

Extraction de primitives 3D pour la reconstruction 3D de bâti milieu urbain à partir de couples d'images satellites haute résolution

chehata@math-info.univ-paris5.fr
nesrine.chehata@ign.fr

Ce travail est réalisé avec l'IGN¹, et le CNES².

Résumé: Ce travail doctoral s'insère dans un projet global d'interprétation de scènes urbaines à partir d'images satellites haute résolution. L'interprétation porte sur deux thèmes, le bâti et les routes. Le projet présenté traite uniquement le thème du bâti. Les algorithmes développés seront adaptés aux données images du futur satellite PLEIADE HR, le successeur de SPOT5. On disposera en entrée de couples d'images satellites panchromatiques, de haute résolution :40-60 cm, ayant un faible rapport Base sur Hauteur B/H=0.1,0.2.

Cette étude détaille trois algorithmes utilisés dans la stratégie globale d'extraction de primitives 3D fiables, qui constitue une étape préliminaire importante pour faciliter la génération d'hypothèses de bâtiments.

Les trois algorithmes utilisent, d'une part, la contrainte photométrique à chaque étape, pour tirer profit du faible rapport B/H du couple et traitent d'autre part les images de façon symétrique.

Cependant, chaque algorithme peut être utilisé indépendamment des autres en fournissant à chaque fois la primitive spécifique en entrée.

Mots Clés: Traitement d'images satellites haute résolution, stéréo, reconstruction 3D de bâtiments, primitives 3D, surfaces planes, délimitation de toits

Abstract : This study is a preliminary operation of a global project which goal is to interpret urban scenes from satellite images in order to recognize and classify roads and buildings. Actually, the research is focused on buildings. We plan to apply these algorithms on the forthcoming satellite data of PLEIADE HR which will follow SPOT5. The input data consists of a panchromatic stereopair of satellite images, at a high resolution of 40-60 cm and a low Base to Height ratio B/H= 0.1,0.2.

Three algorithms are described involving a global strategy for extracting reliable 3D primitives which is a key step and an important preliminary operation to make the generation of building hypothesis easier. All algorithms use photometric constraints at each step to take advantage from low B/H and handle both views in a symmetric way. However, each algorithm can be used independently providing the specific primitives as input data.

Key words: Image Processing, HR satellite images, stereo, 3D building reconstruction, 3D-primitives, planar patches, roofs delimitation

1. Objectifs

Le but du projet est d'interpréter les scènes urbaines à partir des images satellites afin de reconnaître, classifier les routes et les bâtiments et extraire les objets géographiques 3D.

Il n'existe, à ce jour, aucun système qui permet de résoudre le problème de façon automatique. On suppose, dans un premier temps, que le réseau routier est fiable et on traite seulement le thème bâti.

La reconstruction de bâtiments en milieu urbain est un problème extrêmement complexe compte tenu de la diversité des scènes observées. Les approches les plus couramment développées consistent à calculer par des techniques de corrélation [Baillard,97] un Modèle Numérique de Surface qui fournit une représentation maillée en 2D ½ de la scène urbaine. L'extraction de ce MNS est en général suivie d'une étape de segmentation qui permet d'estimer les surfaces planes en utilisant des techniques d'estimation robustes [Fradkin et al.,00].

Cependant, la présence de superstructures (cheminées, chiens assis) ou de motifs répétitifs rendent les

¹ IGN : Institut Géographique National, St Mandé, France.

² CNES : Centre National d'Etudes Spatiales, Toulouse, France
Service QTIS (Qualité et Traitement d'Images Spatiales) .

techniques de corrélation sensibles et l'utilisation de modèles de bâtiments [Fuchs and LeMen,99] [Willuhn and Ade,96] *a priori* de plus en plus importante pour contraindre la reconstruction et permettre une description sémantique de la scène. Dans cette optique, une extraction de primitives 3D fiables est un élément clef et souvent une étape préliminaire pour faciliter la génération d'hypothèses de bâtiments. Ces primitives 3D sont en général des coins [Fischer et al.,98], des plans ou des segments 3D [Fuchs and LeMen,99] [Willuhn and Ade,96].

2. Méthodologie

Pour l'extraction de primitives 3D, on a choisi, dans notre cas, une approche basée sur les segments 3D. Le principe est illustré dans la Figure 1.

Les segments 3D sont obtenus par mise en correspondance des segments 2D dans les deux images. Ensuite l'extraction des plans 3D revient à rechercher des demi-plans pivotant autour des segments 3D extraits. Finalement un algorithme de délimitation des toits par relaxation probabiliste permettra d'obtenir les facettes 3D finales.

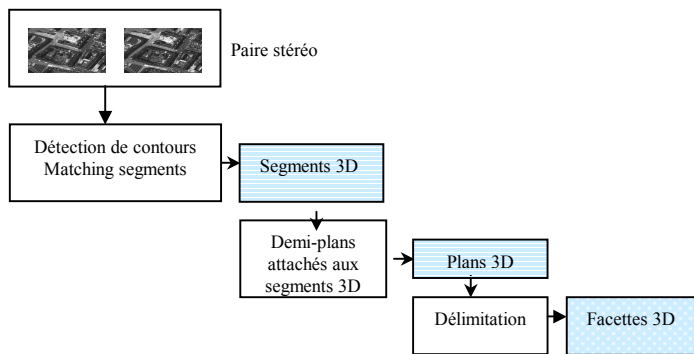


Figure 1. Schéma global de l'extraction de primitives 3D

Les trois étapes seront détaillées dans ce qui suit.

2.1. Extraction de segments 3D

Les contours dans les deux images sont obtenus à l'aide d'un détecteur de Canny Deriche [1] avec une précision subpixelaire. Les lignes finales sont obtenues par régression orthogonale. Dans cette partie, la mise en correspondance des segments 2D et la reconstruction finale des segments 3D seront détaillées.

Afin de réduire la combinatoire, on utilise d'une part la contrainte épipolaire sur les extrémités des segments et une contrainte sur les altitudes Min et Max de chaque point en utilisant en entrée un Modèle Numérique d'Élévation (obtenu par des techniques de corrélation) qui permet de réduire l'espace de recherche.

La contrainte géométrique est réduite à une contrainte de recouvrement. Pour chaque paire de segments satisfaisant la contrainte géométrique ci-dessus, on applique une contrainte photométrique qui prend en compte le voisinage du segment.

En fait, chaque segment est considéré comme une liste de points. La géométrie épipolaire permet d'obtenir une correspondance point à point entre les segments homologues. Les scores de corrélation sont fiables grâce au faible rapport B/H du couple. Le score de la contrainte photométrique est obtenu en moyennant les scores de corrélation individuels de tous les points du segment.

Le score final prend en compte les deux contraintes comme suit :

$$Score_{final} = 0.4 * Score_{geom} + 0.6 * Score_{photo} \quad (1)$$

Un ensemble d'associations consistant est retenu suivant le schéma de « the winner takes all ».

Dans cette approche, les deux images sont traitées de façon symétrique et l'ensemble des correspondances possibles est élagué sous une contrainte de réciprocité qui ne retient que les associations fiables.

Pour la reconstruction 3D des segments, on considère un couple de segments homologues, on projette les extrémités 2D sur la droite 3D correspondante et on applique une stratégie d'union qui consiste à garder les extrémités les plus éloignées. On obtient ainsi le segment 3D final.

Pour valider les segments reconstruits, on utilise un volume de recherche dérivé du MNE par dilatation auquel doivent appartenir les extrémités des segments.

L'apport de cette approche consiste principalement dans l'utilisation de la contrainte photométrique et l'utilisation du MNE pendant l'étape de mise en correspondance et pour valider la reconstruction 3D des segments.

Les segments 3D obtenus permettront de guider par la suite l'extraction des plans 3D.

2.2. Extraction de plans 3D

Le deuxième algorithme permet d'extraire automatiquement des plans 3D attachés aux segments 3D extraits.

Cette méthode essentiellement dérivée de [Baillard,99] a été adaptée à notre contexte. La méthode consiste à extraire des demi-plans fiables définis uniquement par un segment 3D et les scores de similarités de son voisinage entre les deux vues. Cette méthode a été choisie puisqu'elle utilise les deux contraintes géométrique et radiométrique. L'objectif est de déterminer pour chaque segment 3D, à travers deux familles de 1-paramètre $\theta : \pi(\theta)$ et $\pi(\theta')$ s'il y a deux demi-plans attachés ou pas. Si les demi-plans existent, on cherche la meilleure estimation de θ et θ' . Les deux familles

représentent les deux demi-plans de part et d'autre de la référence verticale attachée au segment. Tout d'abord, comme on peut le voir sur la Figure 2, on détermine un ensemble de points texturés au voisinage du segment afin d'avoir une fonction de similarité discriminante. Ainsi, la corrélation sera juste calculée au voisinage de ces points d'intérêt *POI* [Harris 88]. Etant donnée une hypothèse angulaire de θ , le plan $\pi(\theta)$ définit une correspondance point à point entre les deux images. Si le plan est correct, le score de corrélation entre les points correspondants doit être élevé. La fonction de similarité $Sim(\theta)$ doit être sélective et robuste aux point aberrants. Elle est définie comme suit et varie entre [0,1]

$$Sim(\theta) = \sum_{i=1,2} \iint_{POI} Cor^2(x, x') \quad (2)$$

Le terme de corrélation $Cor^2(x, x')$ correspond à la corrélation centrée normalisée entre le point x de la première image et son point correspondant x' dans la deuxième image. Le terme est mis au carré pour donner plus de poids aux scores élevés et pour rendre la fonction de similarité plus sélective.

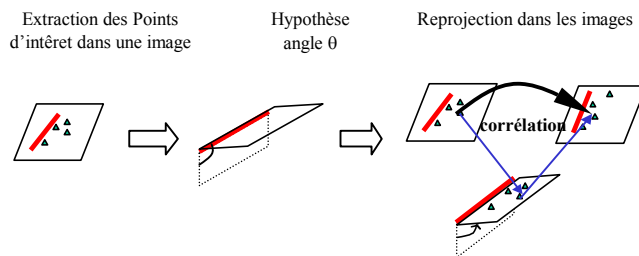


Figure 2. Extraction de demi-plans 3D

La corrélation est discriminante sur des régions texturées d'où l'intérêt d'utiliser des *POI*. Dans notre cas, les *POI* sont calculés dans les deux images et traités de façon symétrique donc le score de corrélation final pour un plan donné est moyenné sur les deux vues. Ceci augmente la fiabilité des plans 3D détectés.

On utilise pour la corrélation des fenêtres adaptatives qui suivent les bordures des toits et les contours extraits. On garde les plans dont le score est supérieur à un seuil préfixé ($S_{th} = 0.5$). L'angle optimal θ renvoie le maximum de la fonction de similarité $Sim(\theta)$ avec $\theta \in [20^\circ, 160^\circ]$ par rapport à une référence verticale. Le processus est itéré pour θ' pour extraire le deuxième demi-plan.

Par la suite, une étape de fusion et d'intersection des plans 3D extraits permettra de grouper les lignes, valider certains segments et retrouver des segments qui n'ont pas été détectés par le premier algorithme.

A ce niveau de l'algorithme, chaque segment a deux hypothèses de demi-plans. On obtient une scène de ce type (cf. Figure 3).

Finalement, la validation de ces hypothèses se fera lors de l'étape de fusion par recouplement d'hypothèses,

c'est-à-dire par validation d'une des hypothèses du segment considéré à l'aide des hypothèses des segments avec lesquels il est en interaction lors de la fusion.

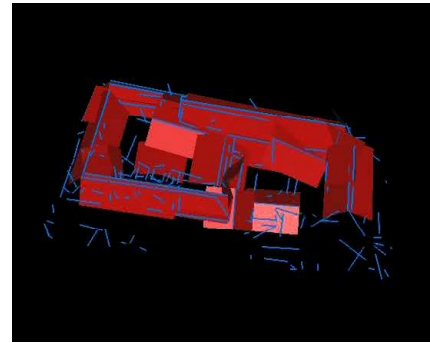


Figure 3. Scène de plans 3D

2.3. Délimitation des toits par relaxation probabiliste

L'étape précédente a permis de déterminer les positions angulaires des demi-plans qui vont servir de support aux pans de toits. En prévision de la phase de fusion où il faudra déterminer les segments qui délimitent le toit, on essaye dans cette partie de déterminer l'étendue de chaque pan de toit, de manière à travailler non plus sur des demi-plans, mais avec des polygones plans positionnés dans l'espace. Pour ce faire on va recourir à la relaxation probabiliste. Etant donné qu'à ce niveau notre demi-plan est positionné, on travaillera plus dans ce qui suit sur des imagerie issues de la reprojection des abords du segment considéré sur le plan $\pi(\theta)$ dans l'espace objet.

La Figure 4 illustre le schéma global de l'algorithme de délimitation des toits.

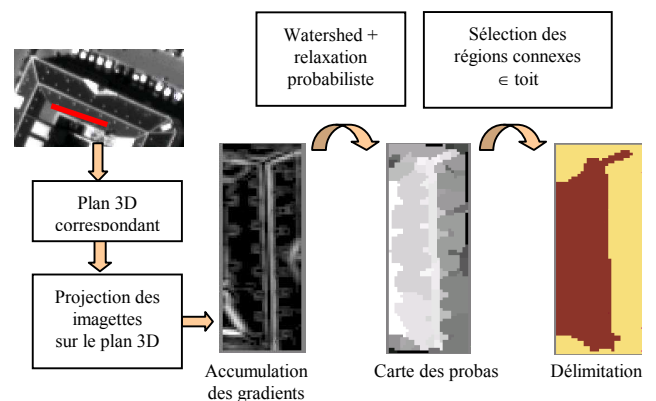
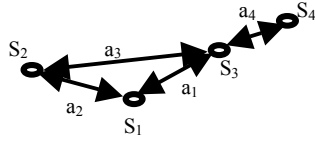


Figure 4 : Schéma global de la délimitation de toits

Tout d'abord, il convient de préciser que la relaxation probabiliste [Fu et Yan,97] s'effectue sur un graphe, c'est-à-dire un ensemble d'objets appelés sommets ayant des caractérisations communes et des relations entre eux matérialisées par des arêtes. Voici un exemple de graphe :



L'idée fondamentale des méthodes de relaxation sur un graphe est de faire évoluer de manière itérative des caractéristiques de chacun des sommets en fonction des caractéristiques de ses voisins, de manière à propager l'information dans le graphe, et ce jusqu'à une position d'équilibre. Dans le cas de la relaxation probabiliste, les sommets sont caractérisés entre autre par des états probabilisés et l'on fait évoluer ces probabilités.

Nous avons utilisé la relaxation probabiliste au sens de Rosenfeld ; on a donc la relation itérative suivante pour la probabilité à l'itération $k+1$ que le sommet i soit dans l'état λ :

$$p_i^{(k+1)}(\lambda) = p_i^{(k)}(\lambda) \frac{1 + s_i^{(k)}(\lambda)}{1 + \bar{s}_i^k}$$

$$s_i^{(k)}(\lambda) = \sum_{j=1}^n d_{ij} \sum_{\lambda'=1}^m c_{ij}(\lambda, \lambda') p_i^{(k)}(\lambda')$$

où :

$$\bar{s}_i^{(k)} = \sum_{\lambda=1}^m p_i^{(k)}(\lambda) s_i^{(k)}(\lambda) \quad (3)$$

avec:

- Un ensemble de m labels : $\{\lambda_i, i=1, \dots, m\}$ pour chaque région
- $c_{ij}(\lambda, \lambda')$ représente la mesure de compatibilité que le sommet i ait le label λ quand le sommet j a le label λ' . Cette compatibilité peut être considérée comme une probabilité conditionnelle qui satisfait la condition suivante :

$$\forall \lambda, \lambda' \quad 0 \leq c_{ij}(\lambda, \lambda') \leq 1$$

$$\sum_{\lambda} c_{ij}(\lambda, \lambda') = 1 \quad (4)$$

- d_{ij} quantifie l'influence de la région j sur la région

i . Elle vérifie $\sum_j d_{ij} = 1$.

Ainsi, lors d'une relaxation probabiliste, trois paramètres sont à régler :

- La fonction d'influence d_{ij}
- Les probabilités conditionnelles c_{ij}
- L'ensemble des probabilités initiales dénotées $p_i^0(\lambda), \lambda=1, \dots, m$

Application à notre problème

Dans notre cas nous avons appliqué la relaxation probabiliste à un graphe d'adjacence, c'est-à-dire un

graphe où les sommets sont des régions et les arêtes indiquent une frontière commune entre deux régions.

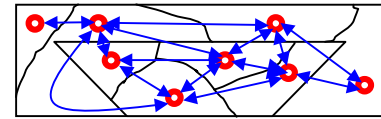


Image des Régions

L'image des régions est obtenue par une segmentation réalisée par Watershed.

Il reste à définir pour ce graphe les états probabilisés, qui sont au nombre de deux : « \in toit » et « \notin toit ».

Les trois paramètres de la relaxation probabiliste ont été réglés de la manière suivante:

Les probabilités initiales d'appartenance au toit ont été fixées à 99% pour les régions adjacentes au segment et à la valeur moyenne de la fonction de similitude sur la région dans les autres cas. La fonction de similarité correspond au score de corrélation entre les deux images re-projetées sur le plan $\pi(\theta)$. Le score de corrélation est pondéré de $\frac{1}{d}$, où d est la distance du

barycentre de la région au segment. En effet, on est sûr que les régions adjacentes au segment appartiennent au toit, et le fait d'avoir une zone de forte probabilité permet de propager cette information aux régions alentours qui sont susceptibles d'appartenir au toit. D'autre part, le score de similitude est lui aussi significatif de l'appartenance au toit, car dans les zones en dehors du toit, la re-projection dans l'espace objet n'est pas pertinente et il y a donc des déformations, d'où un score de corrélation faible, alors que les zones contenues dans le toit sont elles re-projetées correctement, ce qui donne un score de similitude élevé.

Le choix des probabilités conditionnelles n'est pas critique, on choisit de les fixer de la manière suivante :

$$\begin{cases} c_{ij}(\lambda, \lambda) = 0.6 \\ c_{ij}(\lambda, \lambda') = 0.4 \end{cases}$$

avec $\lambda, \lambda' \in \{ " \in \text{toit} ", " \notin \text{toit} " \}$ (5)

Les probabilités conditionnelles aussi proches que possible de 0,5 de manière à favoriser une évolution lente des probabilités, préconisée par la théorie.

La fonction d'influence est, par contre, un paramètre critique. Elle doit satisfaire un certain nombre de conditions préfixées :

- être croissante en fonction de la longueur de la frontière commune entre les deux régions ;
- être décroissante en fonction de la valeur moyenne du gradient le long de la frontière commune aux deux régions. En effet, plus il existe un contour marqué entre deux régions, plus il y a de chance qu'elles n'appartiennent pas au même objet dans l'image ;
- le gradient doit l'emporter sur la longueur commune.

Ainsi, on utilise la fonction d'influence suivante :

$$d_{ij} = \frac{L_{ij} \cdot e^{-\lambda \langle grad \rangle_{ij}}}{\sum_{k=0}^n L_{ik} \cdot e^{-\lambda \langle grad \rangle_{ik}}} \quad (6)$$

où :

- L_{ij} est la longueur de la frontière commune aux deux régions i et j ;
- $\langle grad \rangle_{ij}$ est la valeur moyenne du gradient sur la frontière commune aux régions i et j .
- On utilise $\lambda = -3.5$. Cette valeur permet de délimiter les toits principaux de la scène. Dans les cas tangents, elle tend à renvoyer des toits plus petits qu'en réalité, ce qui est préférable pour la lisibilité de la scène 3D générée.

L'évolution du système conduit à une carte finale des probabilités (Figure 5). Les régions les plus claires ont de fortes probabilités d'appartenir au toit. Finalement, une coupe dans le graphe d'adjacence permet de sélectionner les régions connexes qui ont une probabilité d'appartenir au toit supérieure à un seuil fixé $P_s = 0.75$. La plus grande région sera choisie comme délimitation finale du toit.

3. Résultats

Les premiers résultats sont présentés sur des images simulées par le CNES/QTIS sur le site de Toulouse à 60 cm de résolution (Figure 5,5) ou bien sur des images aériennes avec des faibles B/H=0.2 sur le site d'Amiens à 50 cm (Figure 6).

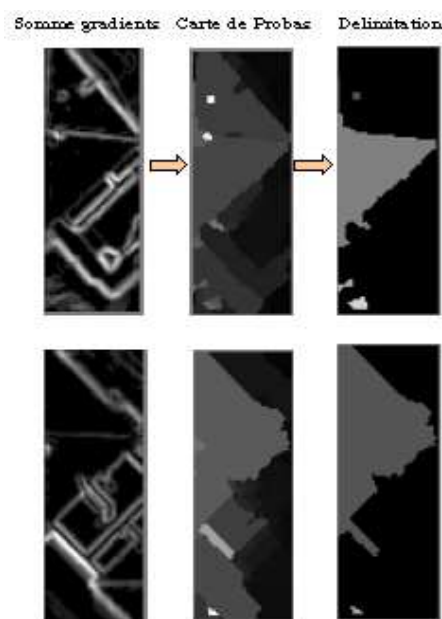


Figure 5 : Délimitation plans

Performance :

Cette approche présente l'avantage d'utiliser la contrainte photométrique à chaque étape mettant à profit le faible rapport B/H de la paire stéréo. Le couple d'images est traité de façon symétrique.

Dans l'approche proposée, la qualité de reconstruction dépend de l'exactitude du segment 3D au départ. S'il y a des segments manquants, des plans entiers risquent d'être manqués.

De plus, un faible rapport B/H entraîne une faible précision de reconstruction 3D. Par cette approche, les erreurs altimétriques sont propagées à chaque étape de l'algorithme général.

Des résultats numériques sont présentés sur le site d'Amiens :

Les segments sont uniquement évalués avec la déviation angulaire par rapport à la ligne de référence correspondante. La longueur du segment n'est pas prise en compte pour l'évaluation.

Un demi-plan π_i est considéré comme correct s'il joint à son segment attaché i un autre segment qui lui est colinéaire. Finalement la délimitation est évaluée visuellement.

Comme on peut le voir dans le tableau 1, à cause du manque de précision de la reconstruction 3D des segments, seulement 29 % des segments sont bien positionnés. C'est le principal inconvénient de la méthode qui propage ensuite les erreurs altimétriques à chaque étape. Par conséquent, le reste des évaluations sera fait uniquement sur les segments considérés comme fiables.

Pour la détection de plans 3D, 66% des plans attachés aux segments fiables sont détectés. Parmi les plans détectés, 73 % sont corrects. Les erreurs sont principalement dues à des façades. En effet, les images re-projetées sur les façades sont très déformées et par conséquent les scores de corrélation ne sont pas fiables.

Finalement les plans corrects sont délimités correctement dans 68 % des cas. Les erreurs de délimitation sont essentiellement dues aux segments dont les longueurs sont surestimées au départ, dans ce cas, certaines régions proches du segment, initialisées comme appartenant au toit, ne le sont pas en réalité. Ce type d'erreur de délimitation peut être résolu par intersections des facettes 3D.

On peut ainsi voir que les deux derniers algorithmes donnent des résultats satisfaisants si les primitives en entrée sont fiables.

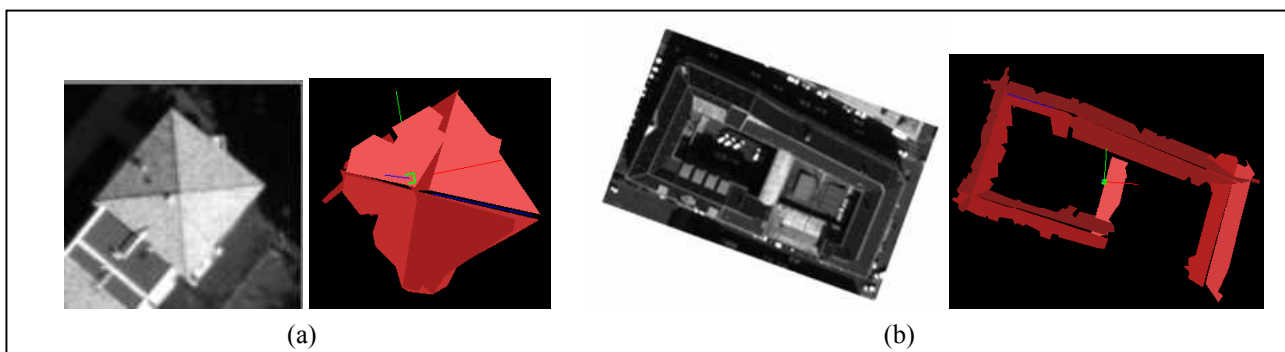


Figure 6: (a)Toulouse 60 cm (b) Amiens 50cm

DétECTION DE SEGMENTS 3D	129 segments		
	37 Corrects (29 %)		89 Faux (71 %)
DétECTION DE PLANS 3D	15 plans attachés manquants (34 %)	30 plans attachés détectés (66%)	7 faux plans détectés
	30 plans détectés		
DÉLIMITATION DE FACETTES 3D	22 plans corrects (73 %)		8 faux plans (27 %)
	22 Bons Plans		
	15 délimitations correctes (68 %)	7 fausses délimitations (32 %)	

Tableau 1. Résultats numériques sur le site d'Amiens 50cm

4. Conclusion

Le travail présenté détaille 3 algorithmes utilisés dans l'extraction de primitives 3D à partir d'un couple d'images satellites HR. Ces algorithmes ont l'avantage de traiter les deux images de façon symétrique mais aussi d'utiliser la contrainte photométrique à chaque étape afin de mettre à profit le faible B/H du couple.

Les trois algorithmes constituent les maillons d'une chaîne globale de reconstruction de facettes 3D, cependant ils sont indépendants les uns des autres et peuvent être utilisés séparément.

Par exemple, l'algorithme de délimitation peut être utilisé pour délimiter a priori les surfaces de corrélation pour la recherche de demi-plans pivotants au lieu d'utiliser des surfaces rectangulaires.

5. Perspectives

La prochaine étape consiste à réaliser la fusion et les intersections de plans 3D. Ceci permettra d'une part de simplifier la scène et de réduire le nombre de primitives et d'autre part de valider les segments 3D détectés précédemment ou de détecter de nouveaux segments par intersections de plans.

L'inconvénient principal de cette approche est la forte imprécision altimétrique de la reconstruction 3D des segments qui est propagée à chaque étape. Pour

s'affranchir de ce problème, on utilisera des primitives 2D plus fiables que les segments. On travaillera sur les segmentations en régions des deux images. La mise en correspondance de ces régions permettra de reconstruire directement des facettes 3D plus fiables

Une étude sera réalisée pour évaluer la précision altimétrique et planimétrique des primitives 3D extraites par cette approche afin de spécifier le modèle 3D final de bâtiments qui peut être extrait dans le contexte satellitaire.

6. Bibliographie

[Alcharroun, 01] " *Positionnement de plans 3D par corrélation multi-vues* ". Stage IGN 2001, pages 12-22.

[Baillard,97] C. Baillard. " *Analyse d'images aériennes stéréoscopiques pour la restitution 3-D des milieux urbains* ". PhD thesis, pages 20-70, ENST, 1997.

[Baillard.C and Zisserman.A, 99] " *Automatic reconstruction of piecewise planar models from multiple views* ". In Proceedings of the Conference on Computer Vision and Pattern Recognition, p 559-565, June 1999

[Fischer et al.,98] A. Fischer, T. Kolbe, F. Lang, A. Cremers, W. Förstner, L. Plümer, and V.Steinlage. "Extracting buildings from aerial images using hierarchical aggregation in 2D and 3D." *CVIU*. 72(2) :163-185. November 1998.

[Fradkin et al., 00] M. Fradkin, H. Maître et M. Roux. "Extraction de bâtiments à partir de plusieurs images aériennes". In *RFIA*, volume III:307-314, Paris, 2000.

[Fu et Yan,97]. Fu, A. et Yan, H. « *A new probabilistic relaxation method based on probability space partition* ». *Pattern recognition*, Vol.30, N°11, Elsevier Science Ltd

[Fuchs and LeMen,99] F. Fuchs and H. LeMen. "Building reconstruction on aerial images through multi-primitive graph matching". In 2nd *IAPR Workshop on Graph-based representations*, Vienna, Austria, 1999.

[Willuhn and Ade,96] W. Willuhn and F. Ade." *A rule-based system for house reconstruction from aerial images*". In *ICPR*, pages 885-889, Août 1996.

7. Publications

[Chehata et al 02] Chehata.N, Desseilligny M-P, Jung.F, Stamon G. "Extraction of 3D primitives from stereopairs of satellite images for automatic reconstruction of buildings" In *IAPR:MVA2002*, p636-639, Décembre2002, Japon.