

Détection des changements d'élévation d'une scène par imagerie satellite stéréoscopique

Cyrielle Guérin, Renaud Binet, Marc Pierrot-Deseilligny

► **To cite this version:**

Cyrielle Guérin, Renaud Binet, Marc Pierrot-Deseilligny. Détection des changements d'élévation d'une scène par imagerie satellite stéréoscopique. RFIA 2012 (Reconnaissance des Formes et Intelligence Artificielle), Jan 2012, Lyon, France. pp.978-2-9539515-2-3. hal-00656563

HAL Id: hal-00656563

<https://hal.archives-ouvertes.fr/hal-00656563>

Submitted on 17 Jan 2012

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

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

Détection des changements d'élévation d'une scène par imagerie satellite stéréoscopique

C. Guérin¹

R. Binet¹

M. Pierrot-Deseilligny²

¹CEA, DAM / ²IGN, MATIS

CEA, DAM, DIF, F-91297 Arpajon, France / IGN, MATIS, 73 avenue de Paris, 94160 Saint Mandé
cyrielle.guerin@cea.fr; renaud.binet@cea.fr; marc.pierrot-deseilligny@ign.fr;

Résumé

Cet article propose une méthode de détection de changements à partir de couples d'images stéréoscopiques très haute résolution. Le but est de mettre en évidence, sur une image labellisée, les changements altimétriques apparus entre deux scènes. Cette méthode s'appuie sur une spatio-triangulation par l'affinage simultané des modèles géométriques de toutes les images de l'étude. Pour les capteurs présentant des défauts de modélisation géométrique, la mise en correspondance des images de chaque couple est effectuée grâce à une recherche bidimensionnelle dans l'espace terrain et image. Les Modèles Numériques de Surface (MNS) issus de chaque couple et comparables à l'échelle du pixel sont soustraits et une classification non supervisée est appliquée à la carte de différence à l'aide d'une régularisation spatiale. Cette technique permet de réduire les fausses alarmes dues au bruit de corrélation tout en gardant une très bonne détection des changements pertinents. Sur les zones testées, on atteint l'objectif de plus de 90% de vrais changements détectés avec un taux de fausses alarmes permettant un gain de temps significatif par rapport à une inspection humaine exhaustive.

Mots Clef

Stéreoscopie, détection de changements, régularisation spatiale, mise en correspondance.

Abstract

We present a change detection methodology from high resolution stereoscopic data. Our objective is to highlight elevation changes appeared between two dates based on the difference of two Digital Surface Models (DSM), with a classified image. A spatio-triangulation is applied to all the images geometrical models. For sensors having significant errors in the geometrical model, the image matching is operated using a bi-dimensional search in image and object space in order to correct residuals errors in the adjustment model. An unsupervised classification is performed using a spatial regularisation criterion and a soft threshold. This technique shows a good efficiency to reduce the number of false alarms due

to the noise induced by mis-matching while keeping a good detection of the true changes. According to the tests we ran, the goal of more than 90% of true changes has been reached with a relatively low rate of false alarms allowing a significant time saving compared to an exhaustive human analysis.

Keywords

Stereoscopy, change detection, stereo matching, spatial regularization.

1 Introduction

Le nombre d'images satellites augmentant plus vite que le nombre d'analystes, l'automatisation des traitements de ces données est devenu indispensable. Dans ce contexte, la détection de changements entre deux scènes est une problématique grandissante dans l'utilisation des images satellites très haute résolution. Elle trouve de très nombreuses applications comme le suivi de sites géologiques (glaciers, zones tectoniques, etc.) ou bien le suivi du bâti pour la mise à jour de bases de données cadastrales [2, 11], ou la cartographie des dégâts après un événement majeur (séisme, tsunami, etc.) afin de localiser rapidement les zones touchées et d'estimer de façon précise les dégâts engendrés [3]. Les changements d'élévation entre deux scènes représentent, notamment dans ce dernier cas, un type de changement très pertinent, car directement liés aux changements de hauteur des bâtiments provoqués par leur destruction, modification ou leur construction. Plus classiquement, les attributs de changements sont basés sur la radiométrie des images [5]. Mais l'intérêt de considérer des changements d'élévation par rapport aux changements radiométriques est de pouvoir s'affranchir des différences radiométriques non pertinentes présentes entre des images de capteurs et/ou de dates différentes dues aux ombres, à la végétation, aux variations saisonnières, etc. C'est pourquoi nous nous intéressons à ces changements particuliers dans un objectif de suivi du bâti.

Ces changements d'élévation peuvent être extraits à partir de couples d'images stéréoscopiques, images d'une même scène acquises quasi simultanément avec un angle de visée différent permettant de remonter à l'élévation des

objets de la scène. L'élévation trouvée et ses précisions altimétrique et planimétrique vont alors dépendre principalement de la résolution des images et de l'angle stéréoscopique du couple [1,4].

Un problème récurrent dans le domaine de la détection de changements, est celui du recalage des images. En effet, sur des images très haute résolution, les objets du sursol visibles et dont l'altitude est inconnue vont générer des problèmes de parallaxe. Il est donc nécessaire de restituer la géométrie de la scène simultanément pour toutes les images afin d'obtenir une cohérence géométrique permettant leur recalage précis [5, 7, 8, 9, 11].

Le but de la chaîne de traitements présentée ici est la création automatique d'une carte géo-localisée des changements d'élévation d'une scène urbaine à partir d'un jeu de couples d'images stéréoscopiques très haute résolution.

Cette chaîne propose tout d'abord une réponse au problème du recalage des images. Puis des Modèles Numériques de Surface (MNS) représentant l'élévation de la surface de la zone sont ensuite générés à partir de chaque couple d'images stéréoscopiques par une méthode de corrélation multi-échelle. Enfin, les changements d'élévation sont extraits grâce à une classification basée sur la différence des MNS.

2 Création de MNS en imagerie spatiale

2.1 Capteurs spatiaux

Aujourd'hui de nombreux capteurs très haute résolution spatiale (THR) tels que QuickBird, Ikonos, WorldView, GeoEye, OrbView, etc., sont capables d'acquisitions stéréoscopiques. Les images peuvent être prises, selon la capacité du satellite, par stéréoscopie latérale (*across track*) ou plus communément en mode avant-arrière le long de la trace (*along track*). Le temps entre les prises de vues pour des satellites tels que WorldView-1 est d'environ une minute, ce qui va générer des difficultés pour des objets mobiles lors de la mise en correspondance.

Les capteurs utilisés sont généralement de type *push-broom*, barrettes linéaires de détecteurs et chaque ligne de l'image est acquise successivement grâce au défilement du paysage dans le plan focal [5]. Afin de déterminer la ligne de visée pour chaque pixel, des informations précises sur la position, l'orientation et la direction de visée sont nécessaires. Or, l'acquisition des images n'est pas parfaite et des défauts apparaissent lors de la prise de vue. Ce sont principalement des défauts d'orientation des capteurs dus aux angles d'attitude du satellite, roulis α_R , tangage α_T et lacet α_L (Figure 1), pouvant engendrer des biais de localisation au sol de 3m pour WorldView-1, et 10m pour OrbView-3 par exemple. Néanmoins, ces défauts sont modélisés et corrigés par l'étape de spatio-triangulation [7] (Paragraphe 2.2). Cependant, d'autres défauts plus

difficiles à modéliser par manque d'informations, ne peuvent pas être corrigés. En effet, le système d'acquisition en barrettes disjointes (Figure 1) complique la modélisation géométrique des images. Lorsque l'image est produite, les segments de chaque ligne sont fusionnés afin de simuler un capteur linéaire dit « parfait » mais des défauts de calibration sur les distances inter-barrettes peuvent subsister, et engendrer ainsi une erreur (Figures 6 et 7). De plus, les variations d'attitude du satellite très haute fréquence temporelle ne sont pas systématiquement restituées dans le modèle de prise de vue.

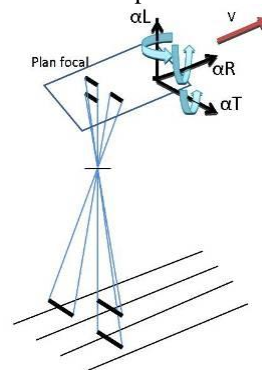


Figure 1: Barrette push-broom et attitudes du satellite

2.2 Spatio-triangulation

La méthode que nous avons développée pour l'analyse des changements d'élévation est illustrée Figure 2.

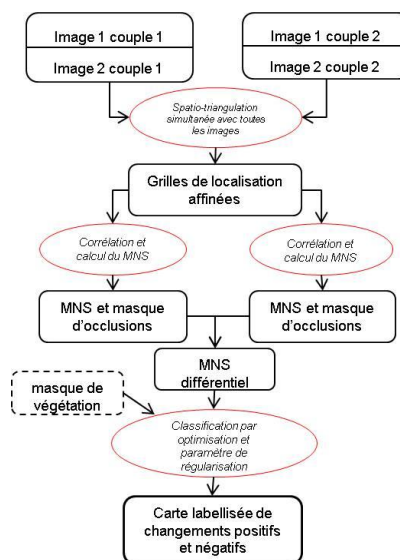


Figure 2: Principe de la méthode de détection de changements

Les techniques classiques en stéréoscopie sont basées sur la recherche des pixels homologues le long des parallaxes entre les deux images afin de calculer l'élévation du point à l'intersection des lignes de visées issues de chaque pixel [9, 11]. La spatio-triangulation a pour objectif de corriger les erreurs d'orientation de la prise de vue afin que les

faisceaux provenant de points homologues s'intersectent au mieux en vue de la création d'un MNS robuste et précis. Son principe est la recherche de paramètres de recalage [3], introduits dans les équations des modèles directs F_1 et inverses F_1^{-1} des images pour résoudre au mieux la divergence entre les coordonnées terrain d'un même point physique. Cet affinage est effectué grâce à l'utilisation de points d'appui ou de liaison. La qualité de la spatio-triangulation influe également sur l'espace de recherche des points homologues. Pour une géométrie parfaite, cette recherche est monodimensionnelle. En réalité, un espace de recherche supplémentaire peut être nécessaire dans certains cas et il est d'autant plus grand que les résidus sont importants.

Les modèles de toutes les images de l'étude sont affinés simultanément. Cela garantit que les images à l'intérieur de chaque couple stéréoscopique et entre les couples soient précisément recalées afin d'obtenir des MNS comparables.

La spatio-triangulation est réalisée avec l'outil Euclide, développé par l'IGN et le CNES, après une recherche exhaustive et automatique de points de liaison dont on contraint la répartition et la qualité sur toutes les images. La solution est qualifiée avec le calcul des résidus des points correspondant à l'écart entre la position réelle des points et la position calculée après le recalage [3].

2.3 Mise en correspondance

La mise en correspondance des images à l'intérieur de chaque couple stéréoscopique est l'étape fondamentale dans la création des MNS [1, 7, 8, 11]. Elle est ici réalisée avec le logiciel MICMAC (Multi-Images Correspondances, Méthodes Automatiques de Corrélation) développé par l'IGN. L'objectif est de retrouver les homologues des pixels sur chaque image afin d'obtenir une carte d'élévation de tous les points de la scène. La recherche monodimensionnelle le long des lignes épipolaires est la technique la plus courante de mise en correspondance [3, 7, 11], l'inconvénient de cette recherche est qu'elle n'admet pas d'erreur dans le modèle affiné de prise de vue, erreur pouvant être conséquente pour certains capteurs.

La méthode de mise en correspondance choisie ici est l'appariement depuis l'espace d'une image maître selon une recherche par parcours de lignes de visées [9]. Son intérêt réside dans le fait que la recherche du pixel homologue est effectuée dans les deux dimensions avec un seul ré-échantillonnage nécessaire. La recherche est d'abord effectuée dans l'espace terrain (parallaxe 1). Pour chaque altitude Z testée et définie selon un pas de recherche fixé, un coefficient de corrélation est calculé entre le pixel de l'image maître et son voisinage et une vignette de même taille centrée sur le pixel correspondant de l'autre image (Figure 2). Un calcul de corrélation dans l'espace image (parallaxe 2) est ensuite effectué selon la

direction orthogonale à la parallaxe 1 autour du pixel de l'autre image pour assurer la recherche bidimensionnelle.

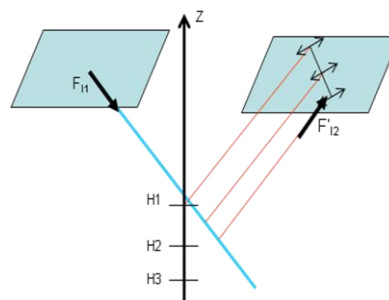


Figure 3: Recherche bidimensionnelle des pixels homologues

La sélection du pixel homologue est alors effectuée par optimisation d'une fonctionnelle globale d'énergie [10] de la forme:

$$E(Z) = \sum A(x, y, Z(x, y)) + \alpha \times F(G(Z)) \quad (1)$$

Avec Z l'altitude recherchée du pixel (l, c) et $A(l, c, Z(l, c))$ le terme d'attache aux données qui mesure la corrélation de vignettes centrées sur le pixel de l'image maître et le pixel de l'image esclave correspondant à l'altitude Z . $F(G(Z))$ est la fonction de régularisation qui dépend des variations de l'altitude Z . α représente le poids accordé à la régularisation par rapport à l'attache aux données. Ce terme est fonction de la connaissance a priori sur la rugosité du terrain pour la parallaxe 1. Sur une scène urbaine, il est choisi assez faible afin de ne pas lisser les variations d'altitudes dues aux bâtiments. Sur la parallaxe 2, ce terme est au contraire important afin de corriger efficacement les erreurs capteurs.

Une stratégie multi-échelle est classiquement utilisée afin de gagner en temps de calcul tout en permettant une recherche sub-pixellique [9, 11].

La mise en correspondance est très sensible aux erreurs de modélisation de la prise de vue. Ainsi, une erreur dans le modèle de l'une des images engendre un décalage des vignettes qui réduit la corrélation et donc le pouvoir discriminant de l'indice de similarité. Une erreur de géométrie peut donc induire des erreurs altimétriques importantes. La perte de corrélation liée à un décalage des vignettes peut se modéliser par la fonction d'auto-corrélation des images. Celle-ci dépend essentiellement de la fonction de transfert de modulation de l'instrument, de la densité spectrale de puissance de l'image et des post-traitements. Pour des images très haute résolution, la perte de corrélation observée est de moins de 15% à 0,5 pixel.

La mise en correspondance effectuée avec les images OrbView-3 montre les avantages certains de la recherche bidimensionnelle. La Figure 4 présente le résultat sur la parallaxe transverse d'après lequel l'erreur calculée peut atteindre 3 pixels engendrant une perte totale de corrélation. La Figure 5 présente la différence calculée entre les parallaxes 1 obtenues par recherche 1D et 2D.

On observe des différences d'altitudes jusqu'à 15m, où les défauts géométriques calculés par la parallaxe 2 sont les plus importants.

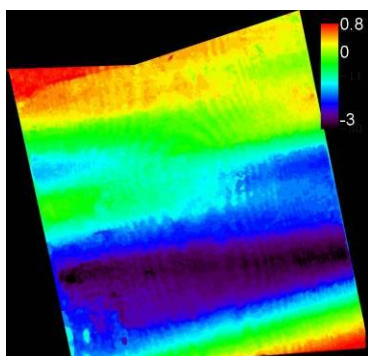


Figure 4 : Parallaxe transverse calculée par recherche 2D en faisceau sur un couple OrbView-3 (échelle en pixels)

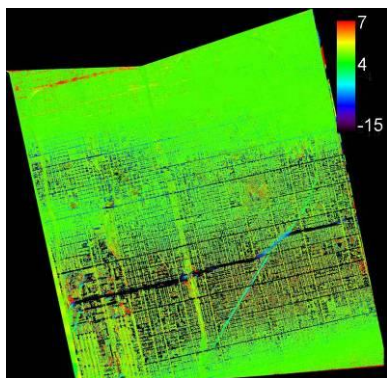


Figure 5 : Différence entre les cartes d'altitudes obtenues avec une mise en correspondance 2D et 1D (échelle en mètres)

Lors du calcul de la parallaxe transverse de WorldView-1, l'erreur liée principalement aux erreurs de fusion des segments en capteur parfait, est comprise entre -0,6 et 0,7 pixel (Figure 6b), la perte de corrélation est alors relativement faible, moins de 25%.

Les images c et d de la Figure 6 montrent les cartes d'élévation calculées respectivement avec une recherche unidimensionnelle et bidimensionnelle afin d'illustrer l'intérêt et l'inconvénient de la recherche 2D. En effet, la route (entourée en jaune sur les figures) présente un défaut d'altitude de près de 3m avec une recherche 1D par rapport à la recherche bidimensionnelle, cette erreur étant non négligeable lorsque l'on souhaite une discrimination des changements dès 2,5m. En revanche, en recherche 2D, on note aussi la présence d'erreurs dans la carte des parallaxes transverses telles que celle entourée en jaune et d'amplitude sub-pixellique (Figure 6b), probablement générée par une texture périodique sur l'image faussant la corrélation. L'erreur altimétrique associée peut être tout aussi gênante pour le calcul du MNS puisqu'elle peut atteindre 2,5m.

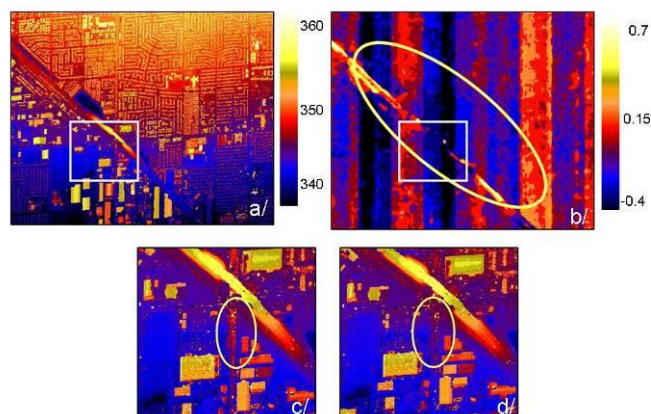


Figure 6: a: carte d'altitudes calculée par géométrie faisceau 2D, parallaxe 1 (en mètres). Les bandes verticales sont dues aux erreurs liées au calcul du capteur parfait. b: carte d'altitudes calculées par géométrie faisceau 2D parallaxe 2 (en pixel). c et d: cartes des altitudes calculées respectivement par géométrie faisceau 1D et par géométrie 2D (en mètres). L'ellipse montre une route d'altitude surestimée par la recherche 1D.

2.4 Qualité des MNS

La précision altimétrique finale des MNS est fonction de la résolution des images et du rapport B/H des couples, B étant la distance entre les deux prises de vue et H l'altitude moyenne du satellite. Un fort B/H permet une bonne précision altimétrique mais va générer de nombreuses parties cachées ou zones d'occlusion par les bâtiments et inversement [1, 4, 5, 12]. Ces zones d'occlusion sont calculées dans un masque lors du ré-échantillonnage des MNS sur une grille régulière géo-référencée et commune pour tous les MNS.

Du fait des inévitables erreurs de mise en correspondance, les MNS calculés présentent des altitudes parfois incohérentes et donc gênantes lors de l'interprétation de la différence des MNS. Comme illustré dans la Figure 6, des artefacts peuvent fausser la corrélation sur des zones de texture périodique ou texturées par du bruit de compression [11]. Des erreurs dites d'adhérence apparaissent aussi au niveau des bordures de bâtiments : le bord contrasté du bâtiment guidant la corrélation lorsque le pixel central de la vignette correspond au sol, engendrant une délocalisation de la bordure. Enfin, les deux scènes stéréoscopiques n'étant pas simultanées, la présence d'objets mobiles, typiquement les véhicules, les nuages, et l'eau, vont aussi générer de nombreuses fausses alarmes.

Certaines de ces erreurs peuvent être en partie filtrées par l'utilisation de masques [7, 8, 11]. Une image multi-spectrale 8 bandes WorldView-2 est utilisée pour le calcul de l'indice de végétation *Normalized Difference Vegetation Index* (NDVI) et le *Normalized Difference Water Index* (NDWI), et tous deux permettent de masquer

les pixels de végétation et d'eau, ces deux classes d'objets étant de moindre intérêt pour le suivi des bâtiments. Les seuils sont choisis empiriquement de manière à préserver au maximum les pixels complémentaires.

Le MNS différentiel obtenu, filtré par les masques d'occlusion et de végétation, permet d'observer, à l'échelle de quelques pixels, un éventuel changement d'élévation (Figure 7c). La discrimination altimétrique possible des changements est liée à l'amplitude du bruit de corrélation des MNS, estimé à environ 1,5m RMS sur ces images. Considérant ce bruit ainsi que l'intérêt opérationnel, le seuil des changements d'intérêt est fixé à une différence de 2,5m, cette valeur étant la hauteur standard d'un étage de bâtiment. Un seuillage à la valeur absolue de 2,5m ne permet pourtant pas d'observer précisément les changements réels de la scène à cause de la présence de trop nombreuses fausses alarmes comme illustré en Figure 8. La détection de changements de petite taille surfacique est ainsi limitée d'une part par la capacité de l'algorithme de création des MNS à discriminer des objets de petite taille (capacité principalement liée à la résolution des images et au gradient d'élévation des objets), et d'autre part par le bruit de corrélation. Tian [11] propose de supprimer une partie de ces fausses alarmes en affectant le minimum des différences d'élévation calculées sur un voisinage du pixel considéré, cette méthode a pour inconvénient de délocaliser les bords et diminue la sensibilité de la détection.

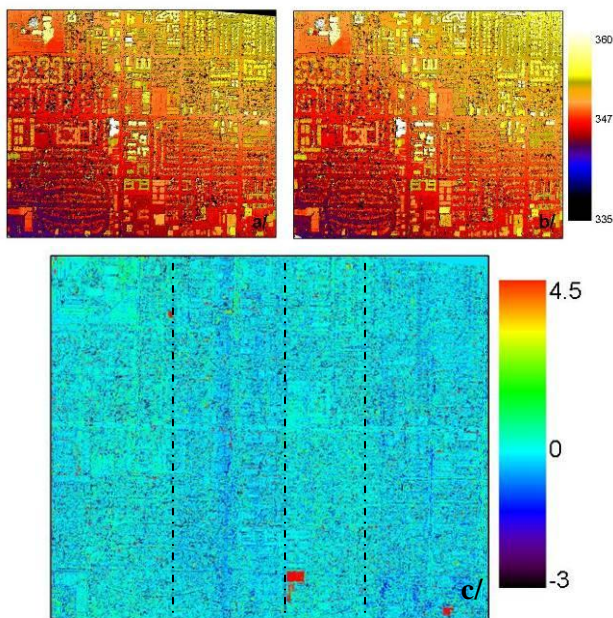


Figure 7 : a et b : MNS WorldView-1 de 2008 et 2011, résolution spatiale de 0,6m, altitudes en mètres. c : Différence des MNS (2011-2008) en mètres. Les bandes verticales montrent les défauts capteurs résiduels.

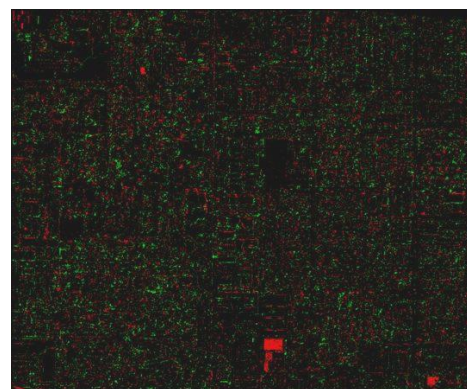


Figure 8 : Seuillage de la différence des MNS (2011-2008). En rouge les pixels montrant une différence de +2,5m, en vert les pixels montrant une différence de -2,5m.

3 Classification par optimisation

Considérant que les changements d'intérêt ont une surface beaucoup plus importante que celle des pixels, et que le bruit de corrélation est de haute fréquence spatiale, nous proposons une méthode de classification non supervisée des pixels de l'image de différence des MNS selon les labels changements positifs – changements négatifs – non changements. C'est une classification spatialement régularisée et basée sur un seuil de détection lié aux différences d'élévation observées. Cette méthode est basée sur le même principe que celui de la mise en correspondance, c'est l'optimisation d'une fonction d'énergie globale avec un critère de régularisation spatiale. En effet, dans une scène urbaine, le taux de changements excède rarement 15% [2] et les changements recherchés sont de type infrastructures donc de taille conséquente. Ainsi, une optimisation avec un critère de régularisation va permettre de lisser la labellisation des pixels dans la direction de l'image parcourue et donc de supprimer les changements de petite taille non pertinents tout en faisant ressortir les changements importants. La fonction d'énergie globale que l'on cherche à minimiser est de la forme :

$$E(D) = E_{diff}(D) + \lambda \times E_{reg}(D) \quad (2)$$

Avec $E_{diff}(D)$ le terme d'attache aux données, λ le poids accordé à la régularisation par rapport à l'attache aux données et $E_{reg}(D)$ le terme de régularisation [6, 7, 9].

L'attache aux données est reliée directement à l'image des différences d'élévation (D). Les valeurs sont normalisées entre 0 et 1 en fonction de l'inverse de la probabilité d'appartenance à chacune des classes, l'objectif étant que l'attache aux données soit la plus faible possible pour privilégier la classe à laquelle appartient réellement le pixel. Cette valeur est calculée pour chaque classe avec une fonction sigmoïde et dépend du seuil, défini à 2,5m, (équation (3)). Le résultat, pour des valeurs de différence de 0 à 8m et pour le calcul de la classe non changement, est illustré en Figure 10.

$$E_{NC}(D) = \frac{1}{1 + e^{-\lambda * abs(D) - seuil}} \quad E_C(D) = 1 - \frac{1}{1 + e^{-\lambda * D - seuil}} \quad (3)$$

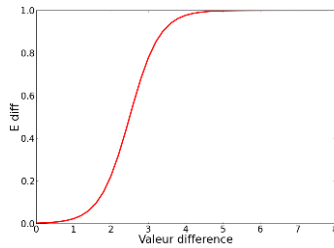


Figure 10 : Attache aux données, calculée pour la classe non changement $E_{NC}(D)$ pour des valeurs d'altitude entre 0 et 8m

Le terme de régularisation dépend, quant à lui, de la variation de classe entre deux pixels consécutifs dans la direction parcourue: il est égal à 0 si il n'y a pas de changement de classe et à 1 pour un changement de classe quelque soit son sens (vers une classe changement ou non changement). Le poids de la régularisation est adaptable à l'image et au résultat recherché : plus il est important, plus la régularisation sera forte par rapport à l'attache aux données, et des changements de petites amplitudes surfacique et altimétrique seront gommés, inversement, plus le poids de la régularisation est faible, plus la sensibilité aux changements augmentera, générant en même temps de nombreuses fausses alarmes.

Afin d'implémenter cette classification, une méthode de programmation dynamique est utilisée. C'est une technique classique d'optimisation monodimensionnelle très utilisée en traitement d'images et notamment dans les problèmes de mise en correspondance [1, 6, 9]. Le principe est la décomposition d'un problème en de nombreux problèmes résolus de façon récursive afin de retrouver la solution du problème original [6]. Dans ce cas, l'optimisation consiste à trouver, pour chaque direction de parcours sur l'image, la solution de labellisation des pixels dans la classe changements positifs, négatifs ou non changement, qui soit la moins coûteuse possible. L'optimisation globale est alors effectuée en optimisant une solution indépendamment dans chacune des huit directions de parcours de l'image selon la Figure 11 [6]. Un pixel de l'image est donc traité huit fois, selon chaque direction, puis les résultats de toutes les sous-optimisations obtenues sont fusionnés pour obtenir une solution de labellisation globale sur l'image.

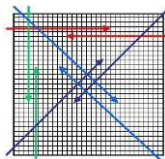


Figure 11 : Directions de lecture de l'image utilisée pour l'optimisation par programmation dynamique.

Cette technique de programmation dynamique nécessite, dans son implémentation, de parcourir chaque direction de lecture une première fois, afin d'obtenir le coût de chaque

labellisation l_n , pour chaque pixel x_i , et en considérant les coûts de labellisation de tous les pixels précédemment calculés dans la direction considérée. On obtient ainsi le coût global de la labellisation $E(l_n)$ de la direction parcourue. Ce coût est défini selon l'équation (4).

$$E(l_n) = \sum_{i=1}^n E_{diff}(x_i) + \sum_{i=1}^{n-1} E_{poids}(x_i, x_{i+1}) \quad (4)$$

Les coûts associés à chaque labellisation de chacun des pixels de l'image et selon la direction considérée sont stockés dans un tableau. Une lecture retour des coûts du tableau permet ensuite de recalculer le coût de chaque label de tous les pixels en utilisant le coût de variation entre le label optimal l_n^* (label ayant le coût le plus faible) calculé précédemment et le label considéré, afin d'obtenir une optimisation globale dans la direction parcourue.

$$l_n(x_i) = \arg \min (E(l_n) + E_{poids}(l_n(x_i), l_n^*(x_{i-1}))) \quad (5)$$

Les tableaux de coût de chaque label calculé selon chaque direction sont alors additionnés et le label final des pixels de l'image est celui dont le coût total est le plus faible.

La carte de changements calculée va dépendre du paramètre λ entré. Ainsi, selon le poids de la régularisation, la carte contiendra un certain taux de fausses alarmes. Certaines peuvent être éliminées grâce aux masques de végétation et d'eau, mais d'autres, causées par exemple par de la végétation non couverte par le masque ou encore par des véhicules, peuvent gêner l'interprétation. Il est néanmoins possible de supprimer les fausses alarmes de superficie très faible, représentant moins de 10 pixels par une opération de morphologie mathématique. Une ouverture avec un élément structurant de taille minimale (3x3 pixels) est alors effectuée sur l'image afin d'éliminer ce bruit.

4 Résultats

Les données utilisées sont des couples stéréoscopiques WorldView-1 à 0,5m de résolution spatiale en bande panchromatique. Deux images sont disponibles, acquises sur la ville de Phoenix, Arizona, en avril 2008 et en juin 2011. Les MNS dérivés de ces couples sont calculés à environ 0,6m de résolution planimétrique et 0,55m de résolution altimétrique. Trois zones test de superficie d'environ 3km² chacune ont été choisies pour effectuer une détection de changements. La zone 1 montre un mélange de bâtiments peu élevés et espacés et de pavillons serrés, la zone 2 contient des bâtiments élevés (jusqu'à 50m) et la zone 3 contient comme la zone 2, de très nombreux bâtiments peu élevés.

La ville de Phoenix est une ville développée et dynamique ce qui permet d'observer des changements de taille variable entre les années 2008 et 2011. Cependant, c'est aussi une zone plutôt riche en végétation, notamment

autour des bâtiments ce qui représente une difficulté pour la recherche précise de changements.

Afin d'être en mesure d'évaluer les résultats de la méthode, les cartes de changements produites ont été étudiées selon les deux critères importants de la détection de changements : la sensibilité et la précision. La sensibilité de la détection représente le nombre de changements trouvés par la méthode par rapport au nombre de changements réels de la scène [2]. Elle est calculée selon l'équation (6) :

$$Sensibilité = \frac{Nb \text{ Changements vrais détectés}}{Nb \text{ Changements réels}} \quad (6)$$

La précision représente le nombre de changements réels détectés par rapport au nombre de changements totaux détectés, il représente le taux de fausses alarmes.

$$Précision = \frac{Nb \text{ Changements vrais détectés}}{Nb \text{ Total de Changements détectés}} \quad (7)$$

Les changements réels contenus dans les zones d'études ont été relevés par une analyse visuelle entre les images et une dizaine de changements ont été relevés sur chaque zone. On considère ici un changement lorsqu'un objet d'une superficie de plus de 20m² et d'une différence d'élévation d'au moins 2,5m (selon le seuil) est apparu ou a disparu dans l'image, les résolutions planimétrique et altimétrique des images ainsi que la présence de nombreuses sources d'erreurs dans les MNS ne permettant pas la recherche de changements plus petits de façon précise.

Les valeurs de sensibilité et précision ont été calculées pour chaque zone et pour différentes valeurs de λ . On considère un changement réel détecté lorsqu'un ou plusieurs pixels labellisés en changements correspondent à une zone de changements sur la scène. De même, on considère comme une fausse alarme un ou plusieurs pixels regroupés correspondant à une région stable de la scène.

Les résultats Figure 12 montrent que le paramètre λ a une très forte influence sur la sensibilité et la précision de la détection. La méthode permet presque toujours de retrouver la totalité des changements à faible régularisation ($\lambda=3$), mais la précision pour de tels résultats est faible avec un taux de fausses alarmes de plus de 90%. Lorsque l'on augmente la régularisation ($\lambda=6$) le taux de fausses alarmes est alors bien meilleur, de 25% à 35% sur les zones 1 et 3 et la sensibilité reste élevée, de 60% à 70%. La zone 2, quant à elle, contenant de hauts bâtiments, demande une régularisation beaucoup plus forte ($\lambda=12$) pour atteindre un taux de fausses alarmes correcte à 50% en maintenant une sensibilité à plus de 60%. En effet, les erreurs de mise en correspondance, d'occlusion et d'ombre sont particulièrement présentes autour de tels bâtiments, générant des fausses alarmes de fortes amplitudes planimétrique et altimétrique, difficiles à corriger. La Figure 13 montre l'effet de la régularisation sur la carte de changements obtenue et l'intérêt du

paramètre de régularisation. En effet, lorsque des changements de grande amplitude sont visés, une régularisation plus forte est définie afin de pouvoir s'affranchir des fausses alarmes, tels que sur la Figure 13, et ainsi d'augmenter la précision.

Cependant, il est à noter que la valeur de la précision ne rend pas totalement compte du résultat puisqu'en réalité, une précision, même faible, représente tout de même un gain de temps certain par rapport à une analyse visuelle de l'image. Un taux de fausses alarmes de 60%, rapportée au nombre de changements de l'image, signifie que près d'un tiers des changements vérifiés par l'opérateur sont de réels changements, les autres faux changements étant vérifiables. Par exemple, sur une image de 3000 x 3000 pixels on dénombre environ 10 changements réels et 30 changements détectés pouvant être invalidés en moins de 2 minutes, tandis qu'une analyse visuelle exhaustive nécessite environ une demi-heure selon l'opérateur soit un gain d'environ un facteur dix.

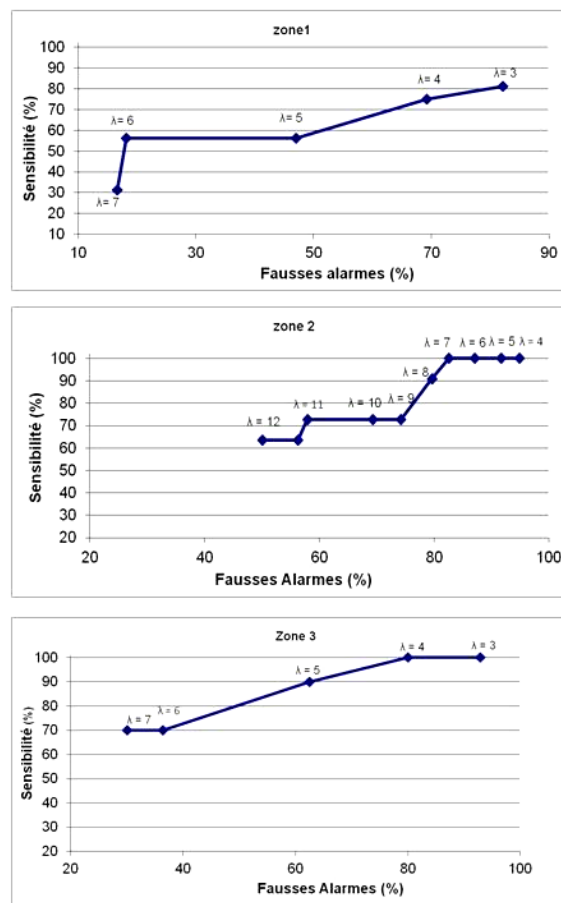


Figure 12 : Résultats de la sensibilité en fonction des fausses alarmes pour les zones 1, 2 et 3

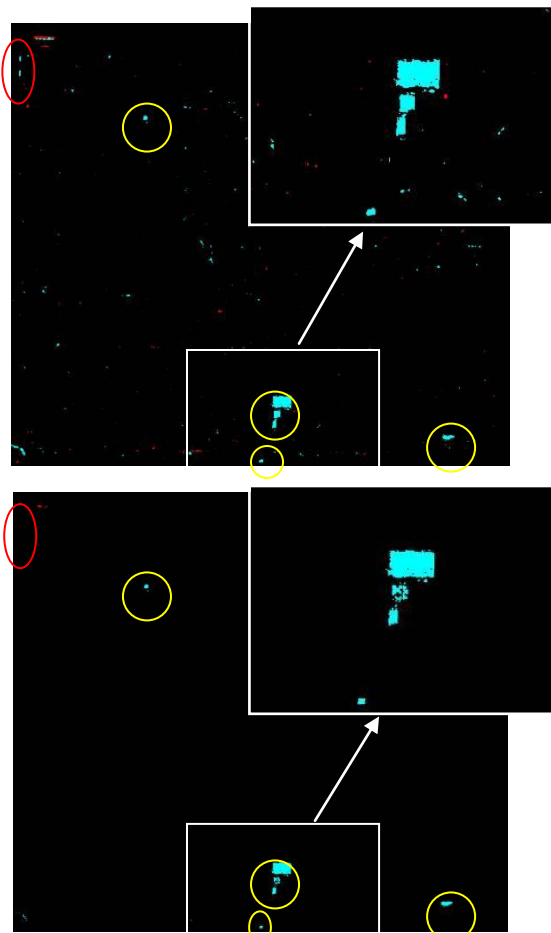


Figure 13 : Cartes de changements obtenues avec différentes valeurs de régularisation : En haut $\lambda=3$, en bas $\lambda=7$. En bleu, les changements positifs et en rouge, les changements négatifs. Les cercles blancs présentent les changements réels détectés et les rouges, les changements réels non détectés.

5 Conclusion et perspectives

La méthode de détection de changements proposée dans cette étude est implémentée afin de pouvoir traiter des couples d'images stéréoscopiques haute résolution de façon générique et automatique. La chaîne de traitements mise en place comprend la restitution simultanée de toutes les images dans une géométrie cohérente commune et la recherche bidimensionnelle des points homologues. Ces deux techniques permettent d'apporter une plus grande robustesse aux défauts géométriques de certains capteurs. L'apport de cette étude réside surtout dans la méthode de classification des cartes de différences obtenues entre les MNS calculés selon des labels changements ou non changements. Cette classification par un critère de régularisation spatiale montre une efficacité satisfaisante de la méthode puisque pour un paramétrage de régularisation donné, le taux de changement réels détectés varie de 60% à 90% tandis que le taux de fausses alarmes est d'environ 60%. Ce dernier résultat est très satisfaisant car il représente en fait un gain de temps d'environ un

facteur dix pour l'analyse des données. Ces performances doivent néanmoins être validées sur d'autres cas d'étude. Dans un contexte opérationnel, notre approche nécessite la comparaison de MNS avant et après événement pour détecter les changements engendrés. Or, si le nombre de données stéréoscopiques augmente, avec notamment l'arrivée prochaine des satellites Pléiades qui seront en mesure d'effectuer de la tri-stéréoscopie, les archives disponibles restent très rares. La création de MNS à partir d'images diachroniques semble alors une piste prometteuse pour résoudre cette difficulté.

Bibliographie

- [1] C. Baillard. Analyse d'images aériennes stéréoscopiques pour la restitution 3D des Milieux Urbains, *Thèse de doctorat, ENST, Lab. Matis*, 1997.
- [2] N. Champion, L. Matikainen, X. Liang, R. Rottensteiner. A test of 2D building change detection methods: Comparison, evaluation and perspectives. *ISPRS Congress*, pages 297-304. 2008.
- [3] A-L. Chesnel, R. Binet. Object oriented assessment of damage due to natural disaster using very high resolution images, *IEEE International Geoscience and Remote Sensing Symposium*, 2007.
- [4] J. Delon, B. Rougé Small baseline stereovision, *Journal of Mathematical Imaging and Vision* vol 3, pp 209-223, 2007.
- [5] Y. Egels, M. Kasser Photogrammétrie numérique, *Edition Paris Hermes Science Publication*, 2001.
- [6] P.F. Felzenszwalb, R. Zabih Dynamic programming and graph algorithms in computer vision, *IAPRS, IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 4, pp. 721-740, 2011.
- [7] H. Hirschmüller, D. Scharstein Evaluation of cost fonctions for stereo matching, *IEEE Conference on Computer Vision and Pattern Recognition*, 2007.
- [8] T. Krauss, M. Lehner, P.Reinartz. Generation of coarse 3D models of urban areas from high resolution satellite images, *International Archives of Photogrammetry*, 2008.
- [9] M. Pierrot-Deseilligny, N. Paparoditis A multiresolution and optimization-based image matching approach: an application to surface reconstruction from Spot5-HRS stereo imagery, *In IAPRS vol 36-1*, 2006.
- [10] F. Rottensteiner, Building change detection from digital surface models and multi-spectral images, *IAPRS, Sensing and Spatial Information Sciences*, vol 36 pp.145-150, 2007.
- [11] J. Tian, H. Chaabouni-Chouayakh, P. Reinartz, T. Krauss, P. d'Angelo, Automatic 3D Change Detection Based on Optical Satellite Stereo Imagery, *ISPRS TC VII Symposium, Vienna*, 2010.