dans ce chapitre, une méthode pour l'extraction tridimensionnelle des bâtiments parallélépipédiques, en monoscopie optique et RSO, à partir des primitives caractéristiques de la présence de ces bâtiments dans les images. Contrairement à la méthode proposée au chapitre précédent, qui repose sur des considérations radiométriques, la méthode développée s'appuie sur la forme des primitives extraites, de sorte que seules des considérations géométriques interviennent. Notons toutefois que la radiométrie de l'image intervient lors de l'extraction des primitives.

Comme au chapitre précédent, nous supposons que les bâtiments sont de forme parallélépipédique et peuvent être décrits au moyen du modèle introduit à la partie 2.2. Ainsi, la méthode développée correspond à une approche *model-based*, conformément à la classification des méthodes d'extraction introduite au chapitre 3. Cette méthode se décompose, comme illustré à la figure 5.1, en quatre étapes :

1. La première étape consiste à extraire les primitives caractéristiques de la présence de bâtiments dans l'image. Nous considérons ainsi les contours de toit et les ombres en imagerie optique, tandis que les échos doubles et les ombres RADAR sont extraits en imagerie RSO. À l'issue de cette première étape, nous disposons d'images binaires au sein desquelles les pixels sont labellisés selon leur appartenance ou non à une primitive donnée. La détection des bâtiments ainsi que l'estimation des paramètres des bâtiments détectés sont alors réalisées directement à partir de ces images binaires ;
2. Lors de la deuxième étape, une détection grossière des bâtiments est réalisée, au moyen d'un test statistique, à partir des images binaires précédemment établies. Pour ce faire, la densité de probabilité de présence des pixels détectés dans les images binaires est tout d'abord établie en simulant, au moyen des méthodes proposées au chapitre 2, les primitives observées dans ces images binaires. La vraisemblance des pixels des images binaires observés dans un voisinage du pixel testé est alors calculée, à l'aide de la densité de probabilité précédente, en supposant que le pixel testé correspond au centre de l'emprise au sol d'un bâtiment. Le seuillage de la carte de vraisemblance ainsi obtenue conduit à une image binaire au sein de laquelle sont identifiés les pixels qui correspondent possiblement au centre de l'emprise au sol d'un bâtiment ;
3. Dans la troisième étape, nous supposons que les pixels identifiés à l'étape précédente correspondent effectivement aux centres de l'emprise au sol de bâtiments. L'orientation (paramètre α) et les dimensions (paramètres w_1 , w_2 et h) de ces bâtiments sont alors estimées par maximisation de vraisemblance à partir des images binaires résultant de l'extraction des primitives. Afin de limiter la complexité calculatoire de cette étape, nous proposons une stratégie d'estimation qui repose sur l'estimation successive, dans des espaces à une dimension, des quatre paramètres précédents, permettant ainsi d'éviter l'estimation conjointe de ces paramètres dans un espace à quatre dimensions ;
4. La dernière étape correspond à une détection fine des bâtiments, qui est réalisée, comme à la seconde étape, au moyen d'un test statistique. La vraisemblance des pixels des images binaires, observés dans un voisinage du pixel testé, est désormais calculée en faisant intervenir

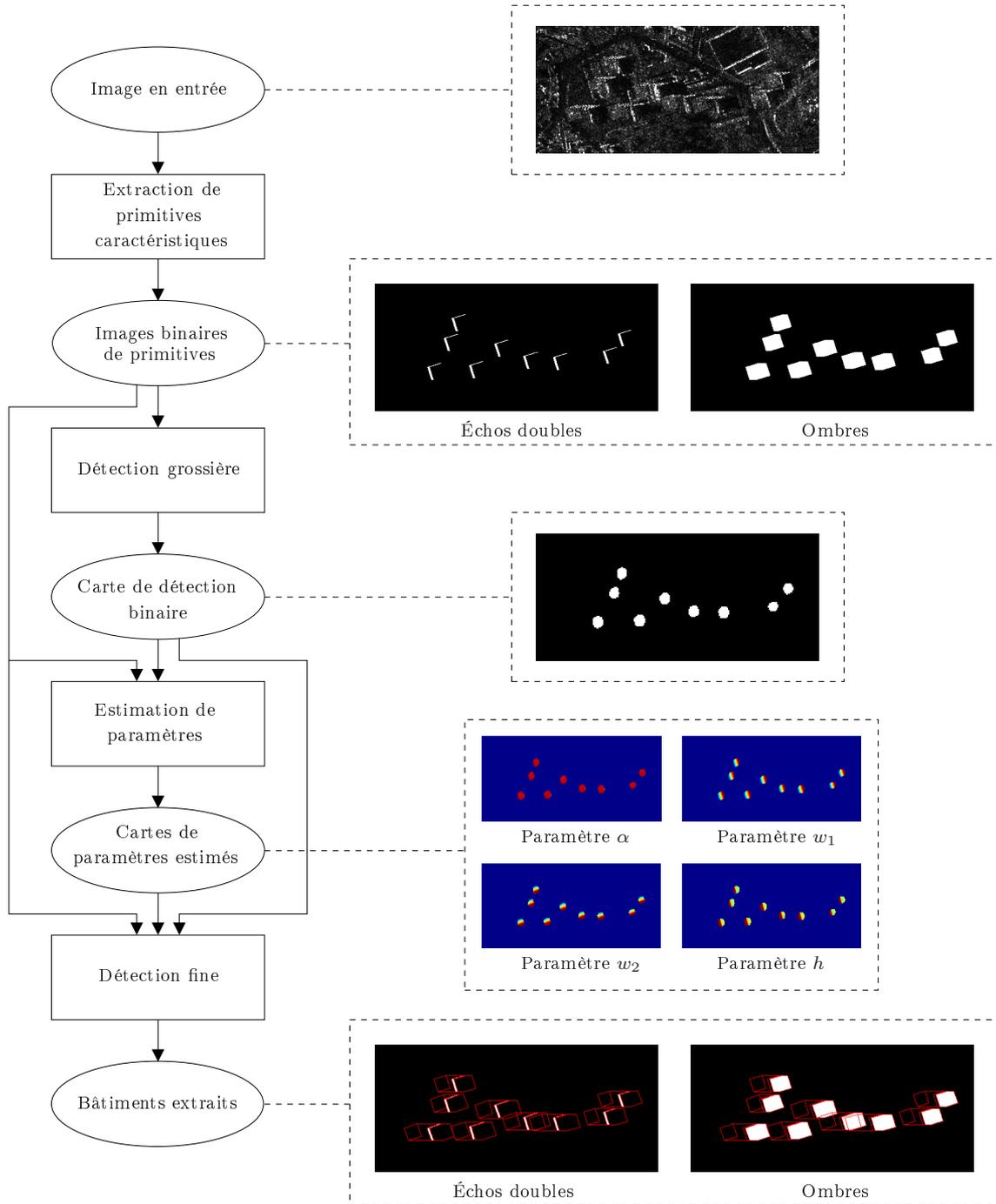


Figure 5.1 – Schéma synoptique de la méthode d'extraction proposée, illustré dans le cas de l'imagerie RSO. Les flèches représentent le flux des données, représentées par des ellipses, entre les différents modules de traitement, représentés par des rectangles.



Figure 5.2 – Image QuickBird sur la région de Marseille, France (©DigitalGlobe).

les paramètres estimés lors de l'étape précédente, ce qui correspond à une vraisemblance généralisée de ces pixels. Une stratégie itérative est alors mise en œuvre à partir de la carte de vraisemblance généralisée obtenue pour extraire les bâtiments les plus vraisemblables. Cette stratégie permet notamment d'éviter les détections multiples d'un même bâtiment.

Les quatre étapes de cette chaîne d'extraction constituent l'objet des parties 5.2, 5.3, 5.4 et 5.5, où elles sont respectivement décrites et illustrées sur données réelles et sur données simulées. Notons également que plusieurs aspects, relatifs à l'implémentation de ces différentes étapes, sont abordés afin de garantir une complexité calculatoire compatible avec une mise en œuvre opérationnelle de la méthode proposée.

5.2 Extraction de primitives

Dans cette première partie, nous décrivons les méthodes mises en œuvre pour extraire les primitives caractéristiques de la présence de bâtiments dans les images optiques et RSO. Comme nous l'avons vu aux sections 3.2.1 et 3.3.1, l'extraction de primitives en imagerie de télédétection constitue une problématique en soi. L'étude de cette problématique sort du cadre de nos travaux, de sorte que nous mettons en œuvre des méthodes relativement simples qui reposent essentiellement sur des approches par seuillage radiométrique. Ces méthodes ont l'avantage d'être peu paramétrées mais présentent des performances limitées. Si les performances de ces méthodes d'extraction permettent d'atteindre, comme nous le verrons par la suite, des résultats d'extraction de bâtiments satisfaisants, il est certain que la mise en œuvre d'approches plus complexes, présentant de meilleures performances, ne peut être que profitable pour les performances globales de la chaîne d'extraction.

5.2.1 Imagerie optique

En imagerie optique, les contours de toit sont extraits au moyen d'un seuillage supervisé de l'image d'amplitude du vecteur gradient, image calculée à partir de l'image initiale I . Ce seuillage, qui fait intervenir un seuil τ_T , conduit à une image binaire, notée B_I^0 , au sein de laquelle les pixels sont identifiés comme appartenant ou non à un contour de toit.

Nous notons L_I^0 l'image binaire correspondant à la vérité terrain relative à l'appartenance des pixels de l'image I à un contour de toit. Nous désignons par B_k et L_k les variables aléatoires discrètes

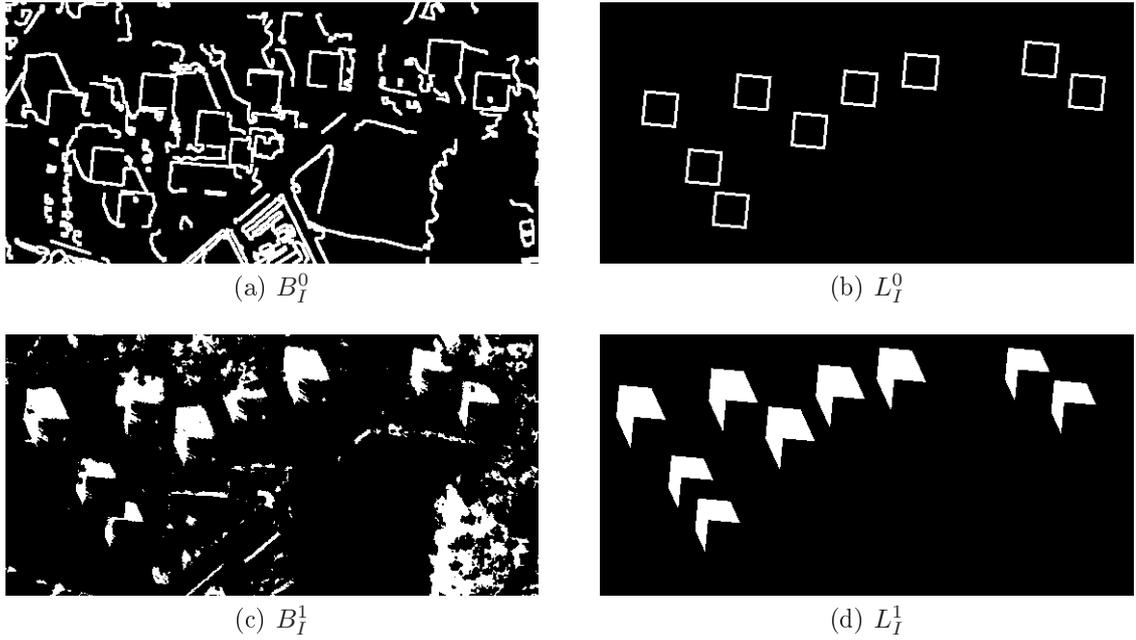


Figure 5.3 – Extraction de primitives en imagerie optique. Les images binaires B_I^0 (a) et B_I^1 (c) résultent respectivement de l'extraction des contours de toit et des ombres à partir de l'image de la figure 5.2. Les images binaires L_I^0 (b) et L_I^1 (d) correspondent aux vérités terrain associées aux images B_I^0 et B_I^1 .

décrivant respectivement les valeurs du pixel k dans les images B_I^0 et L_I^0 , variables aléatoires qui prennent leurs valeurs dans l'ensemble $\{0; 1\}$. Les probabilités de détection p_D^0 et de fausse alarme p_{FA}^0 du détecteur précédemment introduit, probabilités qui caractérisent les performances de ce détecteur, sont alors définies comme suit :

$$p_D^0 = P(B_k = 1 | L_k = 1), \quad (5.1)$$

$$p_{FA}^0 = P(B_k = 1 | L_k = 0). \quad (5.2)$$

La mise en œuvre de cette stratégie d'extraction sur l'image I de la figure 5.2 conduit à l'image binaire B_I^0 présentée à la figure 5.3 (a). Cette image est à comparer à la vérité terrain L_I^0 de la figure 5.3 (b), qui résulte de la simulation des contours de toit des bâtiments définis au tableau 2.1 au moyen de la méthode décrite à la section 2.3.3. Notons que les contours de toit issus de cette méthode sont définis par un ensemble de segments que nous dilatons d'une largeur ϵ_T , fixée empiriquement, de manière à obtenir des contours d'épaisseur non nulle, comparable à l'épaisseur des contours obtenus sur données réelles.

Une estimation des probabilités de détection et de fausse alarme du détecteur peut être réalisée à partir des images B_I^0 et L_I^0 , présentées à la figure 5.3, et conduit aux valeurs suivantes : $p_D^0 = 0.63$ et $p_{FA}^0 = 0.17$.

En ce qui concerne les ombres projetées par les bâtiments, elles sont extraites au moyen d'un seuillage supervisé de l'image I faisant intervenir un seuil noté τ_O .

La mise en œuvre de cette stratégie sur l'image de la figure 5.2 conduit à l'image binaire B_I^1 de la figure 5.3 (c), au sein de laquelle les pixels sont identifiés comme appartenant ou non à l'ombre projetée par un bâtiment. La comparaison de cette image binaire avec la vérité terrain L_I^1 de la figure 5.3 (d), établie au moyen de la méthode de la section 2.3.3, fournit alors une estimation des probabilités de détection et de fausse alarme du détecteur des pixels d'ombre : $p_D^1 = 0.73$ et $p_{FA}^1 = 0.10$.

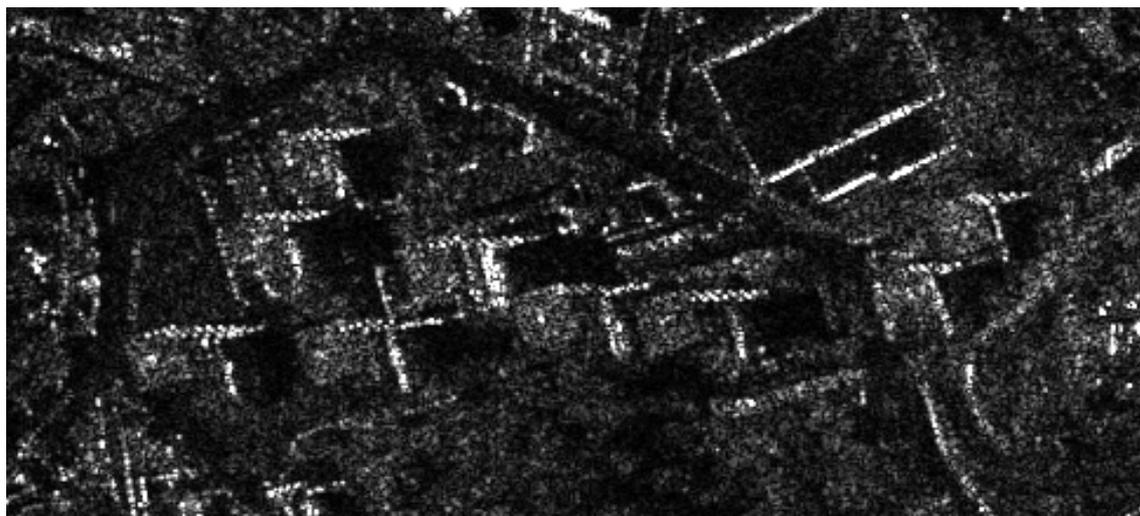


Figure 5.4 – Image TerraSAR-X sur la région de Marseille, France (©Astrium Geo-Information Services).

5.2.2 Imagerie RSO

En imagerie RSO, les échos doubles et les ombres RADAR sont extraits au moyen de seuillages radiométriques impliquant respectivement les seuils τ_{ED} et τ_O , seuillages qui sont réalisés après application d'un filtre anti-speckle visant à débruiter l'image.

Le filtre anti-speckle mis en œuvre est celui développé par Deledalle *et al.* [DDT09]. Ce filtre correspond à une extension du filtre des moyennes non locales, proposé par Buades *et al.* [BCM05], au cas d'un bruit multiplicatif. Le filtre de Deledalle *et al.* présente l'avantage de préserver les contours de l'image, contrairement à certains filtres anti-speckle comme ceux proposés par Lee *et al.* [Lee80] ou Kuan *et al.* [KSSC85] qui ont le défaut de lisser les contours conduisant ainsi à un biais sur la localisation de ces contours. La préservation des contours lors du filtrage anti-speckle est essentielle dans le cadre de l'extraction de bâtiments, afin d'éviter tout biais qui pourrait affecter l'estimation des paramètres des bâtiments.

Cette stratégie d'extraction, appliquée à l'image TerraSAR-X de la figure 5.4, conduit aux images binaires B_I^0 et B_I^1 des figures 5.5 (a) et 5.5 (c). Ces images résultent respectivement de l'extraction des échos doubles et des ombres RADAR. Les images binaires L_I^0 et L_I^1 , présentées aux figures 5.5 (b) et 5.5 (d), ont été simulées à l'aide de la méthode décrite à la section 2.4.3. Ces images de vérité terrain permettent d'estimer les probabilités de détection et de fausse alarme relatives au détecteur des pixels d'échos doubles : $p_D^0 = 0.51$, $p_{FA}^0 = 0.04$, ainsi que celles relatives au détecteur des pixels d'ombre : $p_D^1 = 0.55$ et $p_{FA}^1 = 0.23$.

5.3 Détection grossière des bâtiments

Dans cette partie, nous supposons disposer d'un ensemble d'images binaires, notées B_I^j , qui résultent respectivement de l'extraction des primitives \mathcal{P}^j caractéristiques de la présence de bâtiments dans l'image I , cette image pouvant indifféremment être une image optique ou RSO. À chaque image B_I^j sont associées les probabilités p_D^j et p_{FA}^j qui correspondent aux probabilités de détection et de fausse alarme caractéristiques du détecteur dont est issue l'image B_I^j .

En pratique, les probabilités de détection et de fausse alarme d'un détecteur de primitive peuvent être établies théoriquement, pour un modèle et une puissance de bruit donnés, ou bien empiriquement, de manière similaire à ce qui a été fait à la partie précédente. La détermination théorique de ces probabilités conduit le plus souvent à une sous-estimation de la probabilité de fausse alarme et une

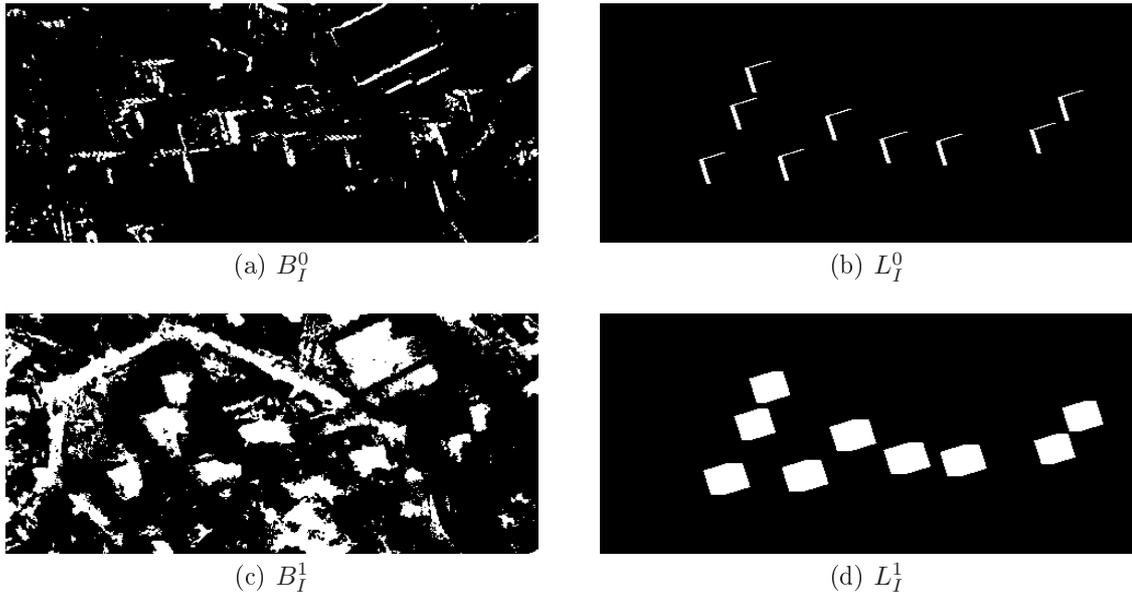


Figure 5.5 – Extraction de primitives en imagerie RSO. Les images binaires B_I^0 (a) et B_I^1 (c) résultent respectivement de l'extraction des échos doubles et des ombres à partir de l'image de la figure 5.4. Les images binaires L_I^0 (b) et L_I^1 (d) correspondent aux vérités terrain associées aux images B_I^0 et B_I^1 .

sur-estimation de la probabilité de détection en raison de l'inadéquation du modèle de bruit avec le bruit présent dans l'image et de la présence dans la scène de structures que nous ne souhaitons pas extraire. Ainsi, la détermination empirique de ces probabilités sur une base d'apprentissage demeure souvent la solution la plus réaliste bien que plus fastidieuse.

L'objectif de cette partie est d'établir une carte de détection grossière des bâtiments à partir des images binaires de primitives B_I^j . Cette carte de détection doit alors permettre de décider, en chaque pixel de l'image I , entre les hypothèses \mathcal{H}_1 et \mathcal{H}_0 qui correspondent respectivement à la présence et à l'absence d'un bâtiment.

En exploitant les méthodes de simulation de la signature géométrique des bâtiments dans une image, nous pouvons déterminer les densités de probabilité de présence des pixels détectés et non détectés dans les images binaires de primitives en présence d'un bâtiment. Nous sommes ainsi en mesure de modéliser statistiquement l'hypothèse \mathcal{H}_1 . L'hypothèse \mathcal{H}_0 n'est, en revanche, pas modélisable, de sorte que le calcul de la carte de détection repose uniquement sur le calcul de la vraisemblance de l'hypothèse \mathcal{H}_1 .

Le seuillage de la carte de détection ainsi établie, qui correspond à un test de significativité statistique, permet alors de définir des régions d'intérêt auxquelles sont restreintes les étapes suivantes de la chaîne d'extraction.

5.3.1 Calcul des densités de probabilité de présence des pixels détectés et non détectés dans les images binaires de primitives en présence d'un bâtiment

En se basant sur le modèle parallélépipédique de bâtiment introduit à la partie 2.2, un bâtiment \vec{b} est repéré dans l'espace terrain \mathcal{E}_T par les coordonnées planimétriques (X_C, Y_C) du centre C de l'emprise au sol de ce bâtiment dans un repère terrain \mathcal{R}_T lié à la Terre. Pour des raisons que nous exposons ultérieurement, le point c , qui résulte de la projection du point C de l'espace terrain vers l'espace image, ne constitue pas toujours le point le plus approprié pour repérer le bâtiment dans l'espace image. Ainsi, nous introduisons un point de référence, noté R dans l'espace terrain et r dans l'espace image, point qui nous permet de repérer un bâtiment \vec{b} dans l'image I .

En imagerie optique, ce point de référence est défini comme le centre du contour du toit d'un bâtiment, de sorte que les coordonnées (X_R, Y_R, Z_R) et (x_r, y_r) de ce point dans l'espace terrain \mathcal{E}_T et dans l'espace image \mathcal{E}_I s'expriment comme suit :

$$(X_R, Y_R, Z_R) = (X_C, Y_C, Z_C + h), \quad (5.3)$$

$$(x_r, y_r) = \mathcal{F}_I^i(X_R, Y_R, Z_R), \quad (5.4)$$

où \mathcal{F}_I^i désigne la fonction de localisation inverse associée à l'image I , h est la hauteur du bâtiment \vec{b} et Z_C correspond à l'altitude du centre de l'emprise au sol de ce bâtiment, altitude qui est fournie par un MNT.

En imagerie RSO, le point de référence est défini comme le centre de l'emprise au sol du bâtiment, de sorte que les points R et r sont confondus respectivement avec les points C et c .

Dans cette section, nous considérons qu'un bâtiment \vec{b} est présent dans l'image I et que les coordonnées (x_r, y_r) du point de référence associé à ce bâtiment, dans l'espace image, sont connues *a priori*.

Par la suite, d désigne un pixel quelconque (pris au hasard) de l'image binaire B_I^j identifié comme appartenant à une primitive \mathcal{P}^j . De même, \bar{d} désigne un pixel quelconque (pris au hasard) de l'image binaire B_I^j identifié comme n'appartenant pas à une primitive \mathcal{P}^j . Nous notons respectivement X_d^j , Y_d^j et $X_{\bar{d}}^j$, $Y_{\bar{d}}^j$ les variables aléatoires correspondant aux coordonnées, dans l'espace image \mathcal{E}_I , du pixel d et du pixel \bar{d} .

Nous établissons, dans cette partie, l'expression des densités de probabilité conjointe des variables aléatoires X_d^j , Y_d^j et $X_{\bar{d}}^j$, $Y_{\bar{d}}^j$, densités de probabilité qui sont notées respectivement p_d^j et $p_{\bar{d}}^j$.

5.3.1.1 Expression à paramètres de bâtiment connus

Nous considérons tout d'abord le cas où les paramètres α , w_1 , w_2 et h du bâtiment \vec{b} présent dans l'image sont connus *a priori*. Les paramètres X_C et Y_C pouvant être calculés à partir des coordonnées (x_r, y_r) du point de référence et de la hauteur h supposés connus, l'ensemble des paramètres du bâtiment \vec{b} sont connus. Ainsi, nous établissons l'expression des densités de probabilité conjointe des variables aléatoires X_d^j , Y_d^j et $X_{\bar{d}}^j$, $Y_{\bar{d}}^j$, conditionnellement au vecteur de paramètres \vec{b} .

Cas d'un détecteur parfait

Lorsque le détecteur permettant d'extraire la primitive \mathcal{P}^j est un détecteur parfait, c'est-à-dire lorsque les probabilités de détection p_D^j et de fausse alarme p_{FA}^j de ce détecteur valent respectivement 1 et 0, la densité de probabilité conjointe des variables aléatoires X_d^j et Y_d^j , conditionnellement au vecteur de paramètres \vec{b} , s'écrit comme suit :

$$\forall (x, y) \in \mathcal{E}_I, p_d^j(x, y | \vec{b}) = \frac{1_{\mathcal{P}_I^j(\vec{b})}(x, y)}{\mathcal{A}(\mathcal{P}_I^j(\vec{b}))}, \quad (5.5)$$

où $\mathcal{P}_I^j(\vec{b})$ correspond à la primitive \mathcal{P}^j résultant de la projection du bâtiment \vec{b} dans l'image I au moyen des algorithmes de simulation de la signature géométrique des bâtiments. $\mathcal{A}(R)$ correspond à l'aire d'une région R définie dans l'espace image \mathcal{E}_I et 1_R désigne la fonction indicatrice sur la région R .

La densité de probabilité de présence $p_d^j(x, y | \vec{b})$ des pixels détectés dans l'image binaire B_I^j correspond donc à une densité de probabilité uniforme sur la région $\mathcal{P}_I^j(\vec{b})$ de l'espace image \mathcal{E}_I . Ainsi, dans le cas d'un détecteur parfait et pour un bâtiment \vec{b} donné, la probabilité d'observer un pixel détecté dans l'image binaire B_I^j est constante non nulle à l'intérieur de la région $\mathcal{P}_I^j(\vec{b})$ et nulle en dehors de cette région.

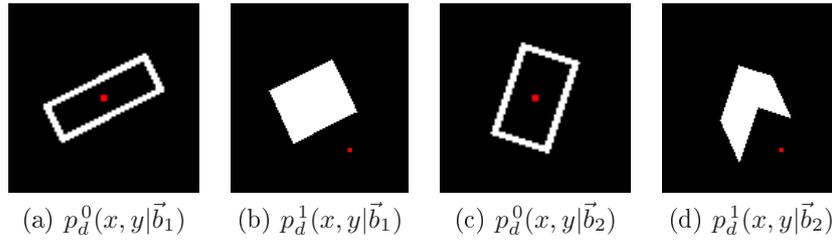


Figure 5.6 – Densités de probabilité de présence des pixels détectés dans les images binaires B_I^0 (a, c) et B_I^1 (b, d) en imagerie optique, conditionnellement aux vecteurs de paramètres $\vec{b}_1 = (X_{C1}, Y_{C1}, 25, 10, 30, 25)$ (a, b) et $\vec{b}_2 = (X_{C2}, Y_{C2}, 70, 15, 25, 20)$ (c, d).

Comme nous l'avons vu au chapitre 2, les effets géométriques liés à la projection des bâtiments de l'espace terrain vers l'espace image ne sont pas uniformes dans l'image. L'équation (5.5) est parfaitement rigoureuse, puisque la non uniformité de ces effets géométriques y est prise en compte par l'intermédiaire de la primitive $\mathcal{P}_I^j(\vec{b})$, qui résulte de la projection du bâtiment \vec{b} dans l'image I , projection qui tient compte de la position (X_C, Y_C) du bâtiment.

Afin de limiter la complexité calculatoire de la méthode proposée, nous supposons dans la suite que les effets géométriques liés à la projection des bâtiments de l'espace terrain vers l'espace image peuvent être considérés comme uniformes dans toute l'image. Cette hypothèse, dont la validité est souvent vérifiée en imagerie satellitaire mais possiblement mise en défaut en imagerie aéroportée, consiste à dire que deux bâtiments ayant la même orientation (α) et les mêmes dimensions (w_1, w_2, h) mais des coordonnées planimétriques (X_C, Y_C) distinctes ont des signatures géométriques identiques dans l'image.

Dans le cadre de cette hypothèse, la densité de probabilité $p_d^j(x, y | \vec{b}')$ calculée pour un bâtiment $\vec{b}' = (X'_C, Y'_C, \alpha, w_1, w_2, h)$ se déduit de la densité de probabilité $p_d^j(x, y | \vec{b})$ calculée pour un bâtiment $\vec{b} = (X_C, Y_C, \alpha, w_1, w_2, h)$ par translation dans l'image I d'un vecteur rr' où r et r' désignent respectivement les points de référence des bâtiments \vec{b} et \vec{b}' . En pratique, les densités de probabilité de présence calculées dans la suite de ce chapitre le seront pour un bâtiment dont le point de référence est situé en (x_0, y_0) , coordonnées du centre de l'image I .

La figure 5.6 présente des exemples de densités de probabilité de présence des pixels détectés dans les images binaires B_I^0 et B_I^1 , résultant respectivement de l'extraction des contours de toit et des ombres en imagerie optique, conditionnellement aux vecteurs de paramètres $\vec{b}_1 = (X_{C1}, Y_{C1}, 25, 10, 30, 25)$ et $\vec{b}_2 = (X_{C2}, Y_{C2}, 70, 15, 25, 20)$, où l'orientation α est exprimée en degrés et les dimensions w_1, w_2, h sont exprimées en mètres. Le point de couleur rouge désigne le point de référence du bâtiment, qui correspond, en imagerie optique, au centre du contour du toit de ce bâtiment. Ces densités de probabilité de présence ont été générées pour une géométrie de prise de vue correspondant à celle de l'image de la figure 5.2.

De même, la figure 5.7 présente des exemples de densités de probabilité de présence des pixels détectés dans les images binaires B_I^0 et B_I^1 , résultant de l'extraction des échos doubles et des ombres en imagerie RSO, conditionnellement aux mêmes vecteurs de paramètres que précédemment. Le point de couleur rouge désigne le point de référence du bâtiment, qui correspond, en imagerie RSO, au centre de l'emprise au sol de ce bâtiment. Ces densités de probabilité de présence ont été générées pour une géométrie de prise de vue correspondant à celle de l'image de la figure 5.4.

Après avoir établi l'expression de la densité de probabilité $p_d^j(x, y | \vec{b})$, nous établissons maintenant l'expression de la densité de probabilité conjointe des variables aléatoires X_d^j, Y_d^j , conditionnellement au vecteur de paramètres \vec{b} , expression qui fait également intervenir la primitive $\mathcal{P}_I^j(\vec{b})$ résultant de la

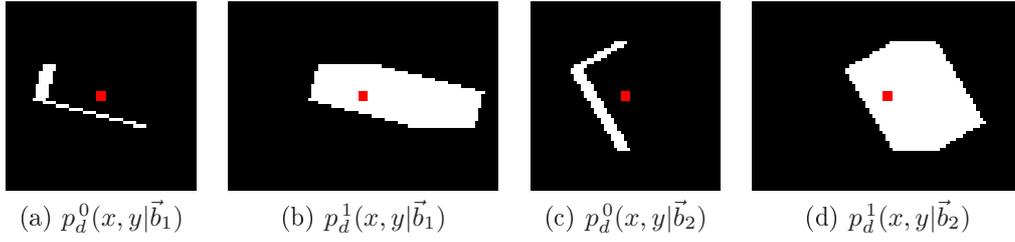


Figure 5.7 – Densités de probabilité de présence des pixels détectés dans les images binaires B_I^0 (a, c) et B_I^1 (b, d) en imagerie RSO, conditionnellement aux vecteurs de paramètres $\vec{b}_1 = (X_{C1}, Y_{C1}, 25, 10, 30, 25)$ (a, b) et $\vec{b}_2 = (X_{C2}, Y_{C2}, 70, 15, 25, 20)$ (c, d).

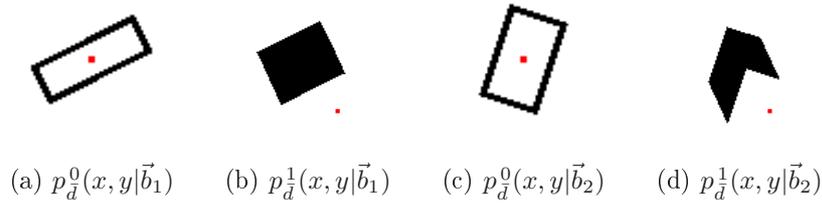


Figure 5.8 – Densités de probabilité de présence des pixels non détectés dans les images binaires B_I^0 (a, c) et B_I^1 (b, d) en imagerie optique, conditionnellement aux vecteurs de paramètres $\vec{b}_1 = (X_{C1}, Y_{C1}, 25, 10, 30, 25)$ (a, b) et $\vec{b}_2 = (X_{C2}, Y_{C2}, 70, 15, 25, 20)$ (c, d).

projection du bâtiment \vec{b} dans l'image I :

$$\forall (x, y) \in \mathcal{E}_I, p_d^j(x, y | \vec{b}) = \frac{1_{\mathcal{E}_I \setminus \mathcal{P}_I^j(\vec{b})}(x, y)}{\mathcal{A}(\mathcal{E}_I \setminus \mathcal{P}_I^j(\vec{b}))}. \quad (5.6)$$

Ainsi, pour un détecteur de primitive parfait, la probabilité d'observer des pixels non détectés dans l'image binaire B_I^j est uniforme sur toute l'image I hormis à l'intérieur de la région $\mathcal{P}_I^j(\vec{b})$ où cette probabilité est nulle.

Des exemples de densités de probabilité conjointe des variables aléatoires X_d^j, Y_d^j sont présentés à la figure 5.8 dans le cas de l'imagerie optique, ainsi qu'à la figure 5.9 dans le cas de l'imagerie RSO.

Par la suite, nous supposons que les dimensions maximales w_1^{max}, w_2^{max} et h^{max} des bâtiments présents dans l'image I sont connues *a priori*, de sorte qu'il est inutile de définir les densités de probabilité $p_d^j(x, y | \vec{b})$ et $p_d^j(x, y | \vec{b})$ sur l'intégralité de l'ensemble \mathcal{E}_I . Nous limitons donc la définition de ces

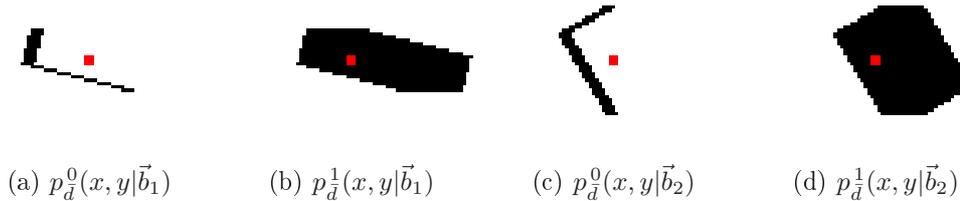


Figure 5.9 – Densités de probabilité de présence des pixels non détectés dans les images binaires B_I^0 (a, c) et B_I^1 (b, d) en imagerie RSO, conditionnellement aux vecteurs de paramètres $\vec{b}_1 = (X_{C1}, Y_{C1}, 25, 10, 30, 25)$ (a, b) et $\vec{b}_2 = (X_{C2}, Y_{C2}, 70, 15, 25, 20)$ (c, d).

densités de probabilité sur un voisinage ν_r^j du pixel de référence r , les dimensions de ce voisinage étant directement reliées aux dimensions maximales de la région $\mathcal{P}_I^j(\vec{b})$ dans l'image I . Les équations (5.5) et (5.6) se réécrivent alors :

$$\forall (x, y) \in \nu_r^j, p_d^j(x, y | \vec{b}) = \frac{1_{\mathcal{P}_I^j(\vec{b})}(x, y)}{\mathcal{A}(\mathcal{P}_I^j(\vec{b}))}, \quad (5.7)$$

$$\forall (x, y) \in \nu_r^j, p_{FA}^j(x, y | \vec{b}) = \frac{1_{\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})}(x, y)}{\mathcal{A}(\nu_r^j \setminus \mathcal{P}_I^j(\vec{b}))}. \quad (5.8)$$

Cas d'un détecteur réel

Nous avons précédemment établi les expressions des densités de probabilité conjointe des variables aléatoires X_d^j, Y_d^j et X_{FA}^j, Y_{FA}^j , conditionnellement au vecteur de paramètres \vec{b} , dans le cas d'un détecteur de primitive parfait, pour lequel $p_D^j = 1$ et $p_{FA}^j = 0$. Nous souhaitons, dans ce paragraphe, établir les expressions de ces densités de probabilité dans le cas d'un détecteur de primitive réel, pour lequel $p_D^j \neq 1$ et $p_{FA}^j \neq 0$.

Plaçons nous tout d'abord dans le cas hypothétique où nous disposons de N images du bâtiment \vec{b} acquises au même instant et par le même capteur, de sorte que nous disposons de N réalisations indépendantes du bruit présent dans l'image I de ce bâtiment. Nous désignons par k un pixel quelconque du voisinage ν_r^j et notons N_k le nombre de fois que ce pixel a été identifié, dans les N images binaires B_I^j associées aux N images I , comme appartenant à la primitive $\mathcal{P}_I^j(\vec{b})$.

La probabilité qu'un pixel détecté au sein du voisinage ν_r^j se trouve en (x_{k_0}, y_{k_0}) , coordonnées d'un pixel k_0 contenu dans le voisinage ν_r^j , est donnée par la relation suivante :

$$p_d^j(x = x_{k_0}, y = y_{k_0} | \vec{b}) = \lim_{N \rightarrow \infty} \frac{N_{k_0}}{\sum_{k \in \nu_r^j} N_k} = \lim_{N \rightarrow \infty} \frac{N_{k_0}}{\sum_{k \in \mathcal{P}_I^j(\vec{b})} N_k + \sum_{k \in \nu_r^j \setminus \mathcal{P}_I^j(\vec{b})} N_k}. \quad (5.9)$$

Pour un pixel k appartenant effectivement à la primitive $\mathcal{P}_I^j(\vec{b})$, le nombre de fois N_k que ce pixel a été détecté dans l'image binaire B_I^j comme appartenant à la primitive $\mathcal{P}_I^j(\vec{b})$ est asymptotiquement relié au nombre N de réalisations du bruit au moyen de la probabilité de détection p_D^j du détecteur de primitive :

$$\lim_{N \rightarrow \infty} N_k = p_D^j N. \quad (5.10)$$

De même, le nombre de fois N_k qu'un pixel k , qui n'appartient pas à la primitive $\mathcal{P}_I^j(\vec{b})$, a été détecté dans l'image binaire B_I^j comme appartenant à la primitive $\mathcal{P}_I^j(\vec{b})$ est asymptotiquement relié au nombre N de réalisations du bruit au moyen de la probabilité de fausse alarme p_{FA}^j du détecteur de primitive :

$$\lim_{N \rightarrow \infty} N_k = p_{FA}^j N. \quad (5.11)$$

Pour un pixel k_0 qui appartient effectivement à la primitive $\mathcal{P}_I^j(\vec{b})$, l'équation (5.9) se réécrit alors :

$$p_d^j(x = x_{k_0}, y = y_{k_0} | \vec{b}) = \frac{p_D^j}{N_{\mathcal{P}_I^j(\vec{b})} p_D^j + N_{\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})} p_{FA}^j}, \quad (5.12)$$

où $N_{\mathcal{P}_I^j(\vec{b})}$ et $N_{\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})}$ désignent respectivement le nombre de pixels contenus dans les régions $\mathcal{P}_I^j(\vec{b})$ et $\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})$. De la même manière, pour un pixel k_0 qui n'appartient pas à la primitive $\mathcal{P}_I^j(\vec{b})$,

l'équation (5.9) se réécrit :

$$p_d^j(x = x_{k_0}, y = y_{k_0} | \vec{b}) = \frac{p_{FA}^j}{N_{\mathcal{P}_I^j(\vec{b})} p_D^j + N_{\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})} p_{FA}^j}. \quad (5.13)$$

Nous déduisons finalement que la densité de probabilité conjointe des variables aléatoires X_d^j et Y_d^j , conditionnellement au vecteur de paramètres \vec{b} , s'écrit comme suit :

$$\begin{aligned} \forall (x, y) \in \nu_r^j, p_d^j(x, y | \vec{b}) &= \frac{p_D^j}{\mathcal{A}(\mathcal{P}_I^j(\vec{b})) p_D^j + \mathcal{A}(\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})) p_{FA}^j} 1_{\mathcal{P}_I^j(\vec{b})}(x, y) \\ &+ \frac{p_{FA}^j}{\mathcal{A}(\mathcal{P}_I^j(\vec{b})) p_D^j + \mathcal{A}(\nu_0^j \setminus \mathcal{P}_I^j(\vec{b})) p_{FA}^j} 1_{\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})}(x, y). \end{aligned} \quad (5.14)$$

Cette densité de probabilité s'écrit comme la somme de deux termes, le premier étant relatif aux vraies détections à l'intérieur de la région $\mathcal{P}_I^j(\vec{b})$ et le second étant relatif à la présence de fausses alarmes en dehors de cette région. Notons enfin que nous retrouvons, pour $p_D^j = 1$ et $p_{FA}^j = 0$, l'expression de la densité de probabilité $p_d^j(x, y | \vec{b})$ que nous avons établie à l'équation (5.5) pour un détecteur de primitive parfait.

Un raisonnement analogue à celui mené précédemment permet d'établir l'expression de la densité de probabilité conjointe des variables aléatoires X_d^j, Y_d^j , conditionnellement au vecteur de paramètres \vec{b} :

$$\begin{aligned} \forall (x, y) \in \nu_r^j, p_d^j(x, y | \vec{b}) &= \frac{(1 - p_D^j)}{\mathcal{A}(\mathcal{P}_I^j(\vec{b})) (1 - p_D^j) + \mathcal{A}(\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})) (1 - p_{FA}^j)} 1_{\mathcal{P}_I^j(\vec{b})}(x, y) \\ &+ \frac{(1 - p_{FA}^j)}{\mathcal{A}(\mathcal{P}_I^j(\vec{b})) (1 - p_D^j) + \mathcal{A}(\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})) (1 - p_{FA}^j)} 1_{\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})}(x, y). \end{aligned} \quad (5.15)$$

5.3.1.2 Expression à position de bâtiment connue

Nous considérons désormais que les paramètres α, w_1, w_2 et h du bâtiment \vec{b} sont inconnus, de sorte que seules les coordonnées (x_r, y_r) du point de référence dans l'espace image sont connues *a priori*. Nous déterminons, dans cette sous-section, les expressions des densités de probabilité conjointe des variables aléatoires X_d^j, Y_d^j et X_r^j, Y_r^j , établies conditionnellement aux coordonnées (x_r, y_r) .

D'après le théorème des probabilités totales, les densités de probabilité $p_d^j(x, y | x_r, y_r)$ et $p_r^j(x, y | x_r, y_r, \alpha, w_1, w_2, h)$ sont reliées par la relation suivante :

$$\forall (x, y) \in \nu_r^j, p_d^j(x, y | x_r, y_r) = \int_{\mathcal{E}_\alpha} \int_{\mathcal{E}_{w_1}} \int_{\mathcal{E}_{w_2}} \int_{\mathcal{E}_h} p_d^j(x, y | x_r, y_r, \alpha, w_1, w_2, h) p(\alpha, w_1, w_2, h) d\alpha dw_1 dw_2 dh, \quad (5.16)$$

où $p(\alpha, w_1, w_2, h)$ désigne la probabilité *a priori* d'observer des bâtiments de paramètres (α, w_1, w_2, h) au sein de l'image I . Dans le cas où une information *a priori*, issue par exemple d'un SIG, est disponible sur un ou plusieurs de ces paramètres, elle peut être introduite au moyen de cette probabilité. Nous supposons ici n'avoir aucune information *a priori* sur l'orientation et les dimensions des bâtiments présents dans l'image. La probabilité *a priori* $p(\alpha, w_1, w_2, h)$ est alors uniforme sur le domaine d'intégration $\mathcal{E}_\alpha \times \mathcal{E}_{w_1} \times \mathcal{E}_{w_2} \times \mathcal{E}_h$.

De la même manière, la densité de probabilité $p_d^j(x, y|x_r, y_r)$ est reliée à la densité de probabilité $p_d^j(x, y|x_r, y_r, \alpha, w_1, w_2, h)$, par la relation suivante :

$$\forall(x, y) \in \nu_r^j, p_d^j(x, y|x_r, y_r) = \int_{\mathcal{E}_\alpha} \int_{\mathcal{E}_{w_1}} \int_{\mathcal{E}_{w_2}} \int_{\mathcal{E}_h} p_d^j(x, y|x_r, y_r, \alpha, w_1, w_2, h) p(\alpha, w_1, w_2, h) d\alpha dw_1 dw_2 dh. \quad (5.17)$$

Dans le cas particulier de l'imagerie RSO, la forme de l'écho double d'un bâtiment ainsi que la position relative de cet écho double par rapport au point de référence, défini comme le centre de l'emprise au sol de ce bâtiment, sont indépendantes de la hauteur du bâtiment. Les densités de probabilité de présence $p_d^0(x, y|x_r, y_r)$ et $p_d^0(x, y|x_r, y_r)$ des pixels détectés et non détectés dans l'image binaire B_I^0 , résultant de l'extraction des échos doubles, peuvent alors se réécrire :

$$\forall(x, y) \in \nu_r^0, p_d^0(x, y|x_r, y_r) = \int_{\mathcal{E}_\alpha} \int_{\mathcal{E}_{w_1}} \int_{\mathcal{E}_{w_2}} p_d^0(x, y|x_r, y_r, \alpha, w_1, w_2) p(\alpha, w_1, w_2) d\alpha dw_1 dw_2, \quad (5.18)$$

$$\forall(x, y) \in \nu_r^0, p_d^0(x, y|x_r, y_r) = \int_{\mathcal{E}_\alpha} \int_{\mathcal{E}_{w_1}} \int_{\mathcal{E}_{w_2}} p_d^0(x, y|x_r, y_r, \alpha, w_1, w_2) p(\alpha, w_1, w_2) d\alpha dw_1 dw_2, \quad (5.19)$$

où le conditionnement des probabilités $p_d^0(x, y|x_r, y_r, \alpha, w_1, w_2, h)$ et $p_d^0(x, y|x_r, y_r, \alpha, w_1, w_2, h)$ par rapport au paramètre h , ainsi que l'intégration par rapport à ce même paramètre, ont été supprimés.

De même, en imagerie optique, la forme du contour du toit d'un bâtiment ainsi que la position relative de ce contour par rapport au point de référence, défini comme le centre du contour du toit de ce bâtiment, sont indépendantes de la hauteur du bâtiment. Ainsi les densités de probabilité de présence $p_d^0(x, y|x_r, y_r)$ et $p_d^0(x, y|x_r, y_r)$ des pixels détectés et non détectés dans l'image binaire B_I^0 , résultant de l'extraction des contours de toit, peuvent se réécrire selon les expressions des équations (5.18) et (5.19). Si le point de référence d'un bâtiment avait été défini comme le centre de l'emprise au sol de ce bâtiment, la position relative du contour du toit de ce bâtiment par rapport au point de référence n'aurait pas été indépendante de la hauteur du bâtiment, en raison des effets de déplacements présentés à la sous-section 2.3.2.2. La suppression du conditionnement des densités de probabilité $p_d^0(x, y|x_r, y_r, \alpha, w_1, w_2, h)$ et $p_d^0(x, y|x_r, y_r, \alpha, w_1, w_2, h)$ par rapport à la hauteur h n'aurait alors pas été possible, ce qui justifie que le point de référence d'un bâtiment ait été défini, en imagerie optique, comme le centre du contour du toit de ce bâtiment.

Étant donné que nous ne disposons pas des expressions analytiques des supports sur lesquels sont définies les fonctions indicatrices intervenant dans les expressions de $p_d^j(x, y|x_r, y_r, \alpha, w_1, w_2, h)$ et de $p_d^j(x, y|x_r, y_r, \alpha, w_1, w_2)$, la densité de probabilité $p_d^j(x, y|x_r, y_r)$ ne peut être calculée que numériquement. L'intégration peut alors être réalisée par somme de Riemann ou bien par méthode de Monte-Carlo.

Ces deux méthodes, mises en œuvre pour le calcul de la densité de probabilité de présence $p_d^0(x, y|x_r, y_r)$ des pixels détectés dans l'image binaire B_I^0 résultant de l'extraction des contours de toit en imagerie optique, conduisent respectivement aux densités de probabilité des figures 5.10 (a) et 5.10 (b). Ces densités de probabilité ont été calculées sur le domaine d'intégration suivant : $\mathcal{E}_\alpha \times \mathcal{E}_{w_1} \times \mathcal{E}_{w_2} = [0; 90] \times [1; 34] \times [1; 34]$, où les angles sont exprimés en degrés et les longueurs sont exprimées en mètres. Le calcul de la densité de probabilité par somme de Riemann est réalisé avec des pas valant respectivement 1 degré, 1,65 m et 1,65 m dans les espaces \mathcal{E}_α , \mathcal{E}_{w_1} et \mathcal{E}_{w_2} , de sorte que le calcul de l'intégrale triple est réalisé sur un ensemble de 40131 points disposés régulièrement dans l'ensemble $\mathcal{E}_\alpha \times \mathcal{E}_{w_1} \times \mathcal{E}_{w_2}$. Le calcul de cette même densité de probabilité au moyen de la méthode

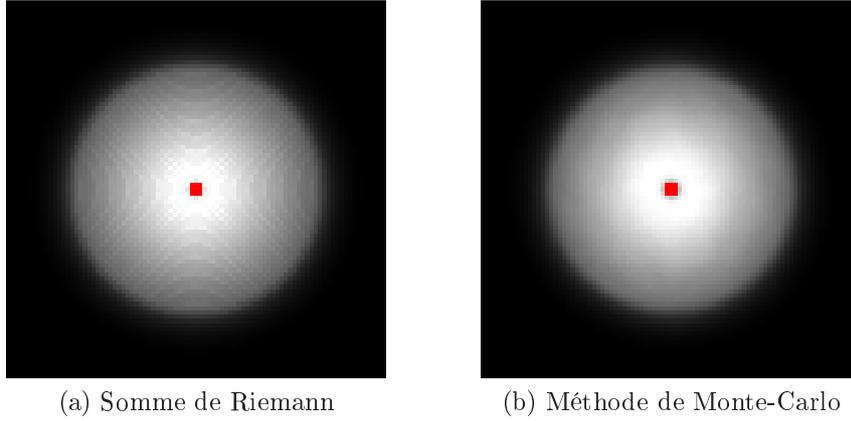


Figure 5.10 – Calcul de la densité de probabilité de présence $p_d^0(x, y|x_r, y_r)$ des pixels détectés dans l'image binaire B_I^0 , résultant de l'extraction des contours de toit en imagerie optique, par somme de Riemann (a) et par méthode de Monte-Carlo (b).

de Monte-Carlo est réalisé avec un tirage, selon une loi uniforme, de 40131 points au sein du domaine d'intégration $\mathcal{E}_\alpha \times \mathcal{E}_{w_1} \times \mathcal{E}_{w_2}$.

Nous pouvons constater à la figure 5.10 que l'intégration par somme de Riemann introduit des artefacts liés au parcours régulier du domaine d'intégration, artefacts qui ne sont pas observés dans le cas d'une intégration par méthode de Monte-Carlo. Lorsque les pas avec lesquels sont parcourus les ensembles \mathcal{E}_α , \mathcal{E}_{w_1} et \mathcal{E}_{w_2} diminuent, les artefacts apparaissant lors de la mise en œuvre de la méthode de Riemann disparaissent. Cependant, l'intégration étant réalisée sur un support tridimensionnel, la diminution des pas d'intégration engendre une explosion de la complexité calculatoire. Ainsi, l'ensemble des densités de probabilité calculées dans cette partie le seront au moyen d'une intégration par méthode de Monte-Carlo.

Nous présentons, à la figure 5.11, les densités de probabilité de présence des pixels détectés et non détectés dans les images binaires de primitives B_I^0 et B_I^1 , résultant respectivement de l'extraction des contours de toit et des ombres en imagerie optique. Ces densités de probabilité, calculées par méthode de Monte-Carlo avec 10^6 tirages, ont été établies pour la géométrie de prise de vue de l'image QuickBird de la figure 5.2.

Nous constatons que les densités de probabilité de présence $p_d^0(x, y|x_r, y_r)$ et $p_d^1(x, y|x_r, y_r)$ présentent logiquement une symétrie de révolution autour du point de référence, étant donné que ces densités de probabilité résultent de l'intégration sur toutes les orientations possibles et que le *spacing* de l'image QuickBird est quasiment le même selon les lignes et les colonnes de cette image. Notons également que la densité de probabilité de présence d'un pixel détecté est d'autant plus élevée que ce pixel est proche du point de référence, ce qui garantit que les bâtiments sont détectés avec un taux de détection similaire quelles que soient leurs dimensions.

En ce qui concerne les densités de probabilité de présence $p_d^1(x, y|x_r, y_r)$ et $p_d^2(x, y|x_r, y_r)$, la dissymétrie de ces densités de probabilité résulte de la prise en compte, par l'intermédiaire des algorithmes de projection de bâtiments, de la direction d'illumination du soleil.

À la figure 5.12, nous présentons les densités de probabilité de présence des pixels détectés et non détectés dans les images binaires B_I^0 et B_I^1 , résultant respectivement de l'extraction des échos doubles et des ombres en imagerie RSO. Ces densités de probabilité ont été calculées par méthode de Monte-Carlo avec 10^6 tirages pour la géométrie de prise de vue de l'image TerraSAR-X de la figure 5.4.

Nous constatons que les pixels détectés dans l'image B_I^0 sont majoritairement situés à gauche du pixel de référence tandis que les pixels détectés dans l'image B_I^1 sont majoritairement localisés à droite de ce point. Ceci est cohérent avec l'acquisition de l'image RSO en visée droite lors d'une

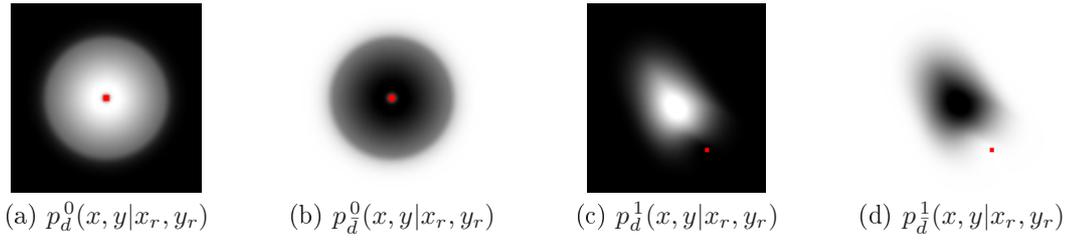


Figure 5.11 – Densités de probabilité de présence des pixels détectés (a, c) et non détectés (b, d) dans les images binaires de primitives B_I^0 (a, b) et B_I^1 (c, d) en imagerie optique.

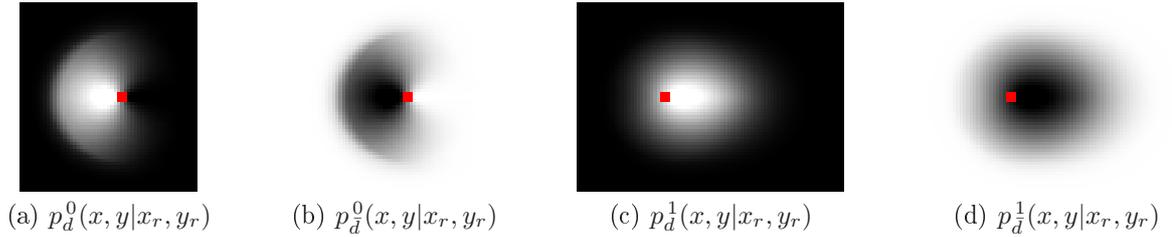


Figure 5.12 – Densités de probabilité de présence des pixels détectés (a, c) et non détectés (b, d) dans les images binaires de primitives B_I^0 (a, b) et B_I^1 (c, d) en imagerie RSO.

passee ascendante.

Pour conclure, nous détaillons l'implémentation mise en œuvre pour calculer les densités de probabilité $p_d^j(x, y|x_r, y_r)$ et $p_{\bar{d}}^j(x, y|x_r, y_r)$. Cette implémentation permet d'obtenir une bonne approximation de ces densités de probabilité en un temps raisonnable. Par souci de clarté, nous exposons uniquement la méthode mise en œuvre pour calculer la densité de probabilité $p_d^j(x, y|x_r, y_r)$, celle mise en œuvre pour le calcul de $p_{\bar{d}}^j(x, y|x_r, y_r)$ étant similaire.

La densité de probabilité $p_d^j(x, y|x_r, y_r)$ est estimée au moyen d'un accumulateur, dont les cases correspondent aux pixels de l'image I , au sein duquel sont comptabilisées les contributions associées à chacun des N_{MC} tirages aléatoires réalisés au sein de l'espace des paramètres lors de la mise en œuvre de la méthode de Monte-Carlo. Pour chacun de ces tirages, nous déterminons la primitive $\mathcal{P}_I^j(\vec{b})$ résultant de la projection du bâtiment \vec{b} dans l'image I . Nous notons alors respectivement $N_{\mathcal{P}_I^j(\vec{b})}$ et $N_{\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})}$ le nombre de pixels contenus dans les régions $\mathcal{P}_I^j(\vec{b})$ et $\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})$.

A priori, la totalité des pixels de l'accumulateur doit être parcourue à chaque tirage, comme indiqué à la figure 5.13 (a). D'après l'équation (5.14), les pixels de couleur gris foncé, pixels contenus dans la région $\mathcal{P}_I^j(\vec{b})$, sont incrémentés d'une quantité égale à $p_D^j / (N_{\mathcal{P}_I^j(\vec{b})} p_D^j + N_{\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})} p_{FA}^j)$ tandis que les pixels de couleur gris clair, pixels n'appartenant pas à la région $\mathcal{P}_I^j(\vec{b})$, sont incrémentés d'une quantité égale à $p_{FA}^j / (N_{\mathcal{P}_I^j(\vec{b})} p_D^j + N_{\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})} p_{FA}^j)$.

Afin d'éviter un parcours systématique de l'ensemble des pixels de l'accumulateur à chaque tirage, opération coûteuse en temps de calcul, nous proposons une méthode permettant de ne parcourir que les pixels de l'accumulateur contenus dans la région $\mathcal{P}_I^j(\vec{b})$, ce qui permet de réduire considérablement les temps de calcul. Cette méthode utilise deux accumulateurs, de manière à séparer les contributions des pixels appartenant à la région $\mathcal{P}_I^j(\vec{b})$ de celles des pixels n'appartenant pas à cette région, comme schématisé à la figure 5.13 (b). Ainsi, pour chaque tirage, les pixels du premier accumulateur contenus dans la région $\mathcal{P}_I^j(\vec{b})$, pixels identifiés par la couleur gris foncé, sont incrémentés d'une quantité égale à $(p_D^j - p_{FA}^j) / (N_{\mathcal{P}_I^j(\vec{b})} p_D^j + N_{\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})} p_{FA}^j)$, tandis que le second accumulateur est incrémenté d'une quantité égale à $p_{FA}^j / (N_{\mathcal{P}_I^j(\vec{b})} p_D^j + N_{\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})} p_{FA}^j)$. Après que tous les tirages ont été effectués, la

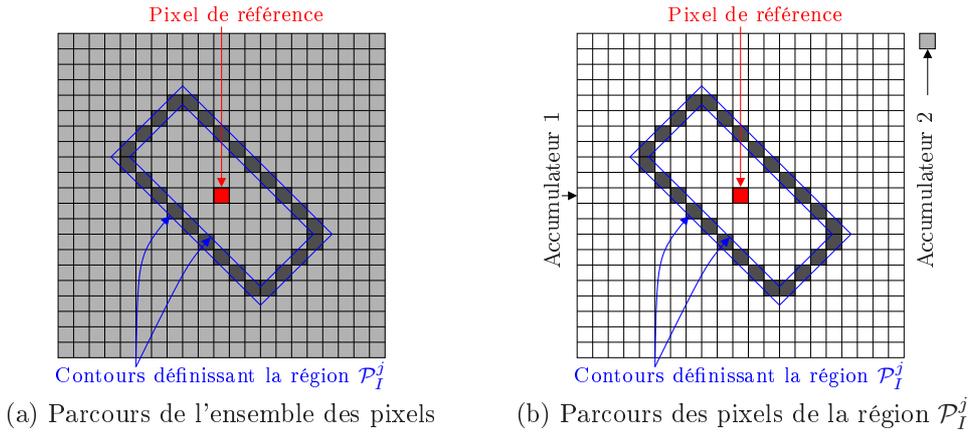


Figure 5.13 – Implémentation du calcul des densités de probabilité de présence des pixels détectés et non détectés dans les images binaires de primitives.

valeur du second accumulateur est ajoutée à chaque pixel du premier accumulateur, ce qui permet d'obtenir un résultat identique à celui résultant d'un parcours systématique de l'ensemble des pixels du premier accumulateur à chaque tirage.

5.3.2 Test de significativité pour la détection des bâtiments

Nous avons établi, à la section précédente, l'expression des densités de probabilité de présence des pixels détectés et non détectés dans les images binaires B_I^j , dans le cas où un bâtiment est présent dans l'image et dont seule la position (x_r, y_r) du point de référence dans cette image est connue.

Dans cette section, nous réalisons, en chaque pixel (x_t, y_t) de l'image binaire B_I^j , un test statistique visant à décider entre les deux hypothèses suivantes :

- \mathcal{H}_1 : « Le pixel (x_t, y_t) correspond au point de référence d'un bâtiment » ;
- \mathcal{H}_0 : « Le pixel (x_t, y_t) ne correspond pas au point de référence d'un bâtiment ».

Lorsque l'hypothèse \mathcal{H}_1 est vérifiée, la répartition des pixels détectés et non détectés au sein de l'image binaire B_I^j , dans un voisinage ν_t^j du pixel testé (x_t, y_t) , est régie par les densités de probabilité de présence p_d^j et $p_{\bar{d}}^j$, densités qui ont été calculées à la section précédente. En revanche, lorsque l'hypothèse \mathcal{H}_0 est vérifiée, nous ne sommes pas en mesure de fournir une modélisation statistique régissant la répartition des pixels détectés et non détectés dans le voisinage du pixel testé. Par conséquent, nous mettons ici en œuvre un test reposant uniquement sur la significativité statistique de l'hypothèse \mathcal{H}_1 .

5.3.2.1 Calcul de la vraisemblance des pixels des images binaires de primitives

Afin de quantifier l'adéquation entre l'image binaire B_I^j et l'hypothèse \mathcal{H}_1 en un pixel (x_t, y_t) de cette image, nous calculons la vraisemblance des pixels de cette image sachant l'hypothèse \mathcal{H}_1 .

En supposant que les pixels détectés et non détectés de l'image B_I^j correspondent respectivement à des réalisations indépendantes des variables aléatoires X_d^j, Y_d^j et $X_{\bar{d}}^j, Y_{\bar{d}}^j$, les log-vraisemblances $\ell_d^j(x_t, y_t)$ et $\ell_{\bar{d}}^j(x_t, y_t)$ des pixels détectés et non détectés dans le voisinage ν_t^j du pixel testé, sachant

l'hypothèse \mathcal{H}_1 , s'écrivent :

$$\forall (x_t, y_t) \in \mathcal{E}_I, \ell_d^j(x_t, y_t) = \sum_{\substack{(x,y) \in \nu_t^j \\ B_I^j(x,y)=1}} \ln p_d^j(x, y | x_t, y_t), \quad (5.20)$$

$$\forall (x_t, y_t) \in \mathcal{E}_I, \ell_{\bar{d}}^j(x_t, y_t) = \sum_{\substack{(x,y) \in \nu_t^j \\ B_I^j(x,y)=0}} \ln p_{\bar{d}}^j(x, y | x_t, y_t). \quad (5.21)$$

Finalement, la log-vraisemblance $\ell^j(x_t, y_t)$ de l'ensemble des pixels de l'image B_I^j contenus dans le voisinage ν_t^j du pixel testé est la somme des log-vraisemblances des pixels détectés et des pixels non détectés :

$$\forall (x_t, y_t) \in \mathcal{E}_I, \ell^j(x_t, y_t) = \ell_d^j(x_t, y_t) + \ell_{\bar{d}}^j(x_t, y_t). \quad (5.22)$$

L'approche qui consisterait à ne considérer que les pixels détectés (ou que les pixels non détectés) pour le calcul de la log-vraisemblance $\ell^j(x_t, y_t)$ conduit à des valeurs incohérentes entre les différents pixels (x_t, y_t) testés, étant donné que le nombre de pixels détectés dans le voisinage ν_t^j varie d'un pixel testé à l'autre. Avec la prise en compte des pixels non détectés, le nombre de pixels intervenant dans le calcul de la log-vraisemblance est le même pour tous les pixels testés, de sorte que les valeurs de cette log-vraisemblance deviennent comparables entre les différents pixels testés.

5.3.2.2 Résultats en imagerie optique

Les figures 5.14 (b) et 5.15 (b) représentent les cartes de log-vraisemblance ℓ^0 obtenues respectivement à partir des images binaires des figures 5.14 (a) et 5.15 (a), images binaires qui résultent de l'extraction des contours de toit dans l'image optique de la figure 5.2. L'image de la figure 5.14 (a) a été obtenue par simulation pour un détecteur dont les probabilités de détection et de fausse alarme valent respectivement $p_D^0 = 0.9$ et $p_{FA}^0 = 0.1$, tandis que l'image de la figure 5.15 (a) a été établie sur données réelles à la section 5.2.1. Les densités de probabilité de présence $p_d^0(x, y | x_r, y_r)$ et $p_{\bar{d}}^0(x, y | x_r, y_r)$, mises en œuvre pour le calcul de ces cartes de log-vraisemblance, ont été générées avec les valeurs p_D^0 et p_{FA}^0 caractéristiques des détecteurs respectifs dont sont issues les images binaires B_I^0 .

De même, les figures 5.16 (b) et 5.17 (b) présentent les cartes de log-vraisemblance ℓ^1 obtenues respectivement à partir des images binaires des figures 5.16 (a) et 5.17 (a), images binaires qui résultent de l'extraction des ombres dans l'image optique de la figure 5.2. L'image de la figure 5.16 (a) a été obtenue par simulation pour un détecteur dont les probabilités de détection et de fausse alarme valent respectivement $p_D^1 = 0.9$ et $p_{FA}^1 = 0.1$, tandis que l'image de la figure 5.17 (a) a été établie sur données réelles à la section 5.2.1. Les densités de probabilité de présence $p_d^1(x, y | x_r, y_r)$ et $p_{\bar{d}}^1(x, y | x_r, y_r)$, mises en œuvre pour le calcul de ces cartes de détection, ont été générées avec les valeurs p_D^1 et p_{FA}^1 caractéristiques des détecteurs respectifs dont sont issues les images binaires B_I^1 .

En comparant les cartes de log-vraisemblance ℓ^0 , obtenues à partir des images binaires B_I^0 réelle et simulée, nous constatons que les données réelles conduisent à des résultats peu satisfaisants. Ceci s'explique par la présence de nombreux contours parasites, dus à la présence d'ombres et d'autres éléments de l'environnement urbain, mais aussi par la forme peu discriminante des densités de probabilité de présence $p_d^0(x, y | x_r, y_r)$ et $p_{\bar{d}}^0(x, y | x_r, y_r)$.

En revanche, les cartes de log-vraisemblance ℓ^1 , obtenues à partir des images binaires B_I^1 réelle et simulée, sont relativement similaires, de sorte que les données réelles conduisent à des résultats satisfaisants aussi bien en termes de détection qu'en termes de fausse alarme.

Un seuillage supervisé des cartes de log-vraisemblance ℓ^0 et ℓ^1 , seuillage qui fait intervenir les seuils τ_{ℓ^0} et τ_{ℓ^1} , conduit finalement aux cartes binaires de détection présentées aux figures 5.18 et 5.19, cartes au sein desquelles sont représentés les pixels de référence des bâtiments présents dans l'image I . Si les résultats obtenus à partir de données simulées sont satisfaisants, ceux obtenus à partir de données

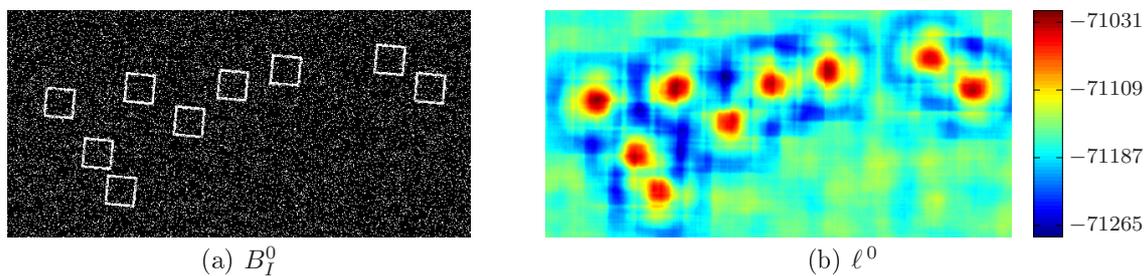


Figure 5.14 – Image binaire B_I^0 simulée (a) et carte de log-vraisemblance ℓ^0 associée (b).

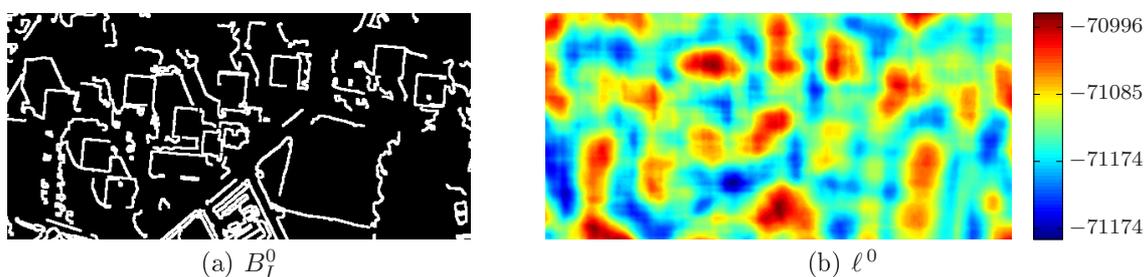


Figure 5.15 – Image binaire B_I^0 réelle (a) et carte de log-vraisemblance ℓ^0 associée (b).

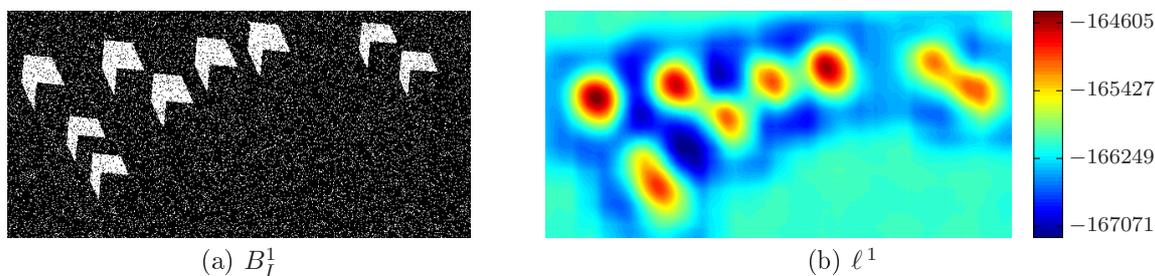


Figure 5.16 – Image binaire B_I^1 simulée (a) et carte de log-vraisemblance ℓ^1 associée (b).

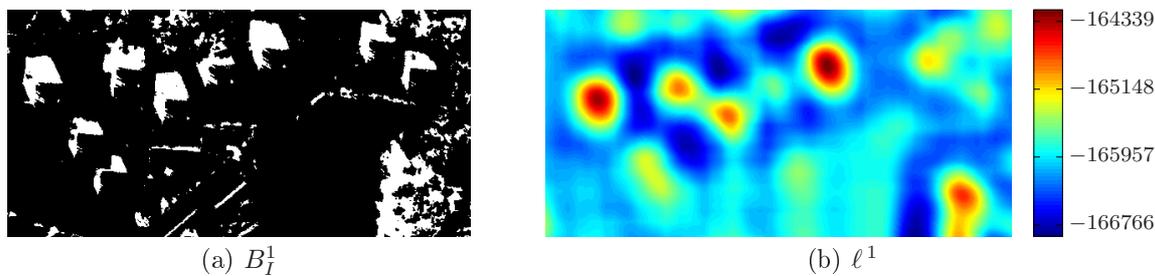


Figure 5.17 – Image binaire B_I^1 réelle (a) et carte de log-vraisemblance ℓ^1 associée (b).

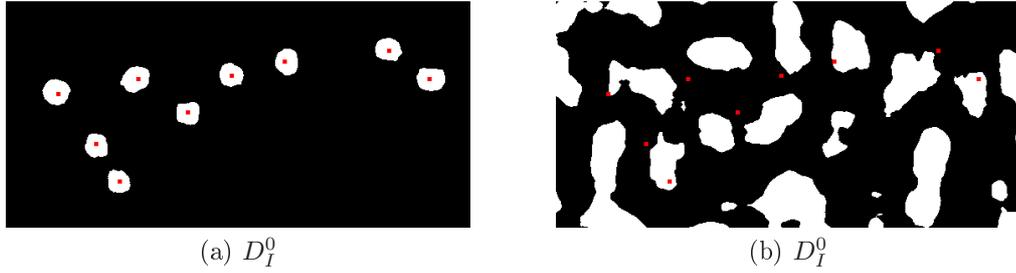


Figure 5.18 – Cartes binaires de détection D_I^0 établies à partir des images binaires de primitives B_I^0 simulée (a) et réelle (b) des figures 5.14 et 5.15.

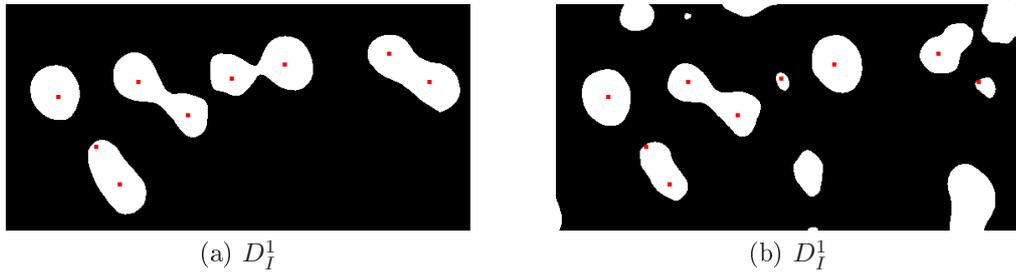


Figure 5.19 – Cartes binaires de détection D_I^1 établies à partir des images binaires de primitives B_I^1 simulée (a) et réelle (b) des figures 5.16 et 5.17.

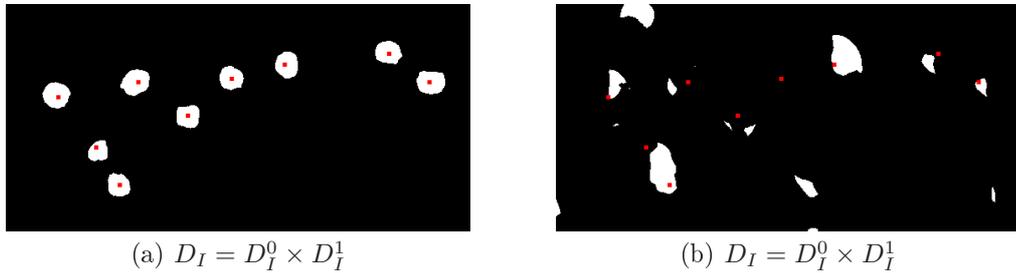
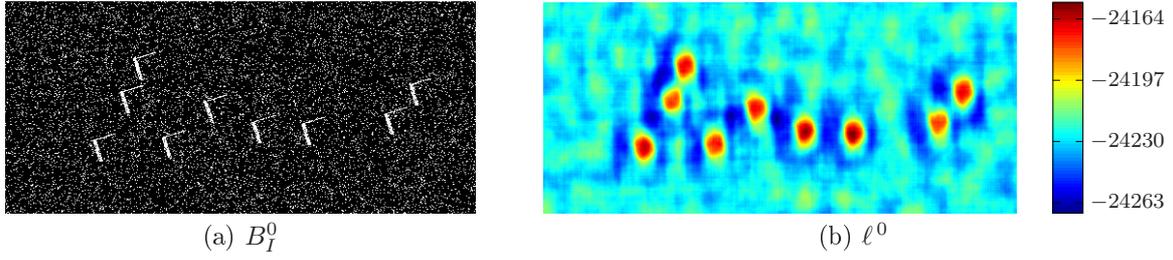
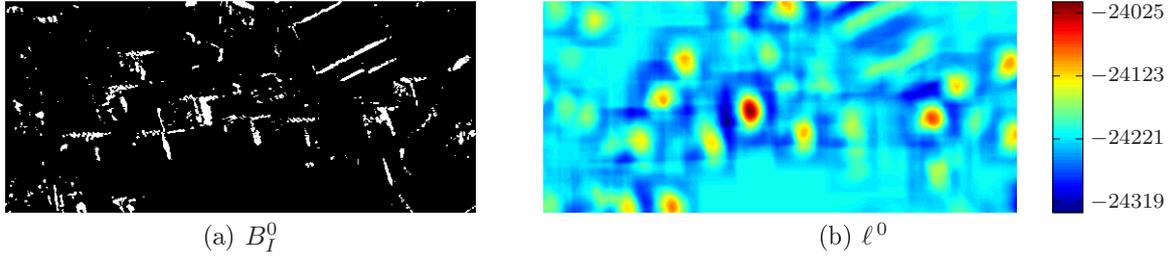


Figure 5.20 – Cartes binaires de détection D_I résultant du produit des cartes binaires de détection D_I^0 et D_I^1 des figures 5.18 et 5.19.

réelles sont mitigés. En effet, la détection à partir de l'image B_I^1 réelle fournit des résultats relativement bons, comparables à ceux obtenus à partir de l'image B_I^1 simulée. En revanche, la détection à partir de l'image B_I^0 réelle est peu satisfaisante, pour les raisons que nous avons précédemment évoquées.

Les images D_I^0 et D_I^1 ainsi établies permettent de définir des régions d'intérêt contenant possiblement les pixels de référence de bâtiments. Une intersection ou une union de ces images peut alors être réalisée afin de fusionner les résultats obtenus indépendamment à partir des images binaires de primitives B_I^0 et B_I^1 . La figure 5.20 montre les images binaires D_I résultant d'une fusion conjonctive des images D_I^0 et D_I^1 obtenues à partir des images B_I^0 et B_I^1 simulées et réelles.

Notons enfin que les expressions des densités de probabilité de présence $p_d^j(x, y|x_r, y_r)$ et $p_d^j(x, y|x_r, y_r)$ des pixels détectés et non détectés dans l'image binaire de primitives B_I^j en présence d'un bâtiment, ont été établies en supposant que seuls pouvaient être détectés, dans le voisinage ν_r^j du point de référence, des pixels de la primitive \mathcal{P}^j de ce bâtiment ainsi que des fausses alarmes. Le cas où sont détectés des pixels appartenant à la primitive \mathcal{P}^j d'un autre bâtiment n'est pas envisagé,

Figure 5.21 – Image binaire B_I^0 simulée (a) et carte de log-vraisemblance ℓ^0 associée (b).Figure 5.22 – Image binaire B_I^0 réelle (a) et carte de log-vraisemblance ℓ^0 associée (b).

de sorte que ces pixels sont traités comme des fausses alarmes. Nous constatons, sur données simulées, que les phénomènes de recouvrement entre bâtiments permettent tout de même de détecter les différents bâtiments, de sorte que les performances limitées de l'approche proposée sur données réelles sont essentiellement liées à la présence de structures parasites dans l'environnement immédiat des bâtiments.

5.3.2.3 Résultats en imagerie RSO

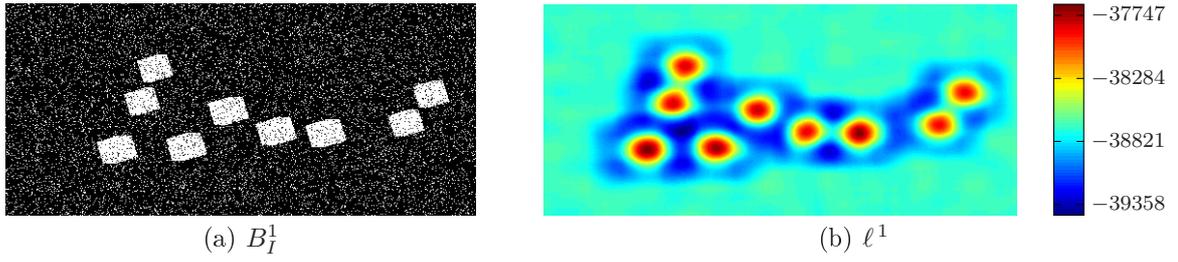
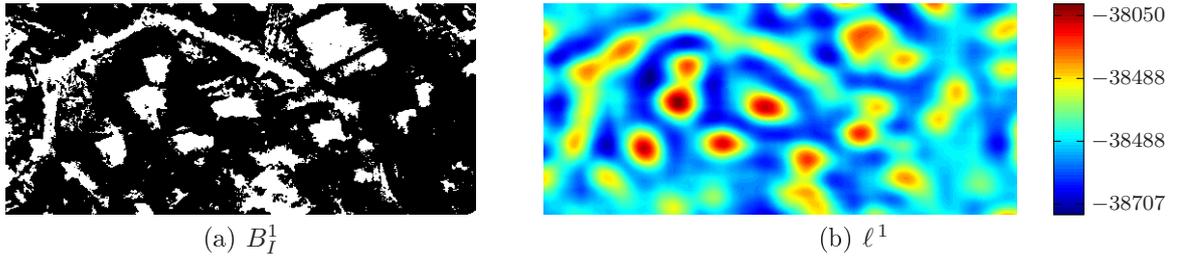
Dans cette sous-section, nous présentons les cartes de log-vraisemblance établies à partir des images binaires résultant de l'extraction des primitives caractéristiques de la présence des bâtiments en imagerie RSO.

Les figures 5.21 (b) et 5.22 (b) présentent les cartes de log-vraisemblance ℓ^0 obtenues respectivement à partir des images binaires des figures 5.21 (a) et 5.22 (a), images binaires qui résultent de l'extraction des échos doubles dans l'image RSO de la figure 5.4. L'image de la figure 5.21 (a) a été obtenue par simulation pour un détecteur dont les probabilités de détection et de fausse alarme valent respectivement $p_D^0 = 0.9$ et $p_{FA}^0 = 0.1$, tandis que l'image de la figure 5.22 (a) a été établie sur données réelles à la section 5.2.2.

De même, les figures 5.23 (b) et 5.24 (b) présentent les cartes de log-vraisemblance ℓ^1 obtenues respectivement à partir des images binaires des figures 5.23 (a) et 5.24 (a), images binaires qui résultent de l'extraction des ombres dans l'image RSO de la figure 5.4. L'image de la figure 5.23 (a) a été obtenue par simulation pour un détecteur dont les probabilités de détection et de fausse alarme valent respectivement $p_D^1 = 0.9$ et $p_{FA}^1 = 0.1$, tandis que l'image de la figure 5.24 (a) a été établie sur données réelles à la section 5.2.2.

Enfin, les figures 5.25 et 5.26 présentent les cartes de détection binaires D_I^0 et D_I^1 résultant du seuillage supervisé des cartes de log-vraisemblance précédentes. Les images D_I de la figure 5.27 sont alors établies par fusion conjonctive de ces images binaires.

En comparant les résultats obtenus sur données réelles et sur données simulées, nous constatons que les cartes de détection établies à partir des données réelles conduisent à de bons résultats en termes

Figure 5.23 – Image binaire B_I^1 simulée (a) et carte de log-vraisemblance ℓ^1 associée (b).Figure 5.24 – Image binaire B_I^1 réelle (a) et carte de log-vraisemblance ℓ^1 associée (b).

de détection mais que le nombre de fausses alarmes est relativement important. Notons, pour finir, que le bâtiment F n'est pas détecté dans l'image binaire D_I^1 en raison de la non détection de l'ombre de ce bâtiment au sein de l'image B_I^1 , ombre qui est masquée par le bâtiment G , comme illustré à la figure 2.17.

5.3.2.4 Implémentation rapide par transformée de Fourier

Dans cette sous-section, nous décrivons les aspects d'implémentation relatifs au calcul de la carte de log-vraisemblance $\ell^j(x_t, y_t)$ dont l'expression est donnée à l'équation (5.22). L'approche qui consiste à parcourir, pour chaque pixel testé, le voisinage ν_t^j de ce pixel présente une complexité calculatoire importante lorsque les dimensions de ce voisinage sont grandes. Les dimensions du voisinage ν_t^j , conditionnées par les valeurs maximales des différents paramètres des bâtiments, sont, dans notre cas, de l'ordre de la centaine de pixels. Ainsi, l'approche précédente s'avère particulièrement coûteuse en temps de calcul.

Nous montrons ici que le calcul de la carte de log-vraisemblance $\ell^j(x_t, y_t)$ peut être réalisé dans l'espace de Fourier, ce qui limite considérablement la complexité calculatoire par rapport à l'approche précédemment évoquée.

Dans le cadre de l'hypothèse d'uniformité des effets géométriques liés à la projection des bâtiments dans les images optiques et RSO, la densité de probabilité de présence $p_d^j(x, y|x_r, y_r)$ des pixels détectés dans l'image B_I^j est reliée à la densité de probabilité $p_d^j(x, y|x_0, y_0)$ par une simple translation :

$$\forall (x, y) \in \mathcal{E}_I, \forall (x_r, y_r) \in \mathcal{E}_I, p_d^j(x, y|x_r, y_r) = p_d^j(x - x_r + x_0, y - y_r + y_0|x_0, y_0), \quad (5.23)$$

où nous rappelons que (x_0, y_0) désigne les coordonnées du pixel situé au centre de l'image I .

En reprenant l'équation (5.20), nous pouvons remarquer que la log-vraisemblance $\ell_d^j(x_t, y_t)$ des pixels détectés dans le voisinage ν_t^j d'un pixel testé (x_t, y_t) peut se réécrire comme la convolution de



Figure 5.25 – Cartes binaires de détection D_I^0 établies à partir des images binaires de primitives B_I^0 simulée (a) et réelle (b) des figures 5.21 et 5.22.

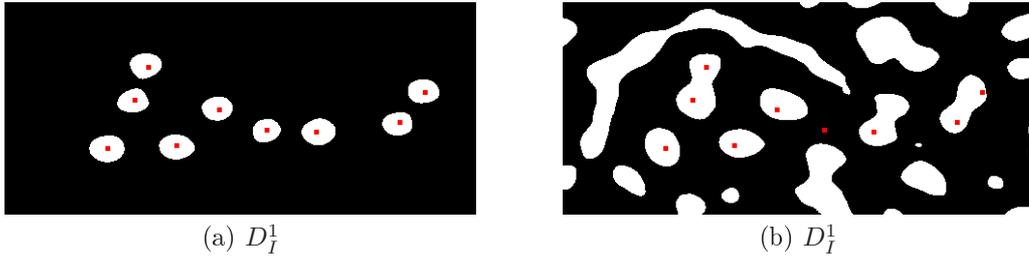


Figure 5.26 – Cartes binaires de détection D_I^1 établies à partir des images binaires de primitives B_I^1 simulée (a) et réelle (b) des figures 5.23 et 5.24.

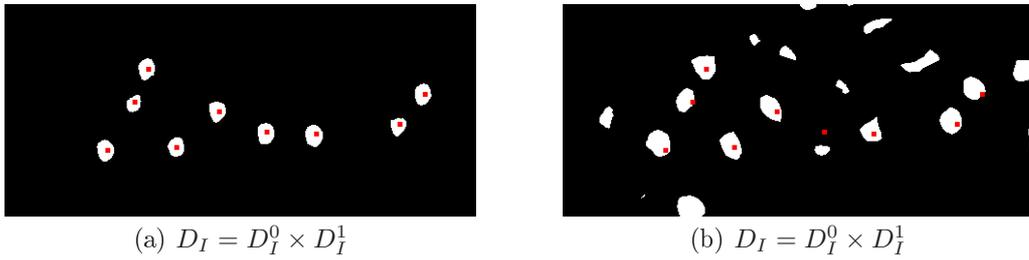


Figure 5.27 – Cartes binaires de détection D_I résultant du produit des cartes binaires de détection D_I^0 et D_I^1 des figures 5.25 et 5.26.

l'image $B_I^j(x, y)$ avec un noyau $f_d^j(x, y) = \ln p_d^j(-x, -y|x_0, y_0)$:

$$\forall (x_t, y_t) \in \mathcal{E}_I, \ell_d^j(x_t, y_t) = \sum_{(x, y) \in \nu_t^j} B_I^j(x, y) \ln p_d^j(x, y|x_t, y_t) \quad (5.24)$$

$$= \sum_{(x, y) \in \nu_t^j} B_I^j(x, y) \underbrace{\ln p_d^j(x - x_t + x_0, y - y_t + y_0|x_0, y_0)}_{f_d^j(x_t - x_0 - x, y_t - y_0 - y)} \quad (5.25)$$

$$\forall (x_t, y_t) \in \mathcal{E}_I, \ell_d^j(x_t, y_t) = [B_I^j * f_d^j](x_t - x_0, y_t - y_0). \quad (5.26)$$

Ainsi, le calcul de cette log-vraisemblance peut être réalisée dans l'espace de Fourier selon la formule suivante :

$$\forall (x_t, y_t) \in \mathcal{E}_I, \ell_d^j(x_t, y_t) = F^{-1} [F[B_I^j]F[f_d^j]](x_t - x_0, y_t - y_0), \quad (5.27)$$

où $F[g]$ et $F^{-1}[g]$ désignent respectivement la transformée de Fourier et la transformée de Fourier inverse de la fonction g .

De manière analogue, la log-vraisemblance $\ell_d^j(x_t, y_t)$ des pixels non détectés dans le voisinage ν_t^j d'un pixel testé (x_t, y_t) peut se réécrire comme la convolution de l'image \bar{B}_I^j avec le noyau $f_d^j(x, y) = \ln p_d^j(-x, -y|x_0, y_0)$, où l'image \bar{B}_I^j est définie comme suit : $\forall(x, y), \in \mathcal{E}_I \bar{B}_I^j(x, y) = 1 - B_I^j(x, y)$. Le calcul de cette log-vraisemblance peut alors être réalisé dans l'espace de Fourier selon l'équation suivante :

$$\forall(x_t, y_t) \in \mathcal{E}_I, \ell_d^j(x_t, y_t) = F^{-1} \left[F[\bar{B}_I^j] F[f_d^j] \right] (x_t - x_0, y_t - y_0). \quad (5.28)$$

Finalement, les cartes de log-vraisemblance $\ell_d^j(x_t, y_t)$ et $\ell_{\bar{d}}^j(x_t, y_t)$ sont établies au moyen de la transformée de Fourier rapide et de la transformée de Fourier inverse rapide. La carte de log-vraisemblance $\ell^j(x_t, y_t)$ résulte alors de la somme des deux cartes précédentes.

5.3.3 Détection grossière à partir des primitives considérées conjointement

À la section précédente, les images binaires D_I définissant les régions d'intérêt ont été obtenues par fusion des images binaires D_I^j et D_I^k établies indépendamment à partir des images binaires B_I^j et B_I^k . Cette approche conduit à des résultats relativement satisfaisants mais nécessite la détermination de deux seuils τ_{ℓ^j} et τ_{ℓ^k} ainsi que la mise en œuvre d'une stratégie de fusion des images binaires D_I^j et D_I^k .

Nous présentons, dans cette section, une extension de la méthode de détection grossière introduite aux sections précédentes, extension qui permet de calculer une unique image de vraisemblance ℓ^{jk} en considérant conjointement deux images binaires B_I^j et B_I^k . Le seuillage de cette image de vraisemblance ℓ^{jk} conduit alors à l'image binaire D_I au sein de laquelle sont définies les régions d'intérêt contenant possiblement le point de référence d'un bâtiment.

5.3.3.1 Calcul des densités de probabilité de présence conjointe des pixels détectés et non détectés dans les images binaires de primitives en présence d'un bâtiment

Nous supposons, dans cette sous-section, qu'un bâtiment \vec{b} est présent dans l'image I et que le point de référence r permettant de repérer ce bâtiment dans l'image, point de coordonnées (x_r, y_r) dans l'espace image \mathcal{E}_I , est connu *a priori*.

Par la suite, nous notons $dd, \bar{d}d, d\bar{d}$ et $\bar{d}\bar{d}$ les couples de pixels constitués d'un pixel de l'image B_I^j et d'un pixel de l'image B_I^k , couples au sein desquels d désigne un pixel détecté tandis que \bar{d} désigne un pixel non détecté.

Nous établissons, dans cette sous-section, l'expression des densités de probabilité de présence des couples de pixels $dd, \bar{d}d, d\bar{d}$ et $\bar{d}\bar{d}$, densités de probabilité qui sont notées respectivement $p_{dd}^{jk}, p_{\bar{d}d}^{jk}, p_{d\bar{d}}^{jk}$ et $p_{\bar{d}\bar{d}}^{jk}$.

Expression à paramètres de bâtiment connus

Nous supposons tout d'abord que les paramètres du bâtiment \vec{b} sont connus et établissons les expressions des densités de probabilité $p_{dd}^{jk}, p_{\bar{d}d}^{jk}, p_{d\bar{d}}^{jk}$ et $p_{\bar{d}\bar{d}}^{jk}$ conditionnellement à ce vecteur de paramètres.

Les couples de pixels dd considérés dans cette section résultent de la sélection équiprobable d'un des pixels détectés de l'image B_I^j puis de la sélection, également équiprobable, d'un des pixels détectés de l'image B_I^k . La sélection des pixels au sein des deux images binaires étant réalisée de manière indépendante, les variables aléatoires X_d^j et Y_d^j , qui correspondent aux coordonnées du premier pixel du couple, sont indépendantes des variables aléatoires X_d^k et Y_d^k , qui correspondent aux coordonnées du second pixel du couple. Ainsi, la densité de probabilité conjointe p_{dd}^{jk} des variables aléatoires X_d^j, Y_d^j, X_d^k et Y_d^k s'écrit comme suit :

$$\forall(x^j, y^j, x^k, y^k) \in \nu_r^j \times \nu_r^k, p_{dd}^{jk}(x^j, y^j, x^k, y^k | \vec{b}) = p_d^j(x^j, y^j | \vec{b}) p_d^k(x^k, y^k | \vec{b}), \quad (5.29)$$

où p_d^j et p_d^k désignent respectivement les densités de probabilité conjointe des variables aléatoires X_d^j , Y_d^j et X_d^k , Y_d^k , densités de probabilité dont les expressions ont été établies à la sous-section 5.3.1.1.

Dans le cas où les images binaires de primitives B_I^j et B_I^k sont obtenues au moyen de détecteurs parfaits, le remplacement, dans l'équation (5.29), des densités de probabilité p_d^j et p_d^k par leurs expressions établies à l'équation (5.7) conduit à l'expression suivante :

$$\forall(x^j, y^j, x^k, y^k) \in \nu_r^j \times \nu_r^k, p_{dd}^{jk}(x^j, y^j, x^k, y^k | \vec{b}) = \frac{1_{\mathcal{P}_I^j(\vec{b})}(x^j, y^j) 1_{\mathcal{P}_I^k(\vec{b})}(x^k, y^k)}{\mathcal{A}(\mathcal{P}_I^j(\vec{b})) \mathcal{A}(\mathcal{P}_I^k(\vec{b}))}. \quad (5.30)$$

Ainsi, en présence du bâtiment \vec{b} , la probabilité que deux pixels, détectés respectivement dans les images binaires B_I^j et B_I^k , soient situés en (x^j, y^j) et (x^k, y^k) est une probabilité uniforme sur le support quadridimensionnel $\mathcal{P}_I^j(\vec{b}) \times \mathcal{P}_I^k(\vec{b})$.

Dans le cas où les images binaires B_I^j et B_I^k sont obtenues au moyen de détecteurs réels, caractérisés respectivement par les probabilités de détection et de fausse alarme p_D^j , p_{FA}^j et p_D^k , p_{FA}^k , la probabilité d'observer des pixels détectés en dehors des primitives $\mathcal{P}_I^j(\vec{b})$ et $\mathcal{P}_I^k(\vec{b})$ est non nulle en raison des fausses alarmes. En remplaçant p_d^j et p_d^k par leurs expressions établies à l'équation (5.14), la densité de probabilité de présence conjointe des pixels détectés dans ces deux images binaires, conditionnellement au vecteur de paramètres \vec{b} , s'écrit :

$$\begin{aligned} \forall(x^j, y^j, x^k, y^k) \in \nu_r^j \times \nu_r^k, p_{dd}^{jk}(x^j, y^j, x^k, y^k | \vec{b}) = \\ \frac{p_D^j p_D^k}{C_{dd}} 1_{\mathcal{P}_I^j(\vec{b})}(x^j, y^j) 1_{\mathcal{P}_I^k(\vec{b})}(x^k, y^k) + \frac{p_D^j p_{FA}^k}{C_{dd}} 1_{\mathcal{P}_I^j(\vec{b})}(x^j, y^j) 1_{\nu_r^k \setminus \mathcal{P}_I^k(\vec{b})}(x^k, y^k) \\ + \frac{p_{FA}^j p_D^k}{C_{dd}} 1_{\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})}(x^j, y^j) 1_{\mathcal{P}_I^k(\vec{b})}(x^k, y^k) + \frac{p_{FA}^j p_{FA}^k}{C_{dd}} 1_{\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})}(x^j, y^j) 1_{\nu_r^k \setminus \mathcal{P}_I^k(\vec{b})}(x^k, y^k), \end{aligned} \quad (5.31)$$

où C_{dd} est une constante de normalisation dont l'expression est donnée par la relation suivante :

$$\begin{aligned} C_{dd} = p_D^j p_D^k \mathcal{A}(\mathcal{P}_I^j(\vec{b})) \mathcal{A}(\mathcal{P}_I^k(\vec{b})) + p_D^j p_{FA}^k \mathcal{A}(\mathcal{P}_I^j(\vec{b})) \mathcal{A}(\nu_r^k \setminus \mathcal{P}_I^k(\vec{b})) \\ + p_{FA}^j p_D^k \mathcal{A}(\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})) \mathcal{A}(\mathcal{P}_I^k(\vec{b})) + p_{FA}^j p_{FA}^k \mathcal{A}(\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})) \mathcal{A}(\nu_r^k \setminus \mathcal{P}_I^k(\vec{b})). \end{aligned} \quad (5.32)$$

Ainsi, la densité de probabilité de présence $p_{dd}^{jk}(x^j, y^j, x^k, y^k | \vec{b})$ s'écrit comme la somme de quatre termes :

- le premier terme correspond au cas où les deux pixels détectés au sein des images binaires B_I^j et B_I^k appartiennent respectivement aux primitives $\mathcal{P}_I^j(\vec{b})$ et $\mathcal{P}_I^k(\vec{b})$ et correspondent par conséquent à deux vraies détections ;
- le second terme correspond au cas où le pixel détecté au sein de l'image B_I^j appartient à la primitive $\mathcal{P}_I^j(\vec{b})$ tandis que le pixel détecté au sein de l'image B_I^k n'appartient pas à la primitive $\mathcal{P}_I^k(\vec{b})$, de sorte que ces pixels correspondent respectivement à une vraie détection et à une fausse alarme ;
- le troisième terme est similaire au second à la différence que le pixel détecté au sein de l'image binaire B_I^j est une fausse alarme et que le pixel détecté dans l'image B_I^k est une vraie détection ;
- le quatrième terme correspond enfin au cas où les deux pixels détectés sont des fausses alarmes.

De manière analogue, les expressions des densités de probabilité de présence p_{dd}^{jk} , p_{dd}^{jk} et p_{dd}^{jk} sont données par les équations suivantes :

$$\forall(x^j, y^j, x^k, y^k) \in \nu_r^j \times \nu_r^k, p_{dd}^{jk}(x^j, y^j, x^k, y^k | \vec{b}) = p_d^j(x^j, y^j | \vec{b}) p_d^k(x^k, y^k | \vec{b}), \quad (5.33)$$

$$\forall(x^j, y^j, x^k, y^k) \in \nu_r^j \times \nu_r^k, p_{dd}^{jk}(x^j, y^j, x^k, y^k | \vec{b}) = p_d^j(x^j, y^j | \vec{b}) p_d^k(x^k, y^k | \vec{b}), \quad (5.34)$$

$$\forall(x^j, y^j, x^k, y^k) \in \nu_r^j \times \nu_r^k, p_{dd}^{jk}(x^j, y^j, x^k, y^k | \vec{b}) = p_d^j(x^j, y^j | \vec{b}) p_d^k(x^k, y^k | \vec{b}). \quad (5.35)$$

Expression à position de bâtiment connue

Nous supposons désormais que seules les coordonnées (x_r, y_r) du point de référence permettant de repérer le bâtiment \vec{b} dans l'image I sont connues. Les expressions des densités de probabilité de présence $p_{dd}^{jk}(x^j, y^j, x^k, y^k | x_r, y_r)$, $p_{\bar{d}\bar{d}}^{jk}(x^j, y^j, x^k, y^k | x_r, y_r)$, $p_{\bar{d}d}^{jk}(x^j, y^j, x^k, y^k | x_r, y_r)$ et $p_{d\bar{d}}^{jk}(x^j, y^j, x^k, y^k | x_r, y_r)$ sont obtenues, comme à la sous-section 5.3.1.2, à partir des expressions de ces densités de probabilité établies conditionnellement au vecteur de paramètres \vec{b} , par intégration sur l'ensemble $\mathcal{E}_\alpha \times \mathcal{E}_{w_1} \times \mathcal{E}_{w_2} \times \mathcal{E}_h$, l'intégration étant réalisée par méthode de Monte-Carlo.

À titre d'exemple, les densités de probabilité $p_{dd}^{jk}(x^j, y^j, x^k, y^k | x_r, y_r, \alpha, w_1, w_2, h)$ et $p_{\bar{d}\bar{d}}^{jk}(x^j, y^j, x^k, y^k | x_r, y_r)$ sont reliées par la relation suivante :

$$\begin{aligned} \forall (x^j, y^j, x^k, y^k) \in \nu_r^j \times \nu_r^k, \quad p_{dd}^{jk}(x^j, y^j, x^k, y^k | x_r, y_r) = \\ \int_{\mathcal{E}_\alpha} \int_{\mathcal{E}_{w_1}} \int_{\mathcal{E}_{w_2}} \int_{\mathcal{E}_h} p_{dd}^{jk}(x^j, y^j, x^k, y^k | x_r, y_r, \alpha, w_1, w_2, h) p(\alpha, w_1, w_2, h) d\alpha dw_1 dw_2 dh, \end{aligned} \quad (5.36)$$

où $p(\alpha, w_1, w_2, h)$ désigne la probabilité *a priori* d'observer des bâtiments de paramètres (α, w_1, w_2, h) au sein de l'image I , probabilité qui est supposée uniforme sur l'ensemble $\mathcal{E}_\alpha \times \mathcal{E}_{w_1} \times \mathcal{E}_{w_2} \times \mathcal{E}_h$.

Les densités de probabilité $p_{dd}^{jk}(x^j, y^j, x^k, y^k | x_r, y_r)$, $p_{\bar{d}\bar{d}}^{jk}(x^j, y^j, x^k, y^k | x_r, y_r)$, $p_{\bar{d}d}^{jk}(x^j, y^j, x^k, y^k | x_r, y_r)$ et $p_{d\bar{d}}^{jk}(x^j, y^j, x^k, y^k | x_r, y_r)$ sont calculées au moyen d'un accumulateur quadridimensionnel au sein duquel sont comptabilisées les contributions des différents tirages réalisés dans le cadre de la méthode de Monte-Carlo. Comme nous pouvons le constater à l'équation (5.31), les densités de probabilité p_{dd}^{jk} , $p_{\bar{d}\bar{d}}^{jk}$, $p_{\bar{d}d}^{jk}$ et $p_{d\bar{d}}^{jk}$, établies conditionnellement au vecteur de paramètres \vec{b} , sont constantes par morceaux sur quatre sous-ensembles constituant une partition de l'ensemble $\nu_r^j \times \nu_r^k$. Ainsi, la méthode mise en œuvre à la sous-section 5.3.1.2 pour éviter un parcours exhaustif de l'accumulateur ne peut plus être appliquée, de sorte que la complexité calculatoire associée au calcul de ces densités de probabilité est conséquente.

Quels que soient les détecteurs mis en œuvre pour l'extraction des primitives, les probabilités de détection p_D^j et p_D^k sont supérieures à 0.5 et les probabilités de fausse alarme p_{FA}^j et p_{FA}^k sont inférieures à 0.5. La densité de probabilité $p_{dd}^{jk}(x^j, y^j, x^k, y^k | \vec{b})$ est donc maximale sur le sous-ensemble $\mathcal{P}_I^j(\vec{b}) \times \mathcal{P}_I^k(\vec{b})$ et minimale sur le sous-ensemble $[\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})] \times [\nu_r^k \setminus \mathcal{P}_I^k(\vec{b})]$. Alors qu'elle peut prendre trois valeurs distinctes sur le sous-ensemble $[\nu_r^j \times \nu_r^k] \setminus [\mathcal{P}_I^j(\vec{b}) \times \mathcal{P}_I^k(\vec{b})]$, nous considérons par la suite que la densité de probabilité $p_{dd}^{jk}(x^j, y^j, x^k, y^k | \vec{b})$ est constante sur ce sous-ensemble. La valeur de cette constante résulte de la moyenne des trois valeurs précédentes pondérées par la mesure des trois sous-ensembles $\mathcal{P}_I^j(\vec{b}) \times [\nu_r^k \setminus \mathcal{P}_I^k(\vec{b})]$, $[\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})] \times \mathcal{P}_I^k(\vec{b})$ et $[\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})] \times [\nu_r^k \setminus \mathcal{P}_I^k(\vec{b})]$.

La densité de probabilité conditionnelle $p_{dd}^{jk}(x^j, y^j, x^k, y^k | \vec{b})$, dont l'expression exacte est donnée à l'équation (5.31), se réécrit alors :

$$\begin{aligned} \forall (x^j, y^j, x^k, y^k) \in \nu_r^j \times \nu_r^k, \quad p_{dd}^{jk}(x^j, y^j, x^k, y^k | \vec{b}) = \\ \frac{p_D^j p_D^k}{C_{dd}} 1_{\mathcal{P}_I^j(\vec{b})}(x^j, y^j) 1_{\mathcal{P}_I^k(\vec{b})}(x^k, y^k) + \frac{(A_{dd}/B_{dd})}{C_{dd}} 1_{[\nu_r^j \times \nu_r^k] \setminus [\mathcal{P}_I^j(\vec{b}) \times \mathcal{P}_I^k(\vec{b})]}(x^j, y^j, x^k, y^k), \end{aligned} \quad (5.37)$$

où la constante de normalisation C_{dd} est inchangée et où A_{dd} et B_{dd} valent :

$$\begin{aligned} A_{dd} = & p_D^j p_{FA}^k \mathcal{A}(\mathcal{P}_I^j(\vec{b})) \mathcal{A}(\nu_r^k \setminus \mathcal{P}_I^k(\vec{b})) + p_{FA}^j p_D^k \mathcal{A}(\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})) \mathcal{A}(\mathcal{P}_I^k(\vec{b})) \\ & + p_{FA}^j p_{FA}^k \mathcal{A}(\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})) \mathcal{A}(\nu_r^k \setminus \mathcal{P}_I^k(\vec{b})), \end{aligned} \quad (5.38)$$

$$\begin{aligned} B_{dd} = & \mathcal{A}(\mathcal{P}_I^j(\vec{b})) \mathcal{A}(\nu_r^k \setminus \mathcal{P}_I^k(\vec{b})) + \mathcal{A}(\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})) \mathcal{A}(\mathcal{P}_I^k(\vec{b})) \\ & + \mathcal{A}(\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})) \mathcal{A}(\nu_r^k \setminus \mathcal{P}_I^k(\vec{b})). \end{aligned} \quad (5.39)$$

Dans le cadre de l'approximation précédente, la densité de probabilité conditionnelle $p_{dd}^{jk}(x^j, y^j, x^k, y^k | \vec{b})$ est désormais constante par morceaux sur deux sous-ensembles, de sorte qu'une stratégie similaire à celle implémentée à la sous-section 5.3.1.2, peut être mise en œuvre. L'intégration sur l'ensemble des paramètres peut alors être réalisée en ne parcourant, pour chaque tirage de Monte-Carlo, que le sous-ensemble $\mathcal{P}_I^j(\vec{b}) \times \mathcal{P}_I^k(\vec{b})$, évitant ainsi un parcours coûteux de l'intégralité de l'ensemble quadridimensionnel $\nu_r^j \times \nu_r^k$.

Avec une approximation similaire à celle adoptée pour le calcul de p_{dd}^{jk} , l'intégration sur l'ensemble des paramètres, nécessaire au calcul de la densité de probabilité $p_{\bar{d}\bar{d}}^{jk}$, peut également être réalisée en ne parcourant, pour chaque tirage de Monte-Carlo, que le sous-ensemble $\mathcal{P}_I^j(\vec{b}) \times \mathcal{P}_I^k(\vec{b})$, ce qui se traduit par un gain important en termes de complexité calculatoire. Pour le calcul des densités de probabilité $p_{\bar{d}\bar{d}}^{jk}$ et $p_{\bar{d}\bar{d}}^{jk}$, la même approximation implique, en revanche, le parcours des sous-ensembles $\mathcal{P}_I^j(\vec{b}) \times [\nu_r^k \setminus \mathcal{P}_I^k(\vec{b})]$ et $[\nu_r^j \setminus \mathcal{P}_I^j(\vec{b})] \times \mathcal{P}_I^k(\vec{b})$, ce qui s'avère relativement coûteux en raison du nombre important d'éléments contenus dans ces sous-ensembles.

Afin de limiter la complexité calculatoire, nous considérons par la suite que les densités de probabilité $p_{\bar{d}\bar{d}}^{jk}$ et $p_{\bar{d}\bar{d}}^{jk}$ sont uniformes sur leur domaine de définition $\nu_r^j \times \nu_r^k$, ce qui permet de s'affranchir du calcul de ces densités de probabilité mais constitue une approximation très forte. En ce qui concerne les densités de probabilité p_{dd}^{jk} et $p_{\bar{d}\bar{d}}^{jk}$, elles sont calculées simultanément, c'est-à-dire que les mêmes tirages de Monte-Carlo sont utilisés pour réaliser l'intégration sur l'ensemble des paramètres.

5.3.3.2 Calcul de la vraisemblance des pixels des images binaires de primitives

Dans cette sous-section, nous souhaitons quantifier l'adéquation de l'hypothèse \mathcal{H}_1 au pixel (x_t, y_t) avec les images binaires de primitives. Contrairement à ce qui a été fait à la section 5.3.2, où l'adéquation de l'hypothèse \mathcal{H}_1 au pixel (x_t, y_t) est quantifiée indépendamment avec les deux images binaires de primitives, nous proposons ici de quantifier conjointement l'adéquation de cette hypothèse avec les images binaires. Pour ce faire, nous calculons la vraisemblance des couples de pixels dd , $\bar{d}\bar{d}$, $d\bar{d}$ et $\bar{d}d$ observés dans le voisinage $\nu_t^j \times \nu_t^k$ du pixel testé.

En supposant que les pixels détectés au sein des images binaires de primitives B_I^j et B_I^k sont des réalisations indépendantes des variables aléatoires X_d^j , Y_d^j , X_d^k et Y_d^k , la log-vraisemblance ℓ_{dd}^{jk} des pixels détectés conjointement dans les voisinages ν_t^j et ν_t^k du pixel testé, sachant l'hypothèse \mathcal{H}_1 , s'écrit :

$$\forall (x_t, y_t) \in \mathcal{E}_I, \ell_{dd}^{jk}(x_t, y_t) = \sum_{\substack{(x^j, y^j) \in \nu_t^j, B_I^j(x^j, y^j)=1 \\ (x^k, y^k) \in \nu_t^k, B_I^k(x^k, y^k)=1}} \ln p_{dd}^{jk}(x^j, y^j, x^k, y^k | x_t, y_t). \quad (5.40)$$

De manière analogue, les log-vraisemblances $\ell_{\bar{d}\bar{d}}$, $\ell_{d\bar{d}}$ et $\ell_{\bar{d}d}$, sachant l'hypothèse \mathcal{H}_1 , s'écrivent comme suit :

$$\forall (x_t, y_t) \in \mathcal{E}_I, \ell_{\bar{d}\bar{d}}^{jk}(x_t, y_t) = \sum_{\substack{(x^j, y^j) \in \nu_t^j, B_I^j(x^j, y^j)=1 \\ (x^k, y^k) \in \nu_t^k, B_I^k(x^k, y^k)=0}} \ln p_{\bar{d}\bar{d}}^{jk}(x^j, y^j, x^k, y^k | x_t, y_t), \quad (5.41)$$

$$\forall (x_t, y_t) \in \mathcal{E}_I, \ell_{d\bar{d}}^{jk}(x_t, y_t) = \sum_{\substack{(x^j, y^j) \in \nu_t^j, B_I^j(x^j, y^j)=0 \\ (x^k, y^k) \in \nu_t^k, B_I^k(x^k, y^k)=1}} \ln p_{d\bar{d}}^{jk}(x^j, y^j, x^k, y^k | x_t, y_t), \quad (5.42)$$

$$\forall (x_t, y_t) \in \mathcal{E}_I, \ell_{\bar{d}d}^{jk}(x_t, y_t) = \sum_{\substack{(x^j, y^j) \in \nu_t^j, B_I^j(x^j, y^j)=0 \\ (x^k, y^k) \in \nu_t^k, B_I^k(x^k, y^k)=0}} \ln p_{\bar{d}d}^{jk}(x^j, y^j, x^k, y^k | x_t, y_t). \quad (5.43)$$

Finalement, la log-vraisemblance des couples de pixels détectés et non détectés dans les voisinages ν_t^j et ν_t^k du pixel t s'écrit comme la somme des log-vraisemblances des couples de pixels dd , $\bar{d}\bar{d}$, $d\bar{d}$ et $\bar{d}d$:

$$\forall (x_t, y_t) \in \mathcal{E}_I, \ell^{jk}(x_t, y_t) = \ell_{dd}^{jk}(x_t, y_t) + \ell_{\bar{d}\bar{d}}^{jk}(x_t, y_t) + \ell_{d\bar{d}}^{jk}(x_t, y_t) + \ell_{\bar{d}d}^{jk}(x_t, y_t). \quad (5.44)$$

Par la suite, nous notons B_{dd}^{jk} , $B_{d\bar{d}}^{jk}$, $B_{\bar{d}d}^{jk}$ et $B_{\bar{d}\bar{d}}^{jk}$ les images quadridimensionnelles définies respectivement sur l'ensemble $\mathcal{E}_I \times \mathcal{E}_I$ comme le produit cartésien des images binaires B_I^j et B_I^k , B_I^j et \bar{B}_I^k , \bar{B}_I^j et B_I^k puis \bar{B}_I^j et \bar{B}_I^k . Nous pouvons montrer, comme à la sous-section 5.3.2.4, que les log-vraisemblances, dont les expressions sont données aux équations (5.40), (5.41), (5.42) et (5.43), peuvent s'écrire comme des convolutions de ces images quadridimensionnelles avec des noyaux dont les expressions font intervenir les logarithmes des densités de probabilité de présence p_{dd}^{jk} , $p_{d\bar{d}}^{jk}$, $p_{\bar{d}d}^{jk}$ et $p_{\bar{d}\bar{d}}^{jk}$. Malheureusement, en raison de la dimension des supports sur lesquels sont définies les opérandes de l'opération de convolution, l'implémentation du calcul de ces log-vraisemblances par transformée de Fourier nécessite l'allocation d'une zone mémoire beaucoup trop importante pour pouvoir être mise en œuvre sur un ordinateur « standard ». Le calcul de ces log-vraisemblances est donc réalisé au moyen d'une implémentation classique de l'opération de convolution, plus coûteuse en complexité calculatoire mais impliquant l'allocation d'une zone mémoire raisonnable.

En pratique, le calcul de la log-vraisemblance ℓ^{jk} au pixel (x_t, y_t) nécessite un parcours exhaustif de l'ensemble $\nu_t^j \times \nu_t^k$, parcours qui s'avère extrêmement coûteux en termes de complexité calculatoire mais qui est cependant indispensable afin de déterminer la nature (dd , $d\bar{d}$, $\bar{d}d$ ou $\bar{d}\bar{d}$) de chacun des couples de pixels observés et ainsi calculer la log-vraisemblance de chacun de ces couples en faisant intervenir la densité de probabilité de présence adaptée (p_{dd}^{jk} , $p_{d\bar{d}}^{jk}$, $p_{\bar{d}d}^{jk}$ ou $p_{\bar{d}\bar{d}}^{jk}$).

Afin de limiter la complexité calculatoire, nous ne faisons plus, par la suite, la distinction entre les couples de pixels $\bar{d}d$, $d\bar{d}$ et $\bar{d}\bar{d}$, de sorte que les couples de pixels observés au sein de l'ensemble $\nu_t^j \times \nu_t^k$ correspondent :

- soit à des couples de pixels détectés, ces couples de vecteurs étant désignés, comme précédemment, par dd ;
- soit à des couples de pixels pour lesquels au moins l'un des deux pixels est un pixel non détecté, couples que nous désignons par la notation $\bar{d}\bar{d}$.

En notant X_{dd}^j , Y_{dd}^j , X_{dd}^k et Y_{dd}^k les variables aléatoires correspondant aux coordonnées des pixels constitutifs d'un couple $\bar{d}\bar{d}$, la densité de probabilité conjointe de ces variables aléatoires, notée $p_{\bar{d}\bar{d}}^{jk}$, s'exprime comme suit :

$$\begin{aligned} \forall (x^j, y^j, x^k, y^k) \in \nu_t^j \times \nu_t^k, p_{\bar{d}\bar{d}}^{jk}(x^j, y^j, x^k, y^k) &= p_{dd}^{jk}(x^j, y^j, x^k, y^k) p(\bar{d}\bar{d}) \\ &+ p_{d\bar{d}}^{jk}(x^j, y^j, x^k, y^k) p(\bar{d}d) + p_{\bar{d}d}^{jk}(x^j, y^j, x^k, y^k) p(d\bar{d}), \end{aligned} \quad (5.45)$$

où $p(\bar{d}\bar{d})$, $p(\bar{d}d)$ et $p(d\bar{d})$ désignent respectivement les probabilités *a priori* que le couple de pixels $\bar{d}\bar{d}$ observé au sein de l'ensemble $\nu_t^j \times \nu_t^k$ corresponde à un couple dd , $\bar{d}d$ ou $d\bar{d}$. En pratique, ces probabilités *a priori* sont calculées à partir des nombres N_d^j , $N_{\bar{d}}^j$, N_d^k et $N_{\bar{d}}^k$ qui correspondent aux nombres de pixels détectés et non détectés au sein des images binaires de primitives B_I^j et B_I^k :

$$p(\bar{d}\bar{d}) = \frac{N_d^j N_{\bar{d}}^k}{N_d^j N_{\bar{d}}^k + N_{\bar{d}}^j N_d^k + N_d^j N_d^k}, \quad (5.46)$$

$$p(\bar{d}d) = \frac{N_{\bar{d}}^j N_d^k}{N_d^j N_{\bar{d}}^k + N_{\bar{d}}^j N_d^k + N_d^j N_d^k}, \quad (5.47)$$

$$p(d\bar{d}) = \frac{N_d^j N_{\bar{d}}^k}{N_d^j N_{\bar{d}}^k + N_{\bar{d}}^j N_d^k + N_d^j N_d^k}. \quad (5.48)$$

Dans le cadre de l'approximation précédente, consistant à ne considérer plus que les couples de pixels dd et $\bar{d}\bar{d}$, la log-vraisemblance des couples de pixels observés dans le voisinage $\nu_t^j \times \nu_t^k$ du pixel testé se réécrit :

$$\forall (x_t, y_t) \in \mathcal{E}_I, \ell^{jk}(x_t, y_t) = \ell_{dd}^{jk}(x_t, y_t) + \ell_{\bar{d}\bar{d}}^{jk}(x_t, y_t), \quad (5.49)$$

où $\ell_{\overline{dd}}^{jk}$ désigne la log-vraisemblance des couples de pixels \overline{dd} et s'exprime comme suit :

$$\begin{aligned} \forall (x_t, y_t) \in \mathcal{E}_I, \ell_{\overline{dd}}^{jk}(x_t, y_t) = & \sum_{\substack{(x^j, y^j) \in \nu_t^j, B_t^j(x^j, y^j)=1 \\ (x^k, y^k) \in \nu_t^k, B_t^k(x^k, y^k)=0}} \ln p_{\overline{dd}}^{jk}(x^j, y^j, x^k, y^k | x_t, y_t) \\ & + \sum_{\substack{(x^j, y^j) \in \nu_t^j, B_t^j(x^j, y^j)=0 \\ (x^k, y^k) \in \nu_t^k, B_t^k(x^k, y^k)=1}} \ln p_{\overline{dd}}^{jk}(x^j, y^j, x^k, y^k | x_t, y_t) \\ & + \sum_{\substack{(x^j, y^j) \in \nu_t^j, B_t^j(x^j, y^j)=0 \\ (x^k, y^k) \in \nu_t^k, B_t^k(x^k, y^k)=0}} \ln p_{\overline{dd}}^{jk}(x^j, y^j, x^k, y^k | x_t, y_t). \end{aligned} \quad (5.50)$$

En remplaçant les log-vraisemblances $\ell_{\overline{dd}}^{jk}$ et $\ell_{\overline{dd}}^{jk}$ par leurs expressions respectives et après un réarrangement des différents termes de la somme, l'équation (5.49) peut finalement se réécrire de la manière suivante :

$$\begin{aligned} \ell^{jk}(x_t, y_t) = & \sum_{(x^j, y^j, x^k, y^k) \in \nu_t^j \times \nu_t^k} \ln p_{\overline{dd}}^{jk}(x^j, y^j, x^k, y^k | x_t, y_t) \\ & + \sum_{\substack{(x^j, y^j) \in \nu_t^j, B_t^j(x^j, y^j)=1 \\ (x^k, y^k) \in \nu_t^k, B_t^k(x^k, y^k)=1}} \ln \left[\frac{p_{\overline{dd}}^{jk}(x^j, y^j, x^k, y^k | x_t, y_t)}{p_{\overline{dd}}^{jk}(x^j, y^j, x^k, y^k | x_t, y_t)} \right]. \end{aligned} \quad (5.51)$$

Nous constatons, à partir de cette dernière équation, que la log-vraisemblance ℓ^{jk} peut être calculée au pixel (x_t, y_t) en sommant deux termes :

- le premier terme correspond à la log-vraisemblance qui serait obtenue si tous les couples de pixels observés au sein de l'ensemble $\nu_t^j \times \nu_t^k$ étaient des couples \overline{dd} . Ce terme, dont la valeur est la même pour tous les pixels testés (x_t, y_t) et dont le calcul est par conséquent réalisé une seule fois, est obtenu en sommant les valeurs de $\ln p_{\overline{dd}}^{jk}$ sur l'intégralité de son support de définition ;
- le second terme correspond à une correction du premier terme, où la valeur de la log-vraisemblance des couples de vecteurs dd est improprement calculée à l'aide de la densité de probabilité $p_{\overline{dd}}^{jk}$. Ces log-vraisemblances sont donc remplacées par des log-vraisemblances calculées à l'aide de la densité de probabilité p_{dd}^{jk} .

Finalement, le calcul de la log-vraisemblance au pixel (x_t, y_t) est réalisé en ne cherchant que les couples de vecteurs dd au sein de l'ensemble $\nu_t^j \times \nu_t^k$, recherche dont la complexité calculatoire est plus faible qu'un parcours exhaustif de l'ensemble $\nu_t^j \times \nu_t^k$ puisque le parcours de l'ensemble ν_t^j n'est réalisé que lorsque qu'un pixel est détecté lors du parcours de l'ensemble ν_t^j .

5.3.3.3 Résultats en imagerie optique

Nous présentons à la figure 5.28 (a) la carte de log-vraisemblance ℓ^{01} établie en considérant conjointement les images binaires de primitive B_I^0 et B_I^1 simulée des figures 5.14 et 5.16. La carte de détection binaire D_I^{01} de la figure 5.28 (b) est alors obtenue par seuillage supervisé de cette carte de log-vraisemblance. Les images de la figure 5.29 présentent des résultats analogues établis à partir des images binaires B_I^0 et B_I^1 réelles des figures 5.15 et 5.17.

Les résultats obtenus sur données simulées permettent de valider l'approche proposée et montrent notamment que les différentes approximations mises en œuvre, lors du calcul des densités de probabilité de présence et de la log-vraisemblance, pour alléger la complexité calculatoire conduisent à des résultats tout à fait satisfaisants.

Malgré la non détection d'un des neuf bâtiments présents dans la scène, en raison des phénomènes de recouvrement entre les bâtiments B et C , les résultats établis à partir des données réelles sont également satisfaisants, en comparaison de ceux établis aux figures 5.15 et 5.17 à partir des mêmes images binaires considérées indépendamment. Ces résultats sont d'autant plus satisfaisants que les

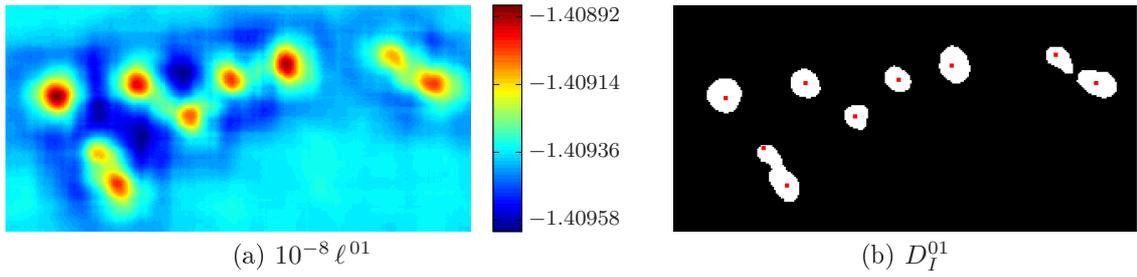


Figure 5.28 – Cartes de log-vraisemblance ℓ^{01} (a) et cartes binaires de détection D_I^{01} (b) établies en considérant conjointement les images binaires B_I^0 et B_I^1 simulées des figures 5.14 et 5.16.

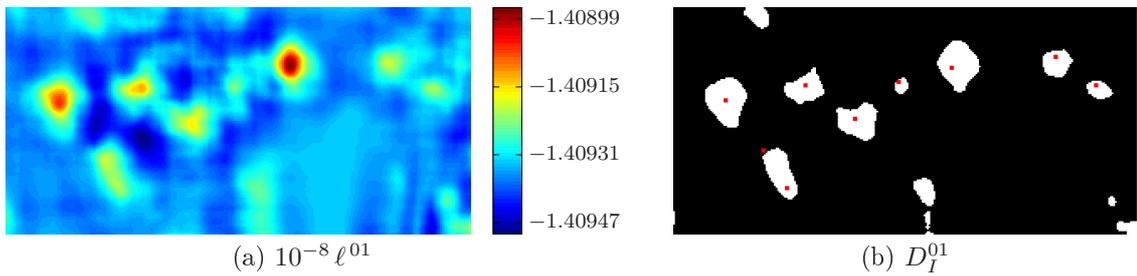


Figure 5.29 – Cartes de log-vraisemblance ℓ^{01} (a) et cartes binaires de détection D_I^{01} (b) établies en considérant conjointement les images binaires B_I^0 et B_I^1 réelles des figures 5.15 et 5.17.

performances des détecteurs au moyen desquels ont été établies les images binaires B_I^0 et B_I^1 sont limitées, ce qui illustre la robustesse de l'approche proposée à une extraction bruitée des primitives caractéristiques des bâtiments.

5.3.3.4 Résultats en imagerie RSO

Pour finir, nous présentons aux figures 5.30 et 5.31 les cartes de log-vraisemblance ℓ^{01} et les cartes binaires de détection D_I^{01} établies en imagerie RSO à partir des images binaires de primitives simulées et réelles des figures 5.21, 5.23, 5.22 et 5.24.

L'interprétation que nous faisons des résultats obtenus est la même que celle faite en imagerie optique à la sous-section précédente. Malgré la présence de fausses alarmes et la non détection d'un des neuf bâtiments présents dans la scène, nous pouvons constater les bonnes performances obtenues lors de la mise en œuvre de la méthode de détection sur données réelles. Ces résultats, qui sont à comparer à ceux obtenus aux figures 5.22 et 5.24, montrent, comme en imagerie optique, la robustesse de l'approche proposée à une extraction bruitée des primitives caractéristiques des bâtiments.

5.4 Estimation des paramètres d'un bâtiment

Nous avons présenté, à la partie précédente, une méthode de détection grossière permettant de définir un ensemble de régions d'intérêt au sein de l'espace image \mathcal{E}_I , régions qui contiennent possiblement les points de référence de bâtiments.

Dans cette partie, nous considérons un pixel (x_r, y_r) de l'image I identifié, lors de l'étape de détection grossière, comme étant possiblement le pixel de référence d'un bâtiment et déterminons les paramètres α , w_1 , w_2 et h de ce bâtiment. Pour ce faire, nous mettons en œuvre une méthode d'inversion de modèle par maximisation de vraisemblance qui s'appuie, comme pour la détection

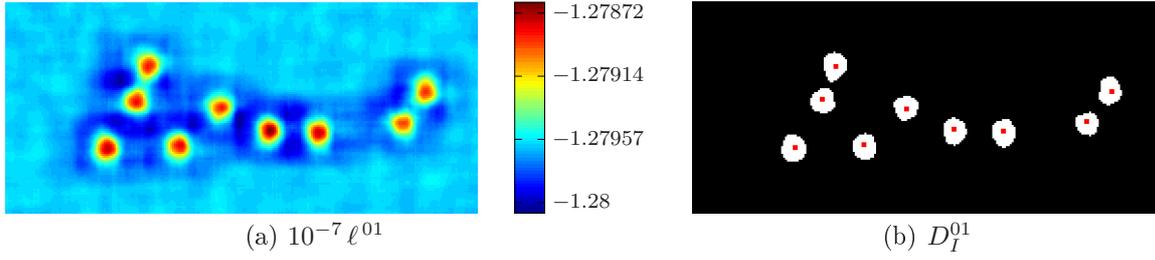


Figure 5.30 – Cartes de détection ℓ^{01} (a) et cartes binaires de détection D_I^{01} (b) établies en considérant conjointement les images binaires B_I^0 et B_I^1 simulées des figures 5.21 et 5.23.

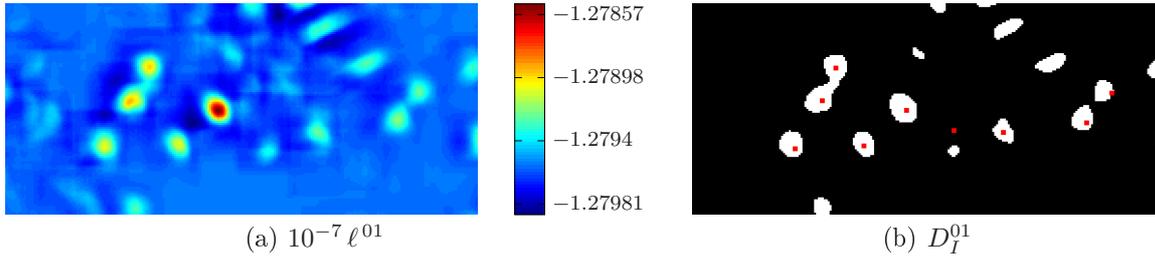


Figure 5.31 – Cartes de détection ℓ^{01} (a) et cartes binaires de détection D_I^{01} (b) établies en considérant conjointement les images binaires B_I^0 et B_I^1 réelles des figures 5.22 et 5.24.

grossière, sur les densités de probabilité de présence des pixels détectés et non détectés dans les images binaires de primitives.

5.4.1 Positionnement du problème

En supposant, comme précédemment, que les pixels détectés dans l'image binaire B_I^j , constituent des réalisations indépendantes des variables aléatoires X_d^j et Y_d^j , la log-vraisemblance ℓ_d^j des pixels détectés dans le voisinage ν_r^j du pixel de référence, pour un jeu de paramètres (α, w_1, w_2, h) donné, s'exprime comme suit :

$$\forall (\alpha, w_1, w_2, h) \in \mathcal{E}_\alpha \times \mathcal{E}_{w_1} \times \mathcal{E}_{w_2} \times \mathcal{E}_h, \ell_d^j(\alpha, w_1, w_2, h) = \sum_{\substack{(x,y) \in \nu_r^j \\ B_I^j(x,y)=1}} \ln p_d^j(x, y | x_r, y_r, \alpha, w_1, w_2, h), \quad (5.52)$$

où $p_d^j(x, y | x_r, y_r, \alpha, w_1, w_2, h)$ désigne la densité de probabilité de présence des pixels détectés dans le voisinage ν_r^j , conditionnellement à l'ensemble des paramètres du bâtiment repéré par le pixel de référence (x_r, y_r) . L'expression de cette densité de probabilité est donnée à l'équation (5.14).

L'estimateur du maximum de vraisemblance consiste à déterminer le jeu de paramètres $(\hat{\alpha}, \hat{w}_1, \hat{w}_2, \hat{h})$ pour lequel la vraisemblance, dont l'expression est donnée à l'équation précédente, est maximale :

$$(\hat{\alpha}, \hat{w}_1, \hat{w}_2, \hat{h}) = \underset{\substack{\alpha \in \mathcal{E}_\alpha, w_1 \in \mathcal{E}_{w_1} \\ w_2 \in \mathcal{E}_{w_2}, h \in \mathcal{E}_h}}{\operatorname{argmax}} \ell_d^j(\alpha, w_1, w_2, h). \quad (5.53)$$

Ainsi, l'estimation des paramètres (α, w_1, w_2, h) par maximisation de vraisemblance correspond à un problème d'optimisation qu'il est, en pratique, relativement complexe de résoudre en un temps raisonnable. En effet, la méthode qui consiste à effectuer une recherche exhaustive dans l'ensemble des paramètres $\mathcal{E}_\alpha \times \mathcal{E}_{w_1} \times \mathcal{E}_{w_2} \times \mathcal{E}_h$ conduit certes au maximum global de la vraisemblance mais implique une complexité calculatoire élevée. Une autre solution consiste à mettre en œuvre, comme au chapitre

précédent, une méthode d'optimisation par descente de gradient, ce qui évite une recherche exhaustive dans l'ensemble des paramètres mais peut conduire à des *maxima* locaux de la vraisemblance.

La méthode que nous proposons dans cette partie repose sur une estimation successive des différents paramètres, de sorte que le problème d'optimisation dans l'ensemble quadridimensionnel $\mathcal{E}_\alpha \times \mathcal{E}_{w_1} \times \mathcal{E}_{w_2} \times \mathcal{E}_h$ est décomposé en quatre problèmes d'optimisation distincts au sein des ensembles unidimensionnels \mathcal{E}_α , \mathcal{E}_{w_1} , \mathcal{E}_{w_2} et \mathcal{E}_h . Ces optimisations peuvent alors être réalisées par recherche exhaustive en un temps raisonnable.

5.4.2 Estimation par maximisation de vraisemblance

Dans cette section, θ désigne indifféremment l'un des paramètres que nous souhaitons estimer et \mathcal{E}_θ correspond à l'ensemble des valeurs admissibles de ce paramètre. Nous notons $\vec{\Phi}_c$ le vecteur, appartenant à l'ensemble \mathcal{E}_{Φ_c} , qui regroupe l'ensemble des paramètres connus, paramètres dont l'estimation a été réalisée avant celle du paramètre θ . De même, le vecteur $\vec{\Phi}_i$, appartenant à l'ensemble \mathcal{E}_{Φ_i} , regroupe l'ensemble des paramètres inconnus, paramètres dont l'estimation est réalisée après celle du paramètre θ .

L'estimation par maximisation de vraisemblance du paramètre θ à partir de l'image binaire B_I^j nécessite que soit connue la densité de probabilité conjointe des variables aléatoires X_d^j et Y_d^j , conditionnellement au paramètre θ et au vecteur de paramètres $\vec{\Phi}_c$. Cette densité de probabilité est obtenue, à partir de la densité de probabilité $p_d^j(x, y|x_r, y_r, \alpha, w_1, w_2, h)$ dont l'expression a été établie à la sous-section 5.3.1.1, par intégration sur l'ensemble des paramètres inconnus \mathcal{E}_{Φ_i} :

$$\forall (x, y) \in \nu_r^j, p_d^j(x, y|x_r, y_r, \theta, \vec{\Phi}_c) = \int_{\mathcal{E}_{\Phi_i}} \underbrace{p_d^j(x, y|x_r, y_r, \theta, \vec{\Phi}_c, \vec{\Phi}_i)}_{p_d^j(x, y|x_r, y_r, \alpha, w_1, w_2, h)} p(\vec{\Phi}_i) d\vec{\Phi}_i, \quad (5.54)$$

où la probabilité *a priori* $p(\vec{\Phi}_i)$ est supposée uniforme sur l'ensemble des paramètres inconnus \mathcal{E}_{Φ_i} . Le calcul de cette densité de probabilité, pour des valeurs données de θ et $\vec{\Phi}_c$, est réalisé, contrairement à ce qui a été fait à la partie 5.3, au moyen d'une méthode de Riemann. Comme nous le voyons par la suite, la dimension des domaines d'intégration n'excède jamais deux, de sorte que la densité de probabilité $p_d^j(x, y|x_r, y_r, \theta, \vec{\Phi}_c)$ peut être calculée avec des pas d'intégration suffisamment petits pour que les artefacts mis en évidence à la sous-section 5.3.1.2 n'apparaissent pas, tout en conservant une complexité calculatoire raisonnable.

Une fois la densité de probabilité de présence $p_d^j(x, y|x_r, y_r, \theta, \vec{\Phi}_c)$ calculée, la vraisemblance des pixels détectés dans le voisinage ν_r^j du pixel de référence, étant données les valeurs des paramètres θ et $\vec{\Phi}_c$, peut être calculée selon l'expression suivante :

$$\forall \theta \in \mathcal{E}_\theta, \ell_d^j(\theta) = \sum_{\substack{(x, y) \in \nu_r^j \\ B_I^j(x, y) = 1}} \ln p_d^j(x, y|x_r, y_r, \theta, \vec{\Phi}_c). \quad (5.55)$$

La valeur $\hat{\theta}$ du paramètre θ pour laquelle cette vraisemblance est maximale, valeur qui est déterminée par recherche exhaustive au sein de l'ensemble \mathcal{E}_θ , correspond alors à l'estimation par maximisation de vraisemblance de ce paramètre :

$$\hat{\theta} = \operatorname{argmax}_{\theta \in \mathcal{E}_\theta} \ell_d^j(\theta). \quad (5.56)$$

5.4.3 Stratégie d'estimation

Nous avons introduit, à la section précédente, la méthode au moyen de laquelle sont estimés les paramètres α , w_1 , w_2 et h d'un bâtiment dont la position (x_r, y_r) du point de référence dans l'image est connue. Dans cette section, nous établissons successivement, en imagerie optique et en imagerie RSO, les stratégies d'estimation qui définissent :

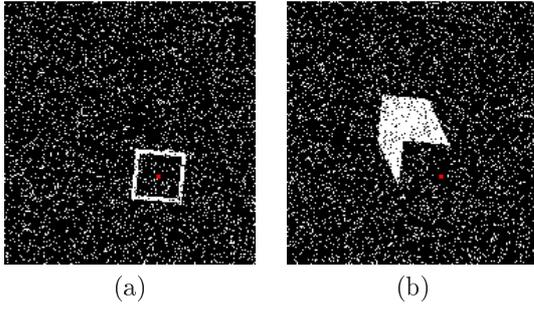


Figure 5.32 – Images binaires B_I^0 (a) et B_I^1 (b) du bâtiment G de la scène d'étude, simulées pour $p_D^0 = 0.9$, $p_{FA}^0 = 0.1$ et $p_D^1 = 0.9$, $p_{FA}^1 = 0.1$.

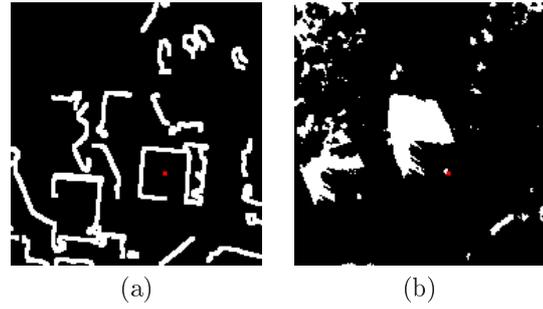


Figure 5.33 – Images binaires B_I^0 (a) et B_I^1 (b) du bâtiment G de la scène d'étude, extraites des images binaires de la figure 5.3.

- pour chaque paramètre, l'image binaire à partir de laquelle est estimé ce paramètre ;
- l'ordre suivant lequel sont estimés les différents paramètres.

La détermination de l'image binaire à partir de laquelle est estimé un paramètre repose sur l'observabilité de ce paramètre au sein des différentes images binaires. À titre d'exemple, l'image binaire B_I^0 résultant de l'extraction des contours de toit en imagerie optique ne contient aucune information sur la hauteur des bâtiments présents dans la scène, de sorte que la hauteur des bâtiments n'est pas observable dans l'image B_I^0 . À l'inverse, la hauteur des bâtiments est observable dans l'image B_I^1 résultant de l'extraction des ombres puisque la longueur de ces ombres est reliée à la hauteur des bâtiments par la relation (2.18).

En ce qui concerne l'ordre suivant lequel sont estimés les paramètres, il est déterminé empiriquement et résulte de l'observation des précisions d'estimation obtenues pour les différentes combinaisons possibles.

Dans les deux sous-sections suivantes, nous présentons les stratégies établies pour l'estimation des paramètres en imagerie optique et en imagerie RSO, stratégies d'estimation qui sont illustrées sur données simulées et sur données réelles.

5.4.3.1 Imagerie optique

En imagerie optique, la stratégie d'estimation mise en œuvre se décompose en trois étapes successives :

1. Estimation successive des paramètres w_1 et w_2 à partir de l'image binaire B_I^0 résultant de l'extraction des contours de toit ;
2. Estimation de l'orientation α du bâtiment à partir de l'image B_I^0 en utilisant les valeurs des paramètres w_1 et w_2 estimées à la première étape ;
3. Estimation de la hauteur h du bâtiment à partir de l'image binaire B_I^1 , la densité de probabilité de présence nécessaire à son estimation étant calculée conditionnellement à l'ensemble des paramètres précédemment estimés.

Dans la suite de cette sous-section, nous détaillons les trois étapes précédentes et mettons en œuvre la stratégie proposée sur les images binaires B_I^0 et B_I^1 des figures 5.32 et 5.33, qui correspondent au bâtiment G de la scène d'étude présentée au chapitre 2.

Les images B_I^0 et B_I^1 de la figure 5.32 ont été établies par simulation pour des détecteurs dont les probabilités de détection et de fausse alarme valent respectivement $p_D^0 = 0.90$, $p_{FA}^0 = 0.10$ et $p_D^1 = 0.90$, $p_{FA}^1 = 0.10$.

Les images B_I^0 et B_I^1 de la figure 5.33 ont été extraites à partir des images binaires des figures 5.3 (a) et 5.3 (c), dont nous rappelons qu'elles sont respectivement caractérisées par les probabilités de détection et de fausse alarme suivantes : $p_D^0 = 0.63$, $p_{FA}^0 = 0.17$ et $p_D^1 = 0.73$, $p_{FA}^1 = 0.10$.

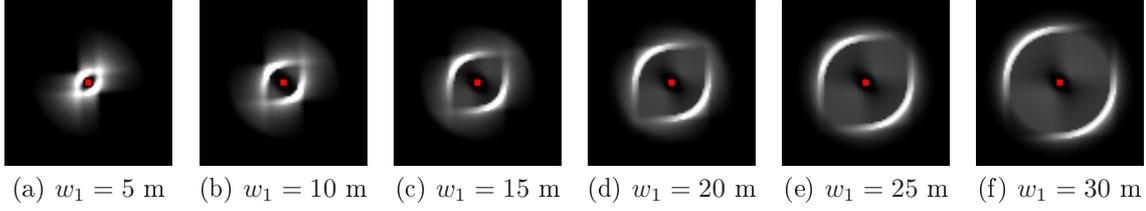


Figure 5.34 – Densités de probabilité de présence conditionnelles $p_d^0(x, y|x_r, y_r, w_1)$ pour l'estimation du paramètre w_1 .

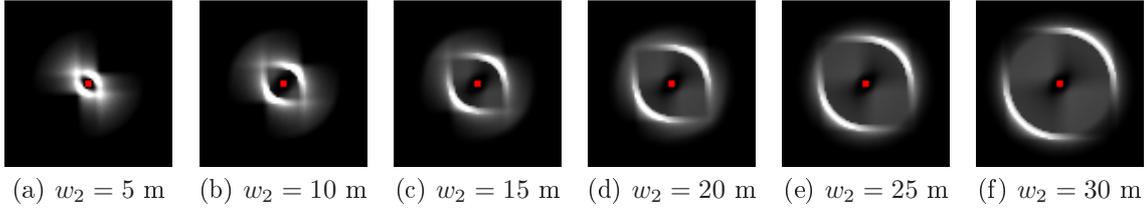


Figure 5.35 – Densités de probabilité de présence conditionnelles $p_d^0(x, y|x_r, y_r, w_2)$ pour l'estimation du paramètre w_2 .

Estimation des paramètres w_1 et w_2

Par la suite, nous désignons les paramètres w_1 et w_2 au moyen des notations w_k et $w_{\bar{k}}$, où $k \in \{1; 2\}$ et \bar{k} est le complémentaire de k dans l'ensemble $\{1; 2\}$.

Étant donné que le paramètre w_k est le premier à être estimé, le vecteur $\vec{\Phi}_c$, introduit à la section 5.4.2, qui regroupe l'ensemble des paramètres précédemment estimés, n'est pas défini. En revanche, le vecteur $\vec{\Phi}_i$ contient les paramètres α , $w_{\bar{k}}$ et h , de sorte que l'équation (5.54) se réécrit comme suit, dans le cas d'une estimation à partir de l'image binaire B_I^0 :

$$\forall k \in \{1; 2\}, \forall (x, y) \in \nu_r^0, p_d^0(x, y|x_r, y_r, w_k) = \int_{\mathcal{E}_\alpha} \int_{\mathcal{E}_{w_{\bar{k}}}} \int_{\mathcal{E}_h} p_d^0(x, y|x_r, y_r, \alpha, w_k, w_{\bar{k}}, h) p(\alpha, w_{\bar{k}}, h) d\alpha dw_{\bar{k}} dh. \quad (5.57)$$

La forme du contour de toit d'un bâtiment ne dépendant pas de la hauteur de ce dernier, nous pouvons supprimer, comme nous l'avons fait à la sous-section 5.3.1.2, le conditionnement de la densité de probabilité $p_d^0(x, y|x_r, y_r, \alpha, w_k, w_{\bar{k}}, h)$ par rapport au paramètre h . L'équation précédente se réécrit alors :

$$\forall k \in \{1; 2\}, \forall (x, y) \in \nu_r^0, p_d^0(x, y|x_r, y_r, w_k) = \int_{\mathcal{E}_\alpha} \int_{\mathcal{E}_{w_{\bar{k}}}} p_d^0(x, y|x_r, y_r, \alpha, w_k, w_{\bar{k}}) p(\alpha, w_{\bar{k}}) d\alpha dw_{\bar{k}}. \quad (5.58)$$

Nous présentons aux figures 5.34 et 5.35 les densités de probabilité de présence $p_d^0(x, y|x_r, y_r, w_1)$ et $p_d^0(x, y|x_r, y_r, w_2)$ nécessaires à l'estimation des paramètres w_1 et w_2 , densités de probabilité qui résultent d'une intégration par méthode de Riemann sur les ensembles $\mathcal{E}_\alpha = [0; 90]$ et $\mathcal{E}_{w_{\bar{k}}} = [5; 30]$ avec des pas d'intégration valant respectivement $\Delta\alpha = 1$ degré, $\Delta w_{\bar{k}} = 0.5$ m. Notons que, dans le cadre de l'hypothèse d'uniformité des effets géométriques liés à la projection des bâtiments dans l'image, ces densités de probabilité ne sont calculées qu'une seule fois pour toute l'image.

Maintenant que nous disposons des densités de probabilité $p_d^0(x, y|x_r, y_r, w_1)$ et $p_d^0(x, y|x_r, y_r, w_2)$, les paramètres w_1 et w_2 peuvent être estimés par maximisation de vraisemblance selon les équations (5.55) et (5.56).

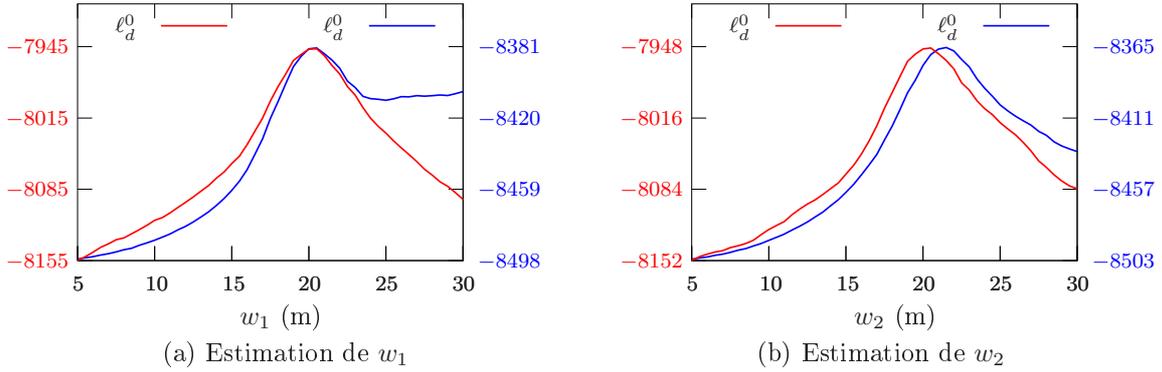


Figure 5.36 – Courbes de log-vraisemblance calculées à partir de l'image binaire B_I^0 simulée de la figure 5.32 (courbes rouges) et de l'image binaire B_I^0 réelle de la figure 5.33 (courbes bleues) pour l'estimation des paramètres w_1 (a) et w_2 (b). Les valeurs estimées des paramètres w_1 et w_2 valent respectivement $\hat{w}_1 = 20.3$ m, $\hat{w}_2 = 20.3$ m pour l'image simulée et $\hat{w}_1 = 20.3$ m, $\hat{w}_2 = 21.4$ m pour l'image réelle.

Nous présentons à la figure 5.36 les courbes de vraisemblance établies à partir des images binaires B_I^0 simulée et réelle des figures 5.32 et 5.33 pour l'estimation des paramètres w_1 et w_2 . Pour le calcul de ces courbes, le point de référence du bâtiment dans l'image est supposé connu et est représenté dans les images B_I^0 par un point de couleur rouge.

Les courbes précédentes résultant d'un calcul de la log-vraisemblance tous les 0.5 m, la moyenne quadratique des erreurs d'estimation liées à l'échantillonnage des courbes est de 0.14 m^1 . Par la suite, la valeur estimée des paramètres correspond à la valeur qui maximise le polynôme d'ordre 2 interpolant le point maximum de la courbe de log-vraisemblance et ses deux voisins immédiats, ce qui permet d'augmenter la précision avec laquelle sont estimés les paramètres.

La méthode proposée conduit finalement à des erreurs d'estimation sur les paramètres w_1 et w_2 qui valent respectivement 0.3 m et 0.3 m dans le cas de l'image simulée puis 0.3 m et 1.4 m dans le cas de l'image réelle, ce qui valide l'approche mise en œuvre pour l'estimation de ces paramètres.

Pour des bâtiments dont l'orientation α est proche de 0 degré ou de 90 degrés, l'approche proposée est cependant mise en défaut, comme l'illustre la figure 5.38 qui représente les courbes de log-vraisemblance calculées à partir de l'image binaire B_I^0 simulée de la figure 5.37 pour l'estimation des paramètres w_1 et w_2 . Chacune des deux courbes présente deux *maxima* localisés approximativement au niveau des valeurs vraies des paramètres w_1 et w_2 , de sorte que l'on observe un phénomène d'ambiguïté entre les paramètres w_1 et w_2 . Alors que la courbe de couleur bleue conduit à une estimation correcte de la valeur du paramètre w_2 , la valeur du paramètre w_1 , estimée au moyen la courbe de couleur rouge, est erronée puisqu'elle est proche de la valeur vraie du paramètre w_2 .

Comme illustré à la figure 5.39, ce phénomène d'ambiguïté entre les paramètres w_1 et w_2 , pour des orientations proches de 0 degré et de 90 degrés, s'explique par la non continuité du modèle paramétrique permettant de décrire les bâtiments en $\alpha = 0$ degré et $\alpha = 90$ degrés. Un bâtiment dont l'orientation est proche de 0 degré ou de 90 degrés peut en effet être décrit par deux jeux de paramètres distincts, pour lesquels les orientations valent respectivement 0 degré et 90 degrés et les valeurs des paramètres w_1 et w_2 sont inversées.

Afin de contrecarrer ce phénomène d'ambiguïté, nous proposons d'estimer les paramètres w_1 et w_2 pour chacune des deux hypothèses \mathcal{H}_α^1 et \mathcal{H}_α^2 suivantes :

- \mathcal{H}_α^1 : « L'orientation α du bâtiment dont le point de référence est situé en (x_r, y_r) appartient à l'intervalle $\mathcal{E}_\alpha^1 = [0; 45[$ » ;

¹L'écart type d'une loi uniforme définie sur le support $[a; b]$ vaut $(b - a)/\sqrt{12}$. La moyenne quadratique des erreurs d'estimation liées à l'échantillonnage vaut donc $0.5/\sqrt{12} = 0.14$ m.

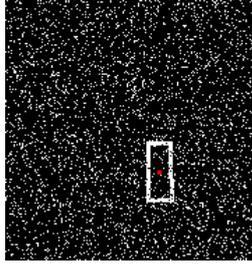


Figure 5.37 – Image binaire B_I^0 simulée pour un bâtiment dont l'estimation des paramètres w_1 et w_2 est sujette à un phénomène d'ambiguïté ($X_C = 703159$ m, $Y_C = 4795086$ m, $\alpha = 89$ degrés, $w_1 = 10$ m, $w_2 = 25$ m, $h = 22$ m).

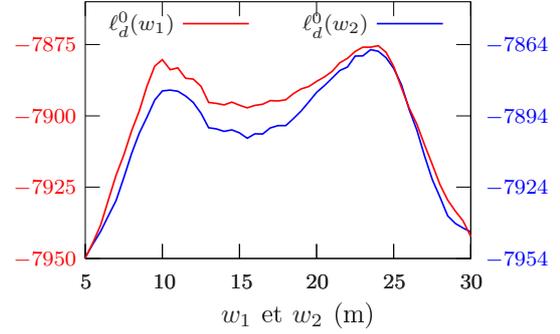


Figure 5.38 – Courbes de log-vraisemblance calculées à partir de l'image binaire B_I^0 simulée de la figure 5.37. Les valeurs estimées des paramètres w_1 et w_2 valent respectivement $\hat{w}_1 = 23.8$ m et $\hat{w}_2 = 23.6$ m.

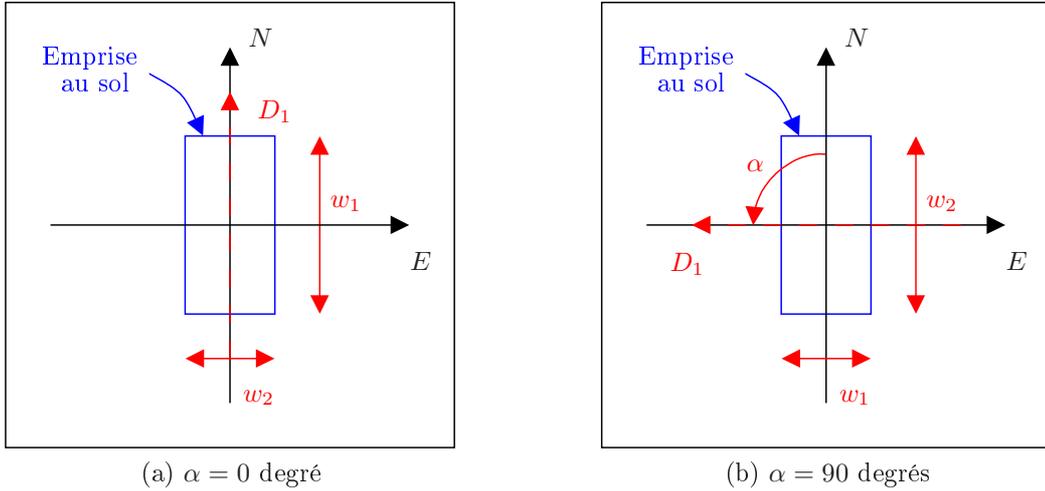


Figure 5.39 – Ambiguïté entre les paramètres w_1 et w_2 pour des bâtiments dont l'orientation α est proche de 0 degré ou 90 degrés. Deux jeux de paramètres distincts, pour lesquels les orientations valent respectivement $\alpha = 0$ degré (a) et $\alpha = 90$ degrés (b) peuvent décrire le même bâtiment, les valeurs des paramètres w_1 et w_2 étant inversées au sein de ces deux jeux de paramètres.

- \mathcal{H}_α^2 : « L'orientation α du bâtiment dont le point de référence est situé en (x_r, y_r) appartient à l'intervalle $\mathcal{E}_\alpha^2 = [45; 90[$ » ;

où les ensembles \mathcal{E}_α^1 et \mathcal{E}_α^2 forment une partition exhaustive de l'ensemble \mathcal{E}_α .

L'estimation du paramètre w_k sous l'hypothèse \mathcal{H}_α^l (avec $l \in \{1; 2\}$) est réalisée au moyen de la densité de probabilité de présence $p_d^0(x, y | x_r, y_r, w_k, \mathcal{H}_\alpha^l)$ dont l'expression est donnée à l'équation suivante :

$$\forall k \in \{1; 2\}, \forall l \in \{1; 2\}, \forall (x, y) \in \nu_r^0, p_d^0(x, y | x_r, y_r, w_k, \mathcal{H}_\alpha^l) = \int_{\mathcal{E}_\alpha^l} \int_{\mathcal{E}_{w_{\bar{k}}}} p_d^0(x, y | x_r, y_r, \alpha, w_k, w_{\bar{k}}) p(\alpha, w_{\bar{k}}) d\alpha dw_{\bar{k}}, \quad (5.59)$$

la seule différence avec l'expression de la densité de probabilité de présence $p_d^0(x, y | x_r, y_r, w_k)$, dont l'expression a été établie à l'équation (5.57), étant le support d'intégration relatif à la variable α .

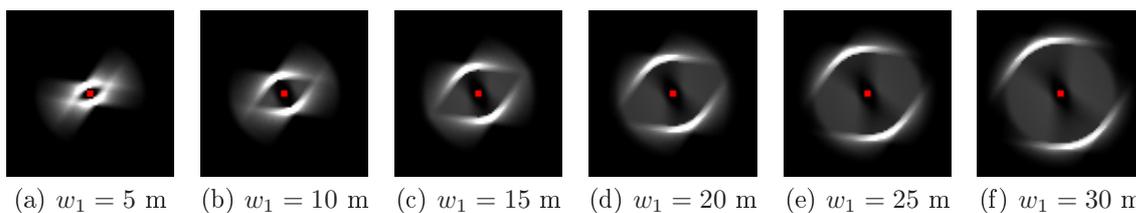


Figure 5.40 – Densités de probabilité de présence conditionnelles $p_d^0(x, y|x_r, y_r, w_1, \mathcal{H}_\alpha^1)$ pour l'estimation du paramètre w_1 .

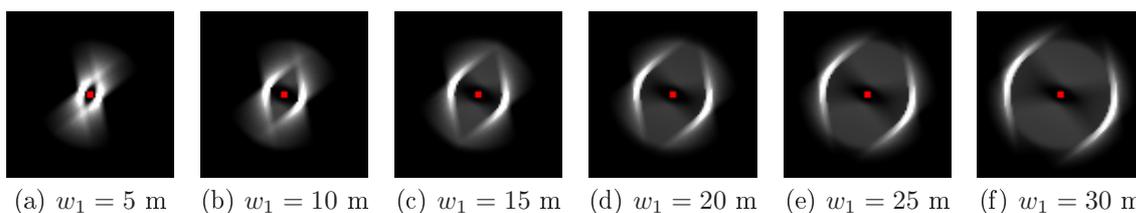


Figure 5.41 – Densités de probabilité de présence conditionnelles $p_d^0(x, y|x_r, y_r, w_1, \mathcal{H}_\alpha^2)$ pour l'estimation du paramètre w_1 .

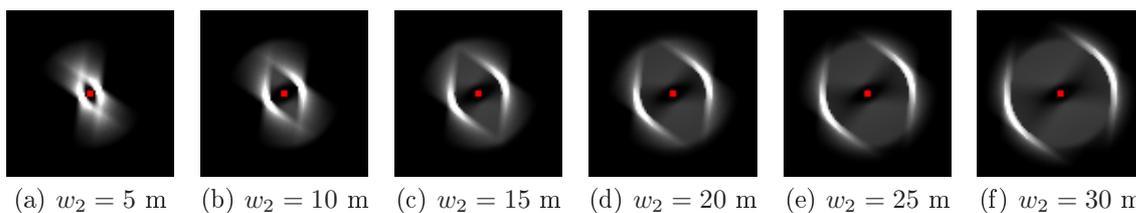


Figure 5.42 – Densités de probabilité de présence conditionnelles $p_d^0(x, y|x_r, y_r, w_2, \mathcal{H}_\alpha^1)$ pour l'estimation du paramètre w_2 .

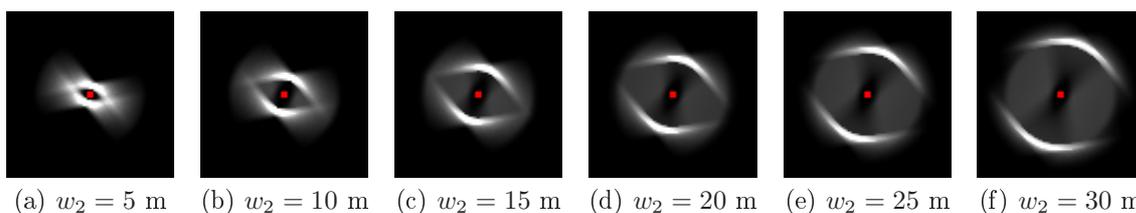


Figure 5.43 – Densités de probabilité de présence conditionnelles $p_d^0(x, y|x_r, y_r, w_2, \mathcal{H}_\alpha^2)$ pour l'estimation du paramètre w_2 .

Les figures 5.40 et 5.41 présentent les densités de probabilité $p_d^0(x, y|x_r, y_r, w_1, \mathcal{H}_\alpha^1)$ et $p_d^0(x, y|x_r, y_r, w_1, \mathcal{H}_\alpha^2)$ au moyen desquelles sont réalisées les estimations du paramètre w_1 pour chacune des deux hypothèses \mathcal{H}_α^1 et \mathcal{H}_α^2 . Les figures 5.42 et 5.43 présentent les densités de probabilité analogues pour l'estimation du paramètre w_2 .

Nous disposons désormais, pour chacun des paramètres w_k , de deux valeurs estimées, notées \hat{w}_k^1 et \hat{w}_k^2 , qui résultent respectivement de la maximisation des log-vraisemblances $\ell_d^0(w_k, \mathcal{H}_\alpha^1)$ et $\ell_d^0(w_k, \mathcal{H}_\alpha^2)$.

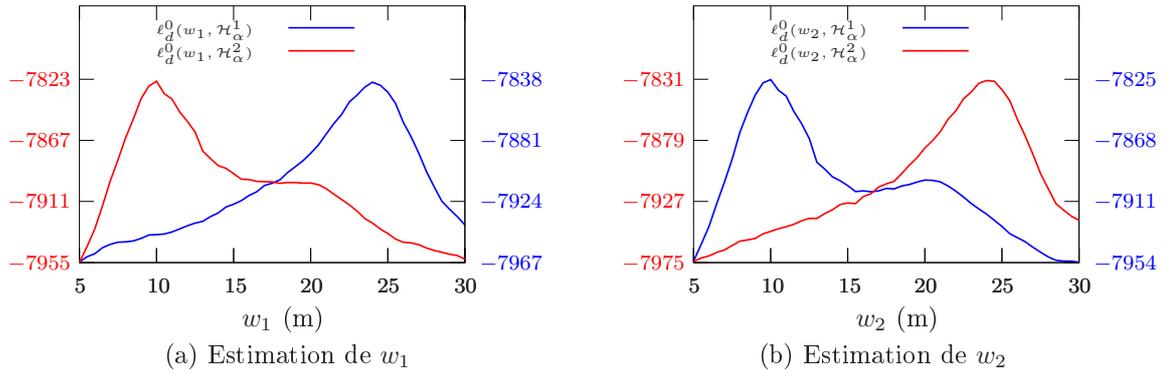


Figure 5.44 – Courbes de log-vraisemblance calculées à partir de l'image binaire B_I^0 simulée de la figure 5.37 (a) pour l'estimation des paramètres w_1 (a) et w_2 (b). Les valeurs estimées des paramètres w_1 et w_2 valent respectivement $\hat{w}_1^0 = 24.1$ m, $\hat{w}_1^1 = 9.9$ m et $\hat{w}_2^0 = 9.9$ m et $\hat{w}_2^1 = 24.2$ m.

calculées comme suit :

$$\forall k \in \{1; 2\}, \forall w_k \in \mathcal{E}_{w_k}, \forall l \in \{1; 2\}, \ell_d^0(w_k, \mathcal{H}_\alpha^l) = \sum_{\substack{(x,y) \in \nu_r^j \\ B_I^0(x,y)=1}} \ln p_d^0(x, y | x_r, y_r, w_k, \mathcal{H}_\alpha^l). \quad (5.60)$$

Finalement, le choix entre les deux couples de valeurs estimées $(\hat{w}_1^1, \hat{w}_2^1)$ et $(\hat{w}_1^2, \hat{w}_2^2)$ est fait en conservant le plus vraisemblable, c'est-à-dire celui pour lequel la somme des log-vraisemblances $\ell_d^0(\hat{w}_1^l, \mathcal{H}_\alpha^l) + \ell_d^0(\hat{w}_2^l, \mathcal{H}_\alpha^l)$ est la plus importante.

La mise en œuvre de cette approche, qui permet de s'affranchir du phénomène d'ambiguïté observé précédemment, conduit aux courbes de log-vraisemblance de la figure 5.44, les courbes bleues et rouges étant relatives respectivement aux hypothèses \mathcal{H}_α^1 et \mathcal{H}_α^2 . Ces courbes, qui ont été calculées à partir de l'image binaire B_I^0 de la figure 5.37, conduisent aux valeurs estimées des paramètres w_1 et w_2 suivantes : $\hat{w}_1^1 = 24.1$ m, $\hat{w}_1^2 = 9.9$ m, $\hat{w}_2^1 = 9.9$ m et $\hat{w}_2^2 = 24.2$ m. Finalement, l'hypothèse retenue est l'hypothèse \mathcal{H}_α^2 , qui correspond à la bonne hypothèse ($\alpha = 89$ degrés pour le bâtiment de la figure 5.37), et les valeurs estimées des paramètres w_1 et w_2 valent donc $\hat{w}_1 = 9.9$ m et $\hat{w}_2 = 24.2$ m. Notons pour finir que si l'hypothèse retenue avait été l'hypothèse \mathcal{H}_α^1 , les valeurs estimées des paramètres w_1 et w_2 auraient été inversées mais seraient restées cohérentes entre elles et l'orientation estimée lors de l'étape suivante aurait été proche de 0 degré, de sorte que la méthode proposée est robuste à une éventuelle erreur lors du choix entre les hypothèses \mathcal{H}_α^1 et \mathcal{H}_α^2 .

De la même manière, la mise en œuvre de cette méthode sur les images des figures 5.32 et 5.33 conduit aux valeurs estimées $\hat{w}_1 = 20.3$ m, $\hat{w}_2 = 20.4$ m pour la première image et $\hat{w}_1 = 20.5$ m, $\hat{w}_2 = 21.5$ m pour la seconde image. Ces valeurs sont quasiment identiques à celles obtenues à partir des courbes de log-vraisemblance de la figure 5.36, ce qui s'explique par le fait que le phénomène d'ambiguïté entre les paramètres w_1 et w_2 n'est pas observé sur ces images pour lesquelles nous rappelons que l'orientation du bâtiment représenté est de 83 degrés.

Estimation du paramètre α

En imagerie optique, l'orientation α est estimée à partir de l'image binaire B_I^0 résultant de l'extraction des contours de toit. Cette estimation est réalisée conditionnellement aux valeurs des paramètres w_1 et w_2 estimées lors de la première étape de la stratégie d'estimation. Ainsi, la densité de probabilité de présence nécessaire à l'estimation du paramètre α est $p_d^0(x, y | x_r, y_r, \alpha, w_1 = \hat{w}_1, w_2 = \hat{w}_2)$, densité de probabilité dont l'expression a été établie à la sous-section 5.3.1.1.

Nous présentons à la figure 5.45 la densité de probabilité $p_d^0(x, y | x_r, y_r, \alpha, w_1 = \hat{w}_1, w_2 = \hat{w}_2)$ calculée pour les valeurs $\hat{w}_1 = 20.3$ m et $\hat{w}_2 = 20.4$ m, valeurs estimées à partir de l'image simulée

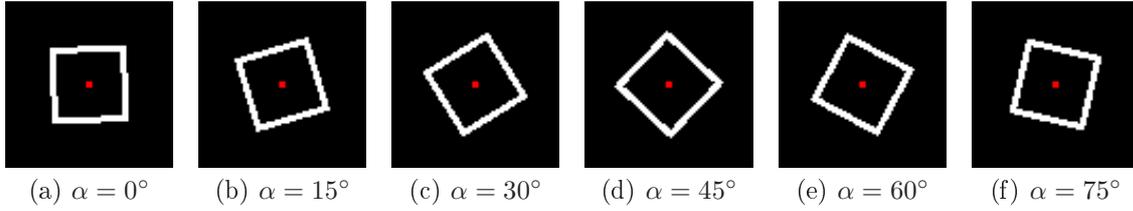


Figure 5.45 – Densités de probabilité de présence conditionnelles $p_d^0(x, y|x_r, y_r, \alpha, w_1 = 20.3, w_2 = 20.4)$ pour l'estimation du paramètre α à partir de l'image binaire B_I^0 de la figure 5.32 (a).

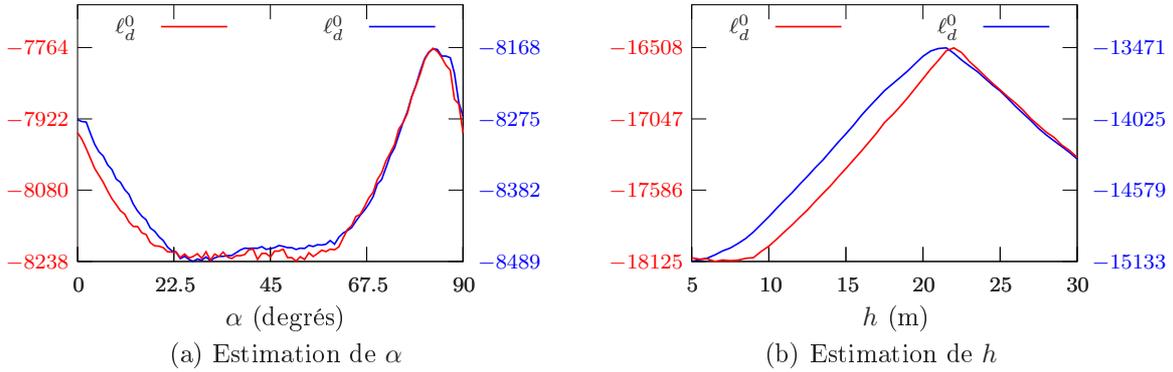


Figure 5.46 – Courbes de log-vraisemblance calculées à partir des images binaires B_I^0 et B_I^1 simulées (courbes rouges) et réelles (courbes bleues) des figures 5.32 et 5.33 pour l'estimation des paramètres α (a) et h (b). Les valeurs estimées des paramètres α et h valent respectivement $\hat{\alpha} = 83.1$ degrés, $\hat{h} = 21.9$ m pour l'image simulée et $\hat{\alpha} = 83.3$ degrés, $\hat{h} = 21.3$ m pour l'image réelle.

de la figure 5.32 (a). À la figure 5.46 (a), la courbe de log-vraisemblance de couleur rouge, qui a été établie à partir de cette même image simulée au moyen de la densité de probabilité de la figure 5.45, conduit à la valeur estimée $\hat{\alpha} = 83.1$ degrés du paramètre α , valeur qui est à comparer à la valeur vraie de ce paramètre qui est de 83.0 degrés.

La courbe de log-vraisemblance de couleur bleue a été établie de manière similaire à partir de l'image réelle de la figure 5.33 (a) et conduit à la valeur $\hat{\alpha} = 84.3$ degrés, valeur qui est, comme en simulation, très proche de la valeur vraie de 83 degrés, validant ainsi l'approche proposée pour l'estimation du paramètre α .

Estimation du paramètre h

La hauteur h est finalement estimée à partir de l'image binaire B_I^1 résultant de l'extraction des ombres. L'estimation est réalisée conditionnellement aux valeurs des paramètres w_1 , w_2 et α estimés lors des deux premières étapes de la stratégie d'estimation. La densité de probabilité de présence nécessaire à cette estimation est donc $p_d^1(x, y|x_r, y_r, \alpha = \hat{\alpha}, w_1 = \hat{w}_1, w_2 = \hat{w}_2, h)$, densité de probabilité dont l'expression a été établie à la sous-section 5.3.1.1 et qui est représentée à la figure 5.47 pour $\hat{\alpha} = 83.1$ degrés, $\hat{w}_1 = 20.3$ m et $\hat{w}_2 = 20.4$ m.

La figure 5.46 (b) représente les courbes de log-vraisemblance calculées à partir des images binaires B_I^1 simulée et réelle des figures 5.32 (b) et 5.33 (b) pour l'estimation du paramètre h . Les valeurs estimées à partir de ces courbes sont respectivement de $\hat{h} = 21.9$ m et $\hat{h} = 21.3$ m pour l'image simulée et pour l'image réelle, valeurs qui sont à comparer à la valeur vraie du paramètre h qui est de 22 m.

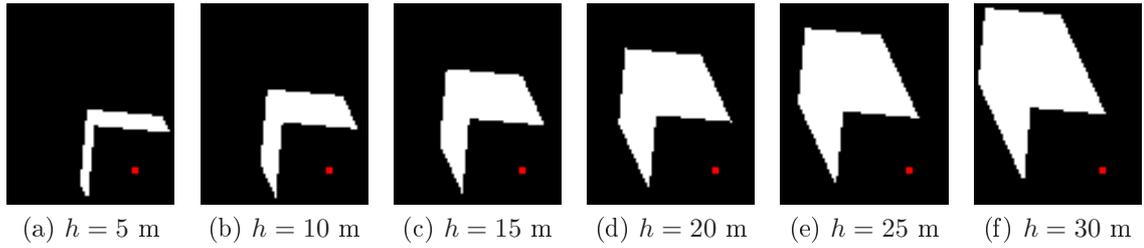


Figure 5.47 – Densités de probabilité de présence conditionnelles $p_d^1(x, y | x_r, y_r, \alpha = 83.1, w_1 = 20.3, w_2 = 20.4, h)$ pour l'estimation du paramètre h à partir de l'image binaire B_I^1 de la figure 5.32 (b).

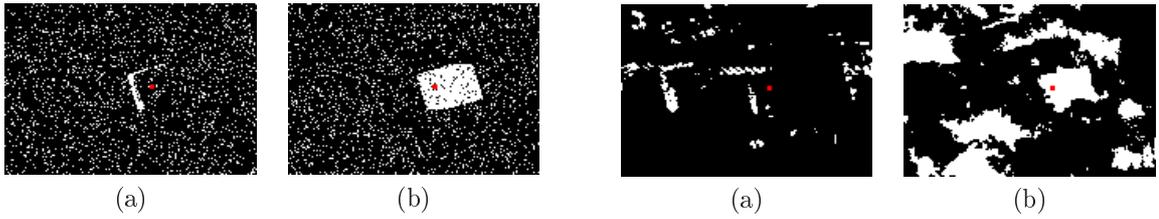


Figure 5.48 – Images binaires B_I^0 (a) et B_I^1 (b) du bâtiment G de la scène d'étude, simulées pour $p_D^0 = 0.9, p_{FA}^0 = 0.1$ et $p_D^1 = 0.9, p_{FA}^1 = 0.1$.

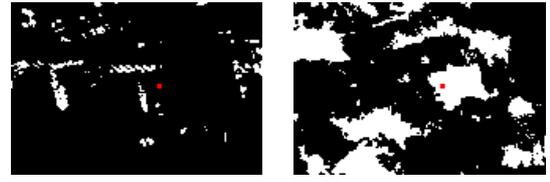


Figure 5.49 – Images binaires B_I^0 (a) et B_I^1 (b) du bâtiment G de la scène d'étude, extraites des images binaires de la figure 5.5.

5.4.3.2 Imagerie RSO

Comme en imagerie optique, nous présentons, dans cette sous-section, la stratégie d'estimation mise en œuvre en imagerie RSO pour estimer les paramètres α, w_1, w_2 et h d'un bâtiment dont les coordonnées du point de référence (x_r, y_r) dans l'image sont connues. Cette stratégie d'estimation est analogue à celle présentée dans le cas de l'imagerie optique : l'ordre d'estimation des paramètres et les images binaires à partir desquelles sont réalisées les estimations sont les mêmes, à la différence que les images binaires B_I^0 et B_I^1 résultent de l'extraction des échos doubles et des ombres RADAR.

Dans la suite de cette sous-section, nous présentons les résultats de la stratégie d'estimation mise en œuvre sur les images simulées et réelles des figures 5.48 et 5.49, qui correspondent au bâtiment G de la scène d'étude.

Les images B_I^0 et B_I^1 de la figure 5.48 ont été établies par simulation pour des détecteurs dont les probabilités de détection et de fausse alarme valent respectivement $p_D^0 = 0.90, p_{FA}^0 = 0.10$ et $p_D^1 = 0.90, p_{FA}^1 = 0.10$.

Les images B_I^0 et B_I^1 de la figure 5.49 ont été extraites à partir des images binaires des figures 5.5 (a) et 5.5 (c), dont nous rappelons qu'elles sont respectivement caractérisées par les probabilités de détection et de fausse alarme suivantes : $p_D^0 = 0.51, p_{FA}^0 = 0.04$ et $p_D^1 = 0.55, p_{FA}^1 = 0.23$.

Estimation des paramètres w_1 et w_2

Le phénomène d'ambiguïté observé en imagerie optique lors de l'estimation des paramètres w_1 et w_2 pour les bâtiments dont l'orientation est proche de 0 degré ou de 90 degrés intervient également en imagerie RSO, puisque ce phénomène résulte de la non continuité du modèle paramétrique décrivant les bâtiments.

Un autre phénomène de discontinuité, spécifique aux images RSO, est également observé pour les bâtiments dont l'orientation est proche de α_P , angle formé par la trace au sol du porteur avec la direction du Nord. Comme l'illustre la figure 5.50, la direction principale du bâtiment est, dans ce cas, parallèle à la trace au sol du porteur et les segments de l'emprise au sol visibles par le capteur pour

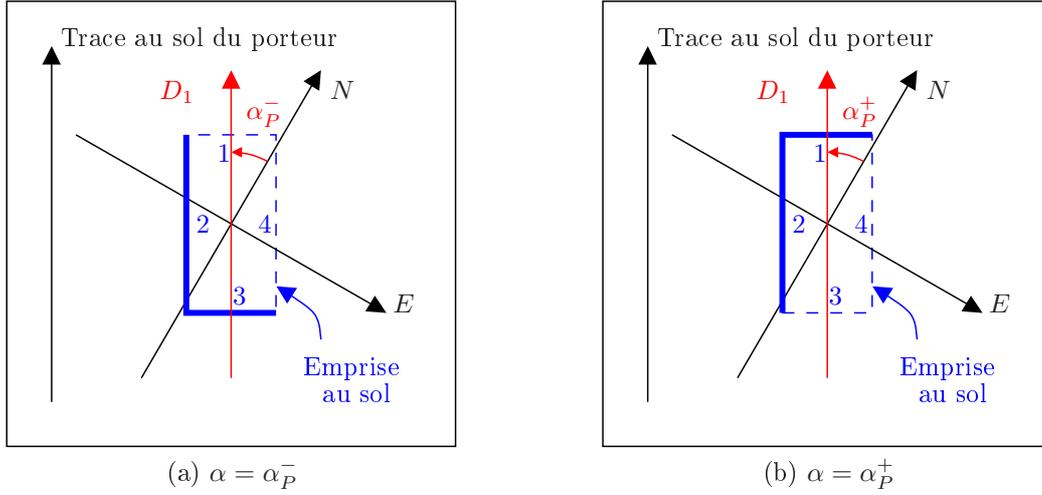


Figure 5.50 – Discontinuité de la forme de l'écho double d'un bâtiment en $\alpha = \alpha_P$. Les segments de l'emprise au sol visibles par le capteur (repérés par des traits pleins) pour $\alpha = \alpha_P^-$ (a) sont différents de ceux visibles pour $\alpha = \alpha_P^+$ (b).

un angle α_P^- sont différents de ceux visibles pour un angle α_P^+ , de sorte que la forme de l'écho double généré par un bâtiment change de manière discontinue en α_P .

En pratique, lorsqu'un bâtiment présente une orientation proche de α_P , seul le segment de l'emprise au sol parallèle à la trace au sol du porteur est illuminé par le capteur et génère un écho double, de sorte que seul un des deux paramètres w_1 ou w_2 peut être estimé à partir de l'image B_I^0 .

La discontinuité de la forme de l'écho double d'un bâtiment en α_P a également des conséquences sur le calcul de la densité de probabilité de présence $p_d^0(x, y | x_r, y_r, w_k)$ au moyen de laquelle est estimé le paramètre w_k . En effet, l'intervalle \mathcal{E}_α sur lequel est intégrée la densité de probabilité $p_d^0(x, y | x_r, y_r, \alpha, w_k, w_{\bar{k}})$ à l'équation (5.58) n'est, *a priori*, pas centré sur la valeur α_P ce qui a pour conséquence de favoriser la configuration où les segments 1 et 2 sont illuminés (configuration 1) par rapport à la configuration où les segments 3 et 4 sont illuminés (configuration 2) ou inversement, alors que ces deux configurations sont *a priori* équiprobables. À titre d'exemple, l'angle α_P vaut 10.3 degrés dans le cas de l'image TerraSAR-X de la figure 5.4, de sorte que la configuration 1 intervient pour les valeurs de α comprises entre 0 degré et 10.3 degrés, tandis que la configuration 2 intervient pour les valeurs de α comprises entre 10.3 degrés et 90 degrés lors de l'intégration sur l'intervalle $\mathcal{E}_\alpha = [0; 90]$. Ceci se traduit par une sur-représentation de la configuration 2 par rapport à la configuration 1. Afin d'éviter la sur-représentation d'une configuration par rapport à une autre, l'intégration sera réalisée sur un intervalle $\mathcal{E}_\alpha = [\alpha_P - 45; \alpha_P + 45]$ centré sur α_P .

L'intégration de la densité de probabilité de présence $p_d^0(x, y | x_r, y_r, \alpha, w_k, w_{\bar{k}})$ par rapport à la variable α étant réalisée sur l'intervalle $\mathcal{E}_\alpha = [\alpha_P - 45; \alpha_P + 45]$, le phénomène d'ambiguïté observé lors de l'estimation des paramètres w_1 et w_2 , intervient désormais pour des bâtiments dont l'orientation est proche de $\alpha_P - 45$ ou $\alpha_P + 45$. Comme en imagerie optique, nous nous affranchissons de ce phénomène en réalisant l'estimation du paramètre w_k conditionnellement aux hypothèses \mathcal{H}_α^1 et \mathcal{H}_α^2 définie comme suit :

- \mathcal{H}_α^1 : « L'orientation α du bâtiment dont le point de référence est situé en (x_r, y_r) appartient à l'intervalle $\mathcal{E}_\alpha^1 = [\alpha_P - 45; \alpha_P[$ » ;
- \mathcal{H}_α^2 : « L'orientation α du bâtiment dont le point de référence est situé en (x_r, y_r) appartient à l'intervalle $\mathcal{E}_\alpha^2 = [\alpha_P; \alpha_P + 45[$ ».

Les densités de probabilité nécessaires à l'estimation des paramètres w_1 et w_2 , conditionnellement aux hypothèses \mathcal{H}_α^1 et \mathcal{H}_α^2 , sont présentées aux figures 5.51, 5.52, 5.53 et 5.54. Ces densités de proba-

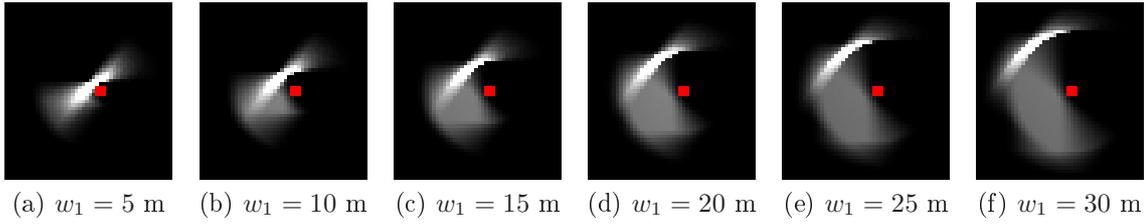


Figure 5.51 – Densités de probabilité de présence conditionnelles $p_d^0(x, y|x_r, y_r, w_1, \mathcal{H}_\alpha^1)$ pour l'estimation du paramètre w_1 .

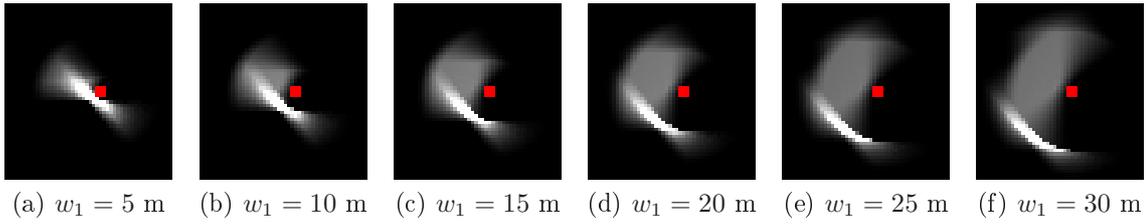


Figure 5.52 – Densités de probabilité de présence conditionnelles $p_d^0(x, y|x_r, y_r, w_1, \mathcal{H}_\alpha^2)$ pour l'estimation du paramètre w_1 .

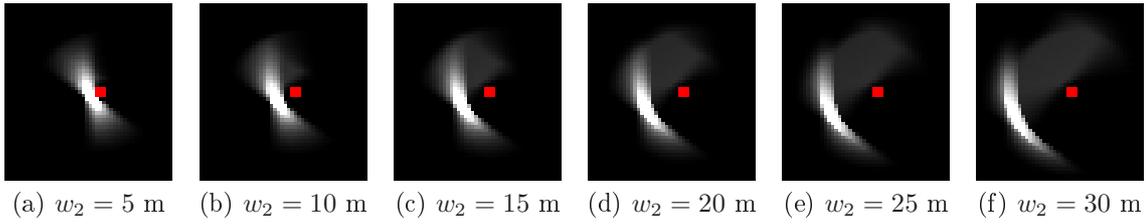


Figure 5.53 – Densités de probabilité de présence conditionnelles $p_d^0(x, y|x_r, y_r, w_2, \mathcal{H}_\alpha^1)$ pour l'estimation du paramètre w_2 .

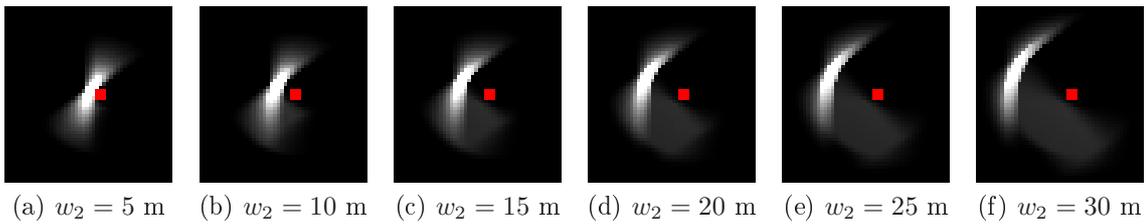


Figure 5.54 – Densités de probabilité de présence conditionnelles $p_d^0(x, y|x_r, y_r, w_2, \mathcal{H}_\alpha^2)$ pour l'estimation du paramètre w_2 .

bilités, qui résultent d'une intégration sur les ensembles \mathcal{E}_α^l et \mathcal{E}_{w_k} , ont été établies pour un détecteur dont les probabilités de détection et de fausse alarme valent respectivement $p_D^0 = 0.9$ et $p_{FA}^0 = 0.1$.

Les courbes de log-vraisemblance de couleur rouge présentées à la figure 5.55 ont été calculées au moyen des densités de probabilité de présence précédentes pour l'estimation des paramètres w_1 et w_2 , conditionnellement aux hypothèses \mathcal{H}_α^1 et \mathcal{H}_α^2 , à partir de l'image binaire B_I^0 simulée de la figure 5.48 (a). Les courbes de couleur bleue ont été établies de manière similaire à partir de l'image réelle de la figure 5.49 (a) au moyen de densités de probabilité calculées pour un détecteur dont les

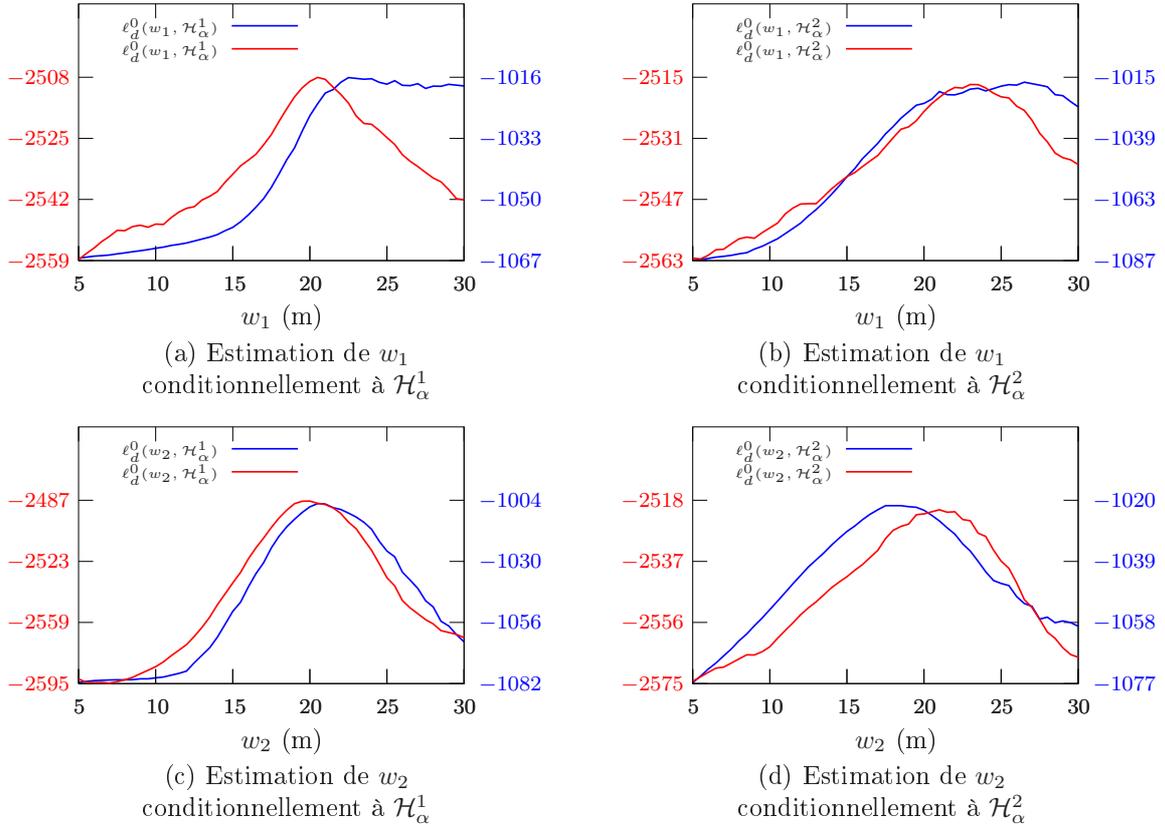


Figure 5.55 – Courbes de log-vraisemblance calculées à partir de l'image binaire B_I^0 simulée de la figure 5.48 (courbes rouges) et de l'image binaire B_I^0 réelle de la figure 5.49 (courbes bleues), conditionnellement aux hypothèses \mathcal{H}_α^1 (a, c) et \mathcal{H}_α^2 (b, d), pour l'estimation des paramètres w_1 (a, b) et w_2 (c, d). Les valeurs estimées des paramètres w_1 et w_2 valent respectivement $\hat{w}_1 = \hat{w}_1^1 = 20.6$ m, $\hat{w}_2 = \hat{w}_2^1 = 19.8$ m pour l'image simulée et $\hat{w}_1 = \hat{w}_1^1 = 22.7$ m, $\hat{w}_2 = \hat{w}_2^1 = 20.7$ m pour l'image réelle.

probabilités de détection et de fausse alarme valent respectivement $p_D^0 = 0.51$ et $p_{FA}^0 = 0.04$. Nous pouvons remarquer que l'estimation du paramètre w_1 , dans le cas de l'image réelle, est perturbée par la présence de balcons sur la façade du bâtiment, balcons dont les échos sont détectés dans l'image binaire B_I^0 ce qui engendre une surestimation du paramètre w_1 .

Estimation du paramètre α

Comme en imagerie optique, l'estimation du paramètre α est réalisée conditionnellement aux paramètres w_1 et w_2 , dont les valeurs ont été estimées à la première étape de la stratégie d'estimation.

Nous présentons à la figure 5.56 la densité de probabilité $p_d^0(x, y|x_r, y_r, \alpha, w_1 = 20.6, w_2 = 19.8)$ au moyen de laquelle est estimé le paramètre α , densité de probabilité qui a été calculée pour un détecteur dont les probabilités de détection et de fausse alarme valent respectivement $p_D^0 = 0.9$ et $p_{FA}^0 = 0.1$.

La courbe de log-vraisemblance de couleur rouge, présentée à la figure 5.57 (a), a été établie à partir de l'image binaire simulée B_I^0 de la figure 5.48 (a) au moyen de la densité de probabilité $p_d^0(x, y|x_r, y_r, \alpha, w_1 = 20.6, w_2 = 19.8)$. La courbe bleue, également présentée à la figure 5.57 (a), a été établie de manière similaire à partir de l'image binaire réelle B_I^0 de la figure 5.49 (a), au moyen de la densité de probabilité $p_d^0(x, y|\alpha, w_1 = 22.7, w_2 = 20.7)$, calculée pour un détecteur dont les probabilités de détection et de fausse alarme valent respectivement $p_D^0 = 0.51$ et $p_{FA}^0 = 0.04$.

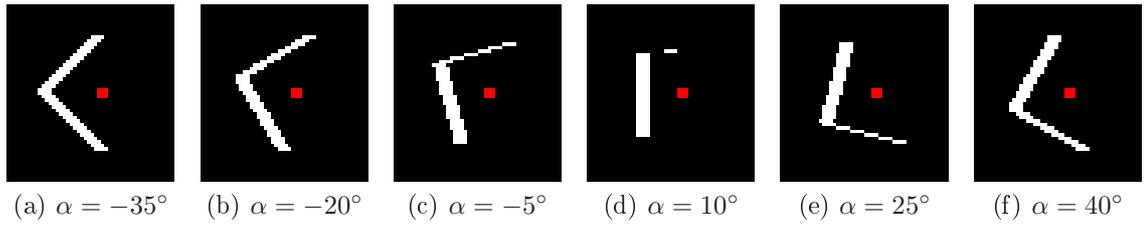


Figure 5.56 – Densités de probabilité de présence conditionnelles $p_d^0(x, y|x_r, y_r, \alpha, w_1 = 20.6, w_2 = 19.8)$ pour l'estimation du paramètre α .

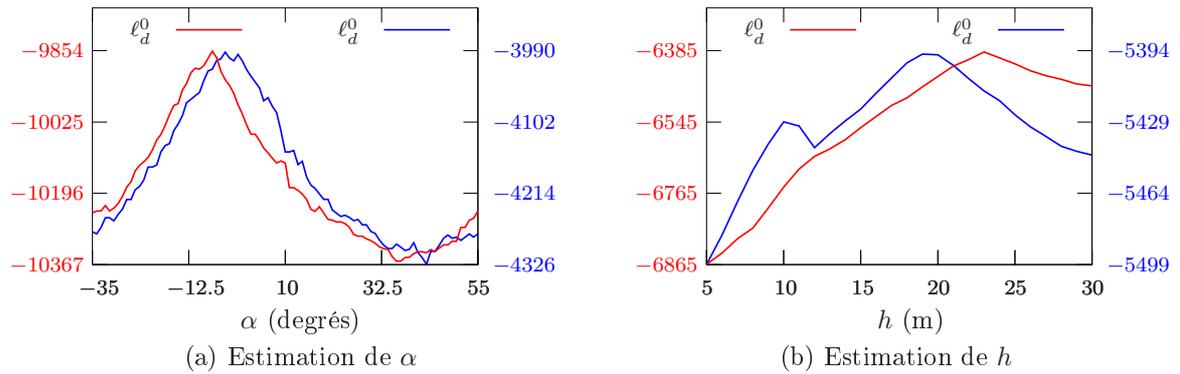


Figure 5.57 – Courbes de log-vraisemblance calculées à partir des images binaires B_I^0 et B_I^1 simulées (courbes rouges) et réelles (courbes bleues) des figures 5.48 et 5.49 pour l'estimation des paramètres α (a) et h (b). Les valeurs estimées des paramètres α et h valent respectivement $\hat{\alpha} = -7.0$ degrés, $\hat{h} = 23.1$ m pour l'image simulée et $\hat{\alpha} = -4.0$ degrés, $\hat{h} = 19.4$ m pour l'image réelle.

Estimation du paramètre h

La troisième et dernière étape de la stratégie d'estimation consiste, comme en imagerie optique, à estimer le paramètre h à partir de l'image binaire B_I^1 , conditionnellement aux valeurs des paramètres w_1 , w_2 et α estimées précédemment.

Nous présentons à la figure 5.58 la densité de probabilité de présence $p_d^1(x, y|x_r, y_r, \alpha = -7.0, w_1 = 20.6, w_2 = 19.8, h)$, densité qui est calculée conditionnellement aux valeurs $\hat{\alpha}$, \hat{w}_1 et \hat{w}_2 estimées à partir de l'image binaire B_I^0 simulée de la figure 5.48 (a). Cette densité a été établie pour un détecteur dont les probabilités de détection et de fausse alarme valent respectivement $p_D^1 = 0.9$ et $p_{FA}^1 = 0.1$.

La courbe de log-vraisemblance de couleur rouge, présentée à la figure 5.57 (b), a été établie à partir de l'image binaire simulée B_I^1 de la figure 5.48 (b), au moyen de la densité de probabilité précédente. De même la courbe bleue a été établie à partir de l'image binaire réelle B_I^1 de la figure 5.49 (b), au moyen de la densité de probabilité $p_d^1(x, y|x_r, y_r, \alpha = -4.0, w_1 = 22.7, w_2 = 20.7, h)$, calculée pour un détecteur dont les probabilités de détection et de fausse alarme valent respectivement $p_D^1 = 0.55$ et $p_{FA}^1 = 0.23$.

Nous constatons, à la figure 5.57 (b), que la courbe de log-vraisemblance établie sur données simulées présente une variation abrupte de sa dérivée au niveau du maximum de la courbe. Ce phénomène apparaît plus nettement encore à la figure 5.60 (a) où la courbe de log-vraisemblance de couleur rouge a été établie à partir d'une image binaire B_I^1 du bâtiment G , simulée pour $p_D^1 = 0.99$ et $p_{FA}^1 = 0.01$. Ce phénomène, que nous expliquons analytiquement par la suite, peut engendrer, pour de faibles valeurs de la probabilité de fausse alarme p_{FA}^1 , un biais lors de l'estimation de la hauteur. En effet, lorsque p_{FA}^1 est faible, l'écart entre les valeurs absolues des dérivées à gauche et à droite du maximum de la courbe est important, ce qui engendre l'introduction d'un biais lors de l'interpolation

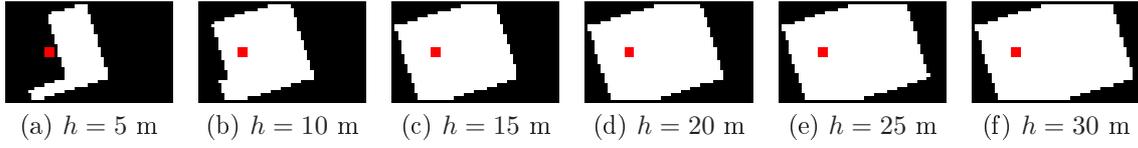


Figure 5.58 – Densités de probabilité de présence conditionnelles $p_d^1(x, y | x_r, y_r, \alpha = -7.0, w_1 = 20.6, w_2 = 19.8, h)$ pour l'estimation du paramètre h .

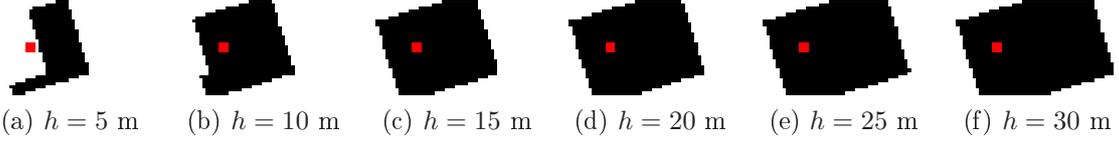


Figure 5.59 – Densités de probabilité de présence conditionnelles $p_d^1(x, y | x_r, y_r, \alpha = -7.0, w_1 = 20.6, w_2 = 19.8, h)$ pour l'estimation du paramètre h .

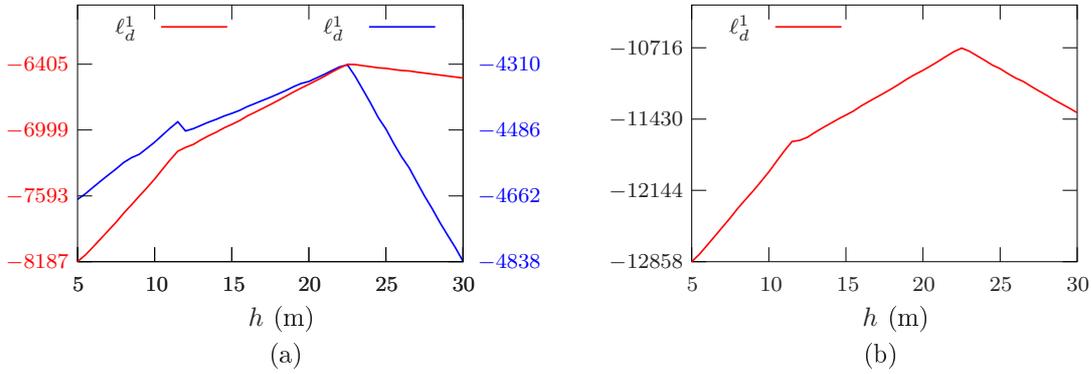


Figure 5.60 – Courbes de log-vraisemblance calculées à partir des pixels détectés et non détectés dans une image binaire B_I^1 du bâtiment G simulée pour $p_D^1 = 0.99$ et $p_{FA} = 0.01$.

de la courbe par un polynôme de degré deux.

Dans le cas de l'estimation du paramètre h à partir de l'image binaire B_I^1 , estimation qui est réalisée conditionnellement aux paramètres α , w_1 et w_2 , l'expression de la log-vraisemblance établie à l'équation (5.55) se réécrit comme suit :

$$\forall h \in \mathcal{E}_h, \ell_d^1(h) = \sum_{\substack{(x,y) \in \nu_r^1 \\ B_I^1(x,y)=1}} \ln p_d^1(x, y | x_r, y_r, \alpha, w_1, w_2, h). \quad (5.61)$$

Conformément à l'équation (5.14), la densité de probabilité $p_d^1(x, y | x_r, y_r, \alpha, w_1, w_2, h)$ qui intervient dans le calcul de la log-vraisemblance $\ell_d^1(h)$ s'écrit :

$$\forall (x, y) \in \nu_r^1, p_d^1(x, y | x_r, y_r, \alpha, w_1, w_2, h) = \frac{p_D^1}{N_{R_O} p_D^1 + (N_{\nu_r^1} - N_{R_O}) p_{FA}^1} 1_{R_O}(x, y) + \frac{p_{FA}^1}{N_{R_O} p_D^1 + (N_{\nu_r^1} - N_{R_O}) p_{FA}^1} 1_{\nu_r^1 \setminus R_O}(x, y), \quad (5.62)$$

où $N_{\nu_r^1}$ et N_{R_O} désignent respectivement le nombre de pixels contenus dans le voisinage ν_r^1 du pixel de référence r et le nombre de pixels contenus dans la région R_O résultant de la projection du bâtiment

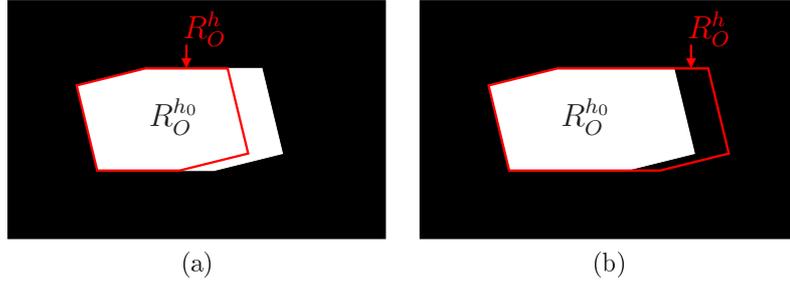


Figure 5.61 – Régions $R_O^{h_0}$ et R_O^h pour le calcul de la log-vraisemblance $\ell_d^1(h)$ à partir d'une image binaire B_I^1 simulée ($p_D^1 = 1, p_{FA}^1 = 0$) pour un bâtiment de hauteur h_0 avec $h_0 \geq h$ (a) et $h_0 \leq h$ (b).

\vec{b} dans l'image I .

Par la suite, nous considérons l'image binaire B_I^1 de la figure 5.61 qui a été simulée pour un bâtiment de hauteur h_0 . La hauteur h_0 est supérieure à $l_a^{max} \tan i^2$, de sorte que ce bâtiment est vu comme un objet semi-occultant pour l'ensemble des positions azimutales du capteur RSO (cf. sous-section 2.4.2.3). Nous souhaitons ici établir l'expression de la log-vraisemblance $\ell_d^1(h)$, calculée à partir de l'image binaire B_I^1 de la figure 5.61, dans le cas où les valeurs estimées des paramètres α , w_1 et w_2 sont égales aux valeurs vraies de ces paramètres.

L'expression de la log-vraisemblance $\ell_d^1(h)$, établie à partir des équations (5.61) et (5.62), est donnée par les équations suivantes :

$$\forall h \in [l_a^{max} \tan i; h_0], \ell_d^1(h) = N_{R_O^h} \ln \left[\frac{p_D^1}{N_{R_O^h} p_D^1 + (N_{\nu_r^1} - N_{R_O^h}) p_{FA}^1} \right] + (N_{R_O^{h_0}} - N_{R_O^h}) \ln \left[\frac{p_{FA}^1}{N_{R_O^h} p_D^1 + (N_{\nu_r^1} - N_{R_O^h}) p_{FA}^1} \right], \quad (5.63)$$

$$\forall h \geq h_0, \ell_d^1(h) = N_{R_O^h} \ln \left[\frac{p_D^1}{N_{R_O^h} p_D^1 + (N_{\nu_r^1} - N_{R_O^h}) p_{FA}^1} \right], \quad (5.64)$$

où les régions $R_O^{h_0}$ et R_O^h , représentées à la figure 5.61, résultent de la projection des bâtiments de paramètres (α, w_1, w_2, h_0) et (α, w_1, w_2, h) dans l'image I . Le premier terme de l'expression (5.63) correspond à la log-vraisemblance des pixels de la région $R_O^{h_0}$ qui appartiennent également à la région R_O^h et qui sont donc considérés comme des vraies détections, tandis que le second terme correspond aux pixels de la région $R_O^{h_0}$ qui n'appartiennent pas à la région R_O^h et qui sont donc considérés comme des fausses alarmes. Dans l'expression (5.64), l'ensemble des pixels de la région $R_O^{h_0}$ appartiennent également à la région R_O^h , de sorte que l'ensemble des pixels détectés de l'image binaire B_I^1 sont considérés comme des vraies détections.

Étant données les expressions de $\ell_d^1(h)$, établies aux équations (5.63) et (5.64), la dérivée partielle de cette log-vraisemblance par rapport à h peut être calculée et s'exprime comme suit :

$$\forall h \in [l_a^{max} \tan i; h_0], \frac{\partial \ell_d^1}{\partial h} = \frac{\partial \ell_d^1}{\partial N_{R_O^h}} \frac{\partial N_{R_O^h}}{\partial h} \quad (5.65)$$

$$\forall h \in [l_a^{max} \tan i; h_0], \frac{\partial \ell_d^1}{\partial h} = \frac{\partial N_{R_O^h}}{\partial h} \left[\ln \left(\frac{p_D^1}{p_{FA}^1} \right) - \frac{N_{R_O^{h_0}} (p_D^1 - p_{FA}^1)}{N_{R_O^h} p_D^1 + (N_{\nu_r^1} - N_{R_O^h}) p_{FA}^1} \right], \quad (5.66)$$

² l_a^{max} et i désignent, conformément aux notations introduites au chapitre 2, la largeur apparente maximale du bâtiment et l'angle d'incidence des ondes RADAR

$$\forall h \geq h_0, \frac{\partial \ell_d^1}{\partial h} = -\frac{\partial N_{R_O^h}}{\partial h} \frac{N_{R_O^{h_0}}(p_D^1 - p_{FA}^1)}{N_{R_O^h} p_D^1 + (N_{\nu_r^1} - N_{R_O^h}) p_{FA}^1}. \quad (5.67)$$

D'après la section 2.4.2, les dimensions de la région R_O selon l'axe azimutal sont indépendantes de la hauteur et sont linéairement liées à cette hauteur selon l'axe radial, de sorte que le terme $\partial N_{R_O^h} / \partial h$ qui intervient dans les équations précédentes peut être considéré comme constant.

Comme le montre l'équation (5.66), les variations de la log-vraisemblance par rapport à la hauteur sont dues, lorsque $h \in [l_a^{max} \tan i; h_0]$, à deux phénomènes antagonistes représentés par les deux termes de la somme entre crochets. Le premier terme, qui est toujours positif puisque $p_D^1 > p_{FA}^1$, traduit le fait qu'un certain nombre de pixels qui étaient considérés comme des fausses alarmes sont, lorsque h augmente, considérés comme des vraies détections ce qui engendre une forte augmentation de la vraisemblance. Le second terme, qui est toujours négatif, traduit la variation de la vraisemblance des pixels considérés comme des vraies détections en raison de l'évolution de la constante de normalisation de la densité de probabilité $p_d^1(x, y | x_r, y_r, \alpha, w_1, w_2, h)$.

Lorsque $h \geq h_0$, l'équation 5.67 montre que seul le second terme de l'équation précédente est conservé, de sorte que la diminution de la log-vraisemblance avec la hauteur est uniquement liée à l'évolution de la constante de normalisation de la densité de probabilité $p_d^1(x, y | x_r, y_r, \alpha, w_1, w_2, h)$.

Afin de remédier au biais qui peut intervenir en raison de la différence des valeurs de la dérivée à gauche et à droite du maximum, différence qui est d'autant plus marquée que p_{FA}^1 est faible et p_D^1 est élevée, nous proposons de prendre en compte l'ensemble des pixels présents dans le voisinage ν_r^1 (les pixels détectés et les pixels non détectés) pour l'estimation du paramètre h .

Une étude similaire à celle menée précédemment permet de montrer que la dérivée de la log-vraisemblance ℓ_d^1 des pixels non détectés par rapport à la hauteur h s'exprime comme suit :

$$\forall h \in [l_a^{max} \tan i; h_0], \frac{\partial \ell_d^1}{\partial h} = \frac{\partial N_{R_O^h}}{\partial h} \frac{(N_{\nu_r^1} - N_{R_O^{h_0}})(p_D^1 - p_{FA}^1)}{N_{R_O^h}(1 - p_D^1) + (N_{\nu_r^1} - N_{R_O^h})(1 - p_{FA}^1)}, \quad (5.68)$$

$$\forall h \geq h_0, \frac{\partial \ell_{\bar{d}}^1}{\partial h} = \frac{\partial N_{R_O^h}}{\partial h} \left[\frac{(N_{\nu_r^1} - N_{R_O^{h_0}})(p_D^1 - p_{FA}^1)}{N_{R_O^h}(1 - p_D^1) + (N_{\nu_r^1} - N_{R_O^h})(1 - p_{FA}^1)} - \ln \left(\frac{1 - p_{FA}^1}{1 - p_D^1} \right) \right]. \quad (5.69)$$

Ainsi, nous constatons à partir des équations (5.68) et (5.69) que les valeurs des dérivées de ℓ_d^1 calculées à gauche et à droite du maximum sont également distinctes, ce qui est confirmé à la figure 5.60 (a) où la courbe bleue correspond à la log-vraisemblance des pixels non détectés, log-vraisemblance qui est calculée à partir d'une image binaire B_I^1 du bâtiment G simulée pour $p_D^1 = 0.99$ et $p_{FA}^1 = 0.01$.

Contrairement à la log-vraisemblance ℓ_d^1 , la log-vraisemblance $\ell_{\bar{d}}^1$ présente une pente plus importante, en valeur absolue, pour les valeurs de h supérieure à h_0 . La prise en compte de l'ensemble des pixels, ce qui est fait en sommant les log-vraisemblances ℓ_d^1 et $\ell_{\bar{d}}^1$, comme illustré à la figure 5.60 (b), permet ainsi d'obtenir un maximum plus marqué et de limiter le biais d'estimation lié à la différence des valeurs des dérivées à gauche et à droite du maximum.

Pour conclure, ce phénomène, que nous avons présenté dans le cas particulier de l'estimation de la hauteur en imagerie RSO, affecte l'estimation de n'importe quel paramètre dès lors que ce dernier est estimé à partir d'une primitive surfacique. L'estimation de la hauteur en imagerie optique est donc également sujette à un biais, de sorte que nous mettons en œuvre la même approche que celle présentée ici afin d'y remédier.

5.5 Détection fine des bâtiments

La méthode que nous avons introduit à la partie 5.3 permet de réaliser une détection grossière des bâtiments à partir des images binaires B_I^0 et B_I^1 , résultant de l'extraction des primitives caractéristiques de la présence de ces bâtiments dans les images optiques ou RSO. Cette méthode de détection

grossière fournit une image binaire D_I au sein de laquelle est défini un ensemble de pixels pouvant possiblement correspondre aux points de référence de bâtiments.

La méthode que nous présentons dans cette partie consiste à réaliser, en chaque pixel (x_t, y_t) de l'image D_I identifié comme étant possiblement le point de référence d'un bâtiment, un test statistique visant à décider entre les deux hypothèses suivantes :

- \mathcal{H}_1 : « Le pixel (x_t, y_t) correspond au point de référence d'un bâtiment » ;
- \mathcal{H}_0 : « Le pixel (x_t, y_t) ne correspond pas au point de référence d'un bâtiment ».

Cette méthode repose, comme à la partie 5.3, sur le calcul de la vraisemblance des pixels des images binaires B_I^j avec l'hypothèse \mathcal{H}_1 , à la différence qu'il s'agit désormais d'une vraisemblance généralisée. Le calcul de cette vraisemblance généralisée implique les paramètres α , w_1 , w_2 et h estimés, en chaque pixel testé, au moyen de la méthode d'inversion de modèle introduite à la partie précédente.

5.5.1 Calcul de la vraisemblance généralisée des pixels des images binaires de primitives

Nous avons vu, à la partie 5.3, que la vraisemblance des pixels des images binaires de primitives B_I^j et B_I^k pouvait être calculée indépendamment à partir de chacune de ces images, conduisant ainsi à deux images de vraisemblance ℓ^j et ℓ^k , ou bien conjointement à partir de ces images, conduisant ainsi à une unique image de vraisemblance ℓ^{jk} . Pour les mêmes raisons qu'à la partie 5.3, nous considérons ici les images conjointement, ce qui revient, dans le cas du calcul de la vraisemblance généralisée, à effectuer une fusion conjonctive des résultats obtenus indépendamment à partir de chacune des images binaires de primitives.

En reprenant l'équation (5.40), la log-vraisemblance généralisée des pixels détectés conjointement dans les voisinages ν_t^j et ν_t^k du pixel testé, sachant l'hypothèse \mathcal{H}_1 , s'écrit :

$$\forall (x_t, y_t) \in \mathcal{E}_I, \ell_{dd}^{jk}(x_t, y_t) = \sum_{\substack{(x^j, y^j) \in \nu_t^j, B_I^j(x^j, y^j)=1 \\ (x^k, y^k) \in \nu_t^k, B_I^k(x^k, y^k)=1}} \ln p_{dd}^{jk}(x^j, y^j, x^k, y^k | x_t, y_t, \hat{\alpha}^t, \hat{w}_1^t, \hat{w}_2^t, \hat{h}^t), \quad (5.70)$$

où $\hat{\alpha}^t$, \hat{w}_1^t , \hat{w}_2^t et \hat{h}^t désignent respectivement les valeurs des paramètres α , w_1 , w_2 et h estimées au moyen de la méthode introduite à la partie précédente, pour un bâtiment dont le point de référence correspond au pixel testé.

D'après l'équation (5.29), la densité de probabilité p_{dd}^{jk} qui intervient dans l'équation précédente s'exprime comme suit :

$$\forall (x^j, y^j, x^k, y^k) \in \nu_t^j \times \nu_t^k, p_{dd}^{jk}(x^j, y^j, x^k, y^k | x_t, y_t, \hat{\alpha}^t, \hat{w}_1^t, \hat{w}_2^t, \hat{h}^t) = p_d^j(x^j, y^j | x_t, y_t, \hat{\alpha}^t, \hat{w}_1^t, \hat{w}_2^t, \hat{h}^t) p_d^k(x^k, y^k | x_t, y_t, \hat{\alpha}^t, \hat{w}_1^t, \hat{w}_2^t, \hat{h}^t). \quad (5.71)$$

En remplaçant la densité de probabilité p_{dd}^{jk} par son expression dans l'équation (5.70), nous obtenons :

$$\begin{aligned} \forall (x_t, y_t) \in \mathcal{E}_I, \ell_{dd}^{jk}(x_t, y_t) = & N_d^k \sum_{\substack{(x^j, y^j) \in \nu_t^j, \\ B_I^j(x^j, y^j)=1}} \ln p_d^j(x^j, y^j | x_t, y_t, \hat{\alpha}^t, \hat{w}_1^t, \hat{w}_2^t, \hat{h}^t) \\ & + N_d^j \sum_{\substack{(x^k, y^k) \in \nu_t^k, \\ B_I^k(x^k, y^k)=1}} \ln p_d^k(x^k, y^k | x_t, y_t, \hat{\alpha}^t, \hat{w}_1^t, \hat{w}_2^t, \hat{h}^t), \end{aligned} \quad (5.72)$$

où N_d^j et N_d^k désignent les nombres de pixels détectés respectivement dans les images binaires B_I^j et B_I^k . Nous voyons donc que la log-vraisemblance calculée en considérant conjointement les pixels détectés au sein des images binaires B_I^j et B_I^k correspond à la somme pondérée des log-vraisemblances de ces pixels considérés indépendamment, la pondération permettant de garantir une contribution

comparable de ces deux log-vraisemblances à la log-vraisemblance globale.

Des expressions similaires à celle de la log-vraisemblance généralisée $\ell_{dd}^{jk}(x_t, y_t)$ peuvent être établies pour les log-vraisemblances généralisées $\ell_{dd}^{jk}(x_t, y_t)$, $\ell_{dd}^{jk}(x_t, y_t)$ et $\ell_{dd}^{jk}(x_t, y_t)$, de sorte que la log-vraisemblance généralisée ℓ^{jk} des pixels des images binaires B_I^j et B_I^k s'écrit finalement :

$$\begin{aligned}
\ell^{jk}(x_t, y_t) = & N_I \sum_{\substack{(x^j, y^j) \in \nu_t^j, \\ B_I^j(x^j, y^j)=1}} \ln p_d^j(x^j, y^j | x_t, y_t, \hat{\alpha}^t, \hat{w}_1^t, \hat{w}_2^t, \hat{h}^t) \\
& + N_I \sum_{\substack{(x^j, y^j) \in \nu_t^j, \\ B_I^j(x^j, y^j)=0}} \ln p_d^j(x^j, y^j | x_t, y_t, \hat{\alpha}^t, \hat{w}_1^t, \hat{w}_2^t, \hat{h}^t) \\
& + N_I \sum_{\substack{(x^k, y^k) \in \nu_t^k, \\ B_I^k(x^k, y^k)=1}} \ln p_d^k(x^k, y^k | x_t, y_t, \hat{\alpha}^t, \hat{w}_1^t, \hat{w}_2^t, \hat{h}^t) \\
& + N_I \sum_{\substack{(x^k, y^k) \in \nu_t^k, \\ B_I^k(x^k, y^k)=0}} \ln p_d^k(x^k, y^k | x_t, y_t, \hat{\alpha}^t, \hat{w}_1^t, \hat{w}_2^t, \hat{h}^t), \tag{5.73}
\end{aligned}$$

où N_I désigne le nombre de pixel de l'image I .

5.5.2 Stratégie de détection des bâtiments à partir de l'image de vraisemblance généralisée

Si l'image de log-vraisemblance généralisée ℓ^{jk} précédemment calculée présente, comme nous le verrons aux sections 5.5.3 et 5.5.4, des pics étroits au niveau des pixels de référence des bâtiments, les pixels voisins des pixels de référence peuvent également présenter des valeurs de vraisemblance généralisée élevées, de sorte qu'une stratégie de détection qui reposerait sur un simple seuillage se traduit par des détections multiples d'un même bâtiment.

Afin d'éviter ces détections multiples, nous mettons en œuvre une stratégie itérative visant à extraire les bâtiments les plus vraisemblables, chaque itération étant décrite par les étapes suivantes :

1. Recherche du maximum de l'image de log-vraisemblance généralisée ℓ^{jk} ;
2. Si la valeur du maximum est supérieure à un seuil de détection $\tau_{\ell^{jk}}$, le pixel pour lequel ce maximum est atteint est considéré comme étant le pixel de référence d'un bâtiment. Dans le cas contraire, nous considérons qu'il n'y a plus de bâtiment à extraire et stoppons la stratégie itérative de détection ;
3. Nous fixons à $-\infty$ la valeur de l'ensemble des pixels de l'image de log-vraisemblance ℓ^{jk} contenus dans l'emprise au sol du bâtiment détecté à l'étape précédente, évitant ainsi les détections multiples.

Notons, pour finir, que cette stratégie de détection évite les détections multiples d'un même bâtiment mais n'interdit toutefois pas la possibilité de recouvrement entre les emprises au sol de deux bâtiments proches.

5.5.3 Résultats en imagerie optique

Nous présentons à la figure 5.62 les cartes de log-vraisemblance généralisée calculées respectivement à partir des images binaires simulées des figures 5.14 (a) et 5.16 (a) et des images binaires réelles des figures 5.15 (a) et 5.17 (a).

La carte de log-vraisemblance obtenue à partir des données simulées présente des pics très étroits, de sorte que l'approche proposée permet de déterminer précisément les pixels de référence des bâtiments présents dans la scène d'étude. La carte de log-vraisemblance calculée sur données réelles présente en revanche des *maxima* peu marqués (bâtiments B , C et F) et dans certains cas

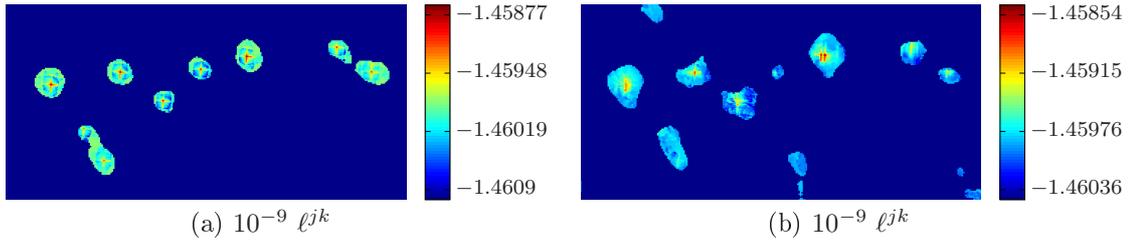


Figure 5.62 – Cartes de log-vraisemblance généralisée établies à partir des images binaires B_I^0 et B_I^1 simulées (a) et réelles (b) des figures 5.14 (a), 5.16 (a), 5.15 (a) et 5.17 (a).

des *maxima* multiples (bâtiments A , E , G), de sorte que la détermination des pixels de référence est logiquement plus délicate que dans le cas des données simulées.

La mise en œuvre de la stratégie de détection itérative à partir de la carte de log-vraisemblance généralisée de la figure 5.62 (a) permet finalement d’extraire les bâtiments dont les paramètres sont donnés au tableau 5.1 et dont les signatures géométriques sont représentées à la figure 5.63 dans les images binaires B_I^0 et B_I^1 . Les résultats obtenus sont excellents puisque les erreurs maximales observées sur les paramètres X_C , Y_C , α , w_1 , w_2 et h sont respectivement de 0.1 m, 0.5 m, 0.6 degrés, 0.6 m, 0.6 m et 0.4 m.

De même, la figure 5.64 et le tableau 5.2 présentent respectivement les signatures géométriques et les paramètres des bâtiments extraits à partir de la carte de log-vraisemblance généralisée de la figure 5.62 (b). En termes de détection, les résultats obtenus sont très bons puisque l’ensemble des bâtiments présents dans la scène ont été détectés alors qu’aucune fausse alarme n’a été détectée. Les performances en estimation sont en revanche plus mitigées, puisque les erreurs maximales observées sur les paramètres X_C , Y_C , α , w_1 , w_2 et h valent respectivement 3.5 m, 5 m, 4.7 degrés, 6.5 m, 10 m et 4 m. Nous rappelons que l’estimation des paramètres w_1 , w_2 et h est réalisée conditionnellement à la position du point de référence des bâtiments, ce qui explique les importantes erreurs observées sur ces paramètres pour les bâtiments A , E , F et G , pour lesquels la position du point de référence est largement erronée. Étant donné les performances très limitées des détecteurs de primitives mis en œuvre, nous pouvons conclure que les résultats obtenus sont relativement bons.

5.5.4 Résultats en imagerie RSO

Dans cette section, nous présentons les résultats de la méthode d’extraction de bâtiments mise en œuvre à partir d’images binaires B_I^0 et B_I^1 résultant de l’extraction des primitives caractéristiques de la présence des bâtiments en imagerie RSO.

Nous présentons tout d’abord, à la figure 5.65, les cartes de log-vraisemblance généralisée calculées respectivement à partir des images binaires B_I^0 et B_I^1 simulées des figures 5.21 (a) et 5.23 (a) et des images binaires B_I^0 et B_I^1 réelles des figures 5.22 (a) et 5.24 (a).

Si la carte de log-vraisemblance généralisée établie à partir des images simulées permet de déterminer correctement la position des points de référence des différents bâtiments présents dans la scène, cette carte présente néanmoins des *maxima* moins marqués que la carte analogue établie en imagerie optique. Ceci s’explique par la forme en « L » des échos doubles, forme au sein de laquelle l’information relative à la position du point de référence est moins redondante qu’elle ne l’est au sein de la forme rectangulaire des contours de toit en imagerie optique.

En ce qui concerne la carte de log-vraisemblance établie à partir des images binaires réelles, elle présente des *maxima* peu marqués ce qui se traduit pour certains bâtiments par une localisation erronée de leur point de référence, comme nous le constatons dans la suite de cette section.

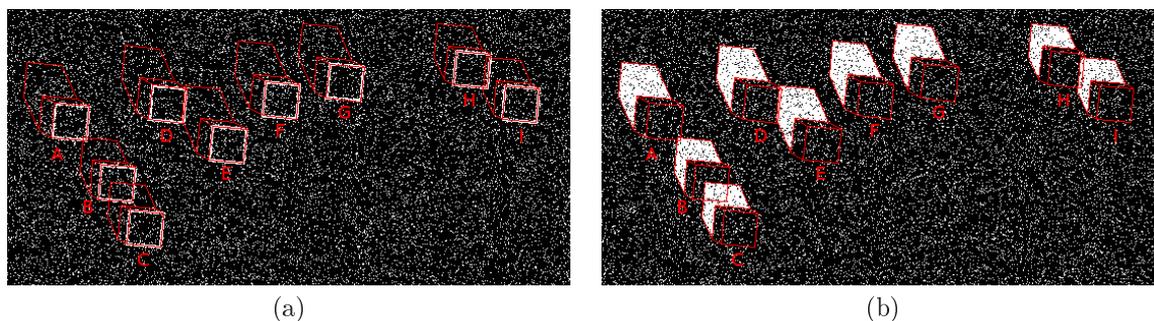


Figure 5.63 – Résultats de la méthode d'extraction sur données simulées en imagerie optique : simulation des signatures géométriques des bâtiments extraits dans les images binaires de primitives B_I^0 (a) et B_I^1 (b) simulées.

Label		X_C (m)	Y_C (m)	α (degrés)	w_1 (m)	w_2 (m)	h (m)
A	Vérité terrain	702995.0	4795064.0	83.0	20.0	20.0	22.0
	Estimation	702995.1	4795064.0	83.3	20.5	20.4	21.9
B	Vérité terrain	703025.0	4795026.0	83.0	20.0	20.0	15.0
	Estimation	703025.1	4795026.0	83.6	20.3	20.5	14.6
C	Vérité terrain	703042.0	4794999.0	83.0	20.0	20.0	15.0
	Estimation	703042.1	4794999.0	83.0	20.4	20.1	15.0
D	Vérité terrain	703053.0	4795074.0	83.0	20.0	20.0	22.0
	Estimation	703053.1	4795074.0	82.3	20.5	20.6	21.3
E	Vérité terrain	703088.5	4795050.0	83.0	20.0	20.0	22.0
	Estimation	703088.5	4795050.0	82.8	20.4	20.2	21.7
F	Vérité terrain	703120.0	4795076.0	83.0	20.0	20.0	22.0
	Estimation	703120.1	4795076.0	82.9	20.3	20.2	21.8
G	Vérité terrain	703159.0	4795086.0	83.0	20.0	20.0	22.0
	Estimation	703159.0	4795085.5	82.7	20.6	20.2	21.9
H	Vérité terrain	703236.0	4795092.0	83.0	20.0	20.0	15.0
	Estimation	703236.0	4795092.0	83.0	20.2	20.1	14.8
I	Vérité terrain	703264.0	4795072.0	83.0	20.0	20.0	15.0
	Estimation	703264.1	4795072.0	83.0	20.1	20.2	15.0

Tableau 5.1 – Résultats de la méthode d'extraction sur données simulées en imagerie optique : valeurs estimées et valeurs vraies des paramètres des neuf bâtiments de la scène d'étude.

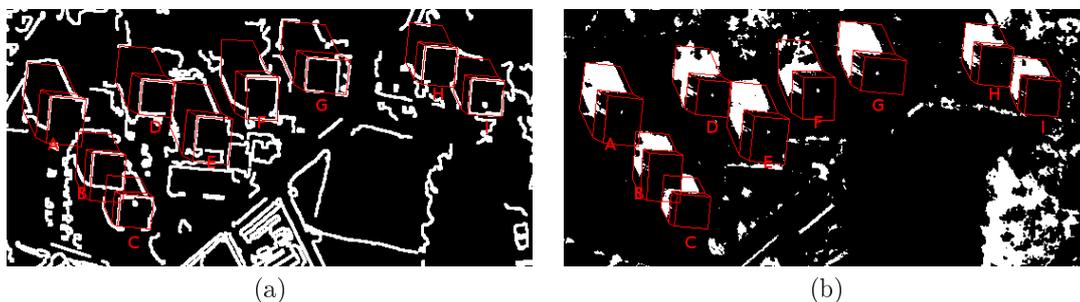


Figure 5.64 – Résultats de la méthode d'extraction sur données simulées en imagerie optique : simulation des signatures géométriques des bâtiments extraits dans les images binaires de primitives B_I^0 (a) et B_I^1 (b) simulées.

Label		X_C (m)	Y_C (m)	α (degrés)	w_1 (m)	w_2 (m)	h (m)
A	Vérité terrain	702995.0	4795064.0	83.0	20.0	20.0	22.0
	Estimation	702995.9	4795059.0	82.1	22.6	30.0	19.6
B	Vérité terrain	703025.0	4795026.0	83.0	20.0	20.0	15.0
	Estimation	703025.3	4795021.5	84.6	20.9	30.0	12.2
C	Vérité terrain	703042.0	4794999.0	83.0	20.0	20.0	15.0
	Estimation	703043.5	4794999.5	83.6	21.2	20.4	11.0
D	Vérité terrain	703053.0	4795074.0	83.0	20.0	20.0	22.0
	Estimation	703053.7	4795073.0	83.3	20.6	22.7	19.3
E	Vérité terrain	703088.5	4795050.0	83.0	20.0	20.0	22.0
	Estimation	703089.1	4795046.5	82.8	22.0	28.6	19.9
F	Vérité terrain	703120.0	4795076.0	83.0	20.0	20.0	22.0
	Estimation	703120.2	4795073.0	87.7	19.6	27.0	21.4
G	Vérité terrain	703159.0	4795086.0	83.0	20.0	20.0	22.0
	Estimation	703162.5	4795086.5	83.0	26.5	21.8	20.7
H	Vérité terrain	703236.0	4795092.0	83.0	20.0	20.0	15.0
	Estimation	703235.9	4795091.0	82.7	21.0	24.5	13.1
I	Vérité terrain	703264.0	4795072.0	83.0	20.0	20.0	15.0
	Estimation	703264.9	4795071.5	84.1	20.8	22.6	12.5

Tableau 5.2 – Résultats de la méthode d'extraction sur données réelles en imagerie optique : valeurs estimées et valeurs vraies des paramètres des neuf bâtiments de la scène d'étude.

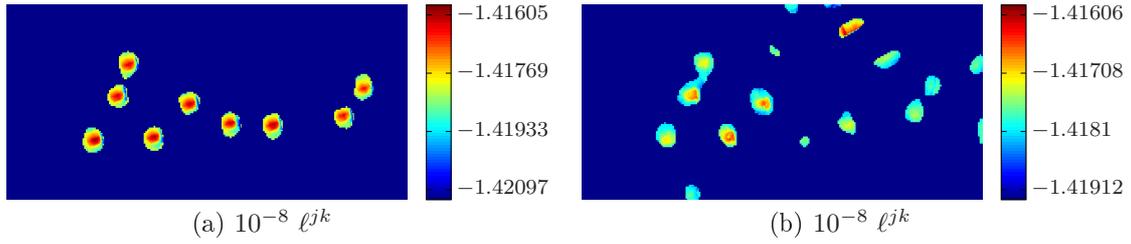


Figure 5.65 – Cartes de log-vraisemblance généralisée établies à partir des images binaires B_I^0 et B_I^1 simulées (a) et réelles (b) des figures 5.21 (a), 5.23 (a), 5.22 (a) et 5.24 (a).

La mise en œuvre de la stratégie de détection, présentée à la section précédente, sur la carte de log-vraisemblance généralisée de la figure 5.65 (a), permet d’extraire les bâtiments dont les paramètres sont donnés au tableau 5.3 et dont les signatures géométriques sont représentées dans les images binaires B_I^0 et B_I^1 à la figure 5.66. Les erreurs maximales, qui sont observées pour les bâtiments D , E et I , valent respectivement 1.7 m, 0.5 m, 0.9 degrés, 2.9 m, 1.0 m et 4.9 m pour les paramètres X_C , Y_C , α , w_1 , w_2 et h . Étant donné l’orientation relative du repère terrain \mathcal{R}_T par rapport au repère image \mathcal{R}_I et l’orientation des bâtiments dans l’image, les erreurs affectant les paramètres X_C , w_1 (longueur du segment le plus fin de l’écho double) et h sont logiquement corrélées.

Pour finir, les images de la figure 5.67 et le tableau 5.4 présentent les résultats d’extraction obtenus à partir de la carte de log-vraisemblance généralisée de la figure 5.65 (b). La stratégie de détection mise en œuvre permet de détecter sept bâtiments sur les neuf présents dans la scène. Elle conduit également à la détection de quatre fausses alarmes, dont une pourrait facilement être supprimée en interdisant les recouvrements entre emprises au sol, puisqu’elle correspond à une détection multiple du bâtiment A . Les autres fausses alarmes sont liées à la présence d’un terrain de sport qui engendre, dans les images binaires résultant de l’extraction des primitives, des structures dont l’agencement et la forme correspondent à ceux d’un bâtiment. En ce qui concerne les performances d’estimation, elles restent relativement limitées, en raison notamment de l’extraction bruitée des primitives et de la difficulté qui en résulte de localiser précisément le point de référence des bâtiments.

5.6 Conclusion

Dans ce chapitre, nous avons développé une seconde méthode pour l’extraction tridimensionnelle des bâtiments parallélépipédiques en monoscopie optique et RSO. Cette méthode réalise une détection-estimation conjointe des bâtiments à partir d’images binaires résultant de l’extraction des primitives caractéristiques de la présence de ces bâtiments dans les images optiques et RSO. La méthode proposée met en œuvre des tests et estimateurs statistiques qui s’appuient sur la vraisemblance géométrique des pixels détectés dans ces images binaires. Le calcul de cette vraisemblance fait intervenir la densité de probabilité de présence des pixels détectés, qui est établie en simulant, au moyen des méthodes développées au chapitre 2, les primitives observées dans les images binaires.

En imagerie optique, comme en imagerie RSO, la méthode proposée fournit des résultats pertinents pour la détection des bâtiments, notamment lorsque les images binaires résultant de l’extraction des primitives sont considérées conjointement.

La méthode proposée pour l’estimation des paramètres des bâtiments, dont le centre de l’emprise au sol est supposé connu, conduit également à de bons résultats. Ces résultats sont d’autant plus satisfaisants que les images binaires sont issues de détecteurs de primitives dont les performances, aussi bien en termes de détection que de fausse alarme, sont relativement limitées, illustrant ainsi la robustesse de cette estimation. La précision avec laquelle sont estimés les paramètres des bâtiments extraits reste cependant relativement limitée en raison de la difficulté de localiser le centre de l’emprise au sol de ces bâtiments.

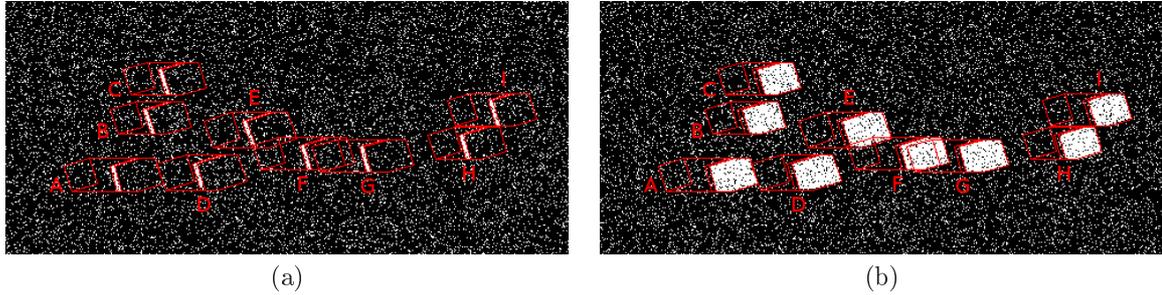


Figure 5.66 – Résultats de la méthode d'extraction sur données simulées en imagerie RSO : simulation des signatures géométriques des bâtiments extraits dans les images binaires de primitives B_I^0 (a) et B_I^1 (b) simulées.

Label		X_C (m)	Y_C (m)	α (degrés)	w_1 (m)	w_2 (m)	h (m)
A	Vérité terrain	702995.0	4795064.0	83.0	20.0	20.0	22.0
	Estimation	702994.9	4795064.0	83.3	20.0	20.1	22.0
B	Vérité terrain	703025.0	4795026.0	83.0	20.0	20.0	15.0
	Estimation	703024.9	4795026.0	83.1	19.5	19.7	16.0
C	Vérité terrain	703042.0	4794999.0	83.0	20.0	20.0	15.0
	Estimation	703041.3	4794999.0	83.5	18.8	20.0	17.0
D	Vérité terrain	703053.0	4795074.0	83.0	20.0	20.0	22.0
	Estimation	703054.3	4795074.5	82.5	22.9	20.9	17.3
E	Vérité terrain	703088.5	4795050.0	83.0	20.0	20.0	22.0
	Estimation	703089.3	4795050.5	82.1	21.5	21.0	18.9
F	Vérité terrain	703120.0	4795076.0	83.0	20.0	20.0	22.0
	Estimation	703119.8	4795076.0	83.8	19.6	19.8	22.9
G	Vérité terrain	703159.0	4795086.0	83.0	20.0	20.0	22.0
	Estimation	703158.6	4795086.0	83.3	19.8	20.2	22.7
H	Vérité terrain	703236.0	4795092.0	83.0	20.0	20.0	15.0
	Estimation	703235.9	4795091.5	83.0	20.0	19.0	15.1
I	Vérité terrain	703264.0	4795072.0	83.0	20.0	20.0	15.0
	Estimation	703262.3	4795072.0	82.6	17.3	20.0	19.9

Tableau 5.3 – Résultats de la méthode d'extraction sur données simulées en imagerie RSO : valeurs estimées et valeurs vraies des paramètres des neuf bâtiments de la scène d'étude.

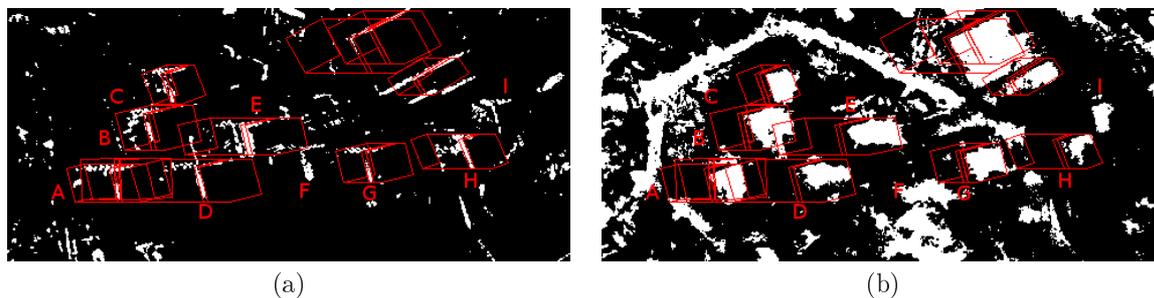


Figure 5.67 – Résultats de la méthode d'extraction sur données simulées en imagerie RSO : simulation des signatures géométriques des bâtiments extraits dans les images binaires de primitives B_I^0 (a) et B_I^1 (b) simulées.

Label		X_C (m)	Y_C (m)	α (degrés)	w_1 (m)	w_2 (m)	h (m)
A	Vérité terrain	702995.0	4795064.0	83.0	20.0	20.0	22.0
	Estimation	702990.9	4795064.0	87.0	13.6	26.1	20.7
B	Vérité terrain	703025.0	4795026.0	83.0	20.0	20.0	15.0
	Estimation	703027.8	4795028.5	86.9	27.4	28.6	13.8
C	Vérité terrain	703042.0	4794999.0	83.0	20.0	20.0	15.0
	Estimation	703043.2	4794998.5	79.8	19.4	24.1	9.2
D	Vérité terrain	703053.0	4795074.0	83.0	20.0	20.0	22.0
	Estimation	703054.8	4795076.0	85.1	23.7	27.3	30.0
E	Vérité terrain	703088.5	4795050.0	83.0	20.0	20.0	22.0
	Estimation	703087.8	4795050.0	87.7	25.4	23.6	30.0
F	Vérité terrain	703120.0	4795076.0	83.0	20.0	20.0	22.0
	Estimation	-	-	-	-	-	-
G	Vérité terrain	703159.0	4795086.0	83.0	20.0	20.0	22.0
	Estimation	703161.1	4795086.5	86.0	24.1	24.8	12.7
H	Vérité terrain	703236.0	4795092.0	83.0	20.0	20.0	15.0
	Estimation	703231.6	4795091.5	77.7	13.3	23.6	22.9
I	Vérité terrain	703264.0	4795072.0	83.0	20.0	20.0	15.0
	Estimation	-	-	-	-	-	-

Tableau 5.4 – Résultats de la méthode d'extraction sur données réelles en imagerie RSO : valeurs estimées et valeurs vraies des paramètres des neuf bâtiments de la scène d'étude.

Comme au chapitre 4, il conviendrait de mettre en œuvre cette méthode sur un ensemble d'images avec des prises de vue variées afin de pouvoir qualifier objectivement les performances de la chaîne d'extraction proposée.

Comme la plupart des méthodes décrites au chapitre 3, la méthode développée repose sur l'extraction des primitives caractéristiques de la présence des bâtiments dans les images optiques et RSO. La méthode se distingue toutefois de ces dernières puisqu'elle ne vectorise pas les primitives extraites mais s'appuie directement sur les images binaires issues de cette extraction de primitives. Ainsi, l'approche proposée reste relativement peu paramétrée et sa mise en œuvre ne nécessite de fixer qu'un nombre restreint de paramètres et de seuils dont le détail est ici rappelé :

- paramètres ϵ_T ou ϵ_{ED} qui correspondent respectivement à la largeur des contours de toit en imagerie optique et à la largeur des échos doubles en imagerie RSO ;
- paramètres w_1^{min} , w_1^{max} , w_2^{min} , w_2^{max} , h^{min} et h^{max} qui désignent les dimensions maximales des bâtiments à extraire ;
- seuils de détection $\tau_{\ell jk}$ impliqués lors des étapes de détection grossière et de détection fine, seuils qui ont été déterminés de manière supervisée.

Aux seuils et paramètres précédents, s'ajoutent également ceux nécessaires à l'extraction des primitives caractéristiques de la présence des bâtiments. Cette extraction de primitives peut être réalisée par une méthode quelconque pour peu qu'elle permette d'établir les images binaires de primitives B_I^j nécessaires à la mise en œuvre de la méthode d'extraction de bâtiments proposée.

Les principales limitations de la méthode développée sont liées, comme au chapitre précédent, à la simplicité du modèle de bâtiment adopté et à la non prise en compte des phénomènes de recouvrement. La mise en œuvre de la méthode avec des modèles paramétriques de bâtiment plus complexe est *a priori* possible, ce qu'il conviendrait de valider par une étude complémentaire. Enfin, la prise en compte des phénomènes de recouvrement est envisageable mais nécessiterait de calculer localement la densité de probabilité de présence des pixels détectés dans les images binaires, engendrant ainsi une complexité calculatoire plus importante.

Conclusion

Travaux réalisés

Après avoir rappelé les notions de base de l'imagerie de télédétection au chapitre 1, nous avons exposé, au chapitre 2, les concepts photogrammétriques et radargrammétriques qui permettent d'expliquer la signature géométrique des bâtiments dans les images de télédétection optiques et RSO. Des algorithmes de projection, au moyen desquels sont simulées ces signatures géométriques, ont alors été établis dans le cas de bâtiments parallélépipédiques.

Les différentes méthodes proposées dans la littérature pour l'extraction des bâtiments en imagerie optique et RSO, méthodes qui s'appuient sur les concepts photogrammétriques et radargrammétriques précédemment introduits, ont ensuite été décrites au chapitre 3. De cet état de l'art, il ressort notamment que la plupart des méthodes développées sont spécifiques à un type de capteur (optique ou RSO) et dépendent de la nature monoscopique ou multiscopique des données.

Nous avons alors proposé, aux chapitres 4 et 5, deux nouvelles méthodes pour l'extraction tridimensionnelle de bâtiments à partir d'une unique image satellitaire très haute résolution, ayant indifféremment été acquise par un capteur optique ou RSO.

Ces deux méthodes reposent sur les algorithmes de projection développés au chapitre 2. En modélisant les effets géométriques liés à la projection des bâtiments, ces algorithmes permettent de prendre en compte la complexité géométrique et radiométrique inhérente aux images haute résolution, complexité qui constitue l'une des principales difficultés de l'analyse de scènes urbaines en imagerie satellitaire.

De plus, chacune des deux méthodes met en œuvre un formalisme unique pour l'extraction des bâtiments en imagerie optique et RSO, ce qui constitue l'une des contributions majeures de ces travaux et laisse entrevoir de nombreuses applications pour l'analyse de scène urbaine dans un contexte multicapteur.

La première méthode proposée s'appuie sur deux critères quantifiant l'adéquation de la signature géométrique d'un bâtiment avec l'image au sein de laquelle cette signature a été simulée.

Plus précisément, le premier critère quantifie l'homogénéité radiométrique des différentes régions constitutives de la signature géométrique d'un bâtiment. Il fait intervenir, pour chacune de ces régions, la vraisemblance généralisée des niveaux de gris des pixels contenus dans cette région. Cette vraisemblance est calculée en supposant que les niveaux de gris des pixels suivent respectivement une loi normale et une loi de Rayleigh au sein des images optiques et des images RSO d'amplitude. Le choix de ces lois a été justifié par des considérations physiques sur la nature du bruit présent dans les images optiques et RSO d'amplitude.

Le second critère mesure l'adéquation des segments constitutifs de la signature géométrique d'un bâtiment avec les bords présents dans l'image. En imagerie optique, où le bruit est additif, le critère repose sur la différence des radiométries moyennes calculées dans des voisinages situés de part et d'autre des différents segments. En imagerie RSO, ce critère fait intervenir le rapport des radiométries

quadratiques moyennes calculées dans ces mêmes voisinages, afin de prendre en compte la nature multiplicative du bruit présent dans ces images.

Après avoir étudié qualitativement et quantitativement le comportement de ces critères sur données simulées et sur données réelles, une stratégie d'optimisation a été établie. Celle-ci repose sur la maximisation successive de ces deux critères, à partir d'un ensemble d'états initiaux répartis uniformément sur l'espace des paramètres. Cette stratégie, qui tire parti de l'attractivité du premier critère et de la sélectivité du second, permet d'estimer les paramètres des bâtiments, en imagerie optique comme en imagerie RSO, avec une précision métrique, dans le cas de données simulées.

Cette première méthode, qui s'apparente à une segmentation locale de l'image contrainte par un modèle de bâtiment, permet d'estimer les paramètres d'un bâtiment. La complexité calculatoire qu'elle implique rend, en revanche, difficile sa mise en œuvre à des fins de détection, bien que celle-ci reste envisageable.

La seconde méthode réalise l'extraction des bâtiments à partir d'images binaires résultant de l'extraction des primitives caractéristiques de ces bâtiments dans les images optiques et RSO. Contrairement à la méthode précédente, qualifiée d'approche radiométrique, la seconde méthode s'appuie uniquement sur la position des pixels détectés dans les images binaires issues de l'extraction des primitives, de sorte que cette seconde méthode est qualifiée de géométrique.

Cette approche géométrique réalise une détection-estimation conjointe des bâtiments présents dans l'image. Elle s'appuie sur le calcul d'une vraisemblance géométrique impliquant la position des pixels détectés dans les images binaires de primitives. Cette vraisemblance géométrique fait intervenir la densité de probabilité de présence des pixels détectés dans les images binaires issues de l'extraction des primitives, densité de probabilité qui est calculée *a priori* pour un modèle de bâtiment donné et pour une géométrie de prise de vue donnée.

Ainsi, la méthode proposée s'appuie sur un formalisme statistique qui offre de nombreux avantages. Ce formalisme permet de s'affranchir de la vectorisation des primitives, évitant ainsi l'introduction d'un certain nombre de paramètres qui sont le plus souvent fixés empiriquement. Ce formalisme permet également de prendre naturellement en compte les performances des détecteurs au moyen desquels les primitives ont été extraites, par l'intermédiaire des probabilités de détection et de fausse alarme de ces détecteurs, ce qui constitue une des originalités de ce travail. Enfin, les informations *a priori*, mêmes partielles, sur l'emplacement et les dimensions des bâtiments présents dans la scène peuvent, lorsqu'elles sont disponibles, être facilement intégrées au sein de ce formalisme.

En termes de performances, l'approche proposée se caractérise par sa robustesse face à une extraction bruitée des primitives, ce qui constitue un des atouts majeurs de cette approche. Sa mise en œuvre sur données simulées et sur données réelles permet d'obtenir des performances satisfaisantes en détection comme en estimation, hormis dans certains cas où les phénomènes d'occlusion entre bâtiments sont importants.

Applications

La principale application de ce travail doctoral est l'extraction automatique ou supervisée de bâtiments en monoscopie optique et RSO, l'objectif étant d'éviter une extraction manuelle qui s'avère souvent fastidieuse et parfois imprécise en raison de la difficulté d'identifier les points devant être saisis dans les images.

Une extraction automatique peut être réalisée au moyen de l'approche géométrique proposée au chapitre 5. Cette méthode nécessite toutefois que soient fixés un certain nombre de seuils pour l'extraction des primitives ainsi que pour les étapes de détection grossière et de détection fine des bâtiments.

Les méthodes développées au chapitre 4 et 5 peuvent également être mises en œuvre dans le cadre d'une extraction supervisée, comme cela est décrit dans le brevet [BMD11]. L'action de l'opérateur est limitée à la saisie approximative du centre du toit du bâtiment, en imagerie optique, ou du centre de son emprise au sol, en imagerie RSO. Le bâtiment peut alors être extrait au moyen

de l'approche radiométrique en maximisant les critères d'adéquation à partir de plusieurs états initiaux, pour lesquels la position du bâtiment correspond au point saisi dans l'image par l'opérateur et les autres paramètres (orientation et dimensions) résultent de tirages aléatoires dans l'espace des paramètres ou d'un parcours régulier de cet espace. L'approche géométrique peut également être mise en œuvre pour extraire ce bâtiment en assimilant l'étape de saisie de l'opérateur à l'étape de détection grossière et en appliquant les autres étapes de la chaîne d'extraction dans un voisinage du point saisi.

Une autre application possible de ce travail est le recalage d'images satellitaires haute résolution par mise en correspondance de bâtiments homologues.

En raison des erreurs affectant les MNT et de la précision limitée avec laquelle sont connus les paramètres d'acquisition d'une image satellitaire (position, vitesse et attitude du satellite), le recalage de deux images au moyen de leurs fonctions de localisation est bien souvent insuffisant pour atteindre la précision requise par les utilisateurs. Ce recalage peut alors être amélioré par la mise en correspondance de points homologues dans les deux images, ces couples de points homologues étant saisis par un opérateur ou bien déterminés automatiquement.

Dans un contexte multicapteur optique-RSO haute résolution, la mise en correspondance, même visuelle, de points homologues entre les images est compliquée, d'une part, par la forte hétérogénéité radiométrique et géométrique entre ces images et, d'autre part, par la nécessité de mettre en correspondance des points au sol. En raison de cette complexité, les méthodes automatiques sont mises en défaut et la saisie supervisée de ces couples de points homologues s'avère particulièrement délicate. Une solution possible, qui a également fait l'objet d'un dépôt de brevet [LRB11], consiste à mettre en correspondance non plus des points homologues mais des bâtiments homologues, bâtiments qui peuvent être extraits selon le procédé décrit précédemment. Cette solution, qui fournit un ensemble de segments homologues au sein du couple d'images à recalcr, présente l'avantage de ne nécessiter de la part de l'utilisateur que la saisie grossière du centre du toit ou de l'emprise au sol des bâtiments homologues.

Liste des travaux

- [BLRD11] E. Barthelet, C. Louis, S. Reynaud, and L. Denise. *Procédé de modélisation de bâtiments à partir d'une image géoréférencée*, 2011. Brevet numéro FR11/02288.
- [BMD11] E. Barthelet, G. Mercier, and L. Denise. Building change detection in a couple of optical and SAR high resolution images. In *IEEE International Geoscience and Remote Sensing Symposium*, pages 2393–2396, July 2011.
- [BMDR12a] E. Barthelet, G. Mercier, L. Denise, and S. Reynaud. A new approach for three-dimensional building extraction in high-resolution monoscopic SAR imagery. In *9th European Conference on Synthetic Aperture Radar*, pages 48–51. VDE, 2012.
- [BMDR12b] E. Barthelet, G. Mercier, L. Denise, and S. Reynaud. Feature based maximum likelihood model inversion for three-dimensional building extraction from single high resolution SAR images. In *IEEE International Geoscience and Remote Sensing Symposium*, pages 3600–3603, July 2012.
- [BMDR12c] E. Barthelet, G. Mercier, L. Denise, and S. Reynaud. *Maximum Likelihood Model Inversion for 3D Building Extraction in High Resolution Optical and SAR Imagery*, July 2012. Séminaire Equipe AYIN, INRIA Sophia Antipolis Méditerranée.
- [LRB11] C. Louis, S. Reynaud, and E. Barthelet. *Procédé de mise en correspondance de plusieurs images associées à des référencements géographiques incohérents*, 2011. Brevet numéro FR 11/02229.

Bibliographie

- [BAE07] BAE Systems. *Trend in Digital Aerial Imaging*, 2007.
- [BB03] A.J. Bennett and D. Blacknell. The extraction of building dimensions from high resolution SAR imagery. In *Proceedings of the International Radar Conference*, pages 182–187, 2003.
- [BCM05] A. Buades, B. Coll, and J.-M. Morel. A non-local algorithm for image denoising. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages 60–65 vol. 2, June 2005.
- [BD62] R. Bellman and S. Dreyfus. *Applied Dynamic Programming*. Princeton University Press, 1962.
- [BDZ12] C. Benedek, X. Descombes, and J. Zerubia. Building development monitoring in multitemporal remotely sensed image pairs with stochastic birth-death dynamics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(1) :33–50, January 2012.
- [BG11] P.T.B. Brett and R. Guida. Bright line detection in cosmo-skymed sar images of urban areas. In *Joint Urban Remote Sensing Event*, pages 253–256, April 2011.
- [BLBG10] D. Brunner, G. Lemoine, L. Bruzzone, and 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, March 2010.
- [BM99] C. Baillard and H. Maître. 3-D 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 Proceedings*, volume 3, pages 287–292, 2000.
- [Can86] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (6) :679–698, November 1986.
- [CBV04] J. Chanussot, J.A. Benediktsson, and M. Vincent. Classification of remote sensing images from urban areas using a fuzzy model. In *IEEE International Geoscience and Remote Sensing Symposium*, pages 556–559, September 2004.
- [CHA10] M. Cetin, U. Halici, and O. Aytakin. Building detection in satellite images by textural features and Adaboost. In *IAPR Workshop on Pattern Recognition in Remote Sensing*, pages 1–4, August 2010.
- [CJDS03] N. Chehata, F. Jung, M.P. Deseilligny, and G. Stamon. A region-based matching approach for 3D-roof reconstruction from HR satellite stereo pairs. In *Digital Image Computing Techniques and Applications*, volume 2, pages 889–898, 2003.
- [CM99] D. Comaniciu and P. Meer. Mean shift analysis and applications. In *The Proceedings of the Seventh IEEE International Conference on Computer Vision*, volume 2, pages 1197–1203, 1999.
- [CON05] F. Cellier, H. Oriot, and J.M. Nicolas. Introduction of the mean shift algorithm in SAR imagery : Application to shadow extraction for building reconstruction. In *Proceedings of EARSeL Workshop*, 2005.

- [CON06] F. Cellier, H. Oriot, and J.-M. Nicolas. Hypothesis Management for Building Reconstruction from High Resolution InSAR Imagery. In *IEEE International Geoscience and Remote Sensing Symposium*, pages 3639–3642, August 2006.
- [CRB99] C. Chesnaud, P. Réfrégier, and V. Boulet. Statistical Region Snake-Based Segmentation Adapted to Different Physical Noise Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(11) :1145–1157, November 1999.
- [CS12] M. Cote and P. Saeedi. Automatic rooftop extraction in nadir aerial imagery of suburban regions using corners and variational level set evolution. *IEEE Transactions on Geoscience and Remote Sensing*, (99) :1–16, 2012.
- [DDT09] C.A. Deledalle, L. Denis, and F. Tupin. Iterative weighted maximum likelihood denoising with probabilistic patch-based weights. *IEEE Transactions on Image Processing*, 18(12) :2661–2672, 2009.
- [Dem67] A.P. Dempster. Upper and lower probability inferences based on a sample from a finite univariate population. *Biometrika*, 54 :515–528, 1967.
- [Den09] L. Denise. SAR Image Exploitation : Introduction to SAR Image Functionalities. Technical report, Thales Communications & Security, Vélizy, France, 2009.
- [Dig06] DigitalGlobe, Longmont, CO, USA. *QuickBird Imagery Products : Product Guide*, 2006. Revision 4.7.1.
- [EAD02] EADS Astrium, Toulouse, France. *SPOT Satellite Geometry Handbook*, 2002.
- [EAD12] EADS Astrium, Toulouse, France. *Pléiades 1 & 2 : General Overview*, 2012.
- [FBB09] A. Ferro, D. Brunner, and L. Bruzzone. An advanced technique for building detection in VHR SAR images. In *Proceedings of SPIE, the International Society for Optical Engineering*, volume 7477, page 27, 2009.
- [FBB10] A. Ferro, D. Brunner, and L. Bruzzone. Building detection and radar footprint reconstruction from single VHR SAR images. In *IEEE International Geoscience and Remote Sensing Symposium*, pages 292–295, 2010.
- [FGI⁺06] G. Franceschetti, R. Guida, A. Iodice, D. Riccio, and G. Ruello. Accuracy of Building Height Estimation from SAR Images. In *IEEE International Geoscience and Remote Sensing Symposium*, pages 3647–3650, August 2006.
- [FGI⁺07] G. Franceschetti, R. Guida, A. Iodice, D. Riccio, G. Ruello, and U. Stilla. Building feature extraction via a deterministic approach : application to real high resolution SAR images. In *IEEE International Geoscience and Remote Sensing Symposium*, pages 2681–2684, 2007.
- [FIR02] G. Franceschetti, A. Iodice, and D. Riccio. A canonical problem in electromagnetic backscattering from buildings. *IEEE Transactions on Geoscience and Remote Sensing*, 40(8) :1787–1801, August 2002.
- [FKL⁺98] A. Fischer, T. H. Kolbe, F. Lang, A. B. Cremers, W. Förstner, L. Plümer, and V. Steinhage. Extracting Buildings from Aerial Images Using Hierarchical Aggregation in 2D and 3D. *Computer Vision and Image Understanding*, 72(2) :185–203, 1998.
- [FMR01] M. Fradkin, H. Maitre, and M. Roux. Building Detection from Multiple Aerial Images in Dense Urban Areas. *Computer Vision and Image Understanding*, 82(3) :181–207, 2001.
- [För94] W. Förstner. A framework for low level feature extraction. In *Lecture Notes in Computer Science*, volume 801, pages 383–394. Springer, 1994.
- [GBR03] Frédéric Galland, Nicolas Bertaux, and Philippe Réfrégier. Minimum description length synthetic aperture radar image segmentation. *IEEE Transactions on Image Processing*, 12(9) :995–1006, 2003.
- [GHS00] P. Gamba, B. Houshmand, and M. Sacconi. Detection and extraction of buildings from interferometric SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 38(1) :611–617, 2000.
- [GIR10] R. Guida, A. Iodice, and D. Riccio. Height Retrieval of Isolated Buildings From Single High-Resolution SAR Images. *IEEE Transactions on Geoscience and Remote Sensing*, 48(7) :2967–2979, 2010.

- [Hei04] A. Hein. *Processing of SAR data : Fundamentals, Signal Processing, Interferometry*. Springer Verlag, 2004.
- [HSD73] R.M. Haralick, K. Shanmugam, and I. Dinstein. Textural Features for Image Classification. *IEEE Transaction on Systems, Man and Cybernetics*, 3(6) :610–621, November 1973.
- [Inf09] Infoterra GmbH. *TerraSAR-X Services - Image Product Guide : Basic and Enhanced Radar Satellite Imagery*, 2009.
- [IS10] M. Izadi and P. Saeedi. Automatic building detection in aerial images using a hierarchical feature based image segmentation. In *International Conference on Pattern Recognition*, pages 472–475, August 2010.
- [IS12] M. Izadi and P. Saeedi. Three-dimensional polygonal building model estimation from single satellite images. *IEEE Transactions on Geoscience and Remote Sensing*, 50(6) :2254–2272, June 2012.
- [JPDM00] H. Jibrini, N. Paparoditis, M.P. Deseilligny, and H. Maitre. Automatic building reconstruction from very high resolution aerial stereopairs using cadastral ground plans. In *International Society for Photogrammetry and Remote Sensing Congress*, pages 16–23, 2000.
- [KE01] M. Kasser and Y. Egels. *Photogrammétrie numérique*. Hermes Science Publications, 2001.
- [KHN00] Zu Whan Kim, A. Huertas, and R. Nevatia. Automatic description of complex buildings with multiple images. In *IEEE Workshop on Applications of Computer Vision*, pages 155–162, 2000.
- [KN04] T. Knudsen and A.A. Nielsen. Detection of buildings through multivariate analysis of spectral, textural, and shape based features. In *IEEE International Geoscience and Remote Sensing Symposium*, volume 5, pages 2830–2833, September 2004.
- [KP09] K. Karantzas and N. Paragios. Recognition-driven two-dimensional competing priors toward automatic and accurate building detection. *IEEE Transactions on Geoscience and Remote Sensing*, 47(1) :133–144, January 2009.
- [KP10] K. Karantzas and N. Paragios. Large-Scale Building Reconstruction Through Information Fusion and 3-D Priors. *IEEE Transactions on Geoscience and Remote Sensing*, 48(5) :2283–2296, May 2010.
- [KS08] A. Katartzis and H. Sahli. A stochastic framework for the identification of building rooftops using a single remote sensing image. *IEEE Transactions on Geoscience and Remote Sensing*, 46(1) :259–271, 2008.
- [KSSC85] D.T. Kuan, A.A. Sawchuk, T.C. Strand, and P. Chavel. Adaptive Noise Smoothing Filter for Images with Signal-Dependent Noise. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7(2) :165–177, March 1985.
- [KWT88] M. Kass, A. Witkin, and D. Terzopoulos. Snakes : Active contour models. *International Journal on Computer Vision*, 1(4) :321–331, 1988.
- [LDZPD06] F. Lafarge, X. Descombes, J. Zerubia, and M. Pierrot-Deseilligny. An automatic building reconstruction method : A structural approach using high resolution satellite images. In *IEEE International Conference on Image Processing*, pages 1205–1208, oct. 2006.
- [Lee80] J.S. Lee. Digital Image Enhancement and Noise Filtering by Use of Local Statistics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2(2) :165–168, March 1980.
- [LHN94] C. Lin, A. Huertas, and R. Nevatia. Detection of buildings using perceptual grouping and shadows. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 62–69, 1994.
- [LL00] M. Lucas and E. Languéno. *Synthèse d'image*. Editions Techniques de l'Ingénieur, 2000.
- [LMP01] J. D. Lafferty, A. McCallum, and F. C. N. Pereira. Conditional random fields : Probabilistic models for segmenting and labeling sequence data. In *Proceedings of the Eighteenth International Conference on Machine Learning*, pages 282–289, San Francisco, CA, USA, 2001. Morgan Kaufmann Publishers Inc.

- [LN95] C. Lin and R. Nevatia. 3-D descriptions of buildings from an oblique view aerial image. In *International Symposium on Computer Vision*, pages 377–382, November 1995.
- [LN98] C. Lin and R. Nevatia. Building detection and description from a single intensity image. *Computer Vision and Image Understanding*, 72(2) :101–121, 1998.
- [LNTL93] A. Lopes, E. Nezry, R. Touzi, and H. Laur. Structure detection and statistical adaptive speckle filtering in SAR images. *International Journal of Remote Sensing*, 14(9) :1735–1758, 1993.
- [LP05] W. Liu and V. Prinet. Building detection from high-resolution satellite image using probability model. In *IEEE International Geoscience and Remote Sensing Symposium*, volume 6, pages 3888–3891, 2005.
- [LTN90] A. Lopes, R. Touzi, and E. Nezry. Adaptive speckle filters and scene heterogeneity. *IEEE Transactions on Geoscience and Remote Sensing*, 28(6) :992–1000, November 1990.
- [Mai01] H. Maître. *Traitement des images de radar à synthèse d’ouverture*. Hermes Science Publication, 2001.
- [May99] H. Mayer. Automatic object extraction from aerial imagery—A survey focusing on buildings. *Computer Vision and Image Understanding*, 74(2) :138–149, 1999.
- [MBM01] E.M. Mikhail, J.S. Bethel, and J.C. McGlone. *Introduction to Modern Photogrammetry*. Wiley New York, NY, 2001.
- [MH80] D. Marr and E. Hildreth. Theory of edge detection. *Proceedings of the Royal Society of London. Series B. Biological Sciences.*, 207(1167) :187–217, 1980.
- [MMB04] C. McGlone, E. Mikhail, and J. Bethel. *Manual of Photogrammetry*. Fifth edition, 2004.
- [MN89] R. Mohan and R. Nevatia. Using perceptual organization to extract 3D structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(11) :1121–1139, November 1989.
- [MS94] J.C. McGlone and J.A. Shufelt. Projective and object space geometry for monocular building extraction. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 54–61, June 1994.
- [MSS02] E. Michaelsen, U. Soergel, and U. Stilla. Grouping Salient Scatterers in InSAR Data for Recognition of Industrial Buildings. In *International Conference on Pattern Recognition*, volume 2, pages 613–616, 2002.
- [MSSD08] E. Michaelsen, U. Stilla, U. Soergel, and L. Doktorski. Extraction of building polygons from SAR images : Grouping and decision-level in the GESTALT System. In *IAPR Workshop on Pattern Recognition in Remote Sensing*, pages 1–4, December 2008.
- [MTN07] V. Le Moigne, F. Tupin, and J. M. Nicolas. Statistical Polygonal Snakes for 3D building reconstruction using High Resolution SAR data. In *Joint Urban Remote Sensing Event*, pages 1–5, 2007.
- [MZ05] S. Muller and D.W. Zaum. Robust building detection in aerial images. In *International Archives of Photogrammetry, Remote Sensing, and Spatial Information Sciences*, volume 36, pages 143–148, August 2005.
- [NB80] R. Nevatia and K. Ramesh Babu. Linear feature extraction and description. *Computer Graphics and Image Processing*, 13(3) :257–269, 1980.
- [NLH97] R. Nevatia, C. Lin, and A. Huertas. A system for building detection from aerial images. In *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, pages 77–86, 1997.
- [NN01] S. Noronha and R. Nevatia. Detection and modeling of buildings from multiple aerial images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(5) :501–518, May 2001.
- [ODZ07] M. Ortner, X. Descombes, and J. Zerubia. Building outline extraction from digital elevation models using marked point processes. *International Journal of Computer Vision*, 72(2) :107–132, 2007.

- [Ori03] H. Oriot. Statistical snakes for building extraction from stereoscopic aerial images. In *International Archives of Photogrammetry Remote Sensing And Spatial Information Sciences*, volume 34, pages 65–72, 2003.
- [PAD00] D. Petit, F. Adragna, and J.D. Durou. The filtering of layover areas in high-resolution IFSAR for the building extraction. In *Proceedings of SPIE, the International Society for Optical Engineering*, volume 4173, pages 230–240, 2000.
- [PIS⁺11] V. Poulain, J. Inglada, M. Spigai, J.-Y. Tourneret, and P. Marthon. High-Resolution Optical and SAR Image Fusion for Building Database Updating. *IEEE Transactions on Geoscience and Remote Sensing*, 49(8) :2900–2910, August 2011.
- [QD04] M. Quartulli and 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.
- [Ris78] J. Rissanen. Modeling by shortest data description. *Automatica*, 14 :465–471, 1978.
- [RJ98] J.A. Richards and X. Jia. *Remote Sensing Digital Image Analysis*. Springer Verlag, Third edition, 1998.
- [SC83] M. Suk and S. M. Chung. A new image segmentation technique based on partition mode test. *Pattern Recognition*, 16(5) :469–480, 1983.
- [SC00] A. Stassopoulou and T. Caelli. Building detection using bayesian networks. *International Journal of Pattern Recognition and Artificial Intelligence*, 14(6) :715–734, 2000.
- [Sha76] G. Shafer. *A mathematical theory of evidence*. Princeton University Press, 1976.
- [Shu96] J.A. Shufelt. Exploiting Photogrammetric Methods for Building Extraction in Aerial Images. In *International Archives of Photogrammetry and Remote Sensing*, volume 31, pages 74–79, 1996.
- [SOG05] E. Simonetto, H. Oriot, and R. Garello. Rectangular building extraction from stereoscopic airborne radar images. *IEEE Transactions on Geoscience and Remote Sensing*, 43(10) :2386–2395, October 2005.
- [SOGC03] E. Simonetto, H. Oriot, R. Garello, and J. Le Caillec. Radargrammetric processing for 3-D building extraction from high-resolution airborne SAR data. In *IEEE International Geoscience and Remote Sensing Symposium*, volume 3, pages 2002–2004, July 2003.
- [SPY06] Zongying Song, Chunhong Pan, and Q. Yang. A Region-Based Approach to Building Detection in Densely Build-Up High Resolution Satellite Image. In *IEEE International Conference on Image Processing*, pages 3225–3228, October 2006.
- [ST10] D. Koc San and M. Turker. Building extraction from high resolution satellite images using hough transform. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Science*, 38, 2010.
- [STD09] H. Sportouche, F. Tupin, and L. Denise. Building extraction and 3D reconstruction in urban areas from high-resolution optical and SAR imagery. In *Joint Urban Remote Sensing Event*, pages 1–11, 2009.
- [STD11] H. Sportouche, F. Tupin, and L. Denise. Extraction and Three-Dimensional Reconstruction of Isolated Buildings in Urban Scenes From High-Resolution Optical and SAR Spaceborne Images. *IEEE Transactions on Geoscience and Remote Sensing*, 49(10) :3932–3946, 2011.
- [Ste98] C. Steger. An unbiased detector of curvilinear structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(2) :113–125, 1998.
- [Tai05] F. Taillandier. Automatic building reconstruction from cadastral maps and aerial images. In *International Archives of Photogrammetry and Remote Sensing*, volume 36, pages 105–110, 2005.
- [TBBD10] O. Tournaire, M. Brédif, D. Boldo, and M. Durupt. An efficient stochastic approach for building footprint extraction from digital elevation models. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(4) :317–327, 2010.

- [TCS+06] A. Thiele, E. Cadario, K. Schulz, U. Thoennessen, and U. Soergel. Building recognition fusing multi-aspect high-resolution interferometric sar data. In *IEEE International Geoscience and Remote Sensing Symposium*, pages 3643–3646, August 2006.
- [TCS+07a] A. Thiele, E. Cadario, K. Schulz, U. Thoennessen, and U. Soergel. Modeling and analyzing insar phase profiles at building locations. In *IEEE International Geoscience and Remote Sensing Symposium*, pages 5053–5056, 2007.
- [TCS+07b] A. Thiele, E. Cadario, K. Schulz, U. Thoennessen, and U. Soergel. Building Recognition From Multi-Aspect High-Resolution InSAR Data in Urban Areas. *IEEE Transactions on Geoscience and Remote Sensing*, 45(11) :3583–3593, November 2007.
- [THC10] A. Thiele, S. Hinz, and E. Cadario. Combining GIS and InSAR data for 3D building reconstruction. In *IEEE International Geoscience and Remote Sensing Symposium*, pages 2418–2421, July 2010.
- [Tis04] C. Tison. *Interférométrie RSO à haute résolution en milieu urbain : application au calcul de MNS urbain*. PhD thesis, 2004. École Nationale Supérieure des Télécommunications, Paris, France.
- [TLB88] R. Touzi, A. Lopes, and P. Bousquet. A statistical and geometrical edge detector for SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 26(6) :764–773, November 1988.
- [TMM+98] F. Tupin, H. Maitre, J.F. Mangin, J.M. Nicolas, and 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.
- [TR04] F. Tupin and M. Roux. 3D information extraction by structural matching of SAR and optical features. *International Archives of Photogrammetry and Remote Sensing*, 35, 2004.
- [TR05] F. Tupin and 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, August 2005.
- [TRW01] TRW Space, Defense & Information Systems, Redondo Beach, Canada. *EO-1/ Hyperion Science Data User's Guide*, 2001.
- [TTM04] C. Tison, F. Tupin, and H. Maitre. Retrieval of building shapes from shadows in high resolution sar interferometric images. In *IEEE International Geoscience and Remote Sensing Symposium*, volume 3, pages 1788–1791, 2004.
- [VD01] C. Vestri and F. Devernay. Using robust methods for automatic extraction of buildings. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 133–138, 2001.
- [WHTS11] J.D. Wegner, R. Hänsch, A. Thiele, and U. Soergel. Building Detection From One Orthophoto and High-Resolution InSAR Data Using Conditional Random Fields. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 4(1) :83–91, March 2011.
- [ZT98] D. Ziou and S. Tabbone. Edge detection techniques-an overview. *International Journal of Pattern Recognition and Image Analysis*, 8(4) :537–559, 1998.