

Un premier pas vers l'extraction de MNS urbains en interférométrie ROS à haute résolution par fusion de détecteurs

Céline Tison^{(a)(b)}, Florence Tupin^(a), Henri Maître^(a)

(a) GET-Télécom Paris UMR 5141
46 rue Barrault
75 013 Paris

Tel : 01 45 81 70 73 – FAX : 01 45 81 37 94

Email : celine.tison@enst.fr, florence.tupin@enst.fr, henri.maître@enst.fr

(b) CNES

Département SI/AR
18 avenue Edouard Belin
31 401 Toulouse Cedex 4

Résumé

Le calcul du MNS (Modèle Numérique de Surface) de villes est un axe majeur de recherche auquel l'imagerie radar en mode interférométrique peut désormais apporter des éléments de réponse, compte tenu de l'amélioration récente de ses résolutions.

Dans cet article, nous proposons une chaîne originale de traitements pour calculer le MNS à partir d'un unique couple interférométrique ROS seulement. Cette méthode globale repose sur l'optimisation conjointe de la hauteur et de la classe de structures détectées à partir d'un champ de Markov défini sur un graphe de régions. Ce dernier est le résultat de la superposition de plusieurs détecteurs (routes, ombres, etc.) qui extraient des primitives de la scène à partir de l'amplitude, la cohérence et l'interférogramme.

Après avoir rappelé le principe de l'interférométrie ROS et ses limites en milieu urbain, nous proposons une analyse des propriétés des images ROS à haute résolution afin de déterminer l'information pertinente qu'elles offrent. Ensuite la méthodologie retenue, comprenant la détection de structures et l'optimisation conjointe classe/hauteur, est décrite. Enfin une analyse de résultats obtenus sur un large ensemble de données réelles conclut cet article.

Mots Clefs : Interférométrie, ROS, milieu urbain, haute résolution, MNS

Abstract

This article addresses the problem of elevation map extraction from high resolution SAR interferograms. Thanks to the improvement of SAR image resolution, it is now possible to study urban areas with such data. As we want to evaluate how precise a DSM (Digital Surface Model) computed from a unique interferometric SAR pair can be, no external data (such as multi incidence images, multi track data or optical image) are added.

Much information can be retrieved from the amplitude image, the interferogram and the coherence image: for instance, road map, building map, and shadow map... We propose different detectors in order to get as many pieces of information as possible to describe the scene. The superimposition of all the detectors defines a region graph processed in a Markovian algorithm. Indeed an optimization is made jointly on the heights and the classes of the region graph.

The complete process has been run on real data. Critical comments of the results help to analyze the inner limits of interferograms and of the algorithm.

Keywords: Interferometry, SAR, urban area, high resolution, MNS

1. Introduction

La moitié de la population mondiale habite désormais dans les villes, ce qui place ces dernières au cœur d'enjeux géopolitiques, économiques et écologiques importants. Ceci renforce le besoin, déjà ancien, de suivre l'évolution des villes via la cartographie. De plus en plus, la troisième dimension est devenue indispensable à l'urbanisme, aux applications militaires ou à la prévention de catastrophes naturelles. L'imagerie optique, grâce à une très bonne résolution spatiale, a été la première à répondre à cette attente par la stéréoscopie. Mais

depuis quelques années, l'imagerie ROS propose des résolutions suffisamment hautes pour étudier l'urbain, ce qui présente des avantages certains de disponibilité car l'imagerie radar n'est pas affectée par les conditions météorologiques.

1.1. L'interférométrie SAR et le milieu urbain

L'imagerie radar est une mesure cohérente complexe (i.e. en phase et en amplitude) de l'onde électromagnétique rétrodiffusée par la scène illuminée. L'interférométrie SAR [Maître et al., 2001] exploite l'information contenue dans la phase de l'image (qui est directement liée au trajet parcouru par

l'onde électromagnétique, donc à la géométrie de la scène) pour calculer l'information de hauteur en chaque pixel. L'interférométrie a prouvé son potentiel sur les zones naturelles à basse résolution, par exemple pour la modélisation tridimensionnelle de l'Etna, mais elle a été peu appliquée au milieu urbain. Plusieurs raisons expliquent cela. Tout d'abord les objets de la ville ne sont accessibles que si l'on dispose d'images à très haute résolution. Par ailleurs l'image radar en milieu urbain est très complexe en raison des réflexions multiples et des recouvrements. Les interférogrammes sont alors très difficiles à dépouiller d'autant que les hypothèses classiquement invoquées pour filtrer la phase (la continuité des hauteurs, l'absence de réflexions multiples) sont en défaut. Enfin, plus la résolution est bonne, plus la nature physique des images doit être prise en compte et plus leurs propriétés diffèrent de celles des images à basse résolution (au niveau statistique par exemple).

Pour s'affranchir du problème de déroulement de phase (levée de l'ambiguïté modulo 2π), l'altitude d'ambiguïté est fixée de façon à ce qu'elle soit supérieure au dénivelé maximal (bâtiments compris) de la scène. Malgré cela, les interférogrammes ne peuvent pas être simplement inversés pour calculer le MNS. Les zones d'ombre et de recouvrement (dues à la géométrie radar), le bruit de l'interférogramme rendent la carte des hauteurs inexploitable (Figure 1 – planche couleur) et il est indispensable de proposer un ensemble de traitements de plus haut niveau qui prendra en compte la cohérence et l'amplitude.

Plusieurs approches ont été proposées pour surmonter ces problèmes. Elles se regroupent en quatre grandes familles, qui sont fondées uniquement sur des méthodes de traitement des images : le « shape from shadow » [Bolter, 2000], la recherche de plans dans le nuage de points 3D obtenus à partir de l'interférogramme [Gamba et al., 2000], la géométrie stochastique [Quartulli et al., 2003] et la reconstruction 3D à partir de zones pré-segmentées [Soergel et al., 2003 ; Petit, 2004]. Dans [Bolter, 2000], l'empreinte des bâtiments est calculée à partir des ombres détectées dans au moins deux (idéalement quatre) passes complémentaires. La hauteur des bâtiments est la moyenne de la hauteur interférométrique sur l'empreinte. Mais il est rare d'avoir autant d'informations et dans de nombreux cas, il faut se contenter d'un unique couple interférométrique. La deuxième méthode [Gamba et al., 2000] est une modélisation par des plans horizontaux de la carte 3D, issue de l'interférogramme. L'optimisation est itérative. De bons résultats sont obtenus sur les bâtiments isolés de grande taille.

Les deux autres méthodes sont plus adaptées à notre contexte car nous souhaitons reconstruire en trois dimensions une ville à partir d'un unique couple

interférométrique. Le principe de la géométrie stochastique [Quartulli et al., 2003] est d'optimiser la position et la taille d'un ou plusieurs modèles de bâtiment en fonction de l'amplitude, la cohérence et l'interférogramme. C'est très prometteur mais cela impose de se restreindre à un petit nombre de modèles de bâtiments pour avoir des temps de calcul raisonnables. Pour cette raison, nous avons choisi de nous diriger vers la dernière approche. En effet, il semble très logique de vouloir d'abord segmenter avant de reconstruire car les attentes 3D sont différentes selon les objets : variations très basses fréquences pour le sol, contour régulier pour les bâtiments, forme sans a priori pour la végétation... Soergel et al. [Soergel et al., 2003] proposent un processus de reconstruction très abouti. L'image d'amplitude est d'abord segmentée. Ensuite les bâtiments sont représentés par des rectangles en introduisant des connaissances a priori sur la géométrie des villes (alignement avec les rues). La segmentation initiale est enrichie par l'utilisation de plusieurs passes ou plusieurs incidences. Un premier MNS est déduit ; il est validé et remis en question en calculant les ombres et zones de recouvrement qu'il implique et en les comparant aux images initiales. Les résultats obtenus sont de bonne qualité dans le cas d'une base de données assez riche.

1.2. Méthode proposée

L'idée générale de segmenter avant de reconstruire est reprise dans notre travail mais nous n'introduisons aucune information autre que celle du couple interférométrique. De plus, nous ne voulons pas contraindre la forme des bâtiments, contrairement à ce qui est proposé dans [Soergel et al., 2003 ; Quartulli et al., 2003]. Notre chaîne est basée sur la fusion de détecteurs, considérés tous au même niveau, afin d'optimiser conjointement la hauteur et la classe d'une région, ce qui est sa principale originalité. Cette approche repose sur l'optimisation d'une énergie markovienne définie sur un graphe de régions. Ce dernier est déterminé par l'intersection de tous les résultats des détecteurs sont : une classification, une carte des coins réflecteurs, l'interférogramme régulé, une carte des ombres, une carte des bâtiments estimés à partir des ombres et une carte des routes.

1.3. Les données

Nous disposons d'un jeu de données interférométriques à haute-résolution sur la ville de Dunkerque contenant à la fois des zones portuaires (industrielles) et des zones urbaines. Ce sont des images acquises par le capteur RAMSES de l'ONERA.

Dans cet article, le principe de l'interférométrie est brièvement rappelé (§ 2.), puis une analyse du contenu des images ROS en milieu urbain est faite. De cette analyse découle le choix et les méthodes de

calcul des détecteurs (§ 4.). L'optimisation de la hauteur et de la classe est également explicitée dans ce paragraphe. L'article se termine par le commentaire de résultats obtenus sur trois zones représentant toute la diversité du milieu urbain (§ 5.).

2. Principe général de l'interférométrie

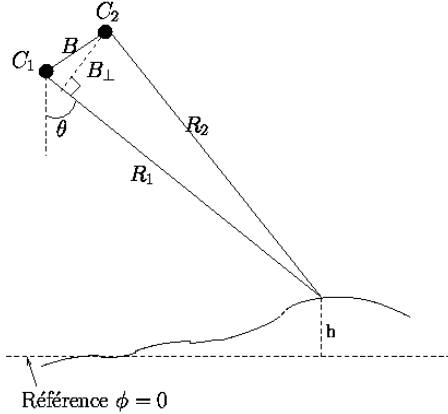


Figure 2 : Configuration géométrique de l'interférométrie

L'interférogramme est l'image des différences de phases de deux images ROS, recalées, sur la même scène. Les deux capteurs (ou le même à deux dates, dans le cas multi-passes) observent la scène avec des angles d'incidence très légèrement différents, ce qui assure un bruit de phase (speckle) identique et une différence de phase non nulle. La configuration géométrique est celle de la Figure 2 dont nous utiliserons les notations.

Le produit hermitien des deux signaux radar complexes s_i est calculé sur une fenêtre L . Sa phase ϕ est la phase interférométrique et son module ρ est la cohérence, indicatrice de la qualité de ϕ :

$$\rho e^{j\phi} = \frac{\sum_{i=1}^L s_1(i) s_2^*(i)}{\sqrt{\sum_{i=1}^L |s_1(i)|^2 \sum_{i=1}^L |s_2(i)|^2}} \quad (1)$$

La phase interférométrique est liée à la différence de chemin optique ($R_2 - R_1$) par :

$$\phi = \frac{2\pi p}{\lambda} (R_2 - R_1) \quad (2)$$

où λ est la longueur d'onde, $p=1$ dans le cas mono-statique (un émetteur et deux récepteurs) et $p=2$ dans le cas bi-statique ou mono-statique ping-pong.

En développant l'Equation (2) au premier ordre, on fait apparaître les contributions dites orbitales ϕ_{orb} et topographiques ϕ_{topo} à la phase [Maître et al., 2001] :

$$\phi = \phi_{orb} + \phi_{topo} \quad (3)$$

$$\phi_{orb} = \frac{2\pi p}{\lambda} \frac{B_{\perp} \Delta R}{R \tan \theta} \quad (4)$$

$$\phi_{topo} = \frac{2\pi p}{\lambda} \frac{B_{\perp} h}{R \sin \theta} \quad (5)$$

ϕ_{orb} , liée à l'éloignement du point au sol par rapport au capteur, se calcule à partir des paramètres physiques et géométriques des capteurs. En chaque point, la seule inconnue est donc ϕ_{topo} ou plutôt l'altitude h dont elle est fonction. Par la suite, nous supposons que l'interférogramme est l'image de la phase topographique (le terme orbital étant retiré au préalable).

L'altitude d'ambiguïté h_{amb} est l'altitude pour laquelle la phase interférométrique a tourné de 2π :

$$h_{amb} = \frac{\lambda R \sin \theta}{p B_{\perp}} \quad (6)$$

Dans le cas de notre jeu de données, h_{amb} est de l'ordre de 200 mètres ce qui signifie qu'un écart de 1m correspond à 0.04 radians. Or le bruit des interférogrammes disponibles est supérieur à 0.04 : il ne sera pas possible d'obtenir une précision métrique.

3. Propriétés des images ROS à haute résolution en milieu urbain

Dans le cas du milieu naturel à basse résolution, les propriétés des images radar sont bien modélisées [Maître et al., 2001]. En revanche, la haute résolution et le milieu urbain restent un domaine d'exploration récent. Nous reprenons dans ce paragraphe des résultats qui nous permettront de décrire les objets (bâtiments, sol, végétation) dans les différents détecteurs.

3.1. Haute résolution et urbain

La Figure 3 présente l'amplitude ROS, la cohérence et l'interférogramme d'un extrait contenant bâtiments, végétation et route. Les trois classes peuvent être distinguées dans l'image d'amplitude : les zones ont des lois statistiques différentes. De plus, les zones d'ombres, qui ont des amplitudes faibles mais proches de celles de la route, apparaissent nettement dans la cohérence. Elles ont en effet une cohérence très faible car il n'y a pas ou très peu de signal de rétrodiffusion. Les coins réflecteurs mur/sol constituent des dièdres qui rétrodiffusent fortement, même lorsqu'ils ne sont pas orientés parallèlement à la trace en raison de la rugosité du sol et/ou du mur [Franceschetti et al., 2002].

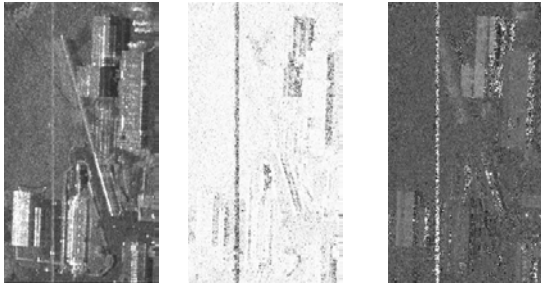


Figure 3 : Amplitude, cohérence et interférogramme (de gauche à droite) en 6 vues sur un extrait du quartier Bayard (images RAMSES, Dunkerque)

En résumé, les bâtiments ont des toits qui sont détectables dans l'image d'amplitude pour la plupart ; ils sont suivis d'une zone d'ombre qui est peu cohérente ; ils sont souvent associés à une réponse linéaire forte associée au coin réflecteur mur/sol.

3.2. Limites intrinsèques aux images

L'analyse des images met à jour des limites intrinsèques :

- l'ambiguïté des zones d'ombre et de recouvrement, qui découlent de la géométrie radar, ne pourra pas être levée avec un seul couple interférométrique,
- la résolution spatiale est insuffisante pour reconstruire les zones pavillonnaires,
- l'altitude d'ambiguïté est trop grande pour avoir une bonne résolution altimétrique et, surtout, distinguer les objets faiblement élevés (par exemple les maisons à un étage).

Nous ne pourrions pas nous affranchir de ces problèmes et ils limiteront la qualité des résultats. Néanmoins les deux derniers sont liés à la base de données dont nous disposons et un autre jeu de données pourrait ne pas connaître de tels problèmes. L'utilisation d'images acquises selon plusieurs traces permet de récupérer de l'information sur l'ensemble de l'image et donc de s'affranchir des zones d'ombre et de recouvrement.

4. Méthodologie

Ce paragraphe expose la méthode employée pour calculer une classification et un MNS de façon conjointe.

4.1. Idée générale

Plusieurs informations complémentaires peuvent être extraites des images d'amplitude (mono-vue), de phase interférométrique et de cohérence. Par exemple, les ombres, les routes, les coins réflecteurs, etc. sont identifiables. Tous ces renseignements se complètent pour construire une classification. La classe est également liée à la hauteur.

Six traitements sont faits sur les images initiales à l'aide de six détecteurs complémentaires pour obtenir six nouvelles images : l'interférogramme filtré, une première classification, une carte des coins réflecteurs, une carte des routes, une carte des ombres et une carte des bâtiments reconstruits à partir des ombres. Toutes sont calculées à partir de l'amplitude, de l'interférogramme, de la cohérence et, éventuellement, de l'un d'entre eux uniquement. Ces résultats sont ensuite fusionnés pour optimiser conjointement la hauteur et la classe de chaque pixel. Ces deux informations sont effectivement liées (un arbre est un objet au-dessus du sol par exemple). Le schéma de la Figure 4 résume le principe général.

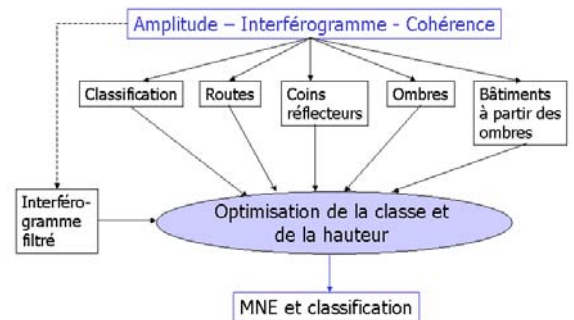


Figure 4 : Schéma général de l'optimisation conjointe de la hauteur et de la classe

L'originalité de la méthode proposée réside dans l'utilisation d'un unique couple interférométrique, l'absence de contrainte sur la forme des bâtiments et l'optimisation conjointe de la hauteur et de la classe. Notons également que le choix des détecteurs peut être remis en question (à la fois du point de vue du nombre et du point de vue de la signification) sans pour autant modifier l'algorithme général. Cette flexibilité peut être très appréciable pour faire évoluer la méthode et pour traiter des bases de données de nature très différente.

Dans la section 4.2., nous décrivons les algorithmes utilisés pour chaque détecteur.

4.2. Les détecteurs

Les détecteurs ont pour but d'extraire des informations structurelles des images. Certains ont été repris de la littérature et légèrement adaptés (détection des routes, des ombres et des coins réflecteurs), d'autres ont été créés spécialement pour ce travail (classification, filtrage, calcul des bâtiments à partir des ombres) et forment des algorithmes originaux par eux-mêmes.

4.2.1. Classification

Elle se fait en deux temps :

- classification à partir de l'amplitude uniquement [Tison et al. 2004 a],

- fusion du résultat précédent avec la cohérence et l'interférogramme [Descombes, 1995].

La classification de l'amplitude est une classification markovienne où les classes sont modélisées par des lois de Fisher (§ 3.1.2.). En effet, dans le cas des images à haute résolution en milieu urbain, il n'est plus possible de faire l'hypothèse de chatoiement développé : le nombre de diffuseurs par cellule est insuffisant, il y a de forts diffuseurs. Il est nécessaire d'utiliser un modèle adapté à ce type d'images et suffisamment souple pour modéliser tous les types de surfaces du milieu urbain. Il a été montré dans [Tison et al., 2004 a] que les lois de Fisher estimées avec la méthode des log-moments répondent parfaitement au problème. L'apprentissage des paramètres des lois est supervisé.

La deuxième étape permet de corriger certaines incertitudes et, surtout, d'introduire les ombres, dont la principale caractéristique est d'avoir une cohérence faible. La fusion est faite par un champ de Markov [Descombes, 1995], les potentiels d'énergie sont décrits par l'utilisateur en fonction de sa connaissance de la scène. L'algorithme est robuste vis-à-vis des paramètres et ces derniers sont les mêmes pour toutes les images d'une même base de données.

La Figure 5 (planche couleur) présente les résultats pour deux étapes, toujours sur la zone de la Figure 3. Tous les bâtiments sont bien détectés et leurs contours sont assez nets. Il existe des confusions au niveau d'arbres placés à côté d'un mur au centre de l'image.

4.2.2. Coins réflecteurs

Les coins réflecteurs sont associés à des lignes très brillantes dans l'image d'amplitude. Ils sont détectés de la façon suivante :

- détection des lignes brillantes par l'algorithme proposé dans [Tupin et al., 1998],
- parmi les points sélectionnés, seuls les pixels au sol (seuillage sur l'interférogramme) et très brillants (seuillage sur l'amplitude) sont conservés, ce qui suppose la définition préalable de deux seuils th_1 et th_2 sur l'amplitude et sur la phase,
- pour chaque zone de bâtiments connexes suffisamment grande (issue de la classification), recherche des meilleures lignes passant par les pixels retenus (transformée de Radon locale).

Les coins réflecteurs les plus importants sont bien détectés (Figure 6). Par contre, dans le cas de bâtiments de géométrie complexe entraînant des coins très petits, la détection n'est pas parfaite. La transformée de Radon peut introduire quelques erreurs en regroupant des pixels appartenant à des

coins proches mais cependant distincts. Ce cas reste heureusement rare. Le choix des deux seuils th_1 et th_2 est dépendant du type de zones traitées.

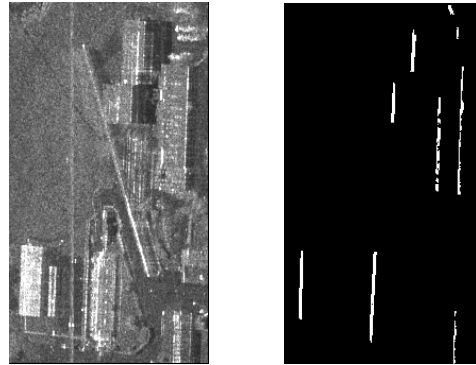


Figure 6 : Résultat de la détection des coins réflecteurs : à gauche image d'amplitude, à droite carte des coins réflecteurs. (images RAMSES, Dunkerque)

4.2.3. Régularisation de l'interférogramme

L'interférogramme multi-vue (ici 6 vues) reste bruité. La phase observée $\bar{\phi}$ s'exprime, en fonction de la phase réelle ϕ et du bruit n :

$$\bar{\phi} = \phi + n \quad (7)$$

Nous avons choisi une régularisation markovienne avec un processus bord [Geman et al., 1992] qui préserve bien les discontinuités du milieu urbain. La probabilité conditionnelle d'avoir une valeur ϕ_s en un site s connaissant $\bar{\phi}_s$ peut alors

s'écrire sous la forme d'une probabilité de Gibbs [Geman et al., 1984], ϕ_s est obtenu en minimisant l'énergie U composée d'un terme d'attache aux données U_{dat}^s et d'un terme de régularisation U_{reg}^s :

$$U(\phi_s | \bar{\phi}_s, \phi_t, t \in V_s, CR) = (1-\beta) U_{dat}^s(\phi_s | \bar{\phi}_s) + \beta \sum_{t \in V_s} U_{reg}^s(\phi_s | \phi_t, t \in V_s, CR) \quad (8)$$

CR est la carte des coins réflecteurs, V_s est l'ensemble des voisins (ici en 8-connexité) du site s . Les coins réflecteurs sont introduits indépendamment du processus bord implicite lié à la fonction de régularisation afin d'assurer que ces structures fines soient parfaitement préservées.

Le terme d'attache aux données au site s est une pénalité quadratique :

$$U_{dat}^s(\phi_s | \bar{\phi}_s) = \frac{(\phi_s - \bar{\phi}_s)^2}{2\pi\sigma_\phi^2} \quad (9)$$

avec l'écart type des phases σ_ϕ estimé par la borne de Cramer-Rao [Maître et al., 2001] :

$$\sigma_\phi = \frac{1 - |\rho|^2}{2L|\rho|^2} \quad (10)$$

L est le nombre de vues et ρ la cohérence interférométrique au pixel s.

Le terme contextuel, au site s, est égal à :

$$U_{reg}^s(\phi_s | \phi_t, t \in V_s, CR) = N_{vois} \sum_k \gamma_{V_s} \Psi \left[\left(D_C^n \phi \right)_k \right] \quad (11)$$

N_{vois} est le nombre de pixels dans le voisinage (8 en 8-connextité), γ_{V_s} est égal à 0 si s et t sont l'un un pixel de coin réflecteur et l'autre non, égal à 1 sinon. D_C^n est une approximation discrète du gradient (n=1) et du laplacien (n=2).

Dans la littérature, la fonction de régularisation Ψ peut prendre plusieurs formes selon le degré de lissage attendu. La forme proposée dans [Geman et al., 1992] a l'avantage d'autoriser des fonctions ϕ à variations linéaires, ce qui peut permettre de préserver les pentes des toits ou de terrain. La forme de la fonction Ψ est :

$$\Psi(u) = - \frac{1}{1 + \left| \frac{u}{\Delta} \right|} \quad (12)$$

Δ est un paramètre qui permet de décrire la pente moyenne attendue. Deux passes de filtrage sont

faites : l'une avec $u = D_C^2(\phi)$ (pour les pentes) et

l'autre avec $u = D_C^1(\phi)$ (pour les surfaces planes). Δ

est d'abord égal à 3 puis à 15.

Plusieurs paramètres Δ et β ont été testés. Les résultats sont robustes vis-à-vis de petites variations.

L'image initiale constitue une bonne initialisation. Donc un algorithme ICM converge assez rapidement vers une solution acceptable. La Figure 7 montre le résultat du filtrage sur la zone de Bayard. L'effet du filtrage est également présenté sur les profils de la Figure 8.

4.2.4. Ombres

Les ombres sont détectées par une segmentation par champs de Markov de l'image d'amplitude avec un grand nombre de classes (15 par exemple). La classe associée à la plus faible amplitude est sélectionnée et des opérations morphologiques sont faites pour retirer les fines structures qui sont principalement associées à la route (érosion).

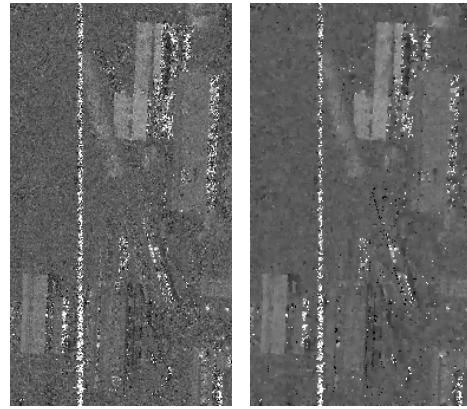


Figure 7 : Interférogramme original (gauche) et filtré (droite) avec $\beta = 0.4$. L'interférogramme filtré est nettement plus lisse mais les bords des bâtiments sont bien préservés ainsi que les coins réflecteurs (images RAMSES, Dunkerque).

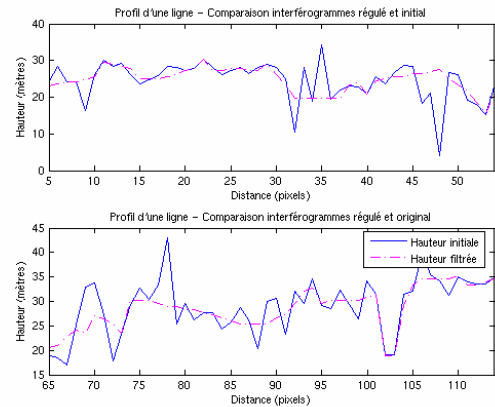


Figure 8 : Profils des interférogrammes filtré et initial pour deux lignes de la Figure 7.



Figure 9 : résultat de la détection des ombres sur l'extrait du quartier Bayard (images RAMSES, Dunkerque).

Les ombres détectées reprennent bien l'ensemble des ombres de l'image (Figures 9 et 13). Notons cependant quelques erreurs au niveau des routes où il est difficile de distinguer la réponse du bitume de celle de l'ombre.

4.2.5. Bâtiments d'après les ombres

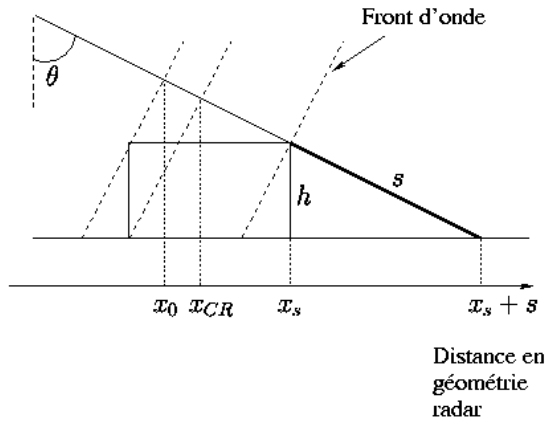


Figure 10 : Configuration géométrique pour le calcul de la position x_0 d'un bâtiment à partir de son ombre s .

Bâtiments isolés et ombres sont reliés par la géométrie de l'acquisition (Figure 10). La taille s de l'ombre est proportionnelle à la hauteur h du bâtiment pour un angle d'incidence θ donné :

$$h = s \cos \theta \quad (13)$$

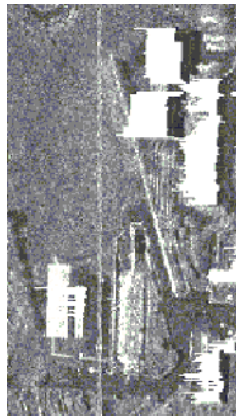


Figure 11 : superposition des bâtiments détectés (en blanc) sur l'image d'amplitude (images RAMSES, Dunkerque).

Dans ce module, nous exploitons ces propriétés afin d'estimer la position d'un bâtiment en fonction de son ombre. Nous faisons l'hypothèse que toutes les ombres détectées au paragraphe précédent sont issues d'un bâtiment et nous nous appuyons sur cette ombre pour détecter ce bâtiment par une classification contextuelle ligne-à-ligne. Pour chacune, le début x_0 du bâtiment est estimé en fonction de l'amplitude et de l'interférogramme. En effet, l'amplitude doit suivre les lois utilisées pour la classification (§ 4.2.1) et les hauteurs extraites de l'interférogramme doivent vérifier l'Equation 13. Le terme d'énergie et la méthode d'optimisation sont explicités dans [Tison et al., 2004 c]. La position de

l'ombre n'est pas remise en cause pour des raisons de temps de calculs.

Le résultat est bon sur la plupart des bâtiments (Figure 11) quand les ombres sont bien détectées. Malheureusement elles ne le sont pas toutes ce qui explique que, sur cette figure, tous les bâtiments ne sont pas identifiés.

4.2.6. Routes

Les routes sont détectées par l'algorithme proposé dans [Lisini et al., 2004]. Une croissance de région contrainte par l'amplitude est appliquée à ce résultat pour avoir une détection de l'ensemble de la route et pas seulement sa partie centrale. La Figure 13 montre le réseau routier détecté pour le quartier Bayard. Il y a quelques confusions entre des ombres et des routes mais le résultat est très proche de la réalité.

4.3. Fusion : optimisation conjointe de la classe et de la hauteur

Six classes, regroupant les grandes classes architecturales des villes, sont recherchées : la route R, la végétation rase V (herbe), la végétation haute A (arbre), les bâtiments B, les coins réflecteurs CR et les ombres O. En parallèle, la hauteur est régularisée.

4.3.1. Définition du graphe de régions

Par la suite, nous notons en majuscule les champs et en minuscule les réalisations des champs.

Une fois toutes les détections effectuées, un détecteur de contours (Sobel) est appliqué à chaque résultat qui est une carte de labels, le plus souvent binaire ; la détection de contours est donc triviale. Ainsi pour chaque détecteur, des régions de valeurs homogènes sont définies. Ensuite, toutes les détections de contours de chaque détecteur sont superposées afin de définir des régions homogènes pour tous les détecteurs. Afin de s'assurer que les contours sont bien fermés, une ligne de partage des eaux est appliquée. On réalise ainsi une partition de l'espace et on définit un graphe d'adjacence. Un vecteur $d^k = [d_1^k, d_2^k, \dots, d_n^k]$ (n étant le nombre de détecteurs) de paramètres est associé à chaque région. La valeur d_i^k est celle de la région k pour l' i ème détecteur.

La phase interférométrique est considérée à part et n'entre pas dans les n détecteurs (bien que l'interférogramme ait servi à calculer les régions). En effet, la carte des hauteurs n'est pas binaire et doit être manipulée de manière indépendante des autres détecteurs. Pour chaque région, la valeur médiane de la phase interférométrique, et donc de la hauteur, est calculée. Par la suite, on travaille avec cette hauteur \bar{h} pour la région entière. Dans la mesure où la détection de contours est aussi appliquée à l'interférogramme filtré, les régions ne peuvent pas contenir de fort gradients de hauteur (sinon elles

seraient scindées en sous régions). L'hypothèse de région plane est donc valide.

4.3.2. Optimisation

Fonction d'énergie

En supposant l'indépendance des régions entre elles et des régions vis-à-vis des hauteurs et des classes, le champ conjoint des hauteurs et des labels (H,L) est markovien et la probabilité $P(H,L|D,\bar{H})$ conditionnelle du champ joint (H,L) sachant le champ des détecteurs (D,\bar{H}) peut être formulée selon un champ de Gibbs. Pour chaque région k, il s'agit de minimiser une énergie U fonction de la classe l_k et de la hauteur h_k en connaissant les valeurs des détecteurs d^k , de la hauteur \bar{h}_k et de son voisinage V_k . L'énergie dépend d'un terme d'attache aux données U_{dat} et d'un terme de régularisation U_{reg} qui sont pondérés par l'aire A_i des régions examinées :

$$U(l_k, h_k | d^k, \bar{h}_k, \{l_j, h_j\}_{j \in V_k}) = (1-\beta)\alpha(A_k)U_{dat}(l_k, h_k | d^k, \bar{h}_k) \quad (14)$$

$$+ \frac{\beta}{n+1} \frac{1}{\sum_{j \in V_k} A_j} \sum_{j \in V_k} A_j U_{reg}(l_k, h_k | l_j, h_j)$$

avec β coefficient de régularisation et α fonction de A_k . Si A_k est grande alors l'influence des voisins doit être moindre que si elle est petite (α est compris entre 1 et 2).

Attache aux données

Le terme d'attache aux données U_{dat} est déterminé par l'utilisateur en fonction de ses connaissances a priori sur la qualité des détecteurs. Pour un jeu d'images, un unique terme d'attache aux données est défini (quel que soit le type de zone observée) par l'utilisateur.

$$U_{dat}(l_k, h_k | d^k, \bar{h}_k) = \sum_{i=1}^n U_D(l_k | d_i^k) + (h_k - \bar{h}_k)^2 \quad (15)$$

Les détecteurs sont binaires ou limités à un petit nombre de valeurs (cas de la classification), il n'y a donc pas trop de valeurs $U_D(l_k | d_i^k)$ à définir par l'utilisateur.

Terme contextuel

Le terme contextuel U_{reg} contient deux contraintes : la première γ est sur la classe (Equation 18) et la seconde ψ sur la hauteur (Equation 17). Les deux contributions sont pondérées par un coefficient ω (égal à 0.3 dans la suite). Mais chacune prend en compte l'autre pour introduire des règles sur la scène.

$$U_{reg}(l_k, h_k | \{l_j, h_j\}_{j \in V_k}) = \gamma(h_k, h_j)(l_k, l_j) + \omega \psi(l_k, l_j)(h_k - h_j) \quad (16)$$

Pour les hauteurs, la régularisation est faite entre deux régions de classe au sol (S ou V) ou entre deux régions de classe B (avec préservation des discontinuités dans ce cas) :

$$\begin{aligned} \psi(l_k, l_j)(h_k - h_j) &= (h_k - h_j)^2 \text{ si } l_j, l_k \in \{S, V\} \\ \psi(l_k, l_j)(h_k - h_j) &= \log(1 + (h_k - h_j)^2) \text{ si } l_j, l_k \in \{B\} \\ \psi(l_k, l_j)(h_k - h_j) &= 0 \text{ sinon} \end{aligned} \quad (17)$$

Pour les classes :

$$\begin{aligned} \text{si } h_k \approx h_j, \gamma(h_k, h_j)(l_k, l_j) &= 0 \text{ si } l_k, l_j \in \{B, CR, O\} \\ \gamma(h_k, h_j)(l_k, l_j) &= \delta(l_k - l_j) \text{ sinon} \\ \text{si } h_k < h_j, \gamma(h_k, h_j)(l_k, l_j) &= c(l_k, l_j) \\ \text{si } h_k > h_j, \gamma(h_k, h_j)(l_k, l_j) &= c(l_j, l_k) \end{aligned} \quad (18)$$

δ est le symbole de Kronecker.

Les valeurs $c(l_k, l_j)$ sont définies par l'utilisateur en prenant en compte la réalité (le sol n'est pas plus haut qu'un bâtiment, l'herbe est au même niveau que le sol, etc.).

Optimisation

L'optimisation est faite par un algorithme ICM initialisé par la classification obtenue au paragraphe 4.2.1 et les hauteurs calculées à partir de l'interférogramme.

Notons que cette méthode est transparente au nombre de détecteurs et à leur signification (les termes d'énergie doivent cependant être ajustés). Si un détecteur n'est pas disponible sur une zone, il peut être ignoré en donnant la même valeur à toutes les régions. L'optimisation aura tout de même lieu.

5. Analyse des résultats

5.1. Test sur trois quartiers distincts

L'algorithme a été testé sur trois zones de 2048x2048 pixels qui représentent toute la diversité de la ville :

- quartier Bayard (Figure 14) : mélange d'immeubles isolés de tailles variées et de zones pavillonnaires, réseau routier géométrique très varié,
- quartier Cheminot (Figure 15) : zone pavillonnaire, réseau routier très rectiligne,
- zone portuaire (Figure 16) : zone industrielle constituée de hangars de grandes tailles, pas de réseau routier.

La Figure 13 propose les résultats de tous les détecteurs, la classification et le MNS issus de

l'optimisation pour le quartier Bayard. Les trois zones ont été traitées avec les mêmes paramètres.

Sur la zone Bayard (Figure 14), l'optimisation permet de lever des ambiguïtés sur les zones d'arbres qui étaient au préalable classées comme bâtiments. Le terme contextuel permet également d'enlever des fausses zones de végétations sur les toits des bâtiments.

Sur la zone industrielle (Figure 16), les résultats sont moins bons car les ombres sont difficiles à détecter ce qui entraîne une déficience dans le détecteur des bâtiments à partir des ombres. En conséquence, la classe « arbre » est prépondérante ce qui n'a pas lieu d'être. Elle pourrait être corrigée si on introduisait une connaissance a priori (l'absence d'arbres dans cette zone). Mais nous ne voulons pas avoir des paramètres d'énergie spécifiques à une zone.

La classification du quartier Cheminot (Figure 15) n'est pas très bonne en raison de la faible résolution spatiale des images. Elle est insuffisante pour traiter des petites maisons individuelles. Cela explique l'aspect très morcelé obtenu.

Pour les trois MNS, par contre, les résultats sont bons. Sauf sur les bâtiments peu élevés (moins de trois mètres) car la résolution altimétrique des images est trop faible.

5.2. Validation des hauteurs

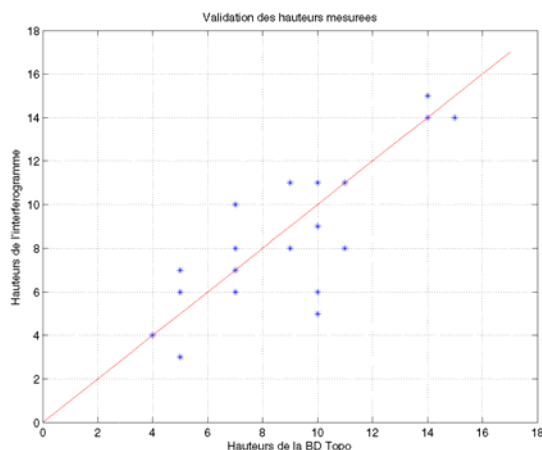


Figure 12 : comparaison de la hauteur estimée à partir de l'interférogramme (en ordonnées (en mètres)) avec la hauteur de la BD Topo© (en abscisses (en mètres)). L'écart avec la droite $y=x$ est relativement faible devant le bruit de l'interférogramme ce qui prouve une bonne adéquation entre la réalité et l'estimation. La comparaison a été faite sur le quartier Bayard.

Les MNS obtenus peuvent être comparés avec la BD Topo© de l'IGN dont nous disposons. La

courbe de la Figure 12 compare les hauteurs moyennes des bâtiments de la BD Topo© avec celles du MNS estimé. La résolution altimétrique des images est faible (§ 2) : un écart de deux mètres est de l'ordre du bruit ! Les résultats obtenus sont donc très proches de la réalité terrain et prouvent la pertinence du MNE calculé.

5.3 Critiques de la méthode

L'algorithme d'optimisation conjointe de la classe et de la hauteur est robuste : de légers changements dans la définition de l'énergie n'affectent pas le résultat final. Les problèmes de robustesse se situent plus au niveau des détecteurs, en particulier de la détection des ombres et des coins réflecteurs qui mettent en jeu des paramètres à régler par l'utilisateur. Mais l'algorithme d'optimisation est indépendant des détecteurs. Il est donc très facile de modifier l'un d'eux pour améliorer le résultat final.

La chaîne proposée comporte certaines limites :

- elle est semi-automatique et certains détecteurs ont plusieurs paramètres à régler,
- elle ne prend pas en compte une approche globale,
- il n'y a pas d'approximation des formes de bâtiments. Dans [Tison et al. 2004 b], une approche par contours actifs est proposée pour calculer la meilleure forme polygonale approximante. Cependant, il n'est pas sûr qu'il soit toujours plus profitable d'approximer la forme des bâtiments. Selon les applications, le résultat par régions peut être plus intéressant car on garde plus de détails sur les toits.

6. Conclusion

L'interférométrie ROS haute résolution est une réponse possible à la reconstruction 3D des bâtiments. La chaîne décrite dans cet article prouve qu'il est possible d'obtenir un MNS précis à partir d'un unique couple interférométrique. La qualité de la reconstruction reste cependant très dépendante de la résolution spatiale des images ROS et de la résolution altimétrique de l'interférogramme. Cela implique un choix particulièrement minutieux des paramètres des capteurs lors de la préparation de la mission de vol.

La méthode développée dans cet article repose sur une optimisation markovienne de la classe et de la hauteur d'un graphe de régions. Ce graphe est obtenu à partir d'un ensemble de détecteurs qui extraient l'information caractéristique de la scène à partir de l'amplitude, de la phase interférométrique et de la cohérence interférométrique. Le résultat final est donc une classification et un MNS. L'ensemble des

détecteurs peut facilement être modifié sans changer la philosophie générale de la méthode.

L'algorithme a été validé sur des images réelles RAMSES sur la ville de Dunkerque. Trois quartiers couvrant une large diversité architecturale sont étudiés. La comparaison avec la BD Topo© de l'IGN sur Dunkerque montre une bonne adéquation entre l'altitude réelle des bâtiments et leur hauteur estimée. La classification est bonne sur le quartier Bayard et comporte des faiblesses sur la zone portuaire et le quartier Cheminot. En effet, sur la dernière zone, les limites de résolution des images sont atteintes. Sur la zone portuaire, les détecteurs de végétation ne sont pas assez performants. Il faut donc améliorer les détecteurs pour réduire la classe «arbre», éventuellement introduire un détecteur dédié à la végétation. Cependant, la végétation a un comportement statistique (au premier et au second ordre) très proche des bâtiments. Il semblerait que la différenciation se fasse plus au niveau géométrique que radiométrique.

Une approche objet, voire de géométrie stochastique, peut ensuite être introduite afin de prendre en compte les relations spatiales entre classes : un bâtiment est proche d'une ombre et d'un coin réflecteur, il est souvent parallèle à la route. Cette approche permettrait de repositionner les objets entre eux et de lever certaines ambiguïtés. Un autre point important à développer est une boucle de validation qui permettrait de remettre en question le MNS calculé en fonction des données d'entrée. Cette boucle, proposée dans [Soergel et al., 2003], améliorerait certainement la classification et le MNS.

Remerciements

Les auteurs remercient le CNES et EADS Defence and Security Systems SA, en particulier Jean-Claude Souyris (CNES) et Vincent Leroy (EADS) pour leur soutien technique et financier. Ils remercient également l'ONERA et la DGA pour les données.

Références

[Bolter, 2000] R. Bolter, **Reconstruction of man-made objects from high resolution SAR images**, In *IEEE Aerospace Conference*, vol. 3, pp 287-292, 2000

[Descombes, 1995] X. Descombes, **A fission and fusion Markovian approach for multi-channel segmentation**, In *IGARSS'95*, vol. 1, pp 124-126, 1995

[Franceschetti et al., 2002] G. Franceschetti, A. Iodice, D. Riccio, **A canonical problem in electromagnetic backscattering from buildings**, *IEEE Trans. Geosci. and Rem. Sens.*, vol. 40, n°8, pp 11-21, 2002

[Gamba et al., 2000] P. Gamba, B. Houshmand, M. Saccani, **Detection and extraction of buildings from interferometric SAR data**, *IEEE Trans. Geosci. and Rem. Sens.*, vol. 32, n° 4, pp 855-865, 2000

[Geman et al., 1992] D. Geman, G. Reynolds, **Constrained restoration and the recovery of discontinuities**, *IEEE Trans. On PAMI*, vol 14, n° 3, pp 367-383, 1992

[Lisini et al., 2004] G. Lisini, C. Tison, D. Cherifi, F. Tupin, P. Gamba, **Improving Road Network Extraction in High Resolution SAR Images by Data Fusion**, In *CEOS'04*, May 2004

[Maître et al., 2001] H. Maître (sous la direction de), **Traitement des images RSO**, Hermès Science, 2001

[Petit, 2004] D. Petit, **Reconstruction du "3D" par interférométrie radar haute résolution**, *PhD thesis*, IRIT, Toulouse, janvier 2004

[Quartulli et al., 2003] M. Quartulli, M. Datcu, **Stochastic modelling for structure reconstruction from high resolution SAR data**, In *IGARSS'03*, vol. 6, pp 4080-4082, July 2003

[Soergel et al., 2003] U. Soergel, U. Thoennessen, U. Stilla, **Iterative building reconstruction from multi-aspect INSAR data**, In *ISPRS working group III/3 workshop on "3D reconstruction from airborne laserscanner and INSAR data"*, vol XXXIV, Part 3/W/3, Germany, 8-10 october 2003

[Tupin et al., 1998] F. Tupin, H. Maître, J.-F. Mangin, J.-M. Nicolas, E. Perchersky, **Detection of linear features in SAR images: application to road network extraction**, *IEEE Trans. Geosci. and Rem. Sens.*, vol 36, n°2, 1998

[Tison et al., 2004 a] C. Tison, J.-M. Nicolas, F. Tupin, H. Maître, **A new statistical model of urban areas in high resolution SAR images for Markovian segmentation**, soumis à *IEEE Trans. Geosci. and Rem. Sens. (IGARSS'03 special issue)*, (à paraître) 2004

[Tison et al., 2004 b] C. Tison, F. Tupin, H. Maître, **Extraction of urban elevation models from high resolution interferometric SAR images**, In *EUSAR 2004*, pp 411-414, Ulm, Germany, 24-27 May 2004

[Tison et al., 2004 c] C. Tison, F. Tupin, H. Maître, **Retrieval of building shapes from shadows in high resolution SAR interferometric images**, In *IGARSS'04*, September 2004 (à paraître)

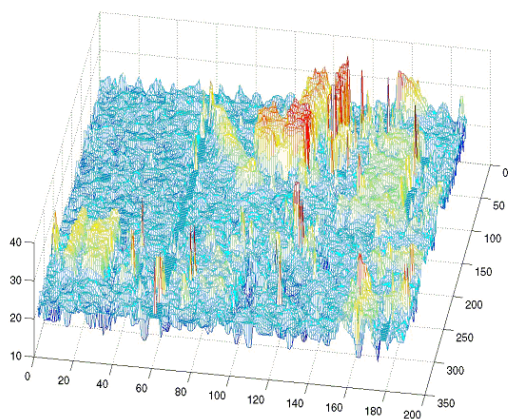


Figure 1 : MNS quasi brut sur un extrait du quartier Bayard, Dunkerque. L'interférogramme est simplement filtré par un filtre médian, les ombres sont masquées et les valeurs extrêmes sont seuillées. Le résultat est néanmoins extrêmement bruité !

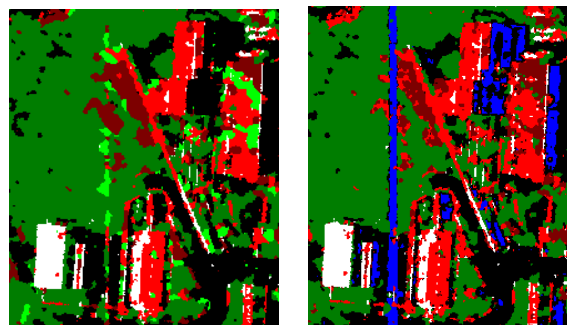


Figure 5 : Résultats de la classification sur l'amplitude uniquement (gauche), puis en fusionnant avec l'interférogramme et la cohérence (droite). Les classes sont : noir=route, vert=végétation, rouge=toit, blanc=fort réflecteur, bleu=ombre. Les différents niveaux de rouge et de vert désignent des sous-classes de bâtiment et de végétation selon leur amplitude moyenne. Après fusion les deux classes de végétation sont regroupées

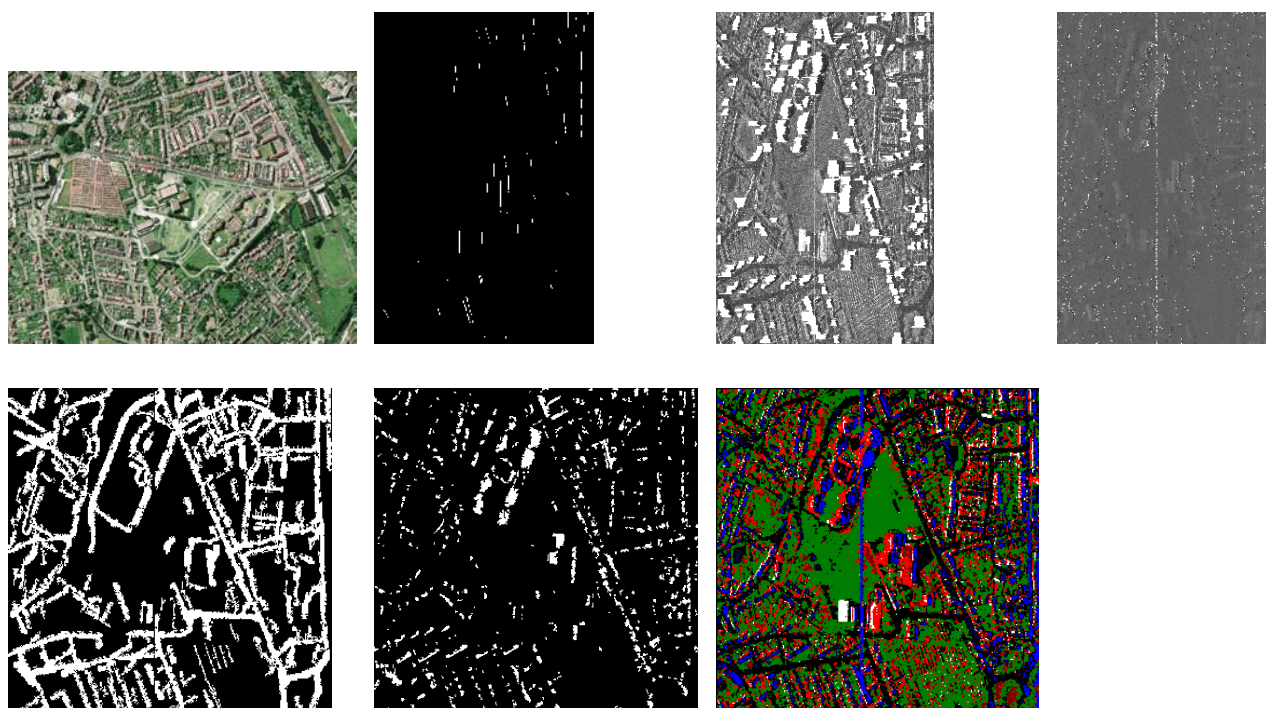


Figure 13 : Résultat des détecteurs sur la zone Bayard. De gauche à droite et de haut en bas : image optique (IGN©), carte des coins réflecteurs, bâtiments détectés à partir des ombres (en blanc) projeté sur amplitude, interférogramme lissé, carte des routes, carte des ombres et classification avec le code couleur de la Figure 4. (images RAMSES, Dunkerque)

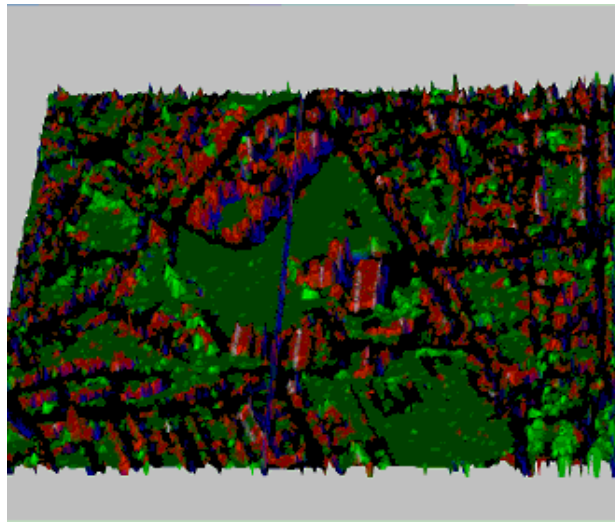


Figure 14 : Résultat de l'optimisation : classification projetée sur l'élévation. Les arbres, classés auparavant comme bâtiments ont bien été discriminés. Les bâtiments apparaissent clairement en relief et le sol est à peu près lisse (au bruit altimétrique près).

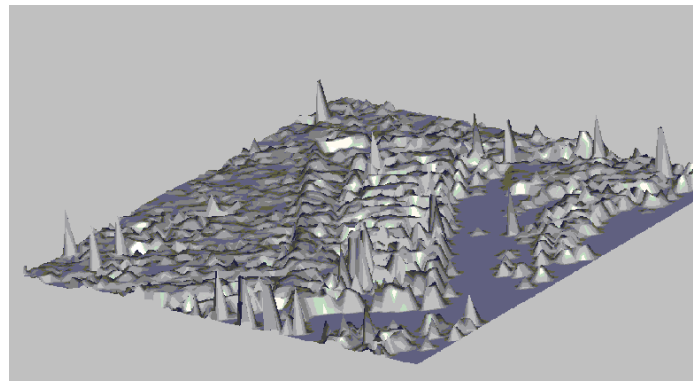
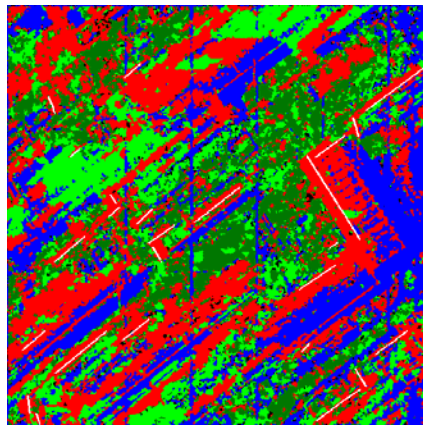
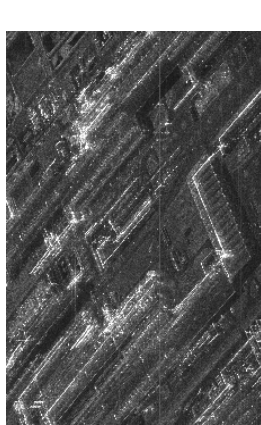


Figure 15 : Résultats sur le quartier industriel (images RAMSES, Dunkerque). De gauche à droite : image d'amplitude multi-vues, classification optimisée (même code couleur que précédemment) et extrait du MNS sur le bâtiment de la droite de l'image. Les détails du toit sont clairement visibles dans le MNS : les cannelures apparaissent. Le MNS est en géométrie radar donc, la réponse du coin réflecteur est marquée par un passage au sol peu après le début du toit.

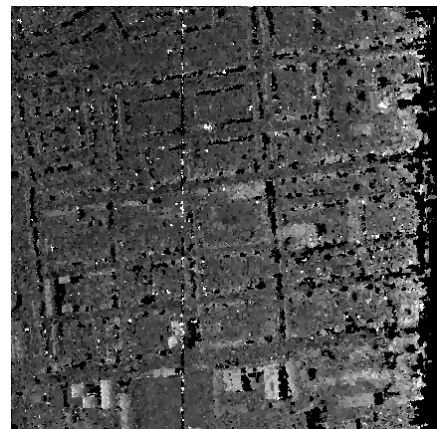
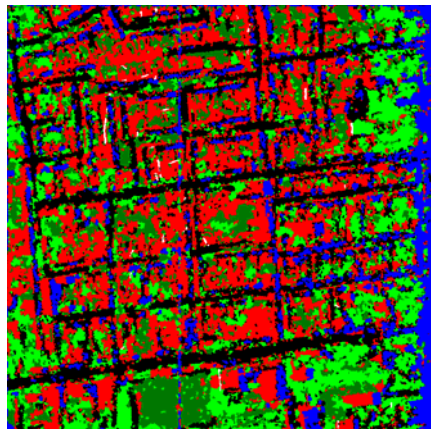
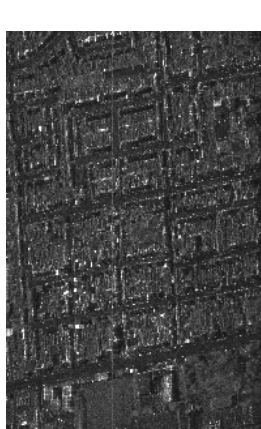


Figure 16 : Résultats sur le quartier Cheminot (images RAMSES, Dunkerque). De gauche à droite : image d'amplitude multi-vues, classification optimisée (même code couleur que précédemment) et MNS de la zone. La classification est mauvaise sur la partie centrale de l'image car les résolutions spatiale et altimétrique sont insuffisantes. Le MNS n'a pas de véritable sens sur ces zones. Par contre, les bâtiments de taille plus importantes sont bien estimés : leurs contours sont nets dans le MNS.