



UNIVERSITE Pierre et Marie Curie – PARIS VI
Ecole Nationale Supérieure des Télécommunications

* * * *

Projet présenté pour l'obtention du diplôme de
Master 2 en Informatique, spécialité : **Imagerie**, parcours : **Traitement
avancé des images**

Par :
Moez BACCOUCHE

Sujet :
**RECALAGE ET FUSION DE DONNEES
OPTIQUES ET RADAR METRIQUES :**
APPLICATION A LA DETECTION DE BATIMENTS

Encadré par :

- | | |
|---------------------|---|
| ➤ Florence TUPIN | -- Professeur à Télécom ParisTech |
| ➤ Hélène SPORTOUCHE | -- Doctorante à Télécom ParisTech,
Thales Communications |

Lieu du stage :



*Département Traitement du Signal et des Images de Télécom ParisTech
Laboratoire de Traitement et de Communication de l'information – UMR 5141*

Année universitaire 2008-2009

Remerciements

Au terme de ce stage, je tiens à exprimer mon immense gratitude à ma responsable de stage, Florence TUPIN, Professeur à Télécom ParisTech, pour la qualité de son encadrement, son expérience dont j'ai eu l'honneur de profiter pendant ces quelques mois, sa disponibilité malgré ses nombreuses responsabilités et pour avoir pris la peine de lire et d'apporter ses corrections à ce rapport.

Je remercie également Hélène SPORTOUCHE, doctorante à Télécom ParisTech / Thales Communications, pour avoir participé à l'encadrement de ce stage et pour avoir su m'aider quand j'en avais besoin.

Je tiens aussi à remercier Jean-Marie NICOLAS, Professeur à Télécom ParisTech, qui, bien qu'il n'apparaisse pas officiellement dans mon encadrement, a fait beaucoup pour ce travail.

Mes remerciements s'adressent également à Nadine POURTHIÉ et Céline TISON du Centre National d'Études Spatiales de Toulouse, qui m'ont accueilli pendant une semaine en mission au sein du CNES et m'ont fait si gentiment profité de leur savoir faire.

Je remercie également tous les membres (chercheurs, doctorants, stagiaires. . .) du département TSI de Télécom ParisTech et tout particulièrement Charles-Alban DELEDALLE, Aymen SHABOU, Gabrielle LEHUREAU, Maria Carolina VANEGAS OROZCO et Dina HADIDI, doctorants au sein du laboratoire pour leur collaboration précieuse et les discussions constructives que nous avons eu pendant la réalisation de ce travail.

Je tiens enfin à remercier l'ensemble des enseignants de la spécialité IMA du Master en informatique de l'Université Pierre et Marie Curie, et plus particulièrement les deux responsables du Master Isabelle BLOCH et Séverine DUBUISSON, ainsi que tous mes collègues étudiants en M2.

Résumé

L'arrivée des nouveaux capteurs optiques et radar, qui permettent d'avoir des images métriques et qui fournissent des informations complémentaires, a relancé, ces dernières années, l'intérêt pour les applications qui permettent l'interprétation automatique des scènes, notamment en milieu urbain.

Ce travail est consacré à la mise au point d'une méthode de détection de bâtiments basée sur l'exploitation conjointe de données optiques et radar, provenant respectivement des capteurs QUICKBIRD et TERRASAR-X.

Pour ce faire, nous avons préalablement recalé les images en les projetant en géométrie radar. Une chaîne de recalage basée sur les fonctions de localisation implémentées sous OTB a été mise au point et améliorée à travers l'exploitation de l'information mutuelle.

Ces données recalées ont ensuite servi à calculer des cartes de score, relatives à des indices, qui traduisent le degré d'appartenance de chaque région à la classe bâtiment. La fusion de ces cartes, combinée à une recherche de rectangles ont permis d'avoir des résultats satisfaisants, mais qui peuvent néanmoins être améliorés en introduisant d'autres indices relatifs à la forme initiale des bâtiments.

MOTS CLÉS : Orfeo ToolBox, recalage et fusion optique / radar, détection de bâtiments.

Table des matières

Table des matières	4
Table des figures	6
1 Introduction générale	10
1.1 Problématique	10
1.2 Objectifs	11
1.3 Organisation du rapport	12
2 État de l’art sur la fusion de données optiques / radar	13
2.1 Introduction	13
2.2 Principes de base de l’imagerie satellitaire optique et radar	13
2.2.1 Imagerie satellitaire optique	13
2.2.2 Imagerie radar à ouverture synthétique (ROS)	14
2.2.2.1 Acquisition des données	15
2.2.2.2 Résolution des images	17
2.2.2.3 Traitement de synthèse	17
2.3 Intérêt de la fusion optique / radar : Etat de l’art	20
2.3.1 Mise en évidence de la complémentarité des images optiques et radar pour l’interprétation des scènes urbaines	20
2.3.2 Fusion optique / radar pour l’étude des bâtiments	21
2.3.2.1 Approche par extraction de primitives sur l’image radar et projection sur l’image optique	21
2.3.2.2 Approche par extraction de primitives sur l’image optique et projection sur l’image radar	21
2.3.2.3 Approche par régions	23
2.3.2.4 Approche par combinaison non hiérarchique d’informations	24
2.3.3 Fusion optique / radar pour l’étude d’autres structures urbaines : exemple des ponts	25
2.4 Conclusion	26
3 Recalage de données optiques / radar métriques	27

3.1	Introduction	27
3.2	Données utilisées et zone d'étude	27
3.2.1	Présentation du capteur optique QUICKBIRD	28
3.2.2	Présentation du capteur radar TERRASAR-X	28
3.2.3	Présentation de la zone d'étude	30
3.3	Modélisation paramétrique de la géométrie de prise de vue	31
3.3.1	Fonction de localisation directe optique	32
3.3.2	Fonction de localisation directe radar	33
3.3.3	Fonctions de localisation inverse	35
3.3.4	Fonctions de localisation implémentées sous OTB	35
3.4	Méthode de recalage proposée	36
3.4.1	Mise en évidence de la notion d'ambiguïté du recalage	36
3.4.2	Schéma synoptique de la méthode de base	37
3.4.3	Amélioration des résultats en exploitant l'information mutuelle	42
3.5	Conclusion	45
4	Détection de bâtiments par fusion de données optiques / radar	47
4.1	Introduction	47
4.2	Méthode proposée	47
4.2.1	Extraction des cartes d'indices	48
4.2.2	Génération des cartes de score	49
4.2.2.1	Cartes de score des ombres optiques et radar	51
4.2.2.2	Carte de score pour les zones brillantes	53
4.2.2.3	Carte de score pour le NDVI	55
4.2.3	Fusion des cartes de score	55
4.2.4	Algorithme de recherche des rectangles	57
4.3	Analyse des résultats	60
4.4	Conclusion	61
5	Conclusion générale	63
A	Données disponibles sur la zone de Toulouse	65
	Bibliographie	68

Table des figures

2.1	Exemples de capteurs couramment utilisés : (a) - Le capteur en fenêtre "Vexcel ULTRACAM" (©Vexcel). (b) - Le capteur ligne par ligne "Leica ADS40 2nd Generation" (©Leica).	14
2.2	Principe de l'acquisition ligne par ligne des détecteurs <i>Push Broom</i> (source [Tebourbi 05]).	15
2.3	Illustration du contenu d'une image panchromatique : approximation (codée sur 8 bits) de la réflectivité des objets au sol (source [Tebourbi 05]).	15
2.4	Géométrie d'acquisition d'un radar à visée latérale (source [Charleux 01]). . .	16
2.5	Compression d'impulsions (source [Charleux 01]).	18
2.6	Antenne synthétique équivalente (source [Charleux 01]).	18
2.7	Cellule de résolution et son pixel correspondant (source [Charleux 01]). . . .	19
2.8	Comparaison entre l'aspect des bâtiments entre l'image optique et radar : (a) - Image radar RAMSES de résolution 50cm, la direction de visée du capteur étant à gauche de l'image. (b) - Image optique de l'IGN de résolution 50cm (source [Tupin 03]).	20
2.9	Principaux résultats présentés dans [Tupin 03] : (a) - Détection des lignes brillantes (en noir) sur l'image radar (b) - En noir : bâtiment détectés par la recherche de rectangles et en blanc : recherche de formes plus complexes. . .	22
2.10	Principaux résultats présentés dans [Sportouche 09a] et [Sportouche 09b] : (a) - Bâtiments détectés sur l'image optique (©DigitalGlobe). (b) - Projection des contours des bâtiments détectés sur l'image radar (©DLR/EADS Astrium). . .	22
2.11	Principaux résultats présentés dans [Wegner 09a] et [Wegner 09b] : (a) - Détection de lignes sur l'image radar superposée à un MNE Laser. (b) - Superposition des lignes détectées sur l'image optique. (c) - Résultat de la méthode décrite dans [Muller 05] appliquée à l'image optique. (d) - Résultat de la fusion. . . .	23
2.12	Primitives extraites sur les images radar et optiques dans [Poulain 09] : (a) - Contours de l'image optique [Canny 86]. (b) - Détection de lignes brillantes sur l'image radar [Tupin 98]. (c) - Masque contenant les ombres de l'image optique. (d) - Masque de végétation (NDVI) calculé à partir de l'image optique multispectrale [Tucker 79]. (e) - Détection de lignes sur l'image optique [Grompone 08]. . .	24

2.13	Bâtiments détectés par la méthode proposée dans [Poulain 09].	25
2.14	Principaux résultats présentés dans [Soergel 07a] et [Soergel 07b] : (a) - Image d'amplitude complexe radar sur laquelle on observe le phénomène des "trois bandes" au niveau des ponts. (b) - Interférogramme correspondant sur lequel on observe le même phénomène. (c) - Image optique recalée en géométrie radar. (d) - Résultat de la fusion : Modèle numérique d'élévation.	26
3.1	QUICKBIRD (source : [DigitalGlobe 09]).	28
3.2	TERRASAR-X (source : [DLR 04]).	28
3.3	Modes d'acquisition de TerraSar-X et leur principe de fonctionnement : (a) - Le mode SpotLight. (b) - Le mode StripMap. (c) : Le mode ScanSar (source [DLR 04]).	29
3.4	Données disponibles sur la zone de Toulouse.	30
3.5	Zone d'étude : Complexe scientifique de Rangueil, Toulouse : (a) - Image TERRASAR-X (©DLR/EADS Astrium). (b) - Image QUICKBIRD panchromatique (©DigitalGlobe). (c) - Image QUICKBIRD multispectrale (©DigitalGlobe).	31
3.6	Intersection de la sphère-distance, du cône-Doppler et de la Terre (source [Polidori 97]).	34
3.7	Mise en évidence de l'ambiguïté du recalage à travers la projection d'un bâtiment : (a) - Cas de l'image optique. (b) - Cas de l'image radar (source : [Lehureau 08]).	36
3.8	Résultat du recalage en utilisant une même valeur d'altitude pour tous les points : (a) - Image radar. (b) - Image optique projetée en géométrie radar. Des points de même couleur ont été rajoutés manuellement pour donner une idée sur le décalage.	38
3.9	Résultat du recalage en utilisant le MNT lors de la projection : (a) - Image radar. (b) - Image optique projetée en géométrie radar.	39
3.10	Estimation des décalages pour chaque image à partir des quatre coins de référence.	39
3.11	Schéma synoptique de la méthode de recalage adoptée.	40
3.12	Résultat du recalage en utilisant le MNT, en estimant les décalages et en les réinjectant lors de la projection : (a) - Image radar. (b) - Image optique projetée en géométrie radar. (c) - Superposition des contours du résultat sur l'image radar.	41
3.13	Résultat du filtrage de l'image radar en utilisant l'algorithme décrit dans [Deledalle 09] : (a) - Image originale. (b) - Image filtrée.	44

3.14	Résultat de la recherche de la translation qui maximise l'information mutuelle : (a) - Représentation 3D des valeurs de l'information mutuelle, maximum atteint pour $t_x = -19$ et $t_y = -10$. (b) - Résultat correspondant : superposition des contours de l'image optique sur l'image radar.	44
3.15	Résultat de la recherche de la similitude qui maximise l'information mutuelle : maximum atteint en appliquant une translation $t_x = -15$ et $t_y = -22$, une rotation d'angle $\theta = 0.2354^\circ$ et une homothétie de facteur $\alpha = 1.0039$	45
4.1	Données recalées utilisées comme entrées pour la méthode de détection de bâtiments : (a) - Image radar filtrée. (b) - Image optique panchromatique. (c) - Bande "rouge" de l'image optique multispectrale. (d) - Bande "proche infra-rouge" de l'image optique multispectrale.	48
4.2	Schéma synoptique de la méthode de détection de bâtiments proposée.	49
4.3	Cartes d'indices extraites à partir des images optiques et radar : (a) - Masque binaire des ombres sur l'image optique. (b) - Masque binaire des ombres sur l'image radar. (c) - Masque binaire des zones brillantes sur l'image radar. (d) - Carte d'indice de végétation NDVI.	50
4.4	Segmentation de l'image optique panchromatique par l'algorithme <i>Mean Shift</i>	50
4.5	Illustration de la génération des cartes de distance orientée pour un objet donné : (a) - Carte de distance morphologique "classique". (b) - Masque de la zone "à 120° de l'objet". (c) - Carte de distance orientée dans cette direction.	51
4.6	Cartes de distance orientée obtenues : (a) - Pour l'ombre optique, direction utilisée : $\alpha = 115^\circ$. (b) - Pour l'ombre radar, direction utilisée : $\alpha = 0^\circ$	52
4.7	Schéma synoptique de la procédure de calcul des scores correspondant aux ombres optiques et radar.	52
4.8	Mise en évidence de l'intérêt de l'utilisation d'une distance en $\exp\left(\frac{distance}{a}\right)$: (a) - Carte de distance orientée obtenue avec une distance linéaire. (b) - Carte de distance orientée obtenue avec une distance en $\exp\left(\frac{distance}{a}\right)$ avec $a = 5$	53
4.9	Carte de score correspondant à : (a) - Ombres optiques. (b) - Ombres radar.	53
4.10	Schéma synoptique de la procédure de calcul des scores correspondant aux zones brillantes.	54
4.11	Carte de score correspondant aux zones brillantes.	54
4.12	Carte de score correspondant à l'indice de végétation.	55
4.13	Chaîne de fusion adoptée.	56
4.14	Résultat de la fusion des cartes de score : (a) - Carte de score finale. (b) - Masque binaire des régions sélectionnées comme étant des bâtiments.	57

4.15	Illustration du principe de l'algorithme de détection de rectangles : (a) - Exemple d'un objet O . (b) - Rectangle englobant O . (c) - Rectangle englobant l'image de O par une rotation (l'angle de la rotation étant $-\theta$, où θ est l'orientation principale de O). (d) - Résultat de l'application de la rotation inverse.	58
4.16	Résultat de l'algorithme de détection de rectangles appliqué sur le masque binaire de la figure 4.14-(b).	59
4.17	Régions du masque binaire de la figure 4.14-(b) qui ont pu être appariées avec des rectangles : Superposition sur l'image optique.	59
4.18	Illustration du critère de décision sur l'appartenance d'un point à la zone "intérieure du rectangle".	60
4.19	Résultat final de la détection de bâtiments : (a) - Carte des bâtiments. (b) - Superposition des contours de la carte des bâtiments sur l'image optique. . . .	61
4.20	Application de la méthode de détection de bâtiments sur une autre zone : (a) - Carte de score finale. (b) - Résultat de l'algorithme de recherche de rectangles. (c) - Carte des bâtiments obtenue. (d) - Superposition de la carte des bâtiments sur l'image optique.	62
5.1	Interférogramme TerraSar-X correspondant au couple 19/12/2007 et 28/08/2008 (un tour de phase correspond à 24.98m) : (a) - Interférogramme complet. (b) - Zoom sur la zone d'étude.	64
A.1	Modèle Numérique de Terrain sur la zone de Toulouse : (a) - Visualisation en niveaux de gris. (b) - Visualisation 3D.	65
A.2	Données QUICKBIRD du 01/04/2002 sur la zone de Toulouse : (a) - Image Panchromatique. (b) - Image multispectrale (©DigitalGlobe).	66
A.3	Données <i>TerraSar-X</i> sur la zone de Toulouse : (a) - Image du 27/11/2007. (b) - Image du 19/12/2007 (©DLR/EADS Astrium).	67

Chapitre 1

Introduction générale

1.1 Problématique

L'une des problématiques actuelles de l'interprétation des images issues des archives satellitaires concerne la multiplication des satellites et des capteurs. La nouvelle génération de capteurs satellitaires radar et optique qui ont des résolutions de l'ordre du mètre soulève de nouveaux défis. Cette amélioration de la résolution nécessite la définition d'outils qui lui soient adaptés. Les différences de contenu pour des résolutions de 10m, 1m et 30cm sont drastiques non seulement pour les données optiques mais plus encore pour les données radar.

Par ailleurs, un des enjeux majeurs de la télédétection est actuellement l'interprétation automatique des scènes urbaines. Avec l'arrivée de ces nouveaux capteurs qui fournissent des images à des résolutions métriques, de nombreuses applications dans ce domaine, tel que la reconstruction 3D en milieu urbain ou encore la détection automatique d'urbanisation, connaissent un nouvel âge d'or. L'un des enjeux actuels est donc l'exploitation conjointe des différentes données disponibles, optiques et radar, pour mettre au point des méthodes automatiques d'interprétation. Pour cela, il est primordial de bien étudier les spécificités physiques de ces deux modalités.

Les systèmes d'imagerie optiques, passifs, mesurent directement le rayonnement réfléchi par la surface terrestre dans le domaine du visible ou de l'infra rouge. Bien que l'utilisation de tels imageurs soit très intéressante car les données fournies sont très proches de la vision qu'aurait un utilisateur, ils ont cependant quelques inconvénients. Le premier concerne leur dépendance vis à vis de la source lumineuse, généralement le soleil, il est alors impossible de faire des acquisitions de nuit. De plus, les longueurs d'ondes utilisées (400 à 800nm) sont très sensibles à la teneur en eau de l'atmosphère, il est donc impossible d'obtenir une image du sol en cas de couverture nuageuse ou de brume.

L'utilisation de systèmes radar (actifs) permet de combler les lacunes de l'imagerie optique. De tels capteurs utilisent le domaine des micro-ondes (1 à 10Ghz), peu sensibles

à l'humidité de l'atmosphère. La qualité des images ne dépend donc pas des conditions météorologiques. Ces systèmes sont actifs, c'est à dire qu'ils disposent de leur propre source d'illumination et sont donc indépendants de l'illumination solaire, ils peuvent ainsi acquérir des images de jour comme de nuit.

Ces différents modes d'observations sont complémentaires et riches en information. Désormais les images satellitaires couvrent la totalité de la planète et un satellite met environ sept heures pour en faire le tour. Les images radar peuvent nous donner des informations malgré les nuages de fumées ou de cendres et les données optiques peuvent nous permettre une interprétation facile.

Néanmoins, cette diversité dans les données est également un facteur de confusion et d'erreur car les différences entre les images, qui font à la fois leur complémentarité, rendent aussi leur exploitation plus difficile. L'interprétation des images fournies par ces capteurs peut être effectuée manuellement, mais devant l'augmentation croissante de la quantité de données à disposition, l'automatisation de l'interprétation est d'un intérêt majeur. Il est donc nécessaire de développer de nouveaux outils automatiques de traitements, adaptés à l'interprétation de ce type d'images. Vu les nombreuses différences au niveau de l'acquisition, une étape cruciale de toute méthode qui exploiterait ces deux données est le recalage des images. Bien que de nombreux travaux aient été proposés auparavant, l'arrivée de la haute résolution induit la nécessité d'avoir de nouvelles méthodes de recalage plus précises.

De plus, les différentes applications liées à l'interprétation automatique des scènes urbaines nécessitent une étape de classification des différents éléments composant ce type de scènes. La détection de bâtiments en particulier est alors une étape clé dont dépend la précision des traitements ultérieurs.

1.2 Objectifs

L'objectif de ce stage est de mettre en place une chaîne de traitement opérationnelle pour la détection de bâtiments à partir d'images optiques QUICKBIRD et radar TERRASAR-X.

On s'intéressera tout d'abord au recalage précis entre les deux types de données. Le département TSI a mené des travaux dans un cadre automatique et en l'absence de connaissances capteurs pour les deux données. Les résultats obtenus sont satisfaisants.

Néanmoins, cette situation est peu réaliste car les images sont délivrées avec des paramètres capteurs qui permettent d'exploiter les équations géométriques régissant les différents systèmes d'acquisition.

La première étape de ce stage sera donc consacrée à la mise en correspondance des données et à l'évaluation de cette mise en correspondance. Cette partie du stage s'appuiera sur l'OTB (Orfeo ToolBox), bibliothèque libre développée par le CNES dans lequel

plusieurs algorithmes de projection ont déjà été intégrés. On étudiera, le cas échéant, la possibilité d'amélioration des résultats de ces algorithmes pour remédier aux erreurs de recalage.

La détection de bâtiments se fera ensuite par fusion des données recalées. Il faudra tenir compte, lors du choix de la règle de fusion, des spécificités qui caractérisent les bâtiments sur chaque type d'images.

1.3 Organisation du rapport

Le présent document s'organise en trois chapitres principaux. Dans un premier temps, un bref état de l'art sur la fusion optique / radar sera effectué pour bien situer ce stage par rapport aux autres travaux. Les principes de base de l'imagerie optique et radar seront aussi présentés afin de mettre en évidence les différences au niveau de l'acquisition.

Le second chapitre présentera la chaîne de recalage proposée ainsi que ses différentes améliorations. Le recalage étant basé sur les paramètres capteur, une brève description de la modélisation paramétrique de la géométrie de prise de vue sera exposée afin d'introduire la notion de « localisation ».

Enfin, dans le dernier chapitre, nous allons présenter la méthode de détection de bâtiments qui exploite conjointement ces données recalées pour générer des cartes de bâtiments sur la zone étudiée. Une approche de fusion de score par régions calculés sur des indices bien définis sera exposée et les résultats seront analysés.

Chapitre 2

État de l’art sur la fusion de données optiques / radar

2.1 Introduction

Dans ce chapitre, nous allons tout d’abord introduire des notions de base en imagerie satellitaire optique et radar afin de bien comprendre les spécificités liées à chaque mode d’imagerie. L’accent sera mis sur les caractéristiques liées à l’interprétation des scènes urbaines en général et aux bâtiments en particulier. Un bref état de l’art, qui servira à mettre en évidence la complémentarité de ces deux sources d’informations pour différentes applications, sera ensuite effectué.

2.2 Principes de base de l’imagerie satellitaire optique et radar

2.2.1 Imagerie satellitaire optique

Les premières images de télédétection optiques remontent à la première guerre mondiale. Les pilotes survolaient les zones ennemies et faisaient des acquisitions en noir et blanc avec des appareils photo analogiques.

De nos jours, les capteurs optiques modernes offrent des acquisitions dans des bandes spectrales très variées. Une large gamme de capteurs aéroportés et satellitaires existe, chacune particulièrement adaptée à un type d’applications. A titre d’exemple d’applications, nous citons l’acquisition de données pour les systèmes d’informations géographiques (GIS) ou bien encore la reconstruction 3D des zones urbaines. Dans le cadre de ce stage, nous nous intéressons uniquement aux capteurs satellitaires numériques. Ce type de capteurs est particulièrement adapté à l’étude des formes, des détails et des contours des objets se trouvant à la surface de la terre.

En général, les bandes spectrales de ces capteurs satellitaires couvrent le domaine visible et le proche infra-rouge (voire l'infra-rouge). Le canal dit “panchromatique”, qui correspond à la bande spectrale autour du visible, permet d’avoir la meilleure résolution. Il est généralement associé à d’autres canaux qui correspondent à des largeurs de bande plus faibles et à des résolutions moindres.

Deux principaux modes d’acquisition existent, et se distinguent par la disposition des cellules CCD sur le capteur. Pour le premier mode, les cellules CCD sont disposées sur des fenêtres rectangulaires et permettent ainsi de faire directement des acquisitions 2D de la scène (cf. figure 2.1-(a)). Ce mode a néanmoins l’inconvénient d’être difficile à mettre en oeuvre sur le plan technique et surtout limite la résolution.



FIGURE 2.1: Exemples de capteurs couramment utilisés : (a) - Le capteur en fenêtre "Vexcel ULTRACAM" (©Vexcel). (b) - Le capteur ligne par ligne "Leica ADS40 2nd Generation" (©Leica).

Le second mode d’acquisition, dit “en peigne” (ou *Push Broom* en anglais), correspond à une disposition linéaire (dite “en barrettes”) des cellules CCD (cf. figure 2.1-(b)). L’acquisition des images se fait alors ligne par ligne par une translation du capteur. La résolution dépend directement de la vitesse du capteur et de la résolution des cellules CCD. La figure 2.2 illustre le principe d’acquisition du mode *Push Broom*.

Pour le cas des images panchromatiques, l’image peut être vue comme une matrice représentant les niveaux de gris associés à la scène (cf. figure 2.3). Chaque valeur de niveau de gris représente une approximation de la réflectivité de l’objet imagé au sol dans la bande spectrale du capteur.

Une image multispectrale, quant à elle, peut être vue comme étant une superposition de plusieurs réponses à des bandes différentes. Chacune des sous bandes peut avoir une taille et une résolution différente selon les cas.

2.2.2 Imagerie radar à ouverture synthétique (ROS)

Cette partie essaie de dresser une brève description des principes généraux de l’imagerie radar à ouverture synthétique (ROS) et a été largement inspirée de l’ouvrage de H. MAÎTRE [Maitre 08].

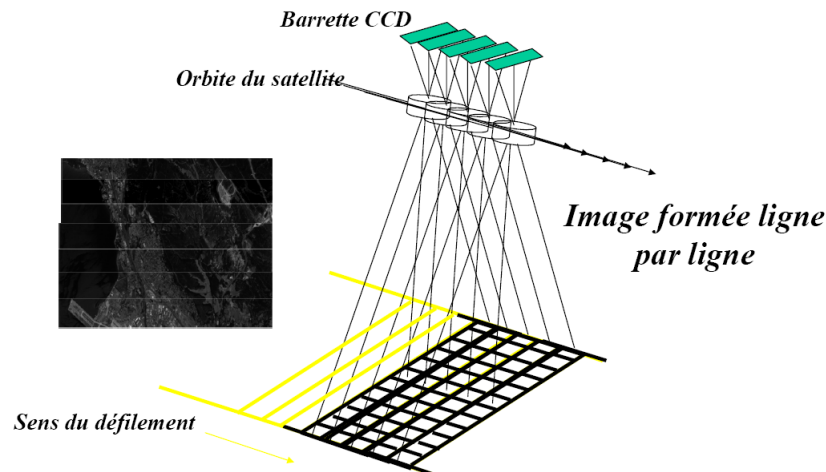


FIGURE 2.2: Principe de l'acquisition ligne par ligne des détecteurs *Push Broom* (source [Tebourbi 05]).

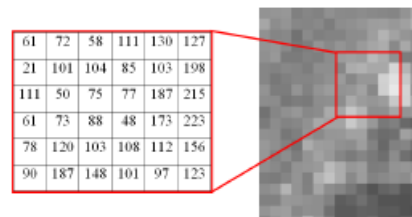


FIGURE 2.3: Illustration du contenu d'une image panchromatique : approximation (codée sur 8 bits) de la réflectivité des objets au sol (source [Tebourbi 05]).

2.2.2.1 Acquisition des données

Les radars utilisés pour l'imagerie satellitaire sont en général des radars à synthèse d'ouverture (ROS ou SAR en anglais *Synthetic Aperture Radar*). Ce type de capteurs à visée latérale illumine une portion de l'espace avec une onde électromagnétique puis reçoit les réponses réfléchies par les objets présents dans cette zone, ce qui permet de détecter leur existence et de déterminer certaines de leurs caractéristiques. En particulier, on exploite la liaison de la mesure radar au temps d'aller-retour de l'onde afin de déduire sa proportionnalité par rapport à la distance radar-cible.

La figure 2.4 montre la géométrie de prise de vue d'un radar imageur. Installé à bord d'un satellite, il envoie latéralement un train d'ondes hyperfréquences. Les échos renvoyés sont enregistrés et l'on obtient ainsi une ligne d'image suivant l'axe des distances. Le satellite se déplace et les trains d'ondes successifs produisent une image du sol par balayage. La longueur d'une image (axe des azimuts) dépend de la durée d'acquisition

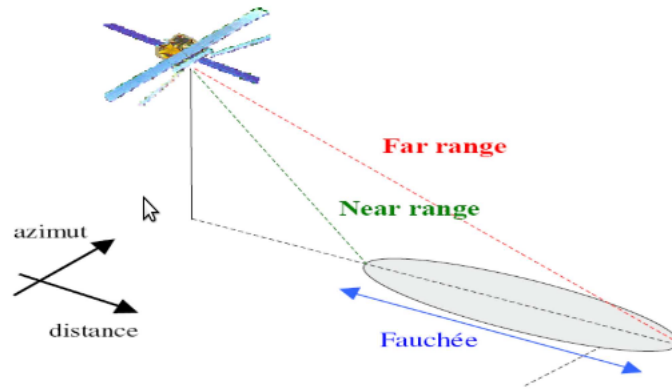


FIGURE 2.4: Géométrie d'acquisition d'un radar à visée latérale (source [Charleux 01]).

tandis que la largeur (axe des distances) est caractérisée par la fauchée du radar. On désigne par *near range* (portée proximale) la distance entre le porteur et le point de l'image le plus proche et par *far range* (portée distale) la distance entre le porteur et le point de l'image le plus éloigné.

En azimuth, le pas inter-pixels est lié à la PRF (*Pulse Repetition Frequency*, soit la fréquence de répétition des ondes émises) et à la vitesse de progression du lobe principal d'éclairement dans la direction azimuthale (la vitesse du porteur). Cette relation se traduit par l'équation :

$$\Delta_{azimuth} = \frac{\nu}{PRF} \quad (2.1)$$

où ν est la norme de la vitesse relative du porteur par rapport au sol dans la direction azimuthale.

En distance, le pas inter-pixels en géométrie radar dépend de la fréquence d'échantillonnage F_e et de la célérité C et s'écrit :

$$\Delta_{distance} = \frac{C}{2 \cdot F_e} \quad (2.2)$$

La distance séparant le radar de la cible est parcourue deux fois (aller/retour), c'est pourquoi on retrouve un facteur deux au dénominateur dans l'équation 2.2. En outre, pour obtenir la taille des pixels projetés au sol, il faut prendre en compte l'angle d'incidence local i de l'onde. Ainsi, le pas d'un pixel en distance varie avec l'angle d'incidence local et devient en géométrie sol :

$$\Delta_{distance}^{sol} = \frac{C}{2 \cdot F_e \cdot \sin(i)} \quad (2.3)$$

2.2.2.2 Résolution des images

Contrairement aux dimensions de l'image, la résolution est complètement indépendante des fréquences d'acquisition (PRF et F_e).

En distance, la résolution est contrainte par la durée de l'impulsion radar τ émise. Pour distinguer les échos de deux cibles ponctuelles, le délai séparant les deux cibles doit être supérieur ou égal à τ , d'où l'équation :

$$\delta_{distance} \simeq C \cdot \frac{\tau}{2} \quad (2.4)$$

Plus τ sera petit et plus la résolution en distance sera fine. Comme il est techniquement difficile de générer des trains d'ondes de durée inférieure à la micro seconde avec une puissance suffisante, la méthode de compression d'impulsions est utilisée (cf paragraphe suivant).

En azimut, la résolution est liée à l'ouverture Δ_θ du faisceau radar et s'écrit :

$$\delta_{azimut} = 2 \cdot D \cdot \Delta_\theta \simeq 2 \cdot D \cdot \frac{\lambda}{L} \quad (2.5)$$

où D est la distance radar-cible et L la longueur de l'antenne.

Ainsi, la résolution en azimut sera supérieure à la dizaine de kilomètres pour les plates-formes satellitaires, et de l'ordre de la centaine de mètres pour les systèmes aéroportés. Pour approcher la résolution des systèmes optiques, il faudrait un angle d'ouverture plus étroit et donc une antenne de plusieurs kilomètres de long.

Afin d'améliorer cette résolution, la technique des radars à synthèse d'ouverture a été mise au point. Cette technique est développée dans le paragraphe suivant.

2.2.2.3 Traitement de synthèse

La partie précédente a mis en évidence les faiblesses du radar en matière de résolution. Deux facteurs limitant ont été mis en exergue : d'une part la technicité requise pour produire une impulsion de courte durée et d'autre part la longueur de l'antenne.

Compression en distance Nous avons vu qu'il fallait produire une impulsion de courte durée et de forte amplitude pour améliorer la résolution en distance et avoir des échos d'amplitude significative. Les études, dont le but était de réduire ce laps de temps, ont abouti au principe de la compression d'impulsions : une impulsion est modulée linéairement en fréquence puis elle est corrélée au signal retourné (l'écho). Des sinus cardinaux se forment, traduisant la position des cibles ponctuelles (cf. figure 2.5).

La résolution en distance devient alors :

$$\delta_{distance}^{comp} \simeq \frac{C}{2 \cdot B} \quad (2.6)$$

où B est la largeur de bande du spectre de l'impulsion émise (et de l'écho reçu).

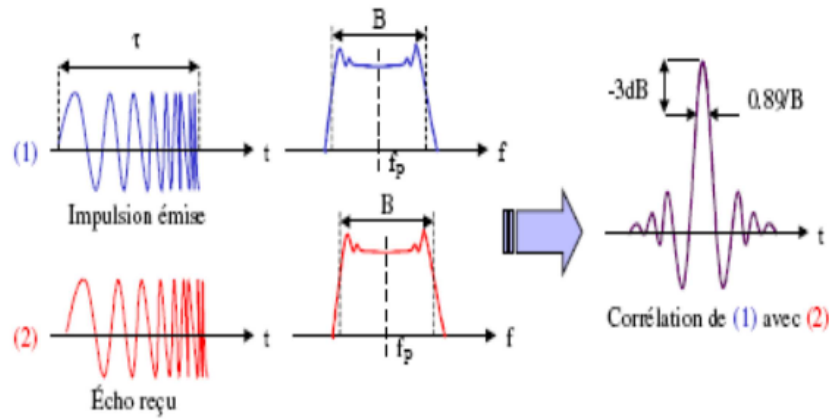


FIGURE 2.5: Compression d'impulsions (source [Charleux 01]).

Compression en azimuth (radar à synthèse d'ouverture) Les radar à ouverture réelle (ROR), émettent et reçoivent des signaux par l'intermédiaire d'une antenne de longueur fixe. Comme il est difficile de transporter une antenne très longue, la résolution des ROR spatiaux est insuffisante pour des applications de télédétection. Afin de s'affranchir de ce problème, les radar à synthèse d'ouverture ont été mis au point. Leur antenne est plus courte mais elle synthétise une longue antenne (cf. figure 2.6).

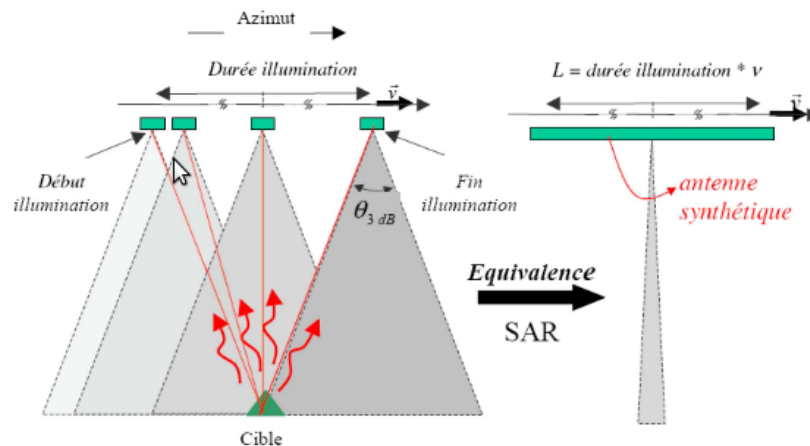


FIGURE 2.6: Antenne synthétique équivalente (source [Charleux 01]).

Au cours de son déplacement, le radar émet plusieurs signaux en direction d'une même cible. L'information concernant la cible se retrouve répartie sur plusieurs lignes à l'origine d'un étalement en azimuth. Le principe de la compression en azimuth consiste donc à isoler l'écho produit par une cible en azimuth au cours du parcours.

Les distances cible-antenne varient au cours de la durée d'illumination. L'écho renvoyé par une même cible a par conséquent une fréquence différente à chaque train d'ondes (fréquence Doppler). C'est cette variation en fréquence du signal qui permet d'extraire l'information étalée en azimuth. Il suffit d'appliquer un traitement similaire à la compression

en distance. Dans ce cas, la largeur de bande du signal produit est celle de l'effet Doppler, elle est donnée par l'équation :

$$B_{Doppler} = \frac{2 \cdot \|\vec{v}\|^2 \cdot \tau}{\lambda \cdot d_0} = \frac{2 \cdot \|\vec{v}\|}{L} \quad (2.7)$$

où L est la longueur de l'antenne, $\vec{v}$ le vecteur vitesse du porteur, τ la durée du train d'ondes, λ la longueur d'onde et d_0 la distance proximale.

Après la compression en azimuth, on obtient une largeur d'impulsion égale à : $\tau' = \frac{1}{B_{Doppler}}$

D'où une nouvelle résolution en azimuth donnée par l'équation :

$$\delta'_{azimut} = \tau' \cdot \|\vec{v}\| = \frac{\|\vec{v}\|}{B_{Doppler}} = \frac{L}{2} \quad (2.8)$$

Pour plus de détails concernant l'imagerie radar à ouverture synthétique, se référer à l'ouvrage de H. MAÎTRE [Maitre 08].

L'image SAR ainsi obtenue est constituée de pixels qui représentent une surface au sol dont les dimensions sont définies dans les directions radiale et azimuthale (cf figure 2.7). Le niveau de gris en un pixel est calculé à partir de l'onde rétrodiffusée par une cellule de résolution .

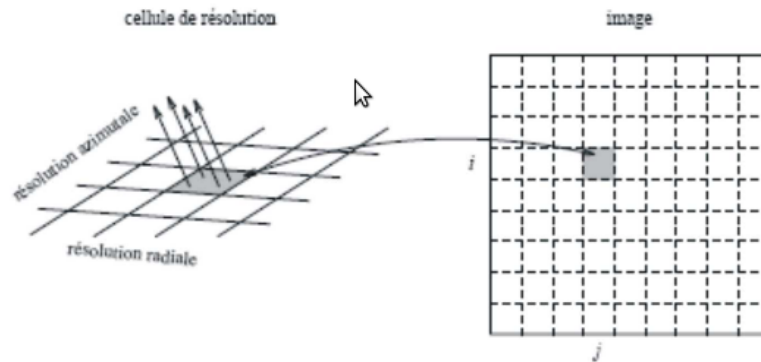


FIGURE 2.7: Cellule de résolution et son pixel correspondant (source [Charleux 01]).

Si on appelle ϵ le champ électro-diffusé par une cellule de résolution (ou bien un pixel), on peut écrire l'équation :

$$\epsilon = A \cdot \exp(j \cdot \varphi) \quad (2.9)$$

où A est l'amplitude du signal et φ sa phase. On peut ainsi en général exploiter deux informations (deux images) fournies par une image radar complexe : l'image d'amplitude et l'image de phase. Dans toute la suite, par abus de langage et dans un cadre qui n'est pas interférométrique, on désignera par "image radar" l'image correspondant à l'amplitude du signal complexe.

2.3 Intérêt de la fusion optique / radar : Etat de l'art

Après avoir présenté les principes de base de l'imagerie optique et radar, nous allons nous intéresser dans cette section à la complémentarité de ces deux sources d'informations. Cette complémentarité sera mise en évidence à travers un bref état de l'art des principaux travaux dans ce domaine.

En effet, plusieurs auteurs ont essayé d'exploiter cette complémentarité pour des applications diverses. Mais depuis quelques années, l'arrivée de la nouvelle génération de capteurs a relancé l'intérêt de la communauté scientifique pour l'étude des scènes urbaines. Faute de pouvoir dresser une liste exhaustive de tous les travaux, nous allons nous intéresser de près à ceux qui se rapprochent beaucoup du sujet de ce stage, et allons les classer en deux sous-sections, selon la structure urbaine étudiée.

2.3.1 Mise en évidence de la complémentarité des images optiques et radar pour l'interprétation des scènes urbaines

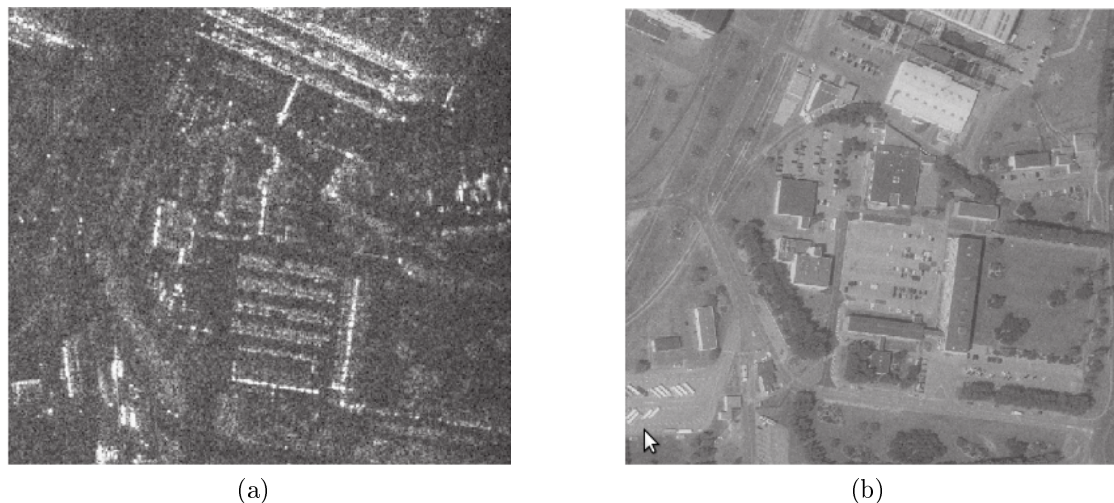


FIGURE 2.8: Comparaison entre l'aspect des bâtiments entre l'image optique et radar : (a) - Image radar RAMSES de résolution 50cm, la direction de visée du capteur étant à gauche de l'image. (b) - Image optique de l'IGN de résolution 50cm (source [Tupin 03]).

La figure 2.8 montre un exemple d'images optique et radar correspondant à la même zone urbaine. L'image radar est très difficile à interpréter sans connaissances a priori, et ce, même pour un photo-interprète humain. En revanche, on observe des lignes brillantes le long de l'interface sol-bâtiment. Ceci est dû aux échos doubles sur les façades orientées dans la direction de visée du capteur [Simonetto 00]. Un autre phénomène très caractéristique des images radar est le repliement (ou *layover*). Ce phénomène se produit surtout pour les objets en hauteur (en l'occurrence les bâtiments dans notre cas) par rapport au sol. Quand l'onde radar arrive sur le sommet de l'objet avant la base, ce sommet

est déplacé vers la direction du capteur, produisant ainsi un aspect visuel sur l'image très caractéristique des objets en hauteur. De plus, quelques toits sont très visibles sur l'image radar (selon le matériau utilisé pour la construction et la taille du toit comparé à la longueur d'onde du capteur), apportant ainsi une information supplémentaire sur la localisation des bâtiments.

L'image optique, quant à elle, permet d'identifier plus facilement les bâtiments car elle est en général très riche en informations de type forme, contours... En revanche, ce type de données est très dépendant des conditions d'acquisition (climat : nuages masquant un bâtiment... , position du soleil : ombres pouvant masquer des éléments de la scène, impossibilité d'avoir des images la nuit...) et ne permet généralement pas de distinguer entre les différents types de bâtiments (toits plats, toits présentant des formes irrégulières...).

Cette complémentarité a été exploitée dans différents travaux. La haute résolution a permis d'approfondir l'étude de deux structures urbaines en particulier, à savoir les bâtiments et les ponts.

2.3.2 Fusion optique / radar pour l'étude des bâtiments

2.3.2.1 Approche par extraction de primitives sur l'image radar et projection sur l'image optique

Dans [Tupin 03], F. TUPIN et M. ROUX proposent une approche de détection de bâtiments en exploitant les informations extraites de deux images de la même zone, l'une optique et l'autre radar (cf. figure 2.8).

L'idée proposée par les auteurs a été d'utiliser l'image radar pour localiser les bâtiments, en se basant sur le principe que ces derniers sont généralement localisés à proximité d'une ligne brillante. La détection de ces lignes a été effectuée par l'algorithme décrit dans [Tupin 98] (cf. figure 2.9-(a)).

Ces lignes sont ensuite exploitées pour définir une zone de recherche, dans laquelle on essaie de trouver des formes rectangulaires (ou des formes plus complexes) en faisant un tri sur des candidats provenant de la vectorisation des contours de CANNY-DERICHE [Canny 86] extraits à partir de l'image optique. Le résultat de la détection est illustré sur la figure 2.9-(b).

2.3.2.2 Approche par extraction de primitives sur l'image optique et projection sur l'image radar

Dans [Sportouche 09a] et [Sportouche 09b], SPORTOUCHE et al. ont proposé une nouvelle approche d'extraction et de reconstruction 3D des bâtiments en utilisant des images optiques et radar à haute résolution. La particularité de ces travaux est qu'ils n'utilisent pour estimer la hauteur des bâtiments, qu'une seule image optique et une autre radar

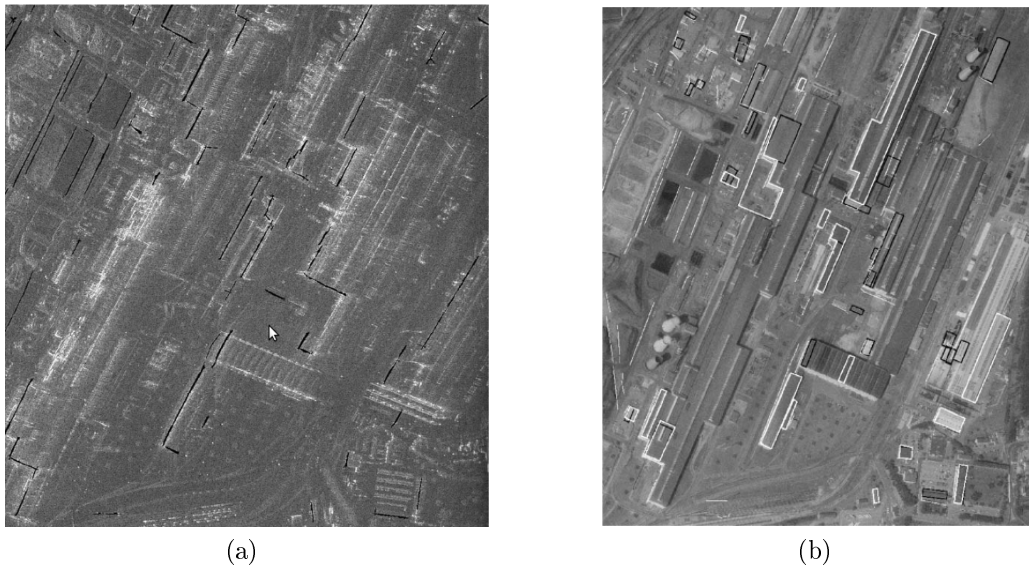


FIGURE 2.9: Principaux résultats présentés dans [Tupin 03] : (a) - Détection des lignes brillantes (en noir) sur l'image radar (b) - En noir : bâtiment détectés par la recherche de rectangles et en blanc : recherche de formes plus complexes.

(par opposition aux autres travaux qui utilisent soit un couple interférométrique d'images radar, soit un couple stéréoscopique d'images optiques).

La première étape de l'algorithme est de localiser les bâtiments sur l'image optique en ayant recours à des outils morphologiques (le profil morphologique différentiel ou *DMP* [Pesaresi 01]) et géométriques (en essayant de faire correspondre les résultats de la détection avec des rectangles). Cette signature (appelée *Building Footprint*) extraite sur l'image optique (cf. figure 2.10-(a)) est ensuite labellisée et projetée en géométrie radar (cf. figure 2.10-(b)).

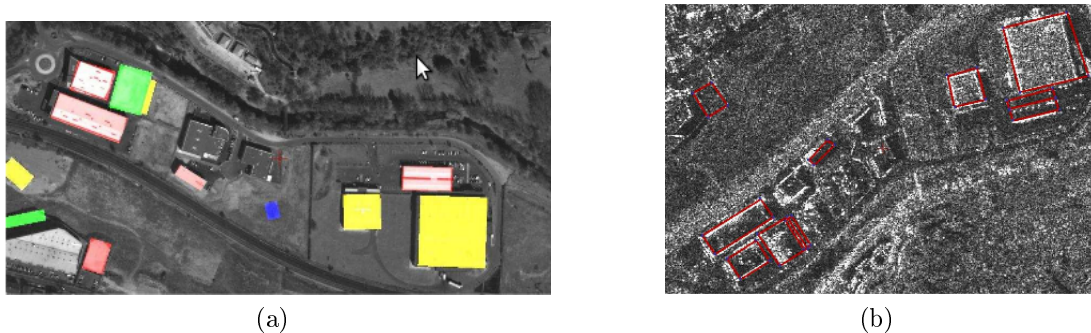


FIGURE 2.10: Principaux résultats présentés dans [Sportouche 09a] et [Sportouche 09b] : (a) - Bâtiments détectés sur l'image optique (©DigitalGlobe). (b) - Projection des contours des bâtiments détectés sur l'image radar (©DLR/EADS Astrium).

Puisque la projection dépend de l'altitude h du point (cf. section 3.3 du chapitre 3), et que la variation de h se traduit comme une translation des objets projetés, plusieurs valeurs de hauteur sont testées pour chaque bâtiment et la valeur retenue est celle qui

maximise deux critères correspondant à deux modèles de réponses des bâtiments. Ceci permet donc d'estimer la hauteur des bâtiments en exploitant les informations fournies par les deux capteurs.

2.3.2.3 Approche par régions

Comme exemple d'autres travaux récents, on peut citer ceux de WEGNER et al., qui présentent dans [Wegner 09a] et [Wegner 09b] une méthode de détection de bâtiments basée sur la fusion optique / radar mais avec une approche "par régions". Comme dans [Tupin 03], la présence d'un bâtiment est caractérisée par celle d'une ou de plusieurs lignes brillantes (les auteurs définissent pour cela la notion de "*corner line*") sur l'image radar (cf. figure 2.11-(a)).

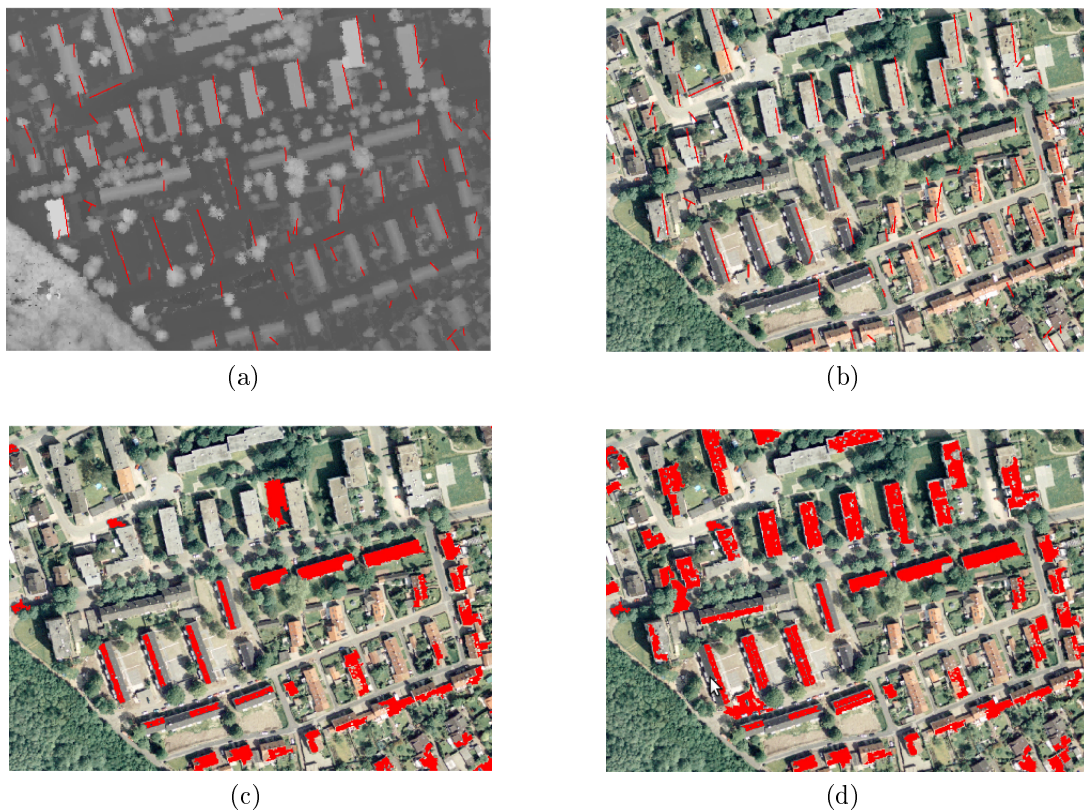


FIGURE 2.11: Principaux résultats présentés dans [Wegner 09a] et [Wegner 09b] : (a) - Détection de lignes sur l'image radar superposée à un MNE Laser. (b) - Superposition des lignes détectées sur l'image optique. (c) - Résultat de la méthode décrite dans [Muller 05] appliquée à l'image optique. (d) - Résultat de la fusion.

Ensuite, un algorithme de détection de bâtiments est appliqué séparément sur l'image optique. L'algorithme choisi par les auteurs est celui décrit dans [Muller 05]. Le résultat est celui illustré par la figure 2.11-(c). La fusion se fait ensuite par des règles simples du type "une région est considérée comme étant un bâtiment si elle a été pré-sélectionnée par l'algorithme décrit dans [Muller 05] et qu'elle a été validée par la présence d'une

ligne brillante d'une certaine longueur". Le résultat de cette détection est illustré sur la figure 2.11-(d).

Cette approche se distingue par rapport aux deux précédentes par le fait qu'elle soit symétrique. En effet, la chaîne de détection proposée attribue la même importance aux deux données, contrairement à [Tupin 03] où l'image radar est prédominante et [Sportouche 09a][Sportouche 09b] où on privilégie l'image optique.

2.3.2.4 Approche par combinaison non hiérarchique d'informations

Dans [Poulain 09], POULAIN et al. proposent une chaîne de mise à jour d'une base de données de bâtiments. La détection de changement se fait en comparant les nouvelles entrées aux anciens bâtiments sauvegardés dans la base de données. L'algorithme proposé pour la détection de bâtiments a la particularité de combiner plusieurs types d'informations, représentées sur la figure 2.12, qui sont extraites aussi bien des données radar qu'optiques, panchromatiques et multispectrales (pour le calcul du NDVI¹).

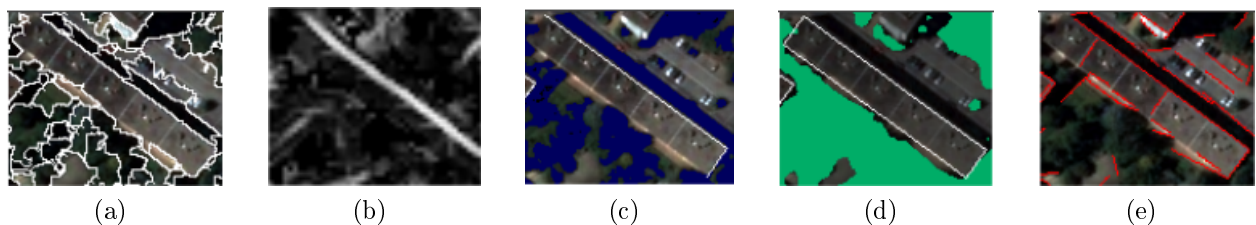


FIGURE 2.12: Primitives extraites sur les images radar et optiques dans [Poulain 09] : (a) - Contours de l'image optique [Canny 86]. (b) - Détection de lignes brillantes sur l'image radar [Tupin 98]. (c) - Masque contenant les ombres de l'image optique. (d) - Masque de végétation (NDVI) calculé à partir de l'image optique multispectrale [Tucker 79]. (e) - Détection de lignes sur l'image optique [Grompone 08].

Les informations fournies par ces primitives ne sont pas assez précises pour permettre de discriminer directement les bâtiments par rapport aux autres classes. En revanche, elles apportent des informations non négligeables sur l'appartenance d'un pixel à la réunion de deux ou plusieurs classes (dont la classe "bâtiment"). Par exemple, le masque NDVI permet de discriminer la végétation du reste de l'image, mais le résultat de l'application d'un tel masque ne permet pas directement de conclure quant à l'existence d'un bâtiment. En effet, le NDVI permet d'extraire les classes "végétation" et "zone urbaine", cette dernière peut être vue comme la réunion de la classe "bâtiment" et de la classe "autres infrastructures" (comme les routes ou les ponts).

De ce fait, ce cas de figure est parfaitement adapté à la théorie des croyances de DEMPSTER-SHAFFER [Smets 94]. En effet, pour prendre en compte l'imprécision des in-

1. NDVI (Normalized Difference Vegetation Index) : indicateur numérique sur la présence de végétation sur une image satellitaire. Il est généralement calculé à partir des bandes rouge et proche infra-rouge d'une image multispectrale.

formations fournies par les primitives, les auteurs proposent un modèle de relations inter-primitives en prenant en compte ce que chaque primitive peut discriminer. Ce modèle peut être résumé comme suit :

- Les lignes brillantes n'apparaissent pas sur tous les bâtiments.
- Les bâtiments ne peuvent pas être de la végétation, mais il existe aussi d'autres objets qui ne le sont pas aussi.
- Chaque objet a un contraste avec son voisinage.
- Les bâtiments et les arbres ont une ombre.
- Les objets qui ont une structure linéaire sont généralement des objets urbains.

Finalement, la décision se fait par une règle de “maximum de plausibilité”, et un exemple de résultat obtenu est illustré sur la figure 2.13.



FIGURE 2.13: Bâtiments détectés par la méthode proposée dans [Poulain 09].

2.3.3 Fusion optique / radar pour l'étude d'autres structures urbaines : exemple des ponts

D'autres travaux ont étudié la complémentarité optique / radar dans le cadre d'autres applications. Par exemple, dans [Soergel 07a] et [Soergel 07b], SOERGEL et al. ont exploité des images optiques aéroportées en les combinant à des interférogrammes² calculés à partir de couples radar pour estimer l'élévation des ponts. Ces infrastructures particulières ont un aspect visuel très caractéristique sur les images radar (cf. figure 2.14-(a)) et donc sur les interférogrammes (cf. figure 2.14-(b)). Ce phénomène, dit des “trois bandes” (ou “*triple-stripe*” en anglais) est dû aux différents rebonds (trois en totalité) que subit l'onde radar quand elle arrive au niveau des interfaces “eau-pont”, “pont-sol” et “pont-eau” [Soergel 07a].

Ce phénomène introduit des erreurs durant la génération du modèle numérique d'élévation, notamment en faussant l'estimation de la hauteur des ponts. D'un autre côté, vu leur géométrie d'acquisition, les images optiques ne présentent pas ce genre de phénomène,

2. Interférogramme : Image de différence de phase entre deux signaux complexes provenant de deux capteurs radar séparés spatialement. Cette technique permet, entre autres, d'estimer l'élévation d'un point imagé. Pour plus de détails, se référer aux deux articles fondateurs de l'interférométrie radar : [Graham 74] et [Zebker 86].

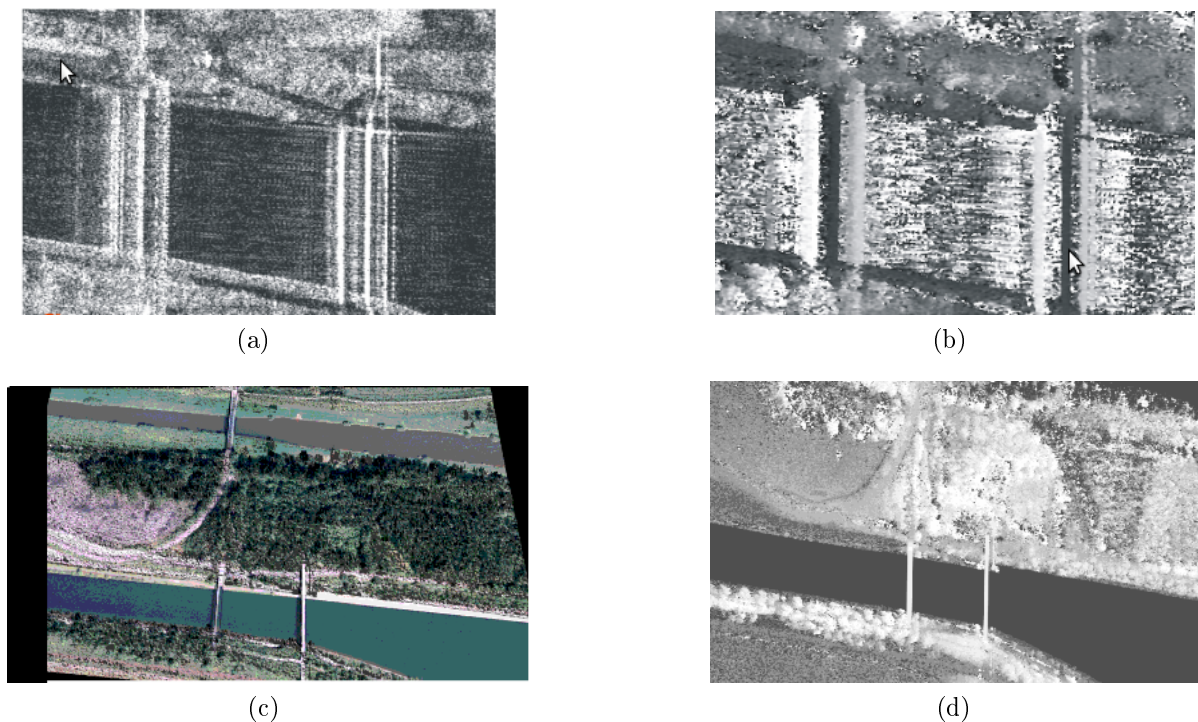


FIGURE 2.14: Principaux résultats présentés dans [Soergel 07a] et [Soergel 07b] : (a) - Image d'amplitude complexe radar sur laquelle on observe le phénomène des “trois bandes” au niveau des ponts. (b) - Interférogramme correspondant sur lequel on observe le même phénomène. (c) - Image optique recalée en géométrie radar. (d) - Résultat de la fusion : Modèle numérique d'élévation.

mais ne contiennent en revanche aucune information de hauteur (cf. figure 2.14-(c)). Cette complémentarité au niveau des informations fournies par les deux types de données est exploitée par les auteurs pour générer un modèle numérique d'élévation (cf. figure 2.14-(d)) corrigé sur lequel les ponts sont parfaitement localisés.

2.4 Conclusion

Ce chapitre a permis, dans un premier temps, de bien comprendre les principes physiques liés à chaque modalité d'imagerie satellitaire. La présentation des travaux liés à la fusion optique / radar en milieu urbain a servi par, la suite, pour mettre en évidence la complémentarité des deux sources, notamment au niveau des bâtiments. Dans le cadre de ce stage, une méthode de détection de bâtiments par fusion de données optiques et radar sera mise au point. Le principe général de cette méthode est assez proche de celui présenté dans [Poulain 09], dans le sens où la détection se fait en combinant plusieurs informations indiquant la présence d'un bâtiment. Pour cela, et comme pour toutes les applications de fusion, des méthodes de recalage précis doivent être mises au point afin de rendre possible l'exploitation simultanée de ces deux sources d'informations. Le chapitre suivant présente la méthode de recalage développée dans le cadre de ce stage.

Chapitre 3

Recalage de données optiques / radar métriques

3.1 Introduction

Après avoir présenté les principes de base de l'imagerie optique et radar et mis en évidence leur complémentarité, nous allons, dans ce chapitre, présenter une méthode de recalage de données optiques / radar que nous avons mis au point. Après avoir présenté les données utilisées et la zone d'étude, nous allons nous intéresser aux fonctions de localisation optiques et radar décrites dans [Tannous 91] et qui serviront de base pour la méthode de recalage proposée. Des améliorations de cette méthode "basique" vont être ensuite introduites pour remédier aux erreurs causées par l'imprécision des données.

3.2 Données utilisées et zone d'étude

Dans le cadre de ce stage, nous avons utilisé des données satellitaires optiques et radar fournies par le CNES (cf. annexe A). Les données optiques sont des images QUICKBIRD datant de 2002 sur la région de Toulouse. Les données radar, quant à elles, sont des images TERRASAR-X datant de 2007 sur la même zone.

Nous allons dans un premier temps présenter ces deux capteurs sans trop entrer dans les détails de leur fonctionnement, mais pour mieux comprendre les spécificités des données utilisées.

Nous allons ensuite présenter la zone sur laquelle l'étude va être faite, et analyser les caractéristiques de cette zone sur les différentes images. Ceci nous servira par la suite pour caractériser la présence d'un bâtiment (cf. chapitre 4).



FIGURE 3.1: QUICKBIRD (source : [DigitalGlobe 09]).

3.2.1 Présentation du capteur optique QUICKBIRD

QUICKBIRD (cf. figure 3.1) est un satellite d'observation de la Terre à haute résolution, lancé en 2001 par un lanceur Delta II à partir de la base de Vandenberg en Californie. Il offre en même temps la plus grande largeur de bande, la plus grande capacité de stockage à bord, et la résolution la plus élevée actuellement disponible pour les satellites commerciaux. QUICKBIRD a été conçu pour imager de grands secteurs, avec un grand degré de précision pour le géoréférencement. Il est capable d'acquérir annuellement des données couvrant plus de 75 millions de km² (plus de trois fois la taille de l'Amérique du Nord). Il collecte des images multispectrales à 2,4 m de résolution et des images panchromatiques à 0,61 m de résolution. Des composés "pan-sharpened" vraies ou fausses couleurs à 70 cm de résolution peuvent être obtenus.

3.2.2 Présentation du capteur radar TERRASAR-X

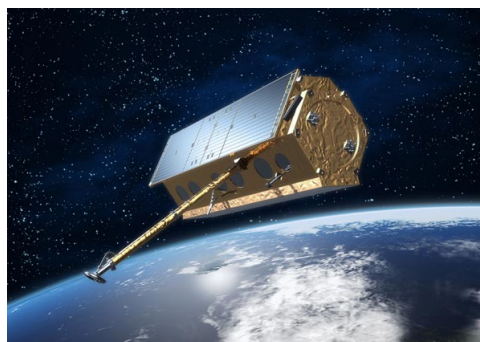


FIGURE 3.2: TERRASAR-X (source : [DLR 04]).

TERRASAR-X (cf. figure 3.2) est un capteur radar satellitaire allemand d'observation de la Terre qui fonctionne en bande X (longueur d'onde : 31 mm, fréquence : 9.6 GHz) et fournit des données haute résolution (jusqu'à 1 m). Il a été lancé le 15 juin 2007 à bord d'une fusée de type Dneper depuis la plate forme Baikonour au Kazakhstan. Les premières images ont commencé à arriver le 19 juin 2007.

Le projet à l'origine de TerraSar-X est né d'une collaboration entre le centre spatial allemand (DLR) et la société EADS Astrium. Les droits d'exploitation ont été confiés à Infoterra.

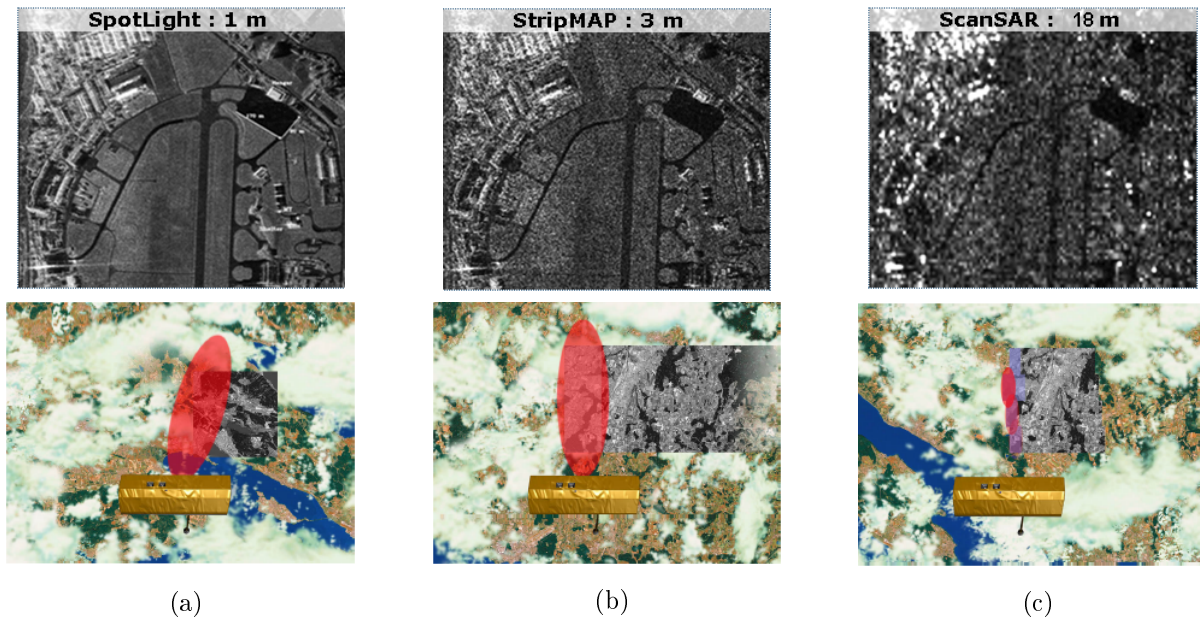


FIGURE 3.3: Modes d'acquisition de TerraSar-X et leur principe de fonctionnement : (a) - Le mode SpotLight. (b) - Le mode StripMap. (c) : Le mode ScanSar (source [DLR 04]).

Le capteur fournit trois types de données suivant leur mode d'acquisition (cf. figure 3.3) :

Le mode SpotLight C'est le mode qui fournit les données d'imagerie radar les mieux résolues du marché. Ce mode met l'accent sur la résolution spatiale (1 m de résolution), au détriment de l'extension de la zone observée (seulement $10 \text{ km} \times 5 \text{ km}$). La figure 3.3-(a) illustre son fonctionnement.

Le mode StripMap C'est un mode plus classique d'imagerie radar. L'antenne vient balayer le sol uniquement grâce au déplacement du satellite. La figure 3.3-(b) illustre son fonctionnement.

Ce mode est présenté comme étant une combinaison optimisée entre haute résolution spatiale et taille de la zone d'acquisition. La fauchée est fixée de l'ordre de 30 km et est donnée principalement par les dimensions de l'antenne (70 cm de largeur). La longueur d'acquisition est principalement limitée par la puissance de la batterie et la température des capteurs.

Le mode ScanSar C'est le mode privilégié pour couvrir de grandes surfaces. Ce mode de fonctionnement permet de couvrir, en tout point du globe, une surface de 100.000 km^2 en une semaine. La figure 3.3-(c) illustre le fonctionnement de ce mode.

3.2.3 Présentation de la zone d'étude

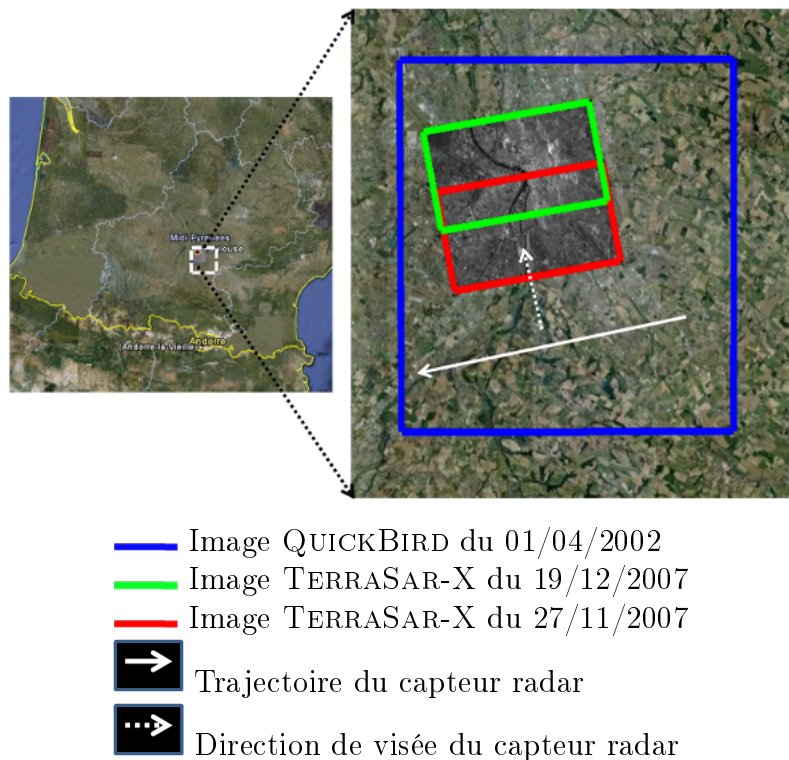


FIGURE 3.4: Données disponibles sur la zone de Toulouse.

Cette étude a été menée sur la zone de Toulouse (cf. figure 3.4), pour laquelle on dispose de données QuickBird, panchromatiques (de résolution 61 cm) et multispectrales (de résolution 2.4 m), et TerraSar-X en mode Spotlight (de résolution 1 m). Pour effectuer les projections, nous avons aussi eu recours à un modèle numérique de terrain (pour donner la valeur de l'altitude de chaque point). Nous avons pour cela utilisé un MNT SRTM¹ téléchargeable gratuitement sur le web. Ce MNT a néanmoins une résolution beaucoup plus faible que les autres images (environ 90 m) et on verra par la suite que cela introduit des erreurs au niveau du recalage.

L'ensemble des données utilisées dans le cadre de ce stage, qui sont illustrées dans l'annexe A, ont été fournies par le CNES dans le cadre d'un projet de collaboration CNES/TSI.

Nous avons choisi d'effectuer cette étude sur la zone représentée sur la figure 3.5. Cette zone est particulièrement intéressante puisque les bâtiments sont très peu contrastés par rapport au reste des éléments constituant la scène. Typiquement, un algorithme de détection de bâtiments basé uniquement sur la segmentation de l'image optique ne suffirait

1. SRTM : *Shuttle Radar Topography Mission* fait référence à des fichiers matriciels et vectoriels topographiques fournis par deux agences états-uniennes : la NASA et la NGA (ex-NIMA). Ces données altimétriques ont été recueillies au cours d'une mission de 11 jours en février 2000 par la navette spatiale Endeavour (STS-99) à une altitude de 233 km en utilisant l'interférométrie radar.

pas à discriminer le bâti du sol, ou du reste de la scène. Ceci va mettre en valeur l'intérêt de la fusion pour ce genre d'applications. D'autant plus que l'image radar est particulièrement riche en informations sur cette zone, notamment à travers la présence de *layover*.

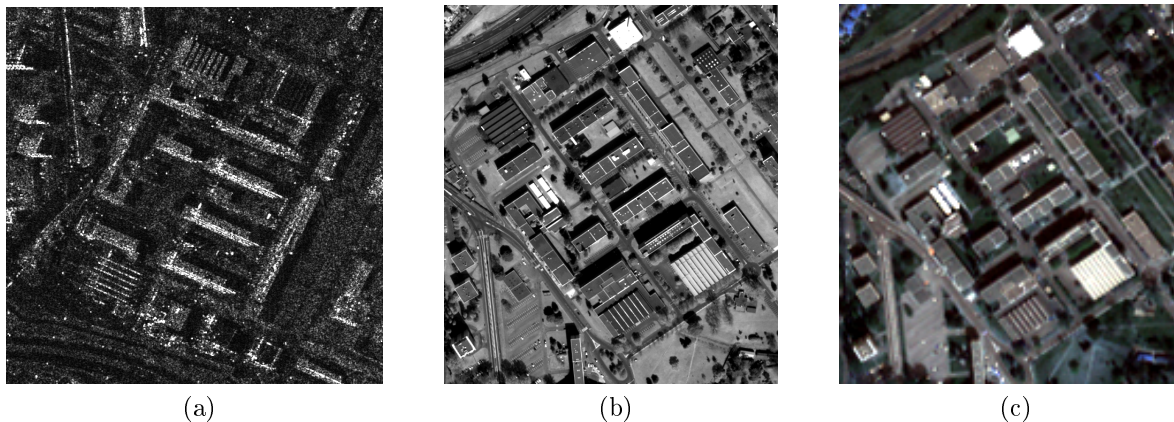


FIGURE 3.5: Zone d'étude : Complexe scientifique de Rangueil, Toulouse : (a) - Image TERRASAR-X (©DLR/EADS Astrium). (b) - Image QUICKBIRD panchromatique (©DigitalGlobe). (c) - Image QUICKBIRD multispectrale (©DigitalGlobe).

Notons néanmoins qu'il sera important de tester nos méthodes sur d'autres zones pour pouvoir généraliser les résultats obtenus.

3.3 Modélisation paramétrique de la géométrie de prise de vue

On trouve dans la littérature trois types de méthodes de recalage optique / radar. Le premier repose sur la mise en correspondance de primitives ou de points d'appui et l'estimation à partir de cet appariement de la transformation géométrique qui permet de projeter une image sur l'autre. Ce type de méthode reste néanmoins limité aux images basse résolution puisque les détails contenus dans les deux types d'images deviennent très difficiles à apparier quand on passe à la haute résolution.

On trouve aussi des méthodes dites iconiques, c'est à dire qui se basent sur l'intensité des pixels, mais qui montrent en général des résultats très peu satisfaisants.

Enfin, lorsqu'on dispose des paramètres d'acquisition des images, on peut formuler de manière paramétrique la géométrie de prise de vue afin de construire ce qu'on appelle des "modèles physiques". Cette modélisation de la prise de vue d'un capteur permet de dériver ce qu'on appelle les "fonctions de localisation" associées.

Ces fonctions, qui relient les coordonnées images aux coordonnées géographiques (et inversement) d'un point M sont couramment appelées "équations de la photogrammétrie".

La localisation dans le sens *image* $\rightarrow$ *sol* est appelée "localisation directe" et est notée G :

$$(p, q) \xrightarrow{G} (X, Y, Z) \quad (3.1)$$

où X , Y et Z sont les coordonnées 3D de la cible à la surface de la terre. Notez bien que certains ouvrages ont recours aux coordonnées géodésiques, c'est à dire la latitude et la longitude.

Les applications où l'image doit être rééchantillonnée dans une géométrie donnée nécessitent que soit connue la fonction de localisation inverse, notée G^{-1} . C'est la fonction qui permet de retrouver les coordonnées images d'un point à partir de ses coordonnées géographiques.

Dans ce qui suit, nous allons détailler les calculs relatifs à ces fonctions de localisation pour les cas optique et radar. Pour plus de détails concernant ces différents calculs, se référer à la thèse de I. TANNOS [Tannous 91] et à d'autres travaux tels que [CNES 05].

3.3.1 Fonction de localisation directe optique

Considérons le cas d'un capteur optique. Chaque pixel de coordonnées (p, q) sur l'image est l'intersection du plan image avec une droite passant par le centre optique $C_0 = (X_C, Y_C, Z_C)$. La fonction de localisation optique est donc simplement une équation traduisant la colinéarité de trois points :

$$\overrightarrow{C_0M} = \begin{bmatrix} X - X_C \\ Y - Y_C \\ Z - Z_C \end{bmatrix} = \alpha \cdot R \cdot \begin{bmatrix} p \\ q \\ -f \end{bmatrix} = \alpha \cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \cdot \begin{bmatrix} p \\ q \\ -f \end{bmatrix} \quad (3.2)$$

où f est la distance focale du capteur optique, R une matrice de rotation 3D (connue) et α un facteur d'homothétie inconnu.

Nous obtenons ainsi un système de trois équations à quatre inconnues (les coordonnées (X, Y, Z) de M ainsi que α). En éliminant ce coefficient, on obtient :

$$\begin{cases} p = -f \cdot \frac{r_{11} \cdot (X - X_C) + r_{12} \cdot (Y - Y_C) + r_{13} \cdot (Z - Z_C)}{r_{31} \cdot (X - X_C) + r_{32} \cdot (Y - Y_C) + r_{33} \cdot (Z - Z_C)} \\ q = -f \cdot \frac{r_{21} \cdot (X - X_C) + r_{22} \cdot (Y - Y_C) + r_{23} \cdot (Z - Z_C)}{r_{31} \cdot (X - X_C) + r_{32} \cdot (Y - Y_C) + r_{33} \cdot (Z - Z_C)} \end{cases} \quad (3.3)$$

Notons que les coordonnées (X_C, Y_C, Z_C) du centre C_0 dépendent du numéro de la ligne pour un capteur de type *Push Broom*.

Le problème donné par l'équation 3.1 étant sous déterminé, il ne peut donc être résolu directement. Il est alors nécessaire d'ajouter une quatrième équation.

Vu qu'on dispose généralement d'un modèle numérique de terrain (MNT) de la zone d'étude, on fait alors l'hypothèse que l'altitude h du point M est connue. Nous ajoutons donc l'équation suivante :

$$Z = h \quad (3.4)$$

Nous verrons néanmoins par la suite qu'à moins de disposer d'un MNT de même résolution que les images (radar ou optiques), ce qui généralement n'est pas le cas, cette hypothèse introduit des erreurs de localisation qu'il faudra corriger en post-traitement.

Finalement, l'équation 3.1 devient :

$$(p, q, h) \xrightarrow{G_{optique}} (X, Y) \quad (3.5)$$

où la fonction $G_{optique}$ s'écrit explicitement de manière analytique :

$$\begin{cases} X = X_C + (h - Z_C) \cdot \frac{r_{11} \cdot p + r_{12} \cdot q - r_{13} \cdot f}{r_{31} \cdot p + r_{32} \cdot q - r_{33} \cdot f} \\ Y = Y_C + (h - Z_C) \cdot \frac{r_{21} \cdot p + r_{22} \cdot q - r_{23} \cdot f}{r_{31} \cdot p + r_{32} \cdot q - r_{33} \cdot f} \end{cases} \quad (3.6)$$

3.3.2 Fonction de localisation directe radar

Pour le cas radar, compte tenue de la visée latérale du capteur, le modèle est évidemment différent de celui développé précédemment. Ce modèle a été introduit dans [Curlander 82] et a été ensuite développé dans [Tannous 91].

Nous supposons la cible $M = (X, Y, Z)$ fixe sur la surface de la Terre. La Terre est modélisée par une ellipsoïde, dont l'équation est donnée par :

$$\frac{x^2}{A^2} + \frac{y^2}{B^2} = 1 \quad (3.7)$$

où A et B sont respectivement le demi grand axe et le demi petit axe de l'ellipsoïde de référence (données connues).

La cible M étant à une altitude h de la surface de la Terre, l'équation précédente devient :

$$\frac{X^2 + Y^2}{(A + h)^2} + \frac{Z^2}{(B + h)^2} = 1 \quad (3.8)$$

Pour une cible donnée, nous noterons r sa distance radiale et t l'instant où cette cible se situe au centre du faisceau. L'axe t correspond donc à l'axe azimutal. De plus, à un pixel de coordonnées (p, q) , on peut associer une distance radiale r et un temps t selon les relations suivantes :

$$t = k \cdot p + t_0 \quad (3.9)$$

$$r = \mu \cdot q + r_0 \quad (3.10)$$

où t_0 est le temps correspondant au début de la scène et où r_0 est la distance correspondant à la première ligne-distance de l'image. k et μ sont des facteurs d'échelle qui correspondent à la taille de la case radar respectivement en azimuth et en distance radiale et valent respectivement N/F_r et C/F_e , où N est le nombre de vues de l'image, C la vitesse de la lumière, F_r la fréquence de répétition et F_e la fréquence d'échantillonnage.

La connaissance de la distance capteur-cible r et de l'instant t correspondant donne la deuxième équation fondamentale de la prise de vue radar. En notant par $\overrightarrow{OS}$ le vecteur position du radar, cette équation s'écrit :

$$r = \|\overrightarrow{SM}\| = \|\overrightarrow{OM} - \overrightarrow{OS}\| \quad (3.11)$$

où $\overrightarrow{OM}$ est la position de la cible à l'instant t (inconnue) et $\overrightarrow{OS}$ la position du capteur (supposée connue). L'équation 3.11 est l'équation de la sphère centrée en S et de rayon r (nous l'appellerons équation "sphère-distance"). A une position S du capteur et à une distance capteur-cible r correspondent plusieurs points sur la surface de la Terre. Ces points sont sur la courbe d'équi-distance, intersection de la sphère-distance et de l'ellipsoïde terrestre (cf. figure 3.6).

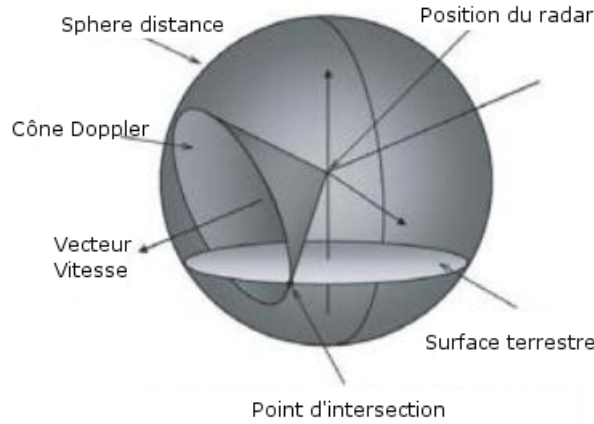


FIGURE 3.6: Intersection de la sphère-distance, du cône-Doppler et de la Terre (source [Polidori 97]).

Le point sur cette courbe qui correspond au pixel (p, q) est obtenu par la connaissance de la visée du radar. Cette visée est donnée par la fréquence Doppler centrale $f_D(p, q)$ estimée lors du traitement de formation de l'image.

L'équation obtenue est la troisième équation fondamentale de la prise de vue radar :

$$f_D(p, q) = \frac{2}{\lambda} \cdot (\overrightarrow{OM} - \overrightarrow{OS}) \cdot (\overrightarrow{V_M} - \overrightarrow{V_S}) \quad (3.12)$$

qui peut aussi s'écrire (en considérant qu'en général la cible est fixe, et que donc $\vec{V}_M = \vec{0}$) :

$$\vec{SM} \cdot \vec{V}_S = \frac{f_D(p, q) \cdot \lambda}{2 \cdot r} \quad (3.13)$$

où $\vec{V}_M$ est la vitesse de la cible à l'instant t et $\vec{V}_S$ celle du satellite.

L'équation 3.13 définit le cône de sommet S , d'axe $\vec{V}_S$ et d'ouverture $2 \cdot \arccos\left(\frac{\lambda \cdot f_D}{2 \cdot \|\vec{V}_S\|}\right)$. On appelle ce cône, "cône-Doppler" (cf. figure 3.6).

L'intersection du cône-Doppler avec la surface de la Terre donne la courbe d'iso-Doppler. L'intersection de ce cube et de la courbe d'équi-distance donne deux points sur la surface de la Terre, la connaissance de la géométrie globale de prise de vue permet d'éliminer l'un de ces points.

Finalement, en supposant qu'on dispose de toutes les informations relatives à la prise de vue (notamment la position et la vitesse du capteur à chaque instant), les équations 3.8, 3.11 et 3.13 forment un système à quatre inconnues, X , Y , Z et h (problème sous-déterminé). En supposant l'altitude h connue (comme pour le cas optique), et qu'en plus, Z et h sont exprimés dans le même référentiel, on se ramène à un problème sur-déterminé qu'on peut résoudre.

3.3.3 Fonctions de localisation inverse

Pour le cas optique aussi bien que pour le radar, la fonction de localisation directe G est non linéaire et donc difficilement inversible. Dans ce cas, la méthode de localisation inverse consiste à prédire la position d'un point de l'image par utilisation d'un polynôme (de degré 4) déduit de la localisation directe.

La position est ensuite affinée par localisation du point au sol et par correction de l'erreur induite par le prédicteur.

3.3.4 Fonctions de localisation implémentées sous OTB

Orfeo Toolbox (OTB) est une librairie C++ de traitement d'images de télédétection basée sur ITK². Le projet a été initié par le Centre National d'Etudes Spatiales (CNES) en 2006 et est resté en développement permanent avec une forte participation de la communauté scientifique. Le but du projet étant de fournir aux chercheurs une boîte à outils qui permette de manipuler les images satellitaires.

OTB a été initialement conçu pour fonctionner avec des images haute-résolution acquises par la constellation ORFEO³ mais depuis, la grande majorité des capteurs a été

2. ITK : *Insight Toolkit* est une bibliothèque C++ *Open Source* de traitement d'images médicales.

3. Pour plus d'informations, se référer au site du programme ORFEO : <http://smc.cnes.fr/PLEIADES/>

prise en compte. En particulier, la version 2.8 (que nous avons utilisée dans le cadre de ce stage) offre des modèles physiques pour une large gamme de capteurs optiques et radar, parmi lesquels QUICKBIRD et TERRASAR-X [CNES 07].

Pour le cas optique, OTB n’offre pas un vrai modèle paramétrique, mais plutôt une version analytique basée sur le modèle RPC (*Rational Polynomial Coefficients*). Ce modèle, très utilisé surtout pour modéliser la géométrie de prise de vue des capteurs optiques, peut être vu comme une approximation polynomiale du modèle réel. Il présente l’avantage d’être plus simple à inverser, mais aussi de n’être dépendant que d’un nombre très réduit de paramètres. Par exemple, pour QUICKBIRD, on fournit avec l’image des “méta-données” parmi lesquelles on peut trouver vingt coefficients RPC qui correspondent aux coefficients a_i et b_i des polynômes modélisant la fonction de transfert.

Cette approximation linéaire a néanmoins l’inconvénient d’être moins précise que le modèle réel, mais aussi utilise des paramètres qui n’ont pas un vrai sens physique pour l’utilisateur (contrairement aux paramètres réels : angle, longueur d’onde, vitesse...).

Pour le cas radar, OTB offre un modèle physique réel utilisant les vrais paramètres de prise de vue stockés dans les méta-données. La localisation inverse se fait donc par des approximations polynomiales (cf. sous section 3.3.3) et est évidemment moins précise que la localisation directe [CNES 07]. On essaiera donc par la suite de projeter l’image optique dans la géométrie radar et non l’inverse, pour exploiter cette propriété.

3.4 Méthode de recalage proposée

3.4.1 Mise en évidence de la notion d’ambiguïté du recalage

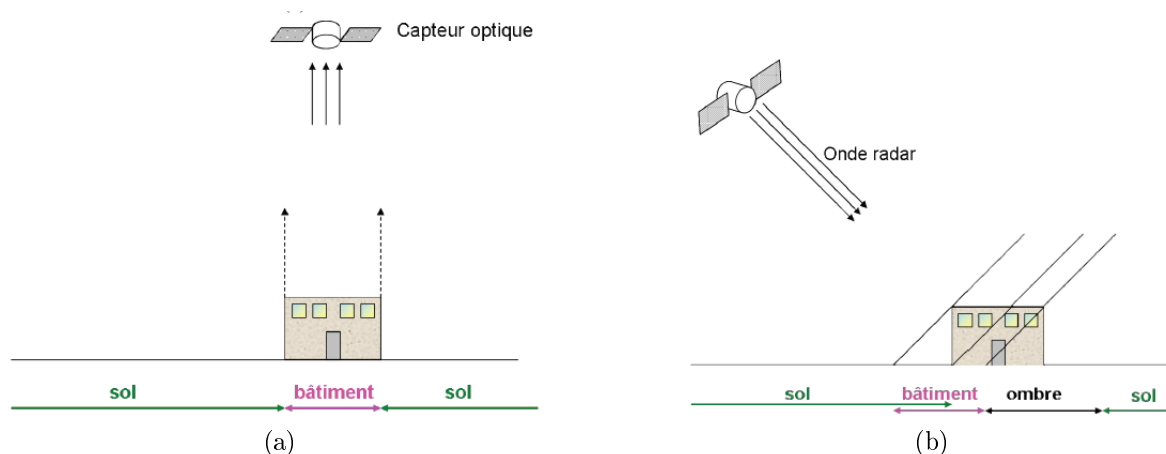


FIGURE 3.7: Mise en évidence de l’ambiguïté du recalage à travers la projection d’un bâtiment : (a) - Cas de l’image optique. (b) - Cas de l’image radar (source : [Lehureau 08]).

Nous avons cité lors du chapitre précédent les différences qui pouvaient exister entre une image radar et une image optique du point de vue de la réponse des différents éléments

constituant une scène urbaine (notamment les bâtiments). Ces différences rendent délicate l'opération de recalage vu qu'une approche classique par appariement de points n'est alors pas envisageable. Les différences peuvent également venir de la géométrie particulière des images. Par exemple, en imagerie radar, en raison de la visée latérale, comme le montre la figure 3.7 qui représente la projection d'un bâtiment. Sur l'image radar, une partie du sol et du bâtiment sont superposés, de plus une ombre apparaît, elle correspond à une zone où le capteur radar ne reçoit aucune information. Ce schéma montre l'ambiguïté du recalage : *“faut-il recaler la projection du sol entre les deux images ou alors la projection du bâtiment ?”*.

La réponse à cette question dépend de l'application visée. Par exemple, pour une application de reconstruction 3D, il est important d'avoir un “bon” recalage au sol, de manière à pouvoir par la suite exploiter l'information de décalage de la réponse des bâtiments sur l'image radar pour retrouver la hauteur de ces derniers [Sportouche 09a][Sportouche 09b].

Dans notre cas, vu l'application souhaitée, à savoir la détection de bâtiments, il est important d'avoir une correspondance parfaite entre les bâtiments sur les deux images, optique et radar. Vu que la localisation de la réponse des bâtiments sur l'image radar dépend de leur hauteur, et que donc, on n'aura pas en général une superposition parfaite entre les deux images, on essaiera de les recaler au mieux, c'est à dire d'avoir le maximum de bâtiments qui soient bien recalés. Typiquement, un bon recalage se traduirait par une superposition relativement bonne des contours optiques sur les lignes brillantes radar.

Notons que sur la zone d'étude que nous avons choisie, les bâtiments ont des hauteurs respectives relativement proches, et que donc, on pourra recaler les images en ayant une superposition des bâtiments relativement bonne.

3.4.2 Schéma synoptique de la méthode de base

Dans cette partie, nous proposons une méthode de recalage basée sur les fonctions de localisation présentées précédemment. Pour les raisons que nous avons évoquées dans le paragraphe précédent, nous avons choisi de projeter les données optiques en géométrie radar. D'autres motivations nous ont aussi incité à faire ce choix. Par exemple, le fait que l'image optique soit de résolution plus fine que l'image radar nous permet de nous passer d'une éventuelle interpolation du résultat du recalage.

Si on note par (p_{rad}, q_{rad}) les coordonnées pixel sur l'image radar, (p_{opt}, q_{opt}) ceux de l'image optique, $(lat, long)$ les coordonnées géodésiques (latitude et longitude) du point, G_{rad} la fonction de localisation directe radar, G_{opt}^{-1} la fonction de localisation inverse optique, projeter l'image optique en géométrie radar revient à trouver pour chaque point (p_{rad}, q_{rad}) de l'image radar, le point (p_{opt}, q_{opt}) de l'image optique tel que :

$$(p_{rad}, q_{rad}) \xrightarrow{G_{rad}} (lat, long) \xrightarrow{G_{opt}^{-1}} (p_{opt}, q_{opt}) \quad (3.14)$$

Nous avons vu auparavant que les fonctions de localisation utilisaient une information d'altitude pour pouvoir projeter les points dans une géométrie de prise de vue donnée. Nous avons dans un premier temps testé ce schéma de fonctionnement en fournissant une même valeur d'altitude qui correspond à l'altitude moyenne de la zone d'étude, à savoir 230m. Le résultat ainsi obtenu est celui illustré sur la figure 3.8.



FIGURE 3.8: Résultat du recalage en utilisant une même valeur d'altitude pour tous les points : (a) - Image radar. (b) - Image optique projetée en géométrie radar. Des points de même couleur ont été rajoutés manuellement pour donner une idée sur le décalage.

Le résultat n'est évidemment pas satisfaisant, mais il permet de mettre l'accent sur l'influence de la précision de h dans le résultat des fonctions de localisation. Il est alors important de fournir pour chaque point une information d'altitude plus précise que le h_{moyen} utilisé précédemment. Pour cela, nous introduisons le modèle numérique de terrain (cf. annexe A) de la zone dans la chaîne de recalage. l'équation 3.14 devient alors :

$$\begin{array}{ccccc}
 (p_{rad}, q_{rad}) & \xrightarrow{G_{rad}} & (lat, long) & \xrightarrow{G_{opt}^{-1}} & (p_{opt}, q_{opt}) \\
 \uparrow & & \uparrow & & \\
 h_{MNT} & & h_{MNT} & &
 \end{array} \quad (3.15)$$

Le résultat ainsi obtenu est celui illustré sur la figure 3.9.

On remarque que même si l'introduction de la donnée d'altitude améliore le résultat de la localisation, la faible résolution du modèle numérique de terrain fait que cette information d'altitude est imprécise. Ceci entraîne des erreurs de localisation qui se traduisent par des erreurs au niveau du recalage.

Afin de remédier à ce problème, nous avons eu recours à une information supplémentaire fournie dans les méta-données. En effet, la plupart des produits de télédétection sont fournis avec des informations concernant les coordonnées géographiques des quatre coins de l'image. Ceci sert en général à localiser la zone couverte par les images.



FIGURE 3.9: Résultat du recalage en utilisant le MNT lors de la projection : (a) - Image radar. (b) - Image optique projetée en géométrie radar.

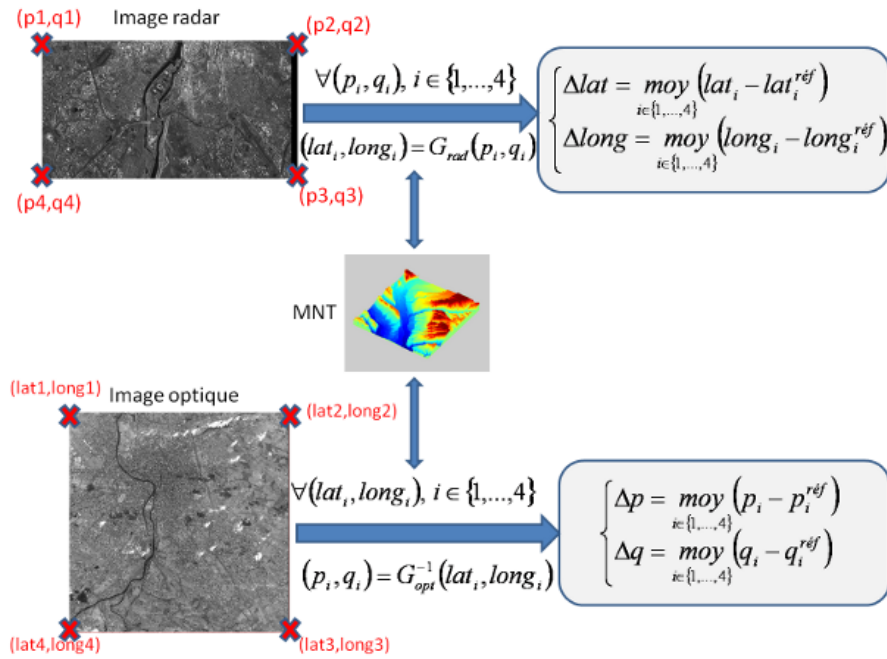


FIGURE 3.10: Estimation des décalages pour chaque image à partir des quatre coins de référence.

Nous avons exploité ces informations pour corriger les erreurs de localisation dues à l'imprécision du MNT. La figure 3.10 montre la méthode adoptée.

Pour chaque coin de l'image radar, on calcule les coordonnées géographiques correspondantes en utilisant la fonction de localisation directe radar. Ces résultats sont ensuite comparés à ceux fournis dans les méta-données et qui servent de référence. Ceci nous permet d'estimer quatre décalages (exprimés en degrés géodésiques) correspondant chacun à un coin. On appelle alors $(\Delta lat, \Delta long)$ la moyenne de ces quatre décalages, et qui correspond à l'erreur moyenne (en degrés) de localisation directe radar.

De la même manière, à partir des quatre coins de l'image optique, on estime un décalage moyen $(\Delta p, \Delta q)$ correspondant à l'erreur moyenne de localisation inverse de l'image optique.

Ces décalages sont ensuite réinjectés dans la chaîne de calcul comme le montre la figure 3.11.

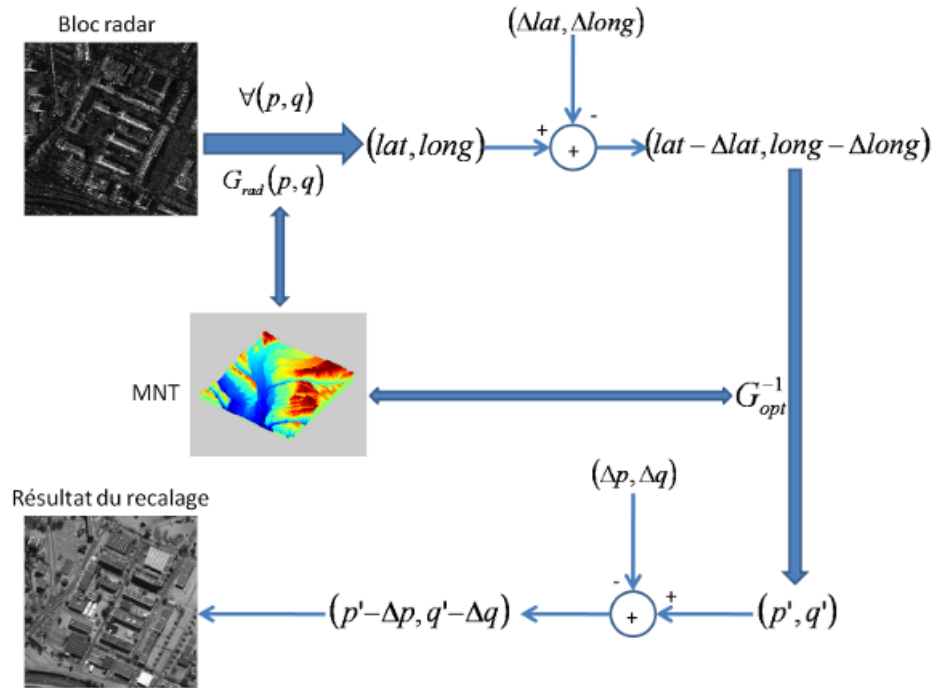


FIGURE 3.11: Schéma synoptique de la méthode de recalage adoptée.

Pour chaque pixel (p, q) de l'image radar, on calcule ces coordonnées géographiques en utilisant la localisation directe radar, dans laquelle on injecte l'altitude h de ce point. On retranche ensuite à ce résultat le décalage $(\Delta lat, \Delta long)$ calculé précédemment. Ceci permet de corriger les erreurs de localisation inverse radar. Ces coordonnées géographiques corrigées sont ensuite prises comme entrée pour la localisation inverse optique. Les coordonnées pixels ainsi obtenus sur l'image optique sont aussi corrigés de la même manière en injectant $(\Delta p, \Delta q)$. Le résultat final ainsi obtenu est illustré sur la figure 3.12.

Afin de mieux analyser la précision du recalage, nous visualisons le résultat de la manière présentée sur la figure 3.12-(c). Pour cela, nous avons extrait les contours de l'image optique (par seuillage du module du gradient) et nous les avons superposé sur l'image radar.

Ceci nous permet de conclure que, même si l'introduction des coins de référence améliore nettement la précision du recalage, des erreurs subsistent et la superposition des deux images n'est pas parfaite.

De plus, le fait de se baser uniquement sur les fonctions de localisation pour faire le recalage fait que les deux images seront bien recalées au sol, mais pas au niveau des bâtiments (cf. sous section 3.4.1).

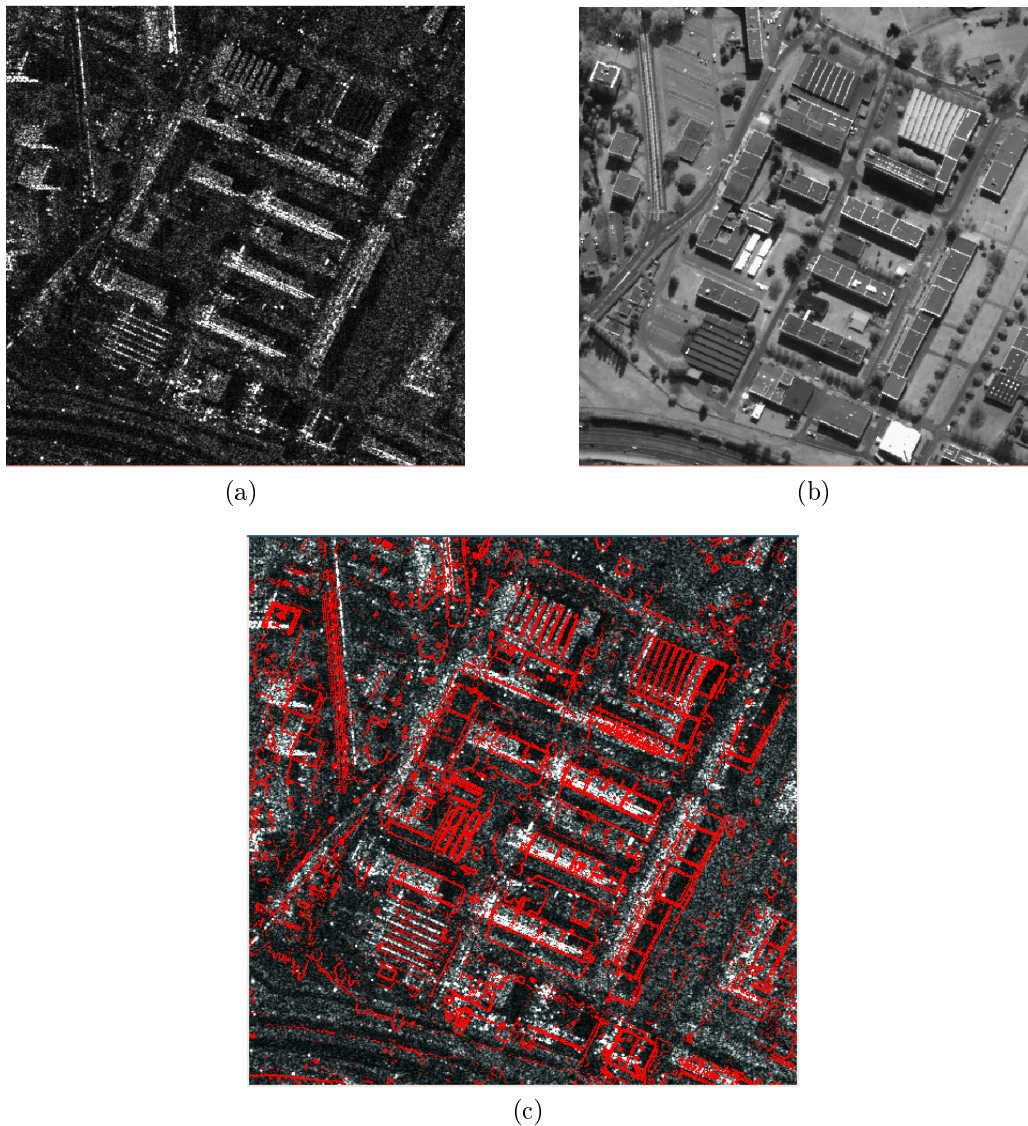


FIGURE 3.12: Résultat du recalage en utilisant le MNT, en estimant les décalages et en les réinjectant lors de la projection : (a) - Image radar. (b) - Image optique projetée en géométrie radar. (c) - Superposition des contours du résultat sur l'image radar.

Pour résoudre ce problème, nous avons dans un premier temps tenté de jouer sur les paramètres d'acquisition radar en considérant qu'ils pouvaient être entâchés d'une certaine erreur. Il est important de rappeler que la localisation optique étant non paramétrique (entièrement basée sur le modèle RPC), les coefficients utilisés lors de la localisation n'ont pas un vrai sens physique, et sont donc difficilement optimisables.

Nous avons pour cela étudié la précision des différents paramètres de prise de vue radar (cette précision est en général fournie dans les méta-données), et avons conclu que les paramètres les plus sensibles sont les paramètres orbitaux.

En effet, nous avons vu que parmi les paramètres pris en compte lors de la modélisation paramétrique de la prise de vue radar, les équations de localisation utilisaient la position et la vitesse du capteur à chaque instant. Pour cela, pour chaque acquisition TerraSar-

X, on fournit avec l'image certains paramètres orbitaux qui permettent de retrouver ces valeurs.

A défaut de fournir un nombre important de mesures de positions et de vitesses, les méta-données contiennent ce qu'on appelle des "vecteurs d'états" et qui sont des valeurs particulières de position et de vitesse qui sont sauvegardées à des intervalles réguliers. Les valeurs à chaque instant sont ensuite retrouvées en effectuant une interpolation de LAGRANGE sur ces points.

Dans le cas de TerraSar-X, les méta-données contiennent douze vecteurs d'état qui contiennent chacun six valeurs, à savoir les positions (X, Y, Z) et les vitesses (V_x, V_y, V_z) du capteur. On fournit aussi la précision de ces mesures qui est de $\epsilon_{pos} = 20cm$ pour les positions et de $\epsilon_{vit} = 5 \cdot 10^{-3}ms^{-1}$ pour les vitesses.

Nous avons tenté de corriger les erreurs de recalage en faisant varier ces valeurs entre $-\epsilon_{pos,vit}$ et $+\epsilon_{pos,vit}$. Nous avons pour cela mis au point un parseur xml qui modifie à chaque itération les valeurs contenues dans le fichier xml des méta-données. Vu le nombre de paramètres sur lesquels se fait l'optimisation (en l'occurrence 72), et à défaut d'avoir recours à des méthodes sophistiquées d'optimisation (type LEVENBERG-MARQUARDT) mais compliquées à mettre en place, nous avons fait l'hypothèse que les erreurs de position et de vitesses sont des erreurs de mesure GPS, et que donc, la même erreur, si elle existe, affecterait toutes les valeurs, de position d'un côté et de vitesse de l'autre.

Nous avons ainsi fait varier les deux erreurs Δ_{pos} et Δ_{vit} et avons calculé à chaque itération un score qui correspond aux moindres carrés sur les quatre coins de référence de l'image radar. Nous avons ainsi trouvé des valeurs de positions et de vitesses du capteur qui sont "meilleures" (au sens des moindres carrés) que les paramètres initiaux. Néanmoins, ces "nouvelles valeurs" n'ont pas d'influence sur le recalage. En effet, même si la localisation radar est devenue ainsi plus précise, la différence de précision est inférieure à un pixel au niveau de l'image optique (vu qu'elle est de résolution supérieure).

Nous avons donc conclu qu'une amélioration de la précision du recalage devait plutôt passer par des raffinements en post-traitement que par des optimisations faites sur les paramètres de prise de vue. Ceci fait l'objet du paragraphe suivant.

3.4.3 Amélioration des résultats en exploitant l'information mutuelle

Dans ce paragraphe, nous allons proposer une méthode de raffinement du recalage obtenu précédemment (cf. figure 3.12). Pour cela, nous allons procéder selon un schéma classique en recalage, à savoir appliquer une transformation géométrique sur l'image optique et calculer une mesure de similarité entre les deux images et garder les paramètres de la transformation qui maximise cette mesure. Vu les différences entre les deux images, une simple corrélation (ou autre mesure du même type) ne convient pas. Une étape cruciale

du recalage est alors le choix de la mesure qui convient.

Dans [Inglada 04], une étude sur les mesures de similarité utilisées pour le recalage multi-modalités a été faite. Il en ressort que la mesure qui convient le mieux au cas optique radar est l'information mutuelle [Viola 97].

L'information mutuelle (IM) est un critère, issu de la théorie de l'information, qui mesure le degré de dépendance au sens probabiliste de deux variables aléatoires. Ce concept de dépendance logique ne doit pas être confondu avec celui de causalité physique, bien qu'en pratique l'un implique souvent l'autre.

Le principal avantage de cette mesure c'est que la forme de dépendance qui existe entre les deux variables aléatoires n'a pas à être spécifiée. Ceci la rend très adaptée au cas des sources multi-modalités. Notons aussi le fait que dans notre cas, vu que les images contiennent principalement des bâtiments, cette mesure devrait donner un bon recalage au niveau des bâtiments et non au niveau du sol, puisque ceux-ci dominent la scène.

Dans le cas des images, l'information mutuelle (selon l'implémentation de Viola-Wells [Viola 97]) s'exprime en fonction de l'entropie. Si on désigne par A et B deux images, par $H(A)$ et $H(B)$ leurs entropies respectives et par $H(A, B)$ l'entropie jointe des deux images, l'information mutuelle de A et B est donnée par :

$$IM(A, B) = H(A) + H(B) - H(A, B) \quad (3.16)$$

Notez bien que si A et B sont indépendantes, $H(A, B) = H(A) + H(B)$ et donc $IM(A, B) = 0$.

Nous avons testé cette mesure sur les deux images illustrées sur la figure 3.12. Nous avons dans un premier temps choisi comme transformation géométrique une translation 2D de paramètres t_x et t_y . Nous avons fait varier ces paramètres entre -20 et $+20$ et avons étudié le comportement de l'information mutuelle mesurée entre les deux images. Néanmoins, nous n'avons pas observé un pic particulier de cette mesure, même si un maximum existe, les valeurs restent très proches les unes des autres.

Ceci est dû au fait que l'image radar est fortement bruitée. En effet, on observe sur l'image un bruit multiplicatif appelé "speckle" qui diminue la dépendance entre les deux images en introduisant des valeurs aléatoires et fausse ainsi les résultats de l'information mutuelle.

Nous avons donc choisi de débruiter cette image radar avant d'appliquer ce processus de raffinement. Plusieurs méthodes de débruitage existent dans la littérature, nous avons opté pour celle décrite dans [Deledalle 09].

Cet algorithme, qui se base sur une estimation au sens du maximum de vraisemblance pondéré dont les poids sont définis via des patches et adaptés au modèle de bruit, convient parfaitement au cas radar puisque des modèles de bruits caractéristiques existent, notamment celui proposé par GOODMAN dans [Goodman 76].

L'application de ce filtre sur l'image radar correspondant à notre zone d'étude donne

le résultat illustré sur la figure 3.13.



FIGURE 3.13: Résultat du filtrage de l'image radar en utilisant l'algorithme décrit dans [Deledalle 09] : (a) - Image originale. (b) - Image filtrée.

En appliquant de nouveau sur l'image filtrée la procédure expliquée précédemment pour la recherche de la translation 2D qui maximise l'information mutuelle, on remarque un pic plus net que pour le cas précédent (cf. figure 3.14). Ce pic correspond à la translation $t_x = -19$ et $t_y = -10$.

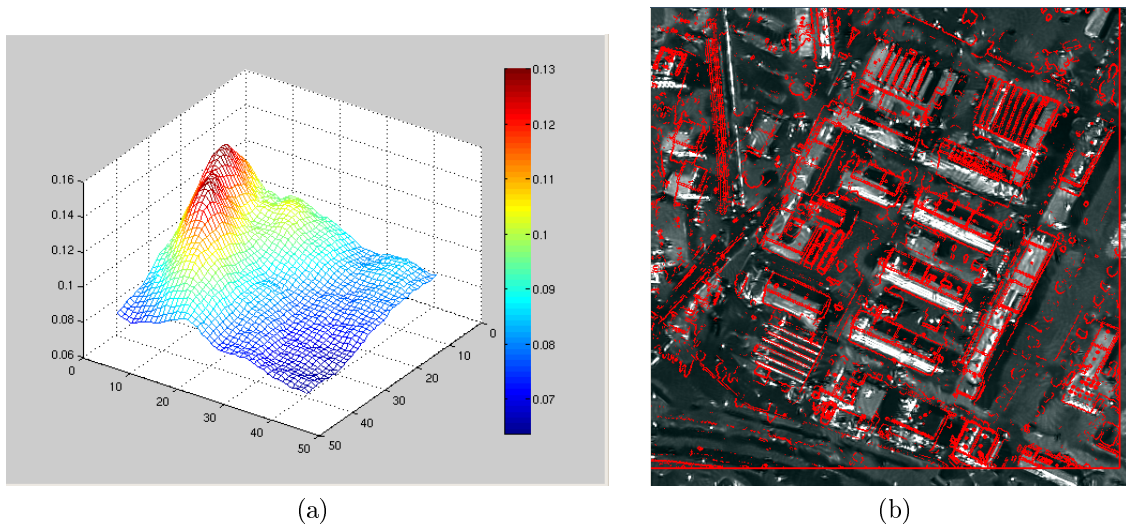


FIGURE 3.14: Résultat de la recherche de la translation qui maximise l'information mutuelle : (a) - Représentation 3D des valeurs de l'information mutuelle, maximum atteint pour $t_x = -19$ et $t_y = -10$. (b) - Résultat correspondant : superposition des contours de l'image optique sur l'image radar.

On remarque que, même si on a une superposition partielle entre les principaux bâtiments, une grande partie de ces derniers reste mal recalée. Ceci est dû au fait qu'une translation ne suffit pas pour bien recalrer les deux images.

Nous avons donc appliqué le même procédé mais en utilisant cette fois une similitude 2D. Vu le nombre important de paramètres de cette transformation (quatre au total, deux pour la translation, un pour la rotation, et un pour l'homothétie, le centre de la rotation étant celui de l'image), une recherche exhaustive serait lourde. Nous avons donc utilisé un algorithme d'optimisation par descente de gradient implémenté sous OTB (ou plus précisément sous ITK : *itk::GradientDescentOptimizer*). Nous avons initialisé les paramètres de la rotation et de l'homothétie à zéro, et ceux de la translation à la valeur trouvée précédemment par la recherche de la translation qui maximise l'information mutuelle. Cela nous permet de nous placer près du maximum local et réduit ainsi le temps de calcul.

Finalement, l'algorithme converge au bout de 83 itérations, et le résultat obtenu est celui illustré sur la figure 3.15. Il correspond aux paramètres de la similitude suivante : une translation $t_x = -15$ et $t_y = -22$, une rotation d'angle $\theta = 0.2354^\circ$ et une homothétie de facteur $\alpha = 1.0039$.

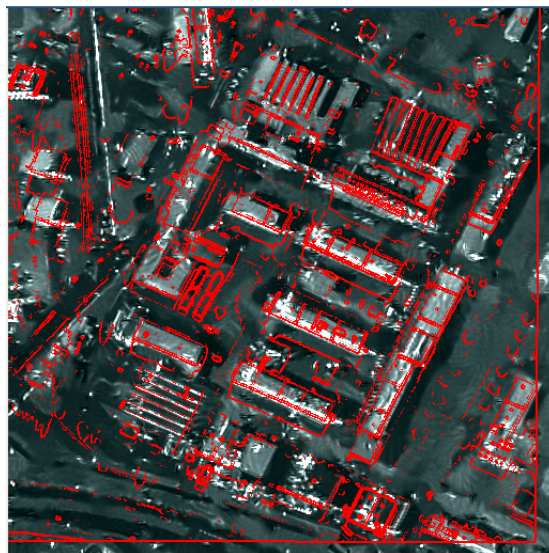


FIGURE 3.15: Résultat de la recherche de la similitude qui maximise l'information mutuelle : maximum atteint en appliquant une translation $t_x = -15$ et $t_y = -22$, une rotation d'angle $\theta = 0.2354^\circ$ et une homothétie de facteur $\alpha = 1.0039$.

Le résultat ainsi obtenu est assez satisfaisant, notamment au niveau de la superposition des bâtiments entre les deux images.

Dans toute la suite, nous allons adopter cette méthode pour recalibrer le reste des données, notamment la bande rouge et proche infra-rouge qui serviront pour le calcul du NDVI (cf. chapitre 4).

3.5 Conclusion

Au cours de ce chapitre, nous avons présenté une méthode de recalage basée sur les modèles physiques implémentés sous OTB et qui correspondent aux capteurs TerraSar-

X et QuickBird. Nous avons constaté que ces modèles ne sont pas précis, et que cette imprécision, combinée à celle du MNT, induisent des erreurs de recalage.

Nous avons donc introduit une chaîne de post-traitements basés sur la mesure de l'information mutuelle afin d'améliorer la précision du recalage. Les résultats ainsi obtenus sont satisfaisants dans la mesure où les données sont assez bien recalées au niveau des bâtiments pour pouvoir les fusionner par la suite.

Le chapitre suivant présente une méthode de détection de bâtiments basée sur la fusion des données optiques et radar précédemment générées.

Chapitre 4

Détection de bâtiments par fusion de données optiques / radar

4.1 Introduction

Ce chapitre présente une méthode de détection de bâtiments que nous avons développée au cours de ce stage. Nous allons dans un premier temps présenter la méthode proprement dite. L’accent sera mis sur la justification du choix de la règle de fusion adoptée. Les résultats seront par la suite analysés et critiqués.

4.2 Méthode proposée

Nous allons proposer dans cette partie une méthode de détection de bâtiments qui exploite les données QuickBird et TerraSar-X dont on dispose et que nous avons recalées avec la méthode présentée précédemment.

Cette méthode prend comme entrée l’image radar filtrée [Deledalle 09], l’image optique panchromatique projetée en géométrie radar (cf. chapitre 3) et les bandes “rouge” et “proche infra-rouge” de l’image optique multispectrale recalée (cf. figure 4.1).

A partir de ces images, nous allons extraire des “cartes d’indices”. Le choix de ces indices est fait de manière à caractériser la présence d’un bâtiment. Ces cartes serviront par la suite de base pour calculer des cartes de score relatives à ces indices. Nous avons choisi une approche “par régions” pour l’attribution des scores, puisque ce type d’approche convient mieux à l’application visée. En effet, il est difficile de caractériser la présence d’un bâtiment au niveau pixelique.

Ces cartes de score sont ensuite fusionnées pour aboutir à une carte de score finale qui attribue des degrés d’appartenance entre 0 et 1 à la classe bâtiments. Cette carte est seuillée pour ne garder que les régions sélectionnées et l’étape finale consiste à effectuer une recherche de rectangles sur ces régions.

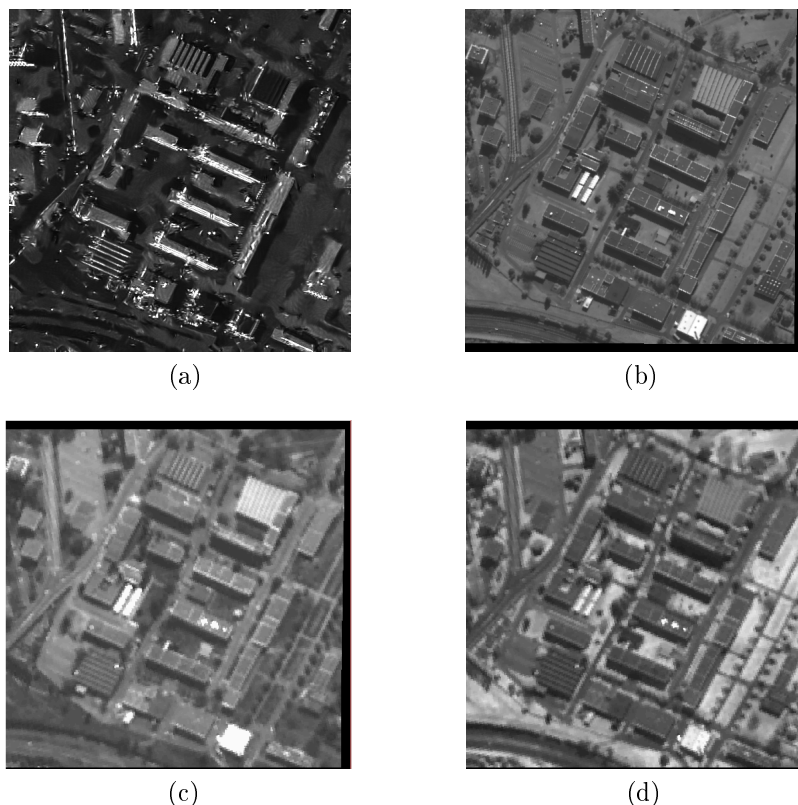


FIGURE 4.1: Données recalées utilisées comme entrées pour la méthode de détection de bâtiments : (a) - Image radar filtrée. (b) - Image optique panchromatique. (c) - Bande “rouge” de l’image optique multispectrale. (d) - Bande “proche infra-rouge” de l’image optique multispectrale.

Le schéma synoptique de cette méthode est illustré sur la figure 4.2.

4.2.1 Extraction des cartes d’indices

Pour définir ces cartes, nous avons besoin de définir quels sont les indices caractérisant la présence d’un bâtiment dans une scène urbaine. Nous avons finalement choisi les indices suivants :

- Présence d’une ombre sur l’image optique.
- Présence d’une ombre sur l’image radar.
- Présence de lignes brillantes ou de *layover* sur l’image radar.
- Absence de végétation.

Nous désignerons dans la suite par “zones brillantes” l’ensemble *layover* et lignes brillantes.

Les cartes d’indices correspondant aux ombres et aux zones brillantes sont des images binaires extraites par simple seuillage (cf. figure 4.3 - (a), (b) et (c)). Les seuils n’étant pas les mêmes pour toutes les images, il serait intéressant d’envisager par la suite une automatisation de cette étape en se basant sur l’histogramme des images, d’autant plus que nous avons observé que ces seuils ne sont pas très sensibles.

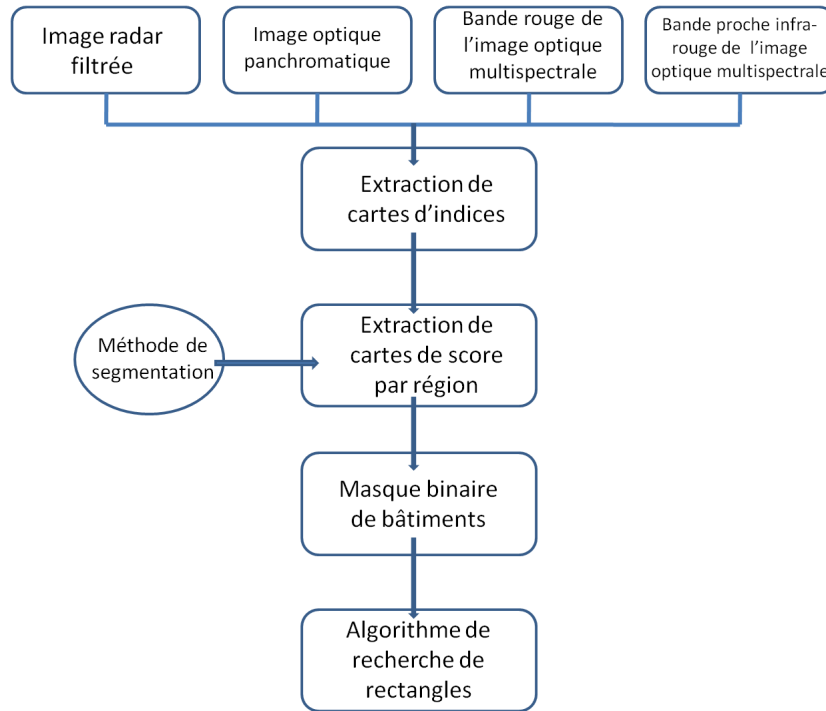


FIGURE 4.2: Schéma synoptique de la méthode de détection de bâtiments proposée.

La présence de végétation est, quant à elle, caractérisée par une carte d'indice de végétation NDVI (cf. figure 4.3) calculée à partir des bandes “rouge” et “proche infra-rouge” de l'image optique multispectrale. Dans [Tucker 79], TUCKER et al. définissent le NDVI comme étant :

$$NDVI = \frac{PIR - R}{PIR + R} \quad (4.1)$$

où PIR désigne la bande “proche infra-rouge” de l'image multispectrale et R la bande rouge.

4.2.2 Génération des cartes de score

Comme nous l'avons mentionné précédemment, nous avons opté pour une approche “par régions” pour l'attribution des scores. Pour cela, une étude concernant les différents algorithmes de segmentation implémentés sous OTB a été faite et nous avons finalement choisi d'utiliser le *Mean Shift* [Comaniciu 02]. Cet algorithme présente l'avantage de donner à l'utilisateur, à travers le choix des paramètres, la possibilité de segmenter l'image d'une manière plus ou moins grossière.

Nous avons choisi les paramètres suivants pour le *Mean Shift* : $\rho_{ng} = 10$, $\rho_{spatial} = 10$ et $\lambda = 1$ afin d'avoir une sur-segmentation de l'image (cf. figure 4.4). Pour plus de détails concernant le choix de ces paramètres, se référer à [Comaniciu 02].

L'idée est que quitte à ce qu'un bâtiment soit divisé en plusieurs régions différentes lors

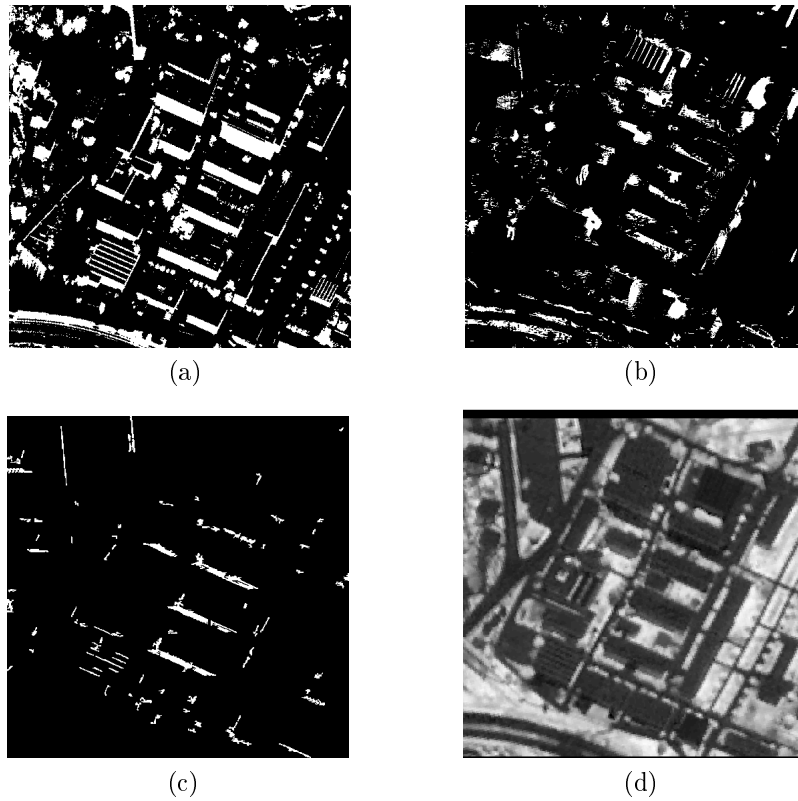


FIGURE 4.3: Cartes d'indices extraites à partir des images optiques et radar : (a) - Masque binaire des ombres sur l'image optique. (b) - Masque binaire des ombres sur l'image radar. (c) - Masque binaire des zones brillantes sur l'image radar. (d) - Carte d'indice de végétation NDVI.



FIGURE 4.4: Segmentation de l'image optique panchromatique par l'algorithme *Mean Shift*.

de la segmentation, l'essentiel c'est que ces régions ne débordent pas sur d'autres éléments de la scène, comme le sol. D'autant plus que sur la zone étudiée, la majorité des bâtiments sont très peu contrastés par rapport à leur voisinage. Même si cette sur-segmentation ne

préserve pas les contours, et donc la forme des bâtiments, la fusion combinée à la recherche de rectangles devrait par la suite permettre de retrouver ces formes.

4.2.2.1 Cartes de score des ombres optiques et radar

Les cartes d'indice des ombres, qui sont des masques binaires obtenus par seuillage des zones sombres sur les deux images, sont exploitées pour calculer des scores (par région) qui traduisent la proximité de chaque région d'une ombre, dans une certaine direction. L'idée est d'exploiter cette connaissance d'une direction privilégiée, qui est celle de l'éclairage du soleil par rapport à la scène pour l'image optique et la direction de visée du capteur pour l'image radar, pour calculer une distance par rapport aux ombres dans cette direction. Notons que, même si ces directions privilégiées sont pour le moment fixées manuellement, nous pouvons envisager dans un second temps des les extraire automatiquement des méta-données, puisqu'elles sont généralement fournies.

Pour calculer ces “cartes de distance orientée”, nous nous sommes basés sur les résultats présentés dans [Bloch 99]. L'idée est de labelliser les objets présents sur les masques des ombres, et pour chaque label, d'effectuer le traitement suivant (cf. figure 4.5) :

- Calculer une carte de distance morphologique “classique” relative à cet objet.
- Définir un masque qui caractérise une position particulière par rapport à cet objet (en l'occurrence la direction privilégiée mentionnée précédemment).
- Masquer la carte de distance morphologique par ce masque de position.

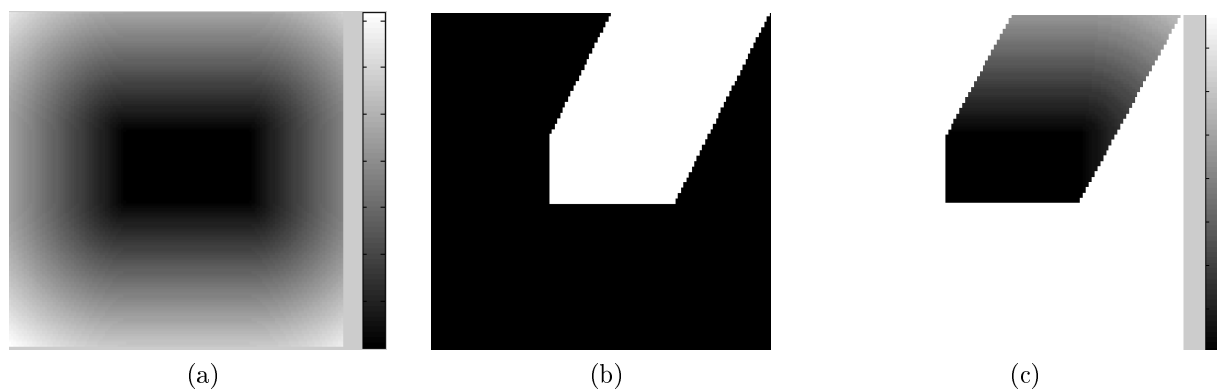


FIGURE 4.5: Illustration de la génération des cartes de distance orientée pour un objet donné : (a) - Carte de distance morphologique “classique”. (b) - Masque de la zone “à 120° de l'objet”. (c) - Carte de distance orientée dans cette direction.

Cette procédure n'est évidemment qu'une approximation des résultats présentés dans [Bloch 99], dans le sens où le calcul de la distance ne se fait pas dans une direction en particulier, mais nous avons estimé que cela serait suffisant pour notre application.

La carte de distance globale (cf. figure 4.6) est ensuite obtenue en fusionnant toutes les cartes correspondant chacune à un objet, et en gardant à chaque fois le minimum de distance en cas de superposition.

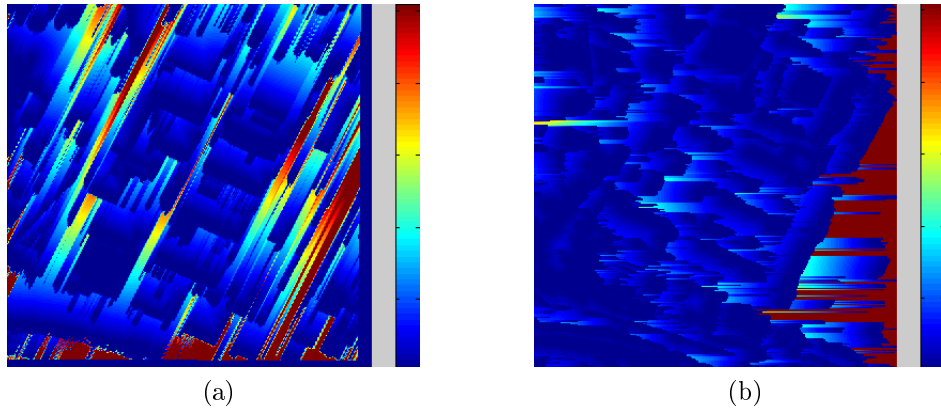


FIGURE 4.6: Cartes de distance orientée obtenues : (a) - Pour l'ombre optique, direction utilisée : $\alpha = 115^\circ$. (b) - Pour l'ombre radar, direction utilisée : $\alpha = 0^\circ$.

Les directions choisies sont de $\alpha_1 = 115^\circ$ pour les ombres optiques et de $\alpha_2 = 0^\circ$ pour les ombres radar.

La génération des cartes de score se fait ensuite suivant la procédure illustrée sur le schéma synoptique de la figure 4.7.

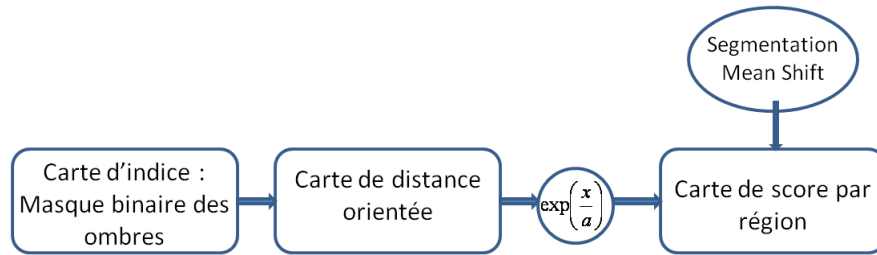


FIGURE 4.7: Schéma synoptique de la procédure de calcul des scores correspondant aux ombres optiques et radar.

La carte de distance est tout d'abord transformée en lui appliquant une fonction $\exp\left(\frac{\text{distance}}{a}\right)$. Ceci permet d'accentuer la notion de proximité par rapport à une ombre donnée en attribuant une distance presque nulle aux régions très proches de l'ombre dans la direction privilégiée, et en discriminant fortement les autres régions au fur et à mesure qu'on s'éloigne. La valeur du paramètre a permet d'accentuer plus ou moins cet effet. Des expérimentations nous ont permis de fixer ce paramètre à $a = 5$. Le résultat de la carte de distance correspondante est celui de la figure 4.8.

Cette carte est ensuite inversée (au sens des niveaux de gris) pour passer d'une distance à un score de degré d'appartenance. Ensuite, pour chaque région segmentée par le *Mean Shift*, on attribue un score qui correspond à la médiane des scores des pixels appartenant à cette région. Ce choix permet d'atténuer l'influence des valeurs extrêmes (*outliers*).

Les cartes de score correspondantes sont illustrées sur la figure 4.9.

On remarque que sur la carte de score des ombres optiques, les régions de fort score correspondent bien aux bâtiments, mais aussi à d'autres éléments de la scène comme les

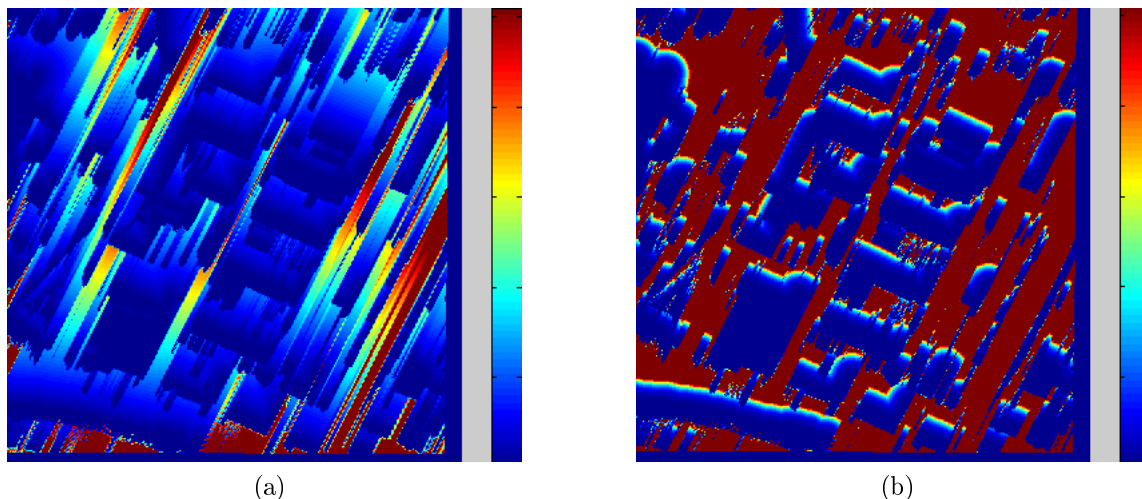


FIGURE 4.8: Mise en évidence de l'intérêt de l'utilisation d'une distance en $\exp\left(\frac{\text{distance}}{a}\right)$: (a) - Carte de distance orientée obtenue avec une distance linéaire. (b) - Carte de distance orientée obtenue avec une distance en $\exp\left(\frac{\text{distance}}{a}\right)$ avec $a = 5$.

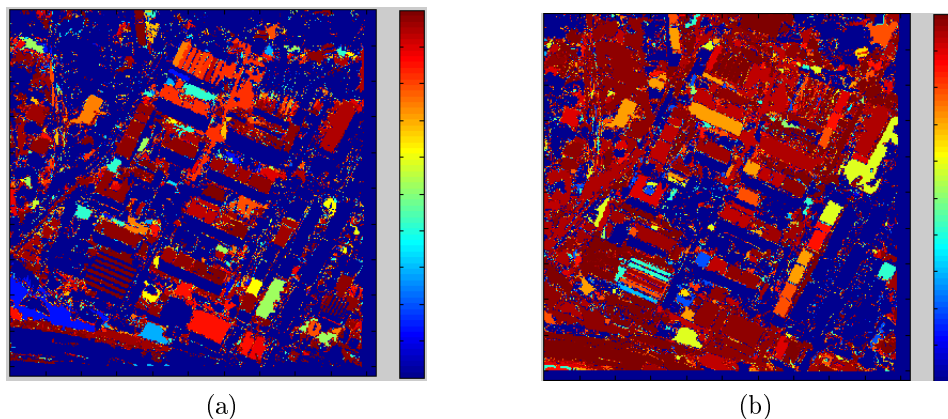


FIGURE 4.9: Carte de score correspondant à : (a) - Ombres optiques. (b) - Ombres radar.

arbres ou les bords de route. Pour l'image radar en revanche, compte tenu du fait que la détection des ombres par seuillage introduit aussi beaucoup d'autres éléments sombres (routes, quelques toits de bâtiments...), il y a beaucoup de fausses alarmes dont on devra tenir compte lors du choix de la règle de fusion.

Néanmoins, les deux cartes de score sont assez complémentaires, ce qui justifie encore une fois l'intérêt de la fusion pour ce genre d'applications.

4.2.2.2 Carte de score pour les zones brillantes

La génération de la carte de score pour les zones brillantes reprend les mêmes étapes que celles décrites précédemment, mais avec quelques modifications (cf. figure 4.10).

En effet, pour ce cas, vu que la position de ces zones brillantes dans la direction de visée du capteur peut être soit devant, soit derrière le bâtiment compte tenu de l'imprécision du recalage (ce dernier cas ne concerne que les lignes brillantes), nous avons calculé deux

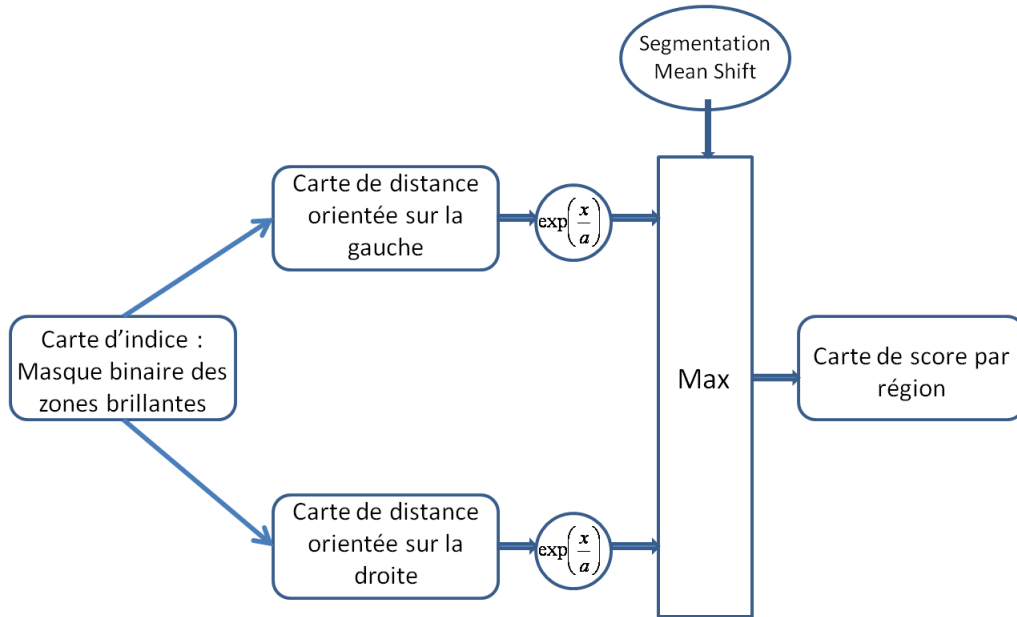


FIGURE 4.10: Schéma synoptique de la procédure de calcul des scores correspondant aux zones brillantes.

cartes de distance orientée, correspondant respectivement aux angles $\alpha_1 = 0^\circ$ et $\alpha_2 = 180^\circ$ (la direction de visée étant l'horizontale sur l'image radar, dans le sens *gauche* $\rightarrow$ *droite*).

Ensuite, après conversion des distances exponentielles en score, pour chaque région segmentée par le Mean Shift, on calcule un “score médian gauche” ($\alpha_1 = 0^\circ$) et un “score médian droit” ($\alpha_2 = 180^\circ$) et nous attribuons à cette région le maximum de ces deux scores. Le résultat obtenu est celui illustré sur la figure 4.11.

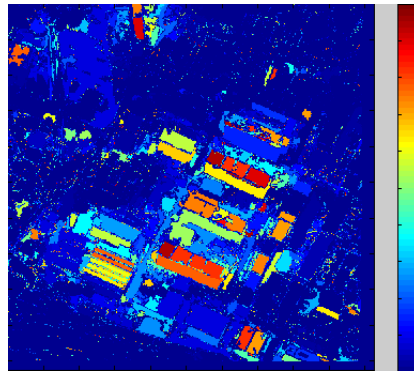


FIGURE 4.11: Carte de score correspondant aux zones brillantes.

Cette carte peut être interprétée de la manière suivante : les régions ayant un score fort correspondent forcément à des bâtiments. En revanche, avoir un score faible n'exclut pas la possibilité que cette région corresponde quand même à un bâtiment. Il faudra par la suite lors du choix des règles de fusion tenir compte de cette particularité.

4.2.2.3 Carte de score pour le NDVI

Comme nous l'avons mentionné auparavant, le NDVI est un indice de végétation codé entre -1 et 1 . Les valeurs proches de 1 indiquent la présence de végétation et inversement.

Avant de calculer la carte de score du NDVI, il faut recoder celui-ci sur 8 bits et lui appliquer une inversion des niveaux de gris. Ensuite, pour chaque région segmentée par le *Mean Shift*, on attribue un score qui correspond, comme pour les cas précédents, à la médiane de tous les scores des pixels appartenant à cette région. Le résultat ainsi obtenu est illustré sur la figure 4.12.

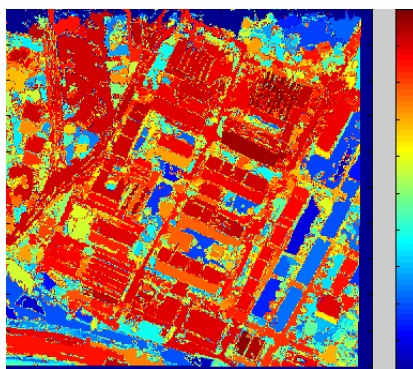


FIGURE 4.12: Carte de score correspondant à l'indice de végétation.

Cette carte indique la présence de végétation (zones de faible score). Cette présence traduit nécessairement l'absence de bâtiment. En revanche, il faut tenir compte de l'existence de zones qui ne soient pas de la végétation mais qui ne sont pas des bâtiments non plus (routes, ponts, parkings...). Il faudra par la suite rajouter cette particularité aux autres citées précédemment et choisir les opérateurs de fusion appropriés.

4.2.3 Fusion des cartes de score

Une fois les cartes de score calculées, il s'agit de les fusionner pour générer une carte de score finale contenant des degrés d'appartenance à la classe bâtiment. Le choix de la règle de fusion est une partie cruciale qui, d'une part nécessite une analyse approfondie des spécificités de chaque indice et des informations qu'on peut en extraire, et d'autre part influe directement sur le résultat final.

Afin de mieux justifier le choix de cette règle, nous allons dans un premier temps rappeler les spécificités de chaque carte de score :

- Carte de score des ombres optique et radar : Ces deux cartes étant complémentaires, nous avons choisi de les fusionner séparément des deux autres. A priori, une somme symétrique devrait convenir à ce cas compte tenu de la variabilité du comportement conjonctif ou disjonctif de cet opérateur en fonction de l'accord ou non des deux cartes.

- Carte de score des zones brillantes : une région ayant un score élevé correspond forcément à un bâtiment, alors qu'un faible score ne suffit pas à conclure à l'absence de ce dernier.
- Carte de score du NDVI : cette carte à un comportement dual de celle des zones brillantes, dans le sens où une région à faible score peut être exclue en tant que bâtiment, mais un score fort n'implique pas nécessairement la présence d'un bâtiment.

Nous avons donc opté pour la chaîne de fusion illustrée sur la figure 4.13.

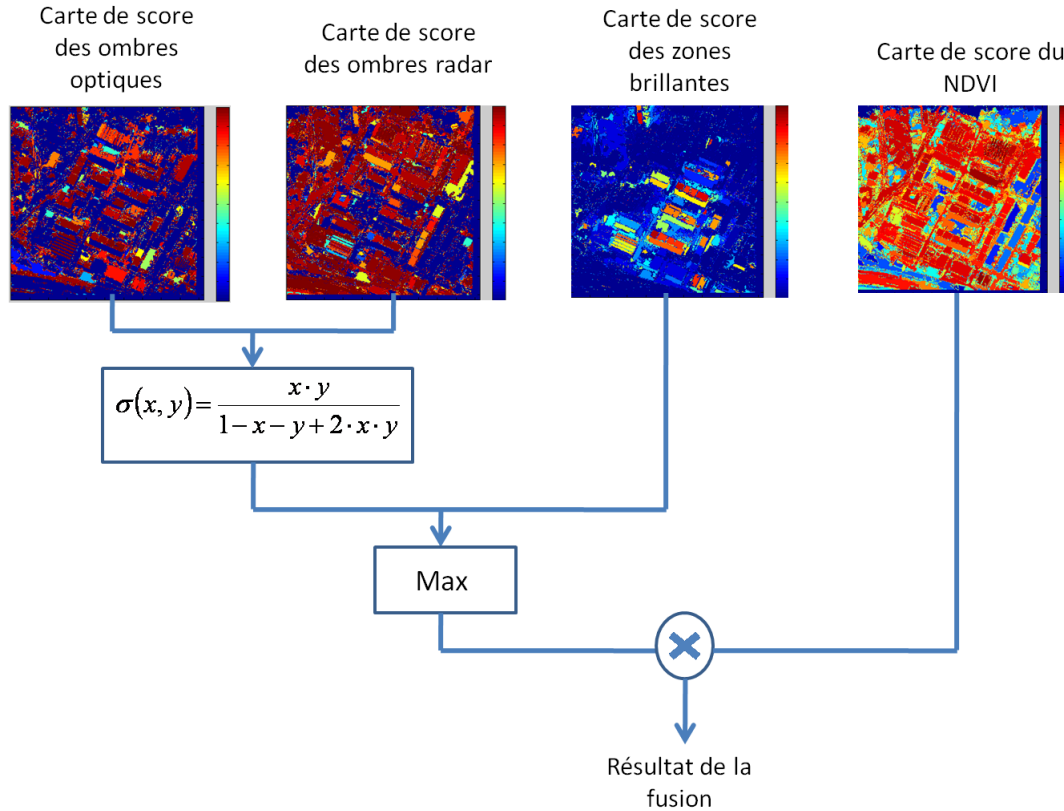


FIGURE 4.13: Chaîne de fusion adoptée.

Les deux cartes de score correspondant aux ombres sont fusionnées par une somme symétrique. Nous avons opté pour l'opérateur $\sigma(x, y) = \frac{x \cdot y}{1 - x - y + 2 \cdot x \cdot y}$ avec des conditions spécifiques pour les cas particuliers, à savoir $\sigma(x, 0) = 0 \forall x$. La sortie de cette première fusion est combinée à la carte de score des zones brillantes en utilisant l'opérateur "max". Ceci vérifie bien les contraintes citées précédemment dans le sens où cette carte sert de validation pour les régions pré-sélectionnées par la première fusion, augmentant ainsi la "sensibilité" (au sens statistique¹) de la méthode de détection. Le NDVI est utilisé à la fin de la chaîne pour pondérer le résultat. Ceci permet de réduire considérablement le taux

1. "Sensibilité" et "spécificité" au sens statistique : la "sensibilité" d'un test est sa capacité à donner un résultat positif lorsque la condition est présente. Elle s'oppose à la "spécificité", qui est la capacité d'un test à donner un résultat négatif lorsque la condition n'est pas présente.

de fausses alarmes améliorant ainsi la “spécificité” de la fusion. Le résultat correspondant à la carte de score finale ainsi obtenue est illustré sur la figure 4.14-(a).

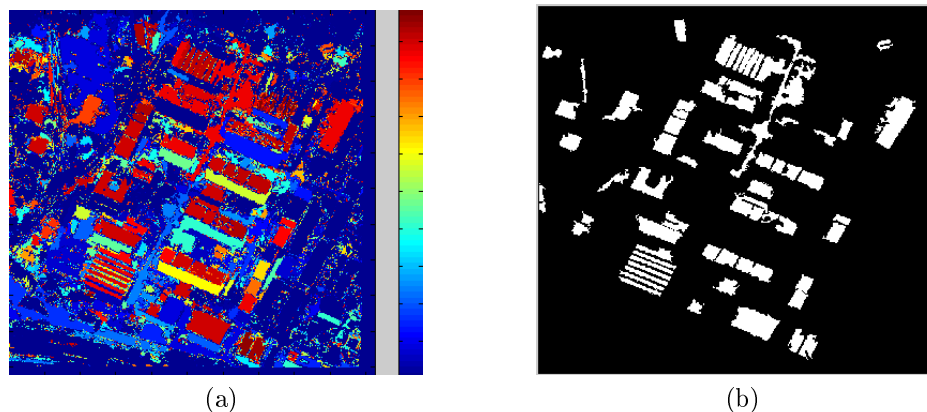


FIGURE 4.14: Résultat de la fusion des cartes de score : (a) - Carte de score finale. (b) - Masque binaire des régions sélectionnées comme étant des bâtiments.

Sur cette carte, les zones de fort score sont bien discriminées par rapport au reste, il est alors facile de les extraire par seuillage. Le seuil étant fixé manuellement mais pourra par la suite être calculé de manière automatique. Une contrainte supplémentaire sur la taille des régions est rajoutée afin d’éliminer celles dont la taille est inférieure à un certain seuil (cf. 4.14-(b)).

Comme prévu, la sur-segmentation introduite par le *Mean Shift* (qui, pour rappel, sert à ce que les régions correspondant à des bâtiments ne débordent pas sur les autres éléments de la scène) fait que le résultat final ne préserve pas la forme géométrique des bâtiments. Une dernière étape consiste donc à rechercher des formes géométriques particulières, en l’occurrence des rectangles, sur le masque binaire de la figure 4.14-(b). Cette partie fait l’objet du paragraphe suivant.

Notons aussi qu’un autre intérêt de l’appariement des régions détectées avec des formes géométriques particulières est la possibilité de sauvegarde des résultats sous forme vectorielle. Ceci est d’autant plus important si on considère que les algorithmes de détection de bâtiments servent en général dans des applications de mise à jour de base de données de scènes urbaines, et que, dans ce genre d’applications, la réduction de la taille des données à stocker est un vrai enjeu actuel.

4.2.4 Algorithme de recherche des rectangles

Afin de retrouver les formes géométriques des bâtiments détectés, nous avons mis au point un algorithme de détection de rectangles. Notons que ceci implique un fort “a priori” sur la forme géométrique des bâtiments dans le sens où on se limite uniquement aux bâtiments rectangulaires. Néanmoins, cette hypothèse pourrait très bien être étendue

par la suite à d'autres formes géométriques particulières (triangles, les formes en T , les formes en $L...$), ce qui pourrait être une étape ultérieure à ce stage.

L'algorithme de détection de rectangles que nous avons mis au point repose sur le principe illustré sur la figure 4.15.

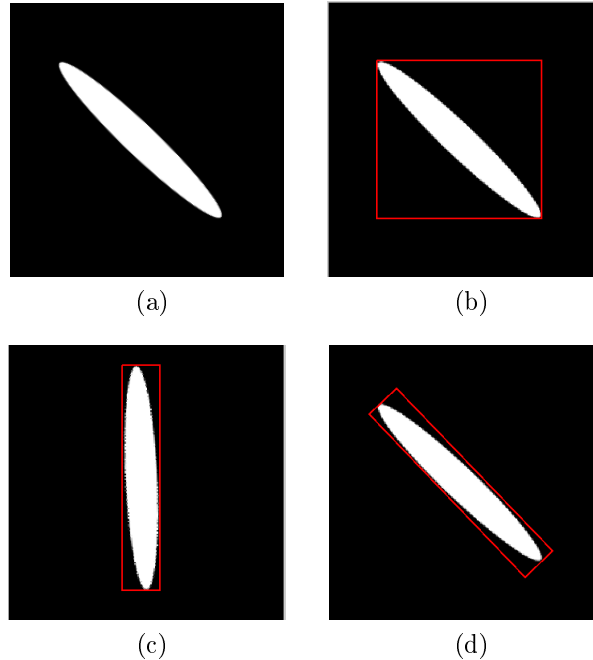


FIGURE 4.15: Illustration du principe de l'algorithme de détection de rectangles : (a) - Exemple d'un objet O . (b) - Rectangle englobant O . (c) - Rectangle englobant l'image de O par une rotation (l'angle de la rotation étant $-\theta$, où θ est l'orientation principale de O). (d) - Résultat de l'application de la rotation inverse.

Soit un objet O ayant une forme géométrique quelconque qu'on veut apparier avec un rectangle, et soit θ l'orientation de cet objet par rapport à l'horizontale. L'idée est que si l'on connaît cette orientation, le rectangle englobant l'objet est simple à retrouver. En effet, il suffit d'appliquer sur cet objet une rotation d'angle $-\theta$ pour qu'il soit orienté suivant l'horizontale. Les quatre coins du rectangle correspondent alors aux extrêma des coordonnées des points de l'objet selon l'horizontale et la verticale.

Puisque dans notre cas l'orientation θ n'est pas connue, il suffit d'effectuer plusieurs rotations en faisant varier l'angle entre 0° et 359° et en gardant l'angle qui minimise le taux de points à l'intérieur du rectangle, mais n'appartenant pas à l'objet. On définit pour cela la notion de "consistance" d'un rectangle R englobant un objet O par :

$$\text{consistance}_O(R) = \frac{\text{Nombre de pixels appartenant à } O}{\text{Nombre de pixels appartenant à } R} \quad (4.2)$$

Ce principe est appliqué sur le masque de la figure 4.14-(b) dont chaque objet connexe a été labellisé et traité à part. Notons que cette procédure de recherche de l'orientation θ

est très coûteuse en temps de calcul. Le résultat de la détection de rectangles est illustré sur la figure 4.16.

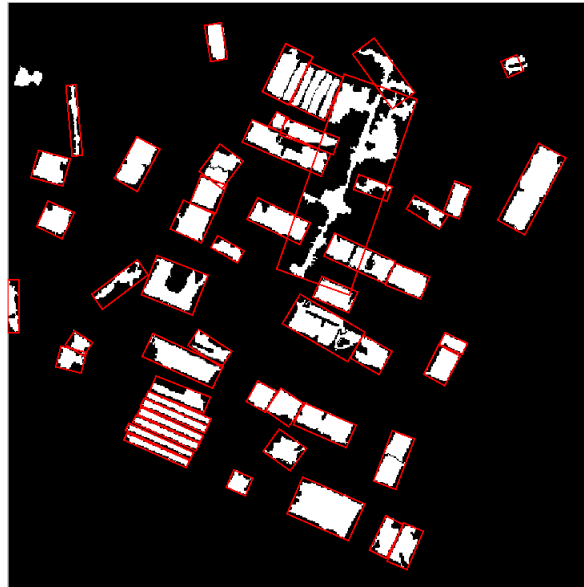


FIGURE 4.16: Résultat de l'algorithme de détection de rectangles appliqué sur le masque binaire de la figure 4.14-(b).

Un seuillage sur la “consistance” permet d’éliminer les objets qui ne peuvent pas être appariés avec des rectangles, et qui correspondent souvent à des fausses alarmes issues de la fusion. Le résultat ainsi obtenu est illustré sur la figure 4.17.

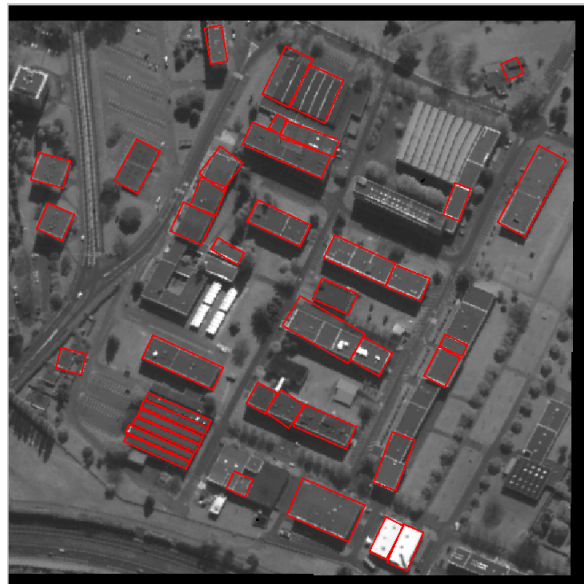


FIGURE 4.17: Régions du masque binaire de la figure 4.14-(b) qui ont pu être appariées avec des rectangles : Superposition sur l'image optique.

On remarque que la sur-segmentation fait qu'un bâtiment est en général divisé en plusieurs objets, et que donc il correspond à plusieurs rectangles adjacents. Une procédure

d'unification des rectangles adjacents a été introduite en post-traitement. L'idée est simple et consiste à construire à partir des rectangles détectés et pré-sélectionnés (cf. figure 4.17) un masque binaire dont les pixels sont mis à 0 s'ils sont à l'intérieur d'un rectangle, et à 1 sinon. Pour décider si un pixel est bien à l'intérieur d'un rectangle, on applique la procédure suivante :

- On sélectionne le rectangle le plus proche du pixel courant en se basant sur sa distance par rapport aux sommets.
- Pour ce rectangle sélectionné, on calcule la somme des distances entre le point considéré et les quatre côtés.
- Le pixel considéré est mis à 1 si cette somme est inférieure à la somme des quatre côtés du rectangle. Par exemple, si on considère le cas illustré sur la figure 4.18, ce point sera mis à 1 si $\sum_{i \in \{1, \dots, 4\}} d_i \leq \sum_{i \in \{1, \dots, 4\}} D_i$.

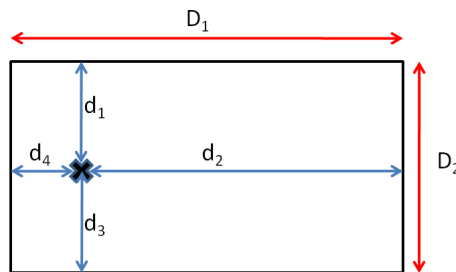


FIGURE 4.18: Illustration du critère de décision sur l'appartenance d'un point à la zone "intérieure du rectangle".

Le résultat ainsi obtenu, et qui constitue le résultat final de la méthode de détection de bâtiments, est illustré sur la figure 4.19.

4.3 Analyse des résultats

Le résultat obtenu (cf. figure 4.19) est assez satisfaisant. En effet, on remarque que, même si certains bâtiments ne sont pas détectés, ceux-ci correspondent en général, soit à des formes géométriques particulières qui ne peuvent pas être appariées avec des rectangles, soit à des zones fortement texturées qui ne sont pas correctement segmentées par le *Mean Shift*.

Une recherche de formes géométriques plus complexes combinée à des méthodes de segmentation plus sophistiquées devraient donc améliorer considérablement les résultats.

Notons aussi que la zone d'étude est particulièrement difficile pour ce genre d'applications vu que les bâtiments sont très peu contrastés par rapport à leur voisinage, les formes géométriques sont assez complexes et les toits sont fortement texturés.

Il est aussi intéressant d'étudier le comportement de cette méthode sur d'autres zones pour que la validation des résultats ne soit pas basée uniquement sur notre zone d'étude.

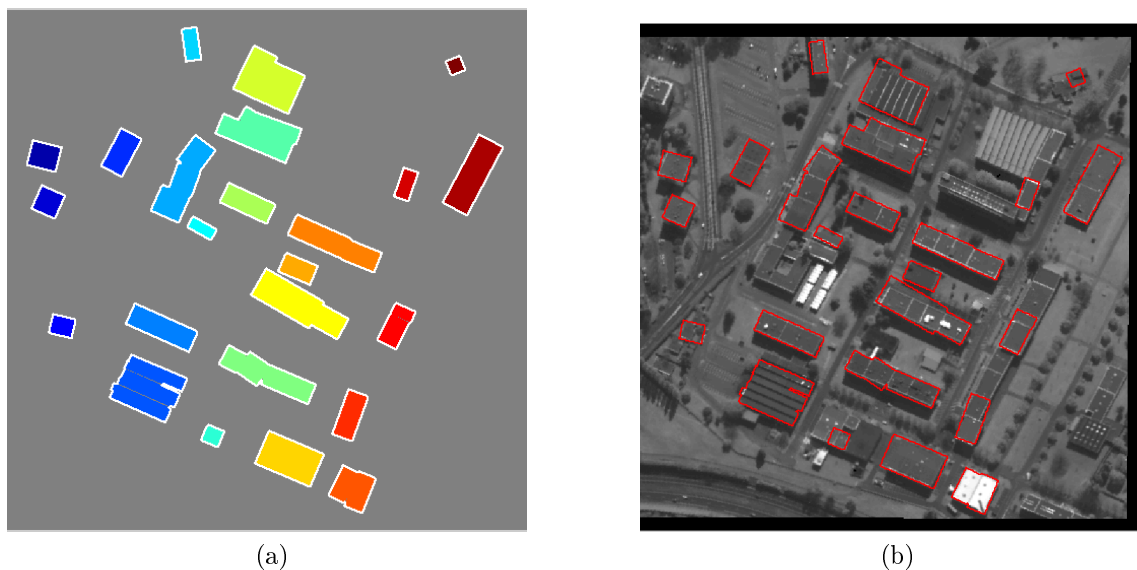


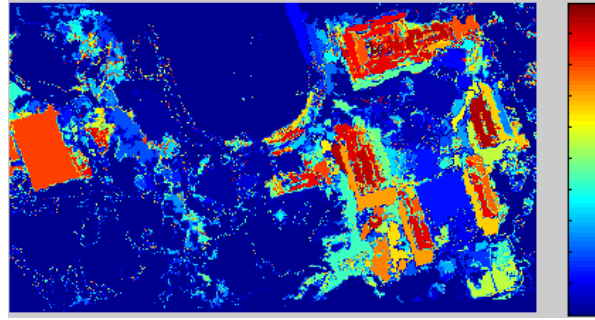
FIGURE 4.19: Résultat final de la détection de bâtiments : (a) - Carte des bâtiments. (b) - Superposition des contours de la carte des bâtiments sur l'image optique.

La figure 4.20 montre un exemple de résultat. Comme pour le cas précédent, la majorité des bâtiments sont détectés, mais leur forme géométrique n'est pas toujours respectée. On observe néanmoins dans cet exemple que l'algorithme arrive à bien détecter un bâtiment dont la forme géométrique est assez complexe (bâtiment en haut à droite de l'image).

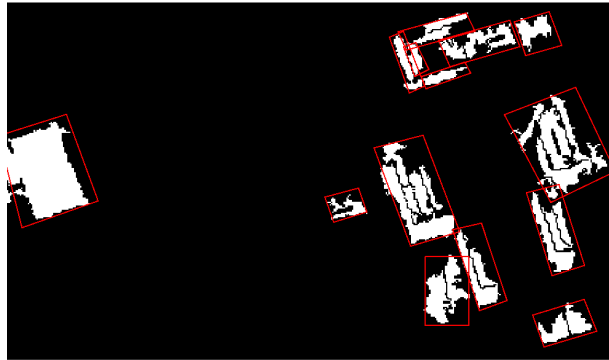
Une étude des performances de cette méthode devrait être envisagée en se basant sur un ensemble de résultats obtenus sur des zones différentes, et sur des scènes urbaines ayant des caractéristiques variées. La comparaison se ferait soit par rapport à une détection manuelle, soit en comparant avec des résultats d'autres méthodes proposées dans la littérature. Ceci permettrait d'étudier le comportement de cette méthode, d'évaluer la précision et le taux des fausses alarmes.

4.4 Conclusion

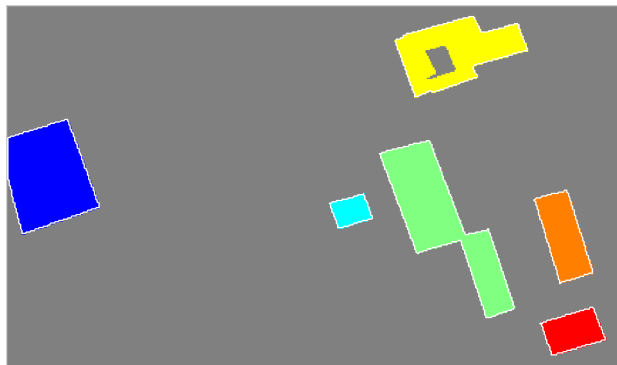
Nous avons présenté, au cours de ce chapitre, une méthode de détection de bâtiments basée sur la fusion des données optiques et radar recalées. Une approche d'attribution de scores par régions, basée sur l'extraction de cartes de distance orientée relatives à des indices, a ainsi été décrite. Les résultats obtenus sont assez satisfaisants surtout si on tient compte de la difficulté liée au choix de la zone. Néanmoins, cette méthode peut encore être améliorée, notamment en remédiant aux erreurs de segmentation introduites par le *Mean Shift*.



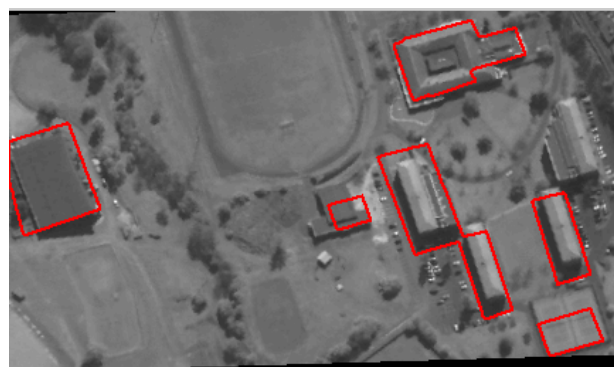
(a)



(b)



(c)



(d)

FIGURE 4.20: Application de la méthode de détection de bâtiments sur une autre zone : (a) - Carte de score finale. (b) - Résultat de l'algorithme de recherche de rectangles. (c) - Carte des bâtiments obtenue. (d) - Superposition de la carte des bâtiments sur l'image optique.

Chapitre 5

Conclusion générale

Dans le cadre de ce stage, nous avons mis au point une chaîne de détection de bâtiments basée sur la fusion d'images optiques et radar, provenant respectivement des capteurs QUICKBIRD et TERRASAR-X.

Une étape de recalage des ces deux données, basée sur les fonctions de localisation implémentées sous OTB, a été présentée. Néanmoins, les erreurs de localisation dues à l'imprécision du MNT et des paramètres capteurs conduisent à un recalage qui n'est pas très précis.

Nous avons, dans un premier temps, essayé de faire varier quelques paramètres de l'acquisition radar afin d'améliorer ces résultats, sans pour autant aboutir à une amélioration satisfaisante. Nous avons donc mis en place, en post-traitement, une méthode de raffinement de ce recalage « grossier » qui consiste à rechercher la similitude 2D qui maximise la mesure de l'information mutuelle entre les deux images. Le résultat ainsi obtenu est satisfaisant dans la mesure où on a une bonne superposition, notamment au niveau des bâtiments, des deux images.

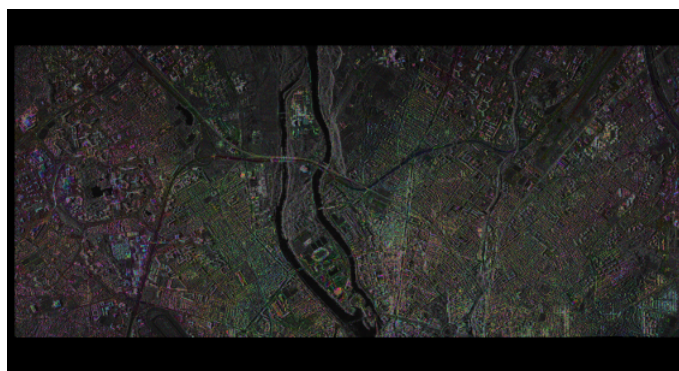
Ces données recalées ont servi par la suite comme entrée pour la méthode de détection de bâtiments. Des cartes d'indices (ombres, zones brillantes et végétation) ont été calculées et exploitées pour extraire des cartes de score par région.

Une règle de fusion, qui prend en compte les spécificités de la signature d'un bâtiment pour chaque indice, a été mise au point, et le résultat de la fusion correspond à un masque binaire des régions pré-sélectionnées comme étant des bâtiments. Ce masque a été enfin apparié avec des rectangles pour aboutir à la carte de bâtiments finale.

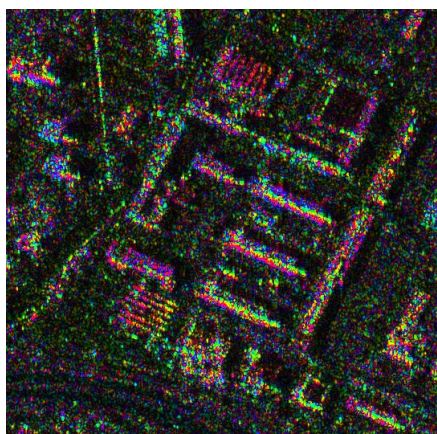
Les perspectives de ce travail sont nombreuses. En effet, on pourrait par exemple introduire d'autres indices dans la règle de fusion pour essayer d'améliorer les résultats. La piste la plus prometteuse semble être la détection des structures linéaires sur l'image optique (par exemple en utilisant la transformée de HOUGH, le détecteur de lignes présenté dans [Grompone 08] ou une simple détection de contours). Ceci permettrait de retrouver plus facilement la forme géométrique des bâtiments qui a été dégradée par la segmentation Mean Shift.

On pourrait aussi étendre l'application à la reconstruction 3D des bâtiments détectés, à travers l'utilisation d'un interférogramme de la zone, d'autant plus qu'on dispose d'un couple interférométrique dont la zone de recouvrement est importante, la base orthogonale est élevée et l'écart temporel est faible.

Dans ce sens, nous avons calculé, durant un court séjour au CNES en utilisant le logiciel DIAPASON, quelques interférogrammes sur la zone étudiée (cf. figure 5.1) afin de retrouver l'information 3D relative aux bâtiments. Néanmoins, la particularité du mode "Spot Light" rend les méthodes classiques de calcul interférométrique insuffisantes [Eineder 08].



(a)



(b)

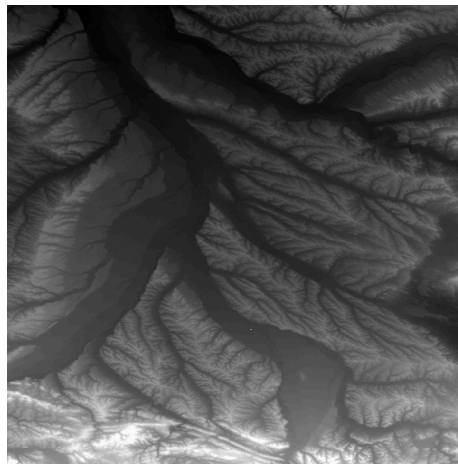
FIGURE 5.1: Interférogramme TerraSar-X correspondant au couple 19/12/2007 et 28/08/2008 (un tour de phase correspond à 24.98m) : (a) - Interférogramme complet. (b) - Zoom sur la zone d'étude.

Par ailleurs le CNES est en train de mettre au point une nouvelle chaîne de calcul intégrée à la nouvelle version de DIAPASON qui devrait résoudre ce problème et permettre d'exploiter les interférogrammes TerraSar-X "Spot Light" .

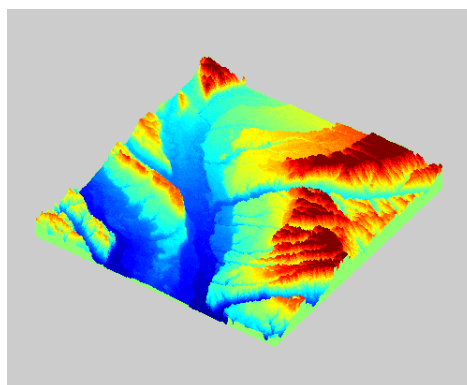
Cette donnée interférométrique pourrait alors être utilisée soit pour améliorer les résultats de la détection de bâtiments (la phase n'étant cohérente que pour les éléments en hauteur, ceci pourrait être exploité comme une information supplémentaire à intégrer dans le processus de fusion), soit pour extraire un modèle 3D des bâtiments détectés.

Annexe A

Données disponibles sur la zone de Toulouse

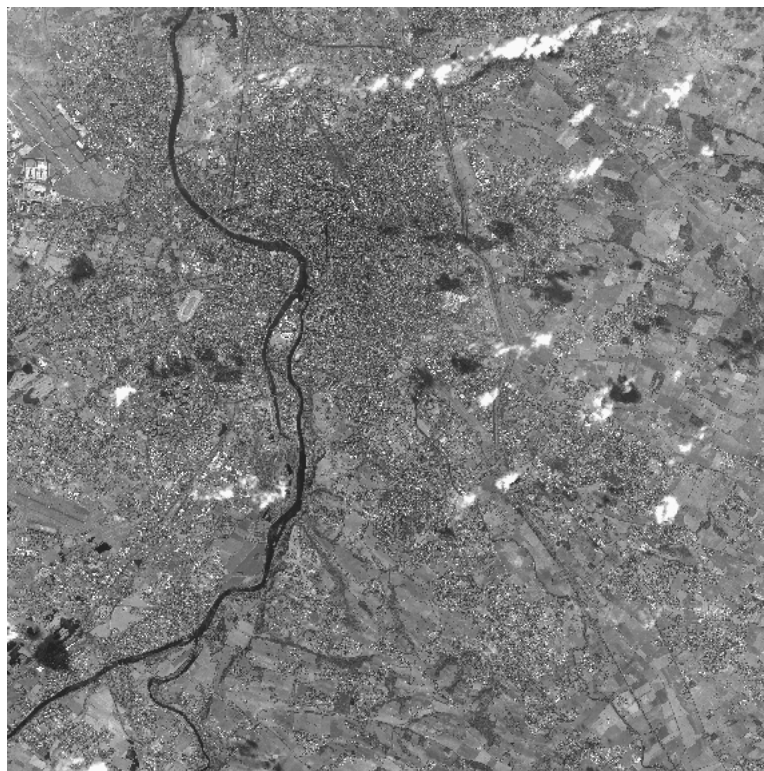


(a)



(b)

FIGURE A.1: Modèle Numérique de Terrain sur la zone de Toulouse : (a) - Visualisation en niveaux de gris. (b) - Visualisation 3D.



(a)



(b)

FIGURE A.2: Données QUICKBIRD du 01/04/2002 sur la zone de Toulouse : (a) - Image Panchromatique. (b) - Image multispectrale (©DigitalGlobe).



(a)



(b)

FIGURE A.3: Données *TerraSar-X* sur la zone de Toulouse : (a) - Image du 27/11/2007. (b) - Image du 19/12/2007 (©DLR/EADS Astrium).

Bibliographie

- [Bloch 99] I. Bloch. *Fuzzy relative position between objects in image processing : a morphological approach*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, no. 7, pages 657–664, 1999.
- [Canny 86] J. Canny. *A computational approach to edge detection*. IEEE Transactions on pattern analysis and machine intelligence, pages 679–698, 1986.
- [Charleux 01] J. Charleux. *Qualité des Modèles Numériques de Terrain pour l'hydrologie, application à la caractérisation du régime de crues des bassins versants*. PhD thesis, Université de marne-la-vallée, France., 2001.
- [CNES 05] CNES. *Recalage d'images optiques et radar par extraction de primitives*. Rapport technique, Etude CNES DCT/SI/AP 04-259, 2005.
- [CNES 07] CNES. The orfeo toolbox software guide, seventh edition, updated for otb-2.8. <http://otb.cnes.fr>, 2007.
- [Comaniciu 02] D. Comaniciu & P. Meer. *Mean shift : A robust approach toward feature space analysis*. IEEE Transactions on pattern analysis and machine intelligence, pages 603–619, 2002.
- [Curlander 82] J.C. Curlander. *Location of spaceborne SAR imagery*. IEEE Transactions on Geoscience and Remote Sensing, pages 359–364, 1982.
- [Deledalle 09] C.A Deledalle, L. Denis & F. Tupin. *Iterative Weighted Maximum Likelihood Denoising with Probabilistic Patch-Based Weights*. IEEE Trans. on Image Processing (to appear), 2009.
- [DigitalGlobe 09] DigitalGlobe. Quickbird imagery products - product guide, revision 5.0, March 2009.
- [DLR 04] DLR. Terrasar-x, ground segment, basic product specification document, March 2004.
- [Eineder 08] M. Eineder, N. Adam, R. Bamler, N. Yague-Martinez & H. Breit. *Spaceborne Spotlight SAR Interferometry With TerraSAR-X*. IEEE Tran-

- sations on Geoscience and Remote Sensing, vol. 47, pages 1524–1535, 2008.
- [Goodman 76] J.W. Goodman. *Some fundamental properties of speckle*. Journal of the optical society of America, vol. 66, no. 11, pages 1145–1150, 1976.
- [Graham 74] L.C. Graham. *Synthetic interferometer radar for topographic mapping*. Proceedings of the IEEE, vol. 62, no. 6, pages 763–768, 1974.
- [Grompone 08] R. Grompone, J. Jakubowicz, J.M. Morel & G. Randall. *On straight line segment detection*. Journal of Mathematical Imaging and Vision, vol. 32, no. 3, pages 313–347, 2008.
- [Inglada 04] J. Inglada & A. Giros. *On the possibility of automatic multisensor image registration*. IEEE Transactions on Geoscience and Remote Sensing, vol. 42, no. 10, pages 2104–2120, 2004.
- [Lehureau 08] G. Lehureau. *Fusion de données optiques et radar à haute résolution en milieu urbain*. Rapport technique, Télécom ParisTech - CNES - Magellium, 2008.
- [Maitre 08] H. Maitre. Processing of synthetic aperture radar images. ISTE Ltd and John Wiley & Sons, Inc, 2008.
- [Muller 05] S. Muller & D.W. Zaum. *Robust building detection in aerial images*. TheInternationalArchivesofPhotogrammetry, Remote Sensing, and Spatial Information Sciences, XXXVI (3/W24), Vienna, Austria, 2005.
- [Pesaresi 01] M. Pesaresi & JA Benediktsson. *A new approach for the morphological segmentation of high-resolution satellite imagery*. IEEE Transactions on geoscience and remote sensing, vol. 39, no. 2, pages 309–320, 2001.
- [Polidori 97] L. Polidori. Cartographie radar. Taylor & Francis, 1997.
- [Poulain 09] V. Poulain, J. Inglada, M. Spigai, J.Y. Tournet & P. Marthon. *Fusion of high-resolution optical and SAR images with vector data bases for change detection*. In IGARSS IEEE International Geoscience & Remote Sensing Symposium, Cape Town, South Africa, July 2009.
- [Simonetto 00] E. Simonetto, H. Oriot & R. Garelo. *Potentiality of High-Resolution SAR Images for Radargrammetric Applications*. In SAR Workshop : CEOS Committee on Earth Observation Satellites ; Working Group on Calibration and Validation, Proceedings of a Conference held 26-29 October 1999, Toulouse, France. Edited by Robert A. Harris and L. Ouwehand. Publisher : Paris : European Space Agency, 2000. ESA-SP, volume 450, 2000.

- [Smets 94] P. Smets. What is Dempster-Shafer's model ? John Wiley & Sons, Inc., New York, USA, 1994.
- [Soergel 07a] U. Soergel, A. Thiele, E. Cadario & U. Thoennesen. *Fusion of High-Resolution InSAR Data and optical Imagery in Scenes with Bridges over water for 3D Visualization and Interpretation*. Proceedings of Urban Remote Sensing Joint Event, vol. 6, 2007.
- [Soergel 07b] U. Soergel, A. Thiele, H. Gross & U. Thoennesen. *Extraction of bridge features from high-resolution InSAR data and optical images*. In Urban Remote Sensing Joint Event, pages 1–6, 2007.
- [Sportouche 09a] H. Sportouche, F. Tupin & L. Denise. *Building detection by fusion of optical and SAR features in metric resolution data*. In IGARSS IEEE International Geoscience & Remote Sensing Symposium, Cape Town, South Africa, July 2009.
- [Sportouche 09b] H. Sportouche, F. Tupin & L. Denise. *Building extraction and 3D reconstruction in urban areas from high-resolution optical and SAR imagery*. In URBAN - IEEE GRSS / ISPRS Joint Workshop on Data Fusion And Remote Sensing over Urban Areas, Shanghai, China, May 2009.
- [Tannous 91] I. Tannous. *Modélisation paramétrique de la géométrie de prise de vue du radar à ouverture synthétique embarqué sur satellite. Application à la superposition d'images radar/optique*. PhD thesis, Université Paris VII, 1991.
- [Tebourbi 05] R. Tebourbi. Cours d'introduction à la télédétection. Ecole supérieure des communications de Tunis, 2005.
- [Tucker 79] C.J. Tucker *et al.* *Red and photographic infrared linear combinations for monitoring vegetation*. Remote Sensing of Environment, vol. 8, no. 2, pages 127–150, 1979.
- [Tupin 98] F. Tupin, H. Maitre, J.F. Mangin, J.M. Nicolas & E. Pechersky. *Detection of linear features in SAR images : application to roadnetwork extraction*. IEEE Transactions on Geoscience and Remote Sensing, vol. 36, no. 2, pages 434–453, 1998.
- [Tupin 03] F. Tupin & M. Roux. *Detection of building outlines based on the fusion of SAR and optical features*. ISPRS Journal of Photogrammetry and Remote Sensing, vol. 58, no. 1-2, pages 71–82, 2003.

- [Viola 97] P. Viola & W.M. Wells III. *Alignment by maximization of mutual information*. International Journal of Computer Vision, vol. 24, no. 2, pages 137–154, 1997.
- [Wegner 09a] J.D. Wegner, S. Auer, A. Thiele & U. Soergel. *Analysis of urban areas combining high-resolution optical and SAR imagery*. In 29th European Association of Remote Sensing Laboratories Symposium, Chania, Greece, June 2009.
- [Wegner 09b] J.D. Wegner, U. Soergel & A. Thiele. *Building extraction in urban scenes from high-resolution InSAR data and optical imagery*. In Urban Remote Sensing Joint Event, 2009.
- [Zebker 86] H.A. Zebker & R.M. Goldstein. *Topographic mapping from interferometer synthetic aperture radar observations*. Journal of Geophysical Research, vol. 91, no. B5, 1986.