

B.3 Reconstruction de Modèles Numériques de Surfaces en zone urbaine à partir de prises de vues aériennes numériques à fort recouvrement

Nicolas Paparoditis, Ibtissam Baghli et Christian Thom

Introduction

Le laboratoire MATIS mène, depuis plusieurs années, des activités de recherche sur la reconstruction automatique de Modèles Numérique de Surface par analyse de couples d'images aériennes stéréoscopiques [BAILLARD 97] [DISSARD et al 99].

Ces MNS sont utiles en tant que tels pour un grand nombre d'applications : utilisation dans des SIG, production de cartes d'inter-visibilité pour l'optimisation du placement d'antennes pour les télécommunications, cartographie des risques (inondations, glissements de terrain, avalanches, etc.), détection automatique de changements pour la mise à jour des bases de données, préparation de missions, réalité virtuelle, etc. Mais sont aussi des données essentielles pour l'élaboration d'autres produits cartographiques : orthophotos, modèles 3-D de bâtiments de type TRAPU [PAPARODITIS et al 98] [FUCHS et al 99] [JIBRINI et al 00], etc.

Pour cette dernière application qui reste un axe prioritaire de recherche du laboratoire MATIS, des images de très grande résolution (< 20 cm) sont nécessaires afin de mettre en évidence l'ensemble des structures tridimensionnelles nécessaires à l'élaboration de la description fine d'un bâtiment [CORD et al 98].

À ces résolutions, et dans un contexte de tissu urbain dense, les parties cachées (zones au pied des bâtiments) sont importantes et une description dense et complète de la scène par analyse stéréoscopique n'est alors pas possible. D'autre part, en zone urbaine, l'analyse stéréoscopique par corrélation rencontre certains problèmes de fiabilité liées aux ambiguïtés de ressemblance des deux images (zones homogènes, surfaces spéculaires, véhicules mobiles, etc.).

L'acquisition de prises de vues à multiples points de vues par un fort recouvrement latéral et longitudinal (un point du paysage est vu dans plus de deux images) permet de résoudre un certain nombre de problèmes. En effet, si le recouvrement est suffisant, il existe toujours un ou plusieurs couples parmi l'ensemble des couples sur lequel un point donné du paysage est visible et où ces problèmes de specularité et de véhicules mobiles n'apparaissent pas.

Des techniques de gestion et d'exploitation de ce type de données ont été traitées abondamment dans la littérature récente [LELOGLU et al 98], [CANU et al 95], [GABET et al 94], [FOURNIER 99] et ont aussi été abordées au laboratoire MATIS dans le cadre d'un stage [MAYER 98].

Dans la majorité des techniques développées, le MNS est reconstruit par une fusion *a posteriori* de tous les MNS élémentaires calculés sur tous ou un sous-ensemble de couples stéréoscopiques par une technique de vote majoritaire (par exemple basé sur la médiane). Ce processus permet de densifier et fiabiliser de manière considérable les résultats quand ceux-ci sont complémentaires et le plus souvent justes sur l'ensemble des couples.

Par contre, le processus d'appariement stéréoscopique élémentaire par corrélation impose l'utilisation de taille de vignettes non négligeables si l'on souhaite des mesures suffisamment fiables. Or la taille des vignettes conditionne directement la qualité morphologique des MNS, i.e. l'aptitude à restituer les pentes, les ruptures de pentes, les microstructures du relief, toutes cruciales pour de nombreuses applications. Plus les vignettes sont grandes et plus les "hautes fréquences" du relief sont lissées, plus les discontinuités sont délocalisées, et plus l'appariement de surfaces en pentes est difficile (du fait de la déformation entre voisinages homologues).

Une solution pour réduire la taille des vignettes est d'améliorer le rapport signal à bruit des images de manière à réduire les ambiguïtés d'appariement : soit naturellement en ayant un système imageur de meilleure qualité, soit artificiellement en augmentant le nombre d'observations. L'originalité de notre travail réside dans l'utilisation d'images de la caméra numérique du Laboratoire OEMI (très bon rapport signal à bruit de 300 environ) avec un fort recouvrement latéral et longitudinal et le développement d'un processus de reconstruction extrêmement simple mais réellement adapté à la nature et à la complémentarité des données.

Le principe de notre approche : appariement multi-images guidé depuis l'espace objet

Un nouveau prototype de corrélateur adapté à ces données a donc été développé en partie dans le cadre d'un stage [BAGHLI 99]. Cet outil permet de construire des MNS, et les ortho-images associées, à partir de données multi-images de la caméra numérique de l'IGN sans passer par un traitement par couple stéréoscopique qui nécessiterait une fusion *a posteriori* de l'ensemble des MNS élémentaires et par conséquent une gestion des raccords et des recouvrements.

Ce prototype est basé sur un concept d'appariement multi-images (corrélation multi-fenêtres) guidé depuis l'espace objet (voir figure B.3.1). Ce prototype suit l'algorithmie suivante. Pour chaque nœud (x,y) de la grille du MNS, on construit un profil de corrélation regroupant tous les coefficients de corrélation calculés pour chacun des z possibles dans un intervalle plausible conditionné par une connaissance *a priori* de la zone (BDZ, BDTopo, etc.).

Pour chaque z donné, on détermine dans l'ensemble des images, les positions (i_k, j_k) hypothétiques correspondantes ($0 < k < n$ où n est le nombre d'images).

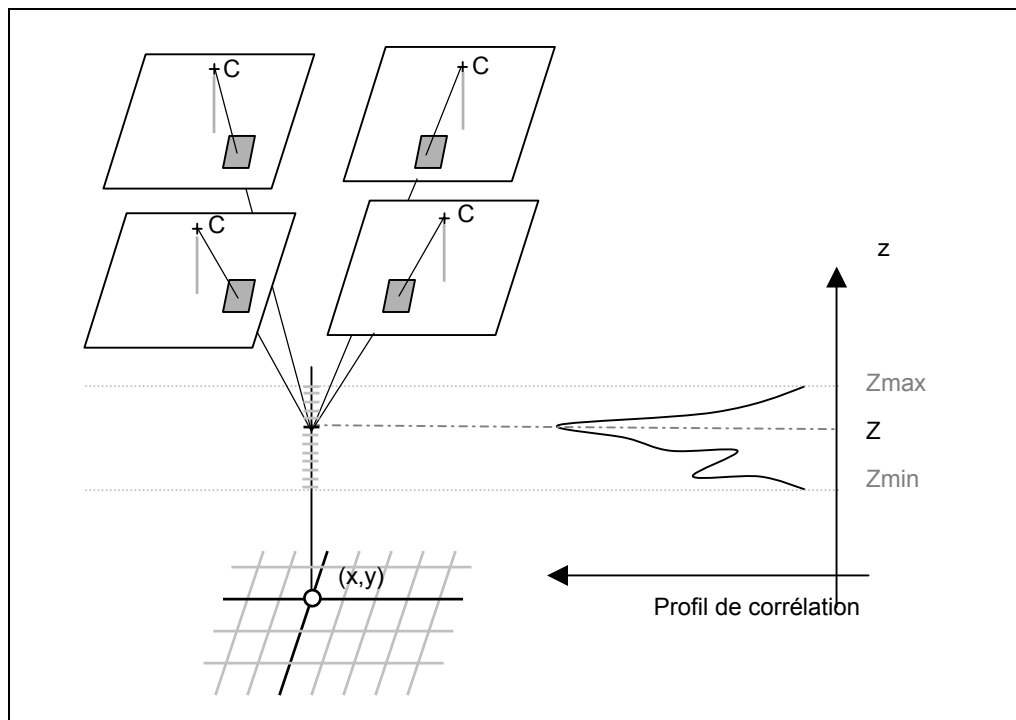


figure B.3.1 : principe de notre appariement multi-vues

La vraisemblance de cette hypothèse est donnée par une mesure directe de la ressemblance de l'ensemble des fenêtres/imagettes centrées sur les (i_k, j_k) . Le Z "estimé" que l'on retient pour un (x, y) est généralement le z pour lequel on atteint la valeur optimale du coefficient de corrélation. De plus, pour ce (x, y) et pour ce Z estimé, on peut calculer directement la radiométrie correspondante dans l'ortho-image à partir de l'ensemble des radiométries observées aux points (i_k, j_k) (en prenant par exemple la moyenne ou la médiane des valeurs).

En pratique, on ne traite pas séquentiellement l'ensemble des nœuds de la grille du MNS. En effet, les (i_k, j_k) ont en réalité des positions subpixelaires. Les imagettes à corrélérer sont donc obtenues par rééchantillonnage des images autour des (i_k, j_k) .

Afin d'éviter la redondance d'opérations de rééchantillonnage pour des points voisins dans le MNS, on rééchantillonne (projette) pour un z donné l'ensemble des images et on calcule pour tous les (x, y) et pour ce z donné, un coefficient de corrélation. Si ce z est le bon, ces images rééchantillonnées doivent être localement superposables et le coefficient de corrélation optimal. Et on recommence itérativement ce processus pour l'ensemble des z .

Un coefficient de ressemblance multi-imagettes

Comment peut-on mesurer la ressemblance d'un ensemble d'imagettes ? L'extension directe du coefficient de corrélation croisé normalisé centré à n imagettes rectangulaires n'est pas possible car celui-ci donne le produit scalaire et donc une information sur l'angle formé par les deux vecteurs de textures V , dont les composantes sont formées par l'ensemble ordonné des radiométries des différents pixels des imagettes.

Nous avons donc défini un coefficient de Corrélation Multi-Image (CMI) adapté :

$$0 \leq \text{MIC}(i_1, j_1, i_2, j_2, \dots, i_n, j_n) = \frac{\text{Var}(\sum_{k=1}^n \vec{v}_k(i_k, j_k))}{\sum_{k=1}^n \text{Var}(\vec{v}_k(i_k, j_k))} \leq n$$

Pourquoi ce coefficient de corrélation ? Si les textures des vignettes images se ressemblent, i.e. les vecteurs de textures sont colinéaires, le critère de similarité CMI est maximal. On remarquera que si $n = 2$, ce coefficient n'est autre que le coefficient de corrélation croisé normalisé centré usuel.

Pondération radiométrique

Vu la stabilité radiométrique des images de la caméra numérique (ce qui n'est pas le cas des clichés scannés) et sous l'hypothèse que la plupart des objets qui décrivent le paysage ont des caractéristiques lambertiennes, nous imposons une contrainte supplémentaire sur la radiométrie " absolue " des voisinages sous la forme d'une fonction de pondération. Cette pondération est nécessaire.

En effet, notre critère de ressemblance " CMI " est " moyenne centrée " ; celui nous donne donc la ressemblance des textures des voisinages et non pas la ressemblance de la radiométrie des voisinages ce qui peut provoquer des appariements aberrants (voir figure B.3.2).

Notre nouvelle mesure de similarité " CMIA " peut s'exprimer de la manière suivante :

$$CMIA (i_1, j_1, \dots, i_n, j_n) = CMI (i_1, j_1, \dots, i_n, j_n) \cdot \exp - \left(\frac{\text{Var}_k(I_k(i_k, j_k))}{k} \right).$$

où Var est la variance, $I(i,j)$ est la radiométrie au pixel (i,j) et k un facteur de normalisation.

L'application de cette fonction de pondération permet d'obtenir un critère de similarité qui caractérise à la fois la ressemblance texturale et la ressemblance radiométrique des voisinages. Celui-ci est donc plus robuste et discriminant (voir figure B.3.2).

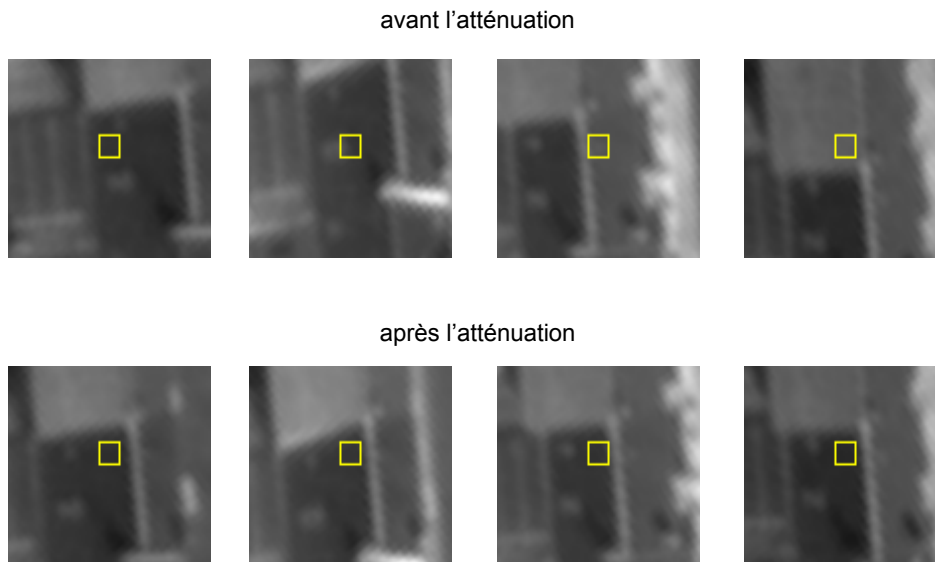


figure B.3.2

Comparaison des homologues trouvés avant et après atténuation. Le petit rectangle représente l'emprise de la fenêtre de corrélation dans l'image. Ici une fenêtre 3 x 3.

Résultats et commentaires

L'augmentation du nombre d'observations permet, d'une part, de réduire de manière considérable les ambiguïtés d'appariement que l'on rencontre en stéréoscopie et par conséquent d'accroître la fiabilité du processus. Elle permet, d'autre part, grâce à la qualité des images de la caméra numérique de l'IGN, d'utiliser des tailles de fenêtres de corrélation de 3 x 3 (voir figure B.3.3).

On peut constater sur les résultats que le MNS est dense, que les micro-structures du relief sont bien restituées, et que les discontinuités sont bien localisées. Toutefois, la petite taille de la vignette est à l'origine d'appariements aberrants dans les zones très homogènes qui n'apparaissent pas avec une fenêtre 5 x 5 (voir figure B.3.4). Par contre, une fenêtre 5 x 5 restitue moins bien les micro-structures du relief et les discontinuités.

Travailler depuis l'espace objet présente de nombreux avantages. On peut travailler de manière transparente sur des images de résolutions différentes et les paramètres s'expriment sous forme métrique. On évite également le rééchantillonnage en aveugle des échantillons 3-D issus de l'appariement stéréoscopique des images afin de constituer une grille MNS régulière dans l'espace objet.

En effet, ces échantillons, issus classiquement d'une mise en correspondance guidée depuis l'espace d'une image, ont par construction une distribution spatiale irrégulière. Enfin, son avantage principal pour notre étude est qu'il permet d'une part de traiter N images de manière naturelle et transparente et d'autre part de construire en même temps le MNS et l'ortho-image associée (voir figure B.3.5).

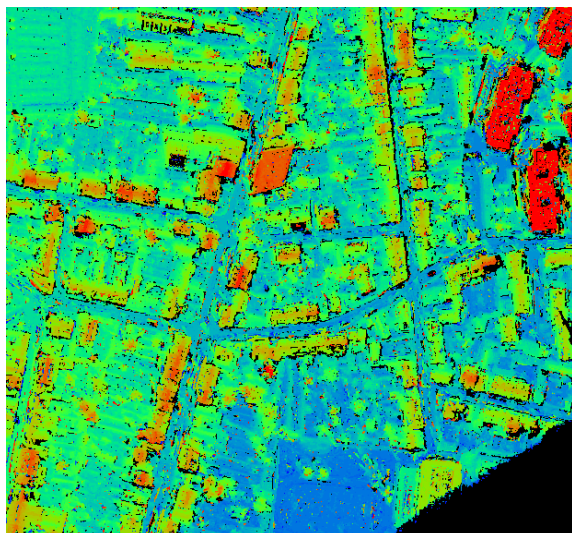


figure B.3.3

MNS avec quatre images et avec une fenêtre 3 x 3.

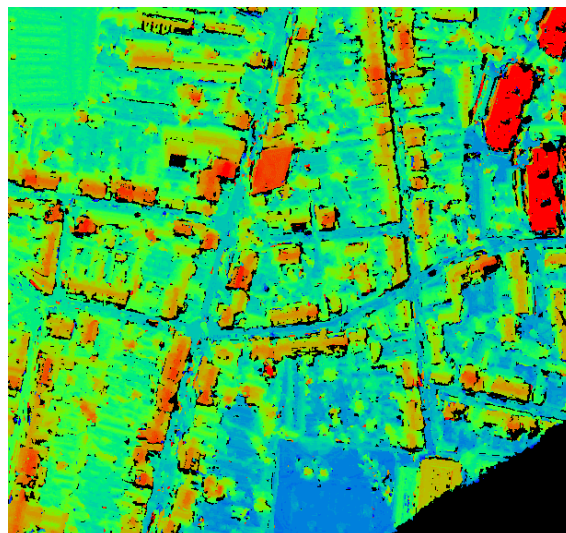


figure B.3.4

MNS avec quatre images et avec une fenêtre 5 x 5.



figure B.3.5

Ortho-image construite avec le MNS de la figure B.3.3 (Mission du Mans avec 40 cm de résolution et un recouvrement latéral et longitudinal de 50%).

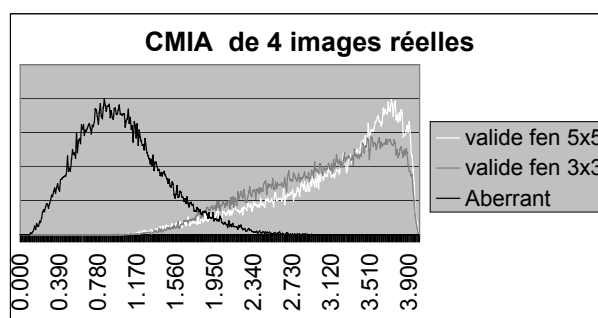


figure B.3.6

Histogramme des coefficients de corrélation.

“ Auto-évaluation ” des mesures et filtrage automatique des points aberrants

Bien entendu, le coefficient de ressemblance multi-images ne donne pas une probabilité mais un bon auto-indicateur de la fiabilité de la mesure. Les figures B.3.7 et B.3.8, page suivante, montrent que les points aberrants du MNS ont des coefficients de corrélation plus faibles. Une expérience simple d'analyse des distributions de coefficients de corrélation le montre de manière plus nette.

Considérons les distributions des coefficients réels trouvés et aberrants. Nous appelons coefficients réels, les coefficients de corrélation issus de notre processus, et coefficients aberrants, les coefficients de corrélation obtenus par le même processus tout en se plaçant délibérément dans un intervalle altimétrique de même amplitude mais ne contenant pas le relief. On peut observer que les deux distributions sont peu mélangées (voir figure B.3.6). D'après l'observation des courbes on peut dire que les points avec des scores de corrélation au dessus de 2,95 sont quasiment sûrs à 100 % et ceux au dessous de 1,0 sont faux à 100%. Entre les deux ?

La relative séparation des distributions dans le cas du multi-vues permet de définir un critère de filtrage valide (voir figure B.3.9, page suivante). On notera que la séparation des deux distributions est d'autant plus importante que le nombre d'images est important. Si l'on se ramène à un traitement stéréoscopique le mélange des distributions est très important ce qui explique la difficulté et l'impossibilité à définir un critère de fiabilité et un seuil de rejet satisfaisant sur la valeur du coefficient de corrélation. Un seuil bas conserve un grand nombre de points aberrants. Un seuil haut rejette un très grand nombre de points valides.

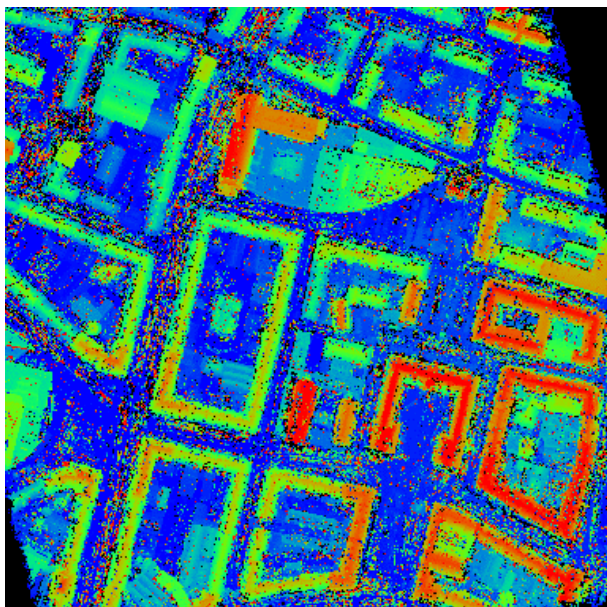


figure B.3.7

MNS brut (sans filtrage) obtenu avec quatre images et avec une fenêtre 3 x 3, (mission sur Amiens avec 40 cm de résolution et un recouvrement latéral et longitudinal de 60%).

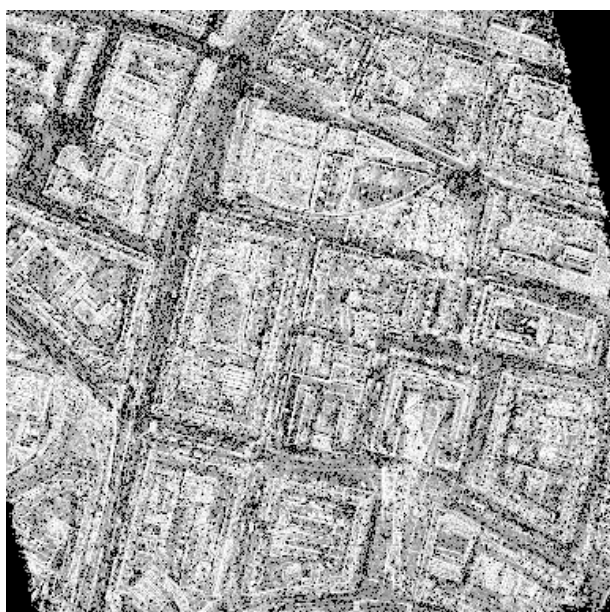


figure B.3.8

Coefficients de corrélation.

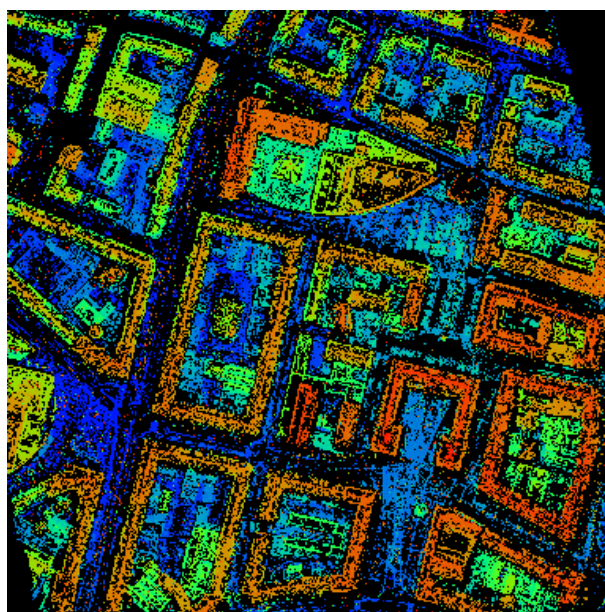


figure B.3.9

MNS filtré avec un seuil de 2,95 sur le CMIA.

Travaux en cours : gestion des zones manquantes

La fonction de pondération radiométrique que nous avons décrite précédemment pénalise bien entendu les zones qui ne sont vues (parties cachées) ou plutôt qui ne sont "bien vues de la même manière" (surfaces spéculaires, véhicules mobiles) que dans un sous-ensemble d'images.

Notre processus d'appariement tend volontairement à restituer uniquement les points du paysage vus de façon assez semblable sur les n images simultanément. Aussi, il restera des trous de corrélation sur un certain nombre de zones : zones homogènes, parties masquées dans une ou plusieurs images, zones liées aux véhicules mobiles, surface ayant donné lieu à des réflexions spéculaires, etc.

Aussi notre stratégie générale sera de compléter progressivement et hiérarchiquement (par rétroaction) le MNS à partir des mesures les plus sûres (à l'aide du filtrage) afin d'éviter la propagation des erreurs dans les traitements.

Pour restituer les zones homogènes on peut recommencer le même processus (corrélation, pondération, filtrage) avec des tailles de vignettes plus grandes (5 x 5 et 7 x 7). De manière à compléter les zones encore manquantes, on applique localement la même stratégie mais sur un sous-ensemble déterministe d'images.

Pour les parties cachées, on détermine à l'aide des échantillons sûrs du MNS et d'un algorithme de z-buffer, le sous-ensemble d'images sur lequel un point du MNS est vu et on applique le même processus que précédemment mais limité à ce sous-ensemble d'images. Pour les surfaces spéculaires et véhicules mobiles, on prend la médiane de tous les coefficients calculés sur tous les sous ensembles possibles de $n - 1$ images. Le traitement hiérarchique de ces zones est un travail en cours.

Conclusion

Ce nouveau prototype de corrélateur, quoique très simple en termes de mise en œuvre et de stratégie, donne des résultats très prometteurs avec des propriétés particulièrement intéressantes pour la production d'ortho-images, pour la détection de micro-structures du relief, et pour la reconstruction du bâti. D'une part, les MNS restitués sont denses, fiables, précis et reproduisent correctement la morphologie du relief (discontinuités, pentes importantes, ruptures de pentes, micro textures 3-D, etc.). D'autre part, on possède pour chacune des mesures du MNS, un indicateur de la "fiabilité" qui peut être utilisé soit pour une rétro-action directe sur nos algorithmes et une vraie stratégie hiérarchique, soit pour mettre en évidence des zones "incertaines", afin que d'autres traitements qui s'appuient sur les MNS puissent en tenir compte.

Perspectives

Jusqu'à présent dans cette étude, le processus de détermination altimétrique a été indépendant pour chaque (x,y). La valeur retenue pour chacun de ces processus correspond au maximum de corrélation du profil. Bien que la multi-corrélation permette de faire sortir distinctement les bons maxima la plupart du temps, il arrive que la bonne solution se trouve dans les maxima secondaires.

Une des perspectives sera de conserver pour chaque processus, un ensemble de valeurs de Z pour lesquelles on observe des maxima locaux du coefficient de corrélation et on utilisera des techniques d'optimisation locales ou globales (relaxation, programmation dynamique, ou autres) dans l'espace objet de manière à étudier la cohérence spatiale de ces maxima dans un voisinage de (x,y) afin de déterminer une "meilleure" solution pour chaque mesure.

Références

- [BAGHLI 99] I. Baghli : *Reconstruction de Modèles Numérique d'Élévation à partir d'images aériennes multi-vues panchromatique*, rapport de stage, DEA de traitement du signal et des images, ENSEA-UCP, septembre 1999.
- [CANU et al 95] D. Canu, N. Ayache & J.-A. Sirat : *Accurate and robust stereovision with a large number of aerial images*, SPIE, pp. 152- 160, Paris, 1995.
- [FOURNIER 99] J. Fournier : *Analyse 3-D multi-vues de scènes urbaines en imagerie aérienne haute résolution*, rapport de stage, DEA de traitement du signal et des images, ENSEA-UCP, septembre 1999.
- [FUCHS et al 99] F. Fuchs, H. Le Men : *Building reconstruction on aerial images through multi-primitive graph matching*, in Proceedings of the 2nd IAPR workshop on graphbased representations, laboratoire MATIS, IGN SR 990023/C-COM, 1999.
- [GABET et al 94] L. Gabet, G. Giraudon & L. Renouard : *Construction automatique de Modèles Numériques de Terrain à haute résolution*, in Bulletin SFPT n° 135, 1994.
- [LELOGLU et al 98] U.M. Leloglu, M. Roux & H. Maître : *Dense urban DEM with three or more high resolution aerial images*, ISPRS Symposium on GIS, Between Vision and applications, Stuttgart (D), septembre 1998.
- [MAYER 98] D. Mayer : *Fusion de résultats de corrélation*, rapport de stage, DEA Sciences de l'Information Géographique, ENSG-ENST-Université de Marne-la-Vallée, laboratoire MATIS, IGN SR 9800 /C-STA, septembre 1998.
- [OKOTUMI et al 91] M. Okotumi and T. Kanade : *Multiple Baseline Stereo*, IEEE transactions on Pattern Analysis and Machine Intelligence, vol. 15, n° 4, avril 1993.

Publications

- [CORD et al 98] M. Cord, N. Paparoditis & M. Jordan : *Dense, reliable and depth discontinuity preserving DEM computation from H.R. V. images*, in Proceedings of ISPRS Commission II symposium, Data Integration : Systems and techniques, vol. 32, n° 2, Cambridge (UK), laboratoire MATIS, IGN SR 980040/C-COM, juillet 1998.
- [JIBRINI et al 00] H. Jibrini, M. Pierrot Deseilligny, N. Paparoditis & H. Maître : *Reconstruction 3-D de bâtiments à partir de données cadastrales vectorisées et d'images aériennes*, RFIA'2000, pp. 299-306, laboratoire MATIS, IGN SR 00-002/C-COM-HJ, février 2000.
- [PAPARODITIS et al 98] N. Paparoditis, M. Cord, M. Jordan and J.-P. Cocquerez : *Building detection and reconstruction from mid and high resolution images*, in Journal of Computer Vision and Image Understanding, vol. 72, n° 2, pp. 122-142, laboratoire MATIS, IGN SR 980039/C-ART, novembre 1998.

