

Reconstruction automatique de bâtiments à partir d'images aériennes : une approche probabiliste générique.

Automatic reconstruction of buildings from aerial images : a generic probabilistic approach.

F. Taillandier¹

R. Deriche²

¹ Institut Géographique National - MATIS

² INRIA - Projet Odyssée

¹ Institut Géographique National. 2/4 avenue Pasteur. 94165 SAINT-MANDE Cédex

² INRIA. 2004 Route des Lucioles - BP 93 - 06902 - Sophia-Antipolis Cédex

e-mail : ¹franck.taillandier@ign.fr, ²Rachid.Deriche@inria.fr

Résumé

Nous présentons une approche originale complète de reconstruction automatique de bâtiments à partir d'images aériennes reposant sur une modélisation générique des bâtiments sous la forme de polyèdres sans surplomb. Le système utilise un arrangement de plans généré par l'intersections de primitives planes verticales et non verticales. L'extraction des formes acceptables dans cet arrangement est faite par recherche des cliques maximales sur un graphe de compatibilité de facettes, étendant en cela les travaux de [1] au cas général présentant des plans verticaux. Le choix du meilleur modèle repose quant à lui sur une approche bayésienne permettant une intégration souple d'observations de types différents. La complexité du modèle est naturellement prise en compte dans ce schéma et des contraintes a priori assurent l'introduction d'une connaissance externe dans le choix de la meilleure forme polyédrique. Des résultats sont présentés pour montrer la validité et l'intérêt de cette approche générique qui mise sur la complémentarité des primitives 3D et sur l'introduction de connaissances externes par le biais des contraintes a priori.

Mots Clefs

images aériennes, polyèdres, reconstruction 3D, reconstruction de bâtiments, formalisation bayésienne.

Abstract

An original approach for automatic building reconstruction from aerial images is presented. The algorithm uses a very generic modelling of buildings as polyhedral form with no overhang in which external knowledge is introduced from a priori constraints on base primitives. The system starts from plane 3D primitives, vertical or not, to

build a planes arrangement from which acceptable forms of buildings are extracted through maximal cliques search, thus extending the work of [1] to the cases where vertical planes may be present. The choice of model is done through bayesian formulation that enables different kinds of observations to be taken into account. Model complexity and a priori constraints on base primitives are also naturally introduced to balance caricature needs and data adequacy. Results on real data show the validity of the approach.

Keywords

aerial images, polyhedral volumes, 3D reconstruction, building reconstruction, bayesian formulation.

1 Introduction

1.1 Contexte et Problématique

La reconstruction de bâtiments à partir d'images aériennes est un problème extrêmement complexe. L'ensemble des approches se heurtent à plusieurs problématiques parfois contradictoires : la nécessité de pouvoir reconstruire n'importe quel bâtiment et donc la généralité de l'approche et l'obligation de faire intervenir dans le processus des notions de simplicité de forme (généralisation) bien connues des cartographes pour éliminer les petits détails (chiens assis, lucarnes, cheminées...) au profit de structures plus importantes. L'objectif d'un tel système de reconstruction est donc d'obtenir une caricature d'un bâtiment générique. En outre, la plupart des approches utilisent des primitives 3D pour inférer la forme du bâtiment. Aux notions de généralisation et de généralité s'ajoutent alors les problèmes d'échecs inévitables des détecteurs (sur ou sous détection) dont le système doit tenir compte. Pour pallier ces erreurs et assurer une certaine généralisation, il paraît nécessaire d'introduire une information externe, soit sous forme de

modèles de bâtiments ce qui peut réduire la généralité de l'approche, soit sous forme de contraintes sur les primitives.

1.2 État de l'Art

Plusieurs auteurs [2, 3, 4] ont choisi de ne pas limiter a priori les formes à reconstruire. La plupart de ces approches utilisent un seul type de primitives (segments [2], coins [3] ou plans [4]) et gèrent difficilement les cas d'erreurs ou de sous-détection. Un autre axe de recherche utilise au contraire des modèles de bâtiments pour restreindre la bibliothèque de formes à reconstruire [5, 6]. L'apport d'une information externe forte permet de contrer les problèmes de sous-détections de primitives (complétion de formes basées sur les modèles) ou de sur-détections (élimination de primitives générant des formes qui ne sont pas dans la bibliothèque). [5] n'utilise qu'un seul type de primitives (coins) pour générer les hypothèses de bâtiments et l'ensemble des images pour choisir le bon modèle, [6] cherche au contraire à tirer parti de la complémentarité des primitives (planes, linéaires...) pour inférer les modèles et valide son choix en fonction d'un Modèle Numérique d'Élévation (MNE). Ces deux stratégies cherchent à contrer les problèmes de manque de généralité inhérente à leurs approches modèles par des "règles" heuristiques dans le cas de [5] ou basées sur des grammaires de graphe chez [6]. Ces stratégies prometteuses sont malgré tout limitées à des formes relativement simples et il paraît difficile de générer l'ensemble des formes visibles en zone urbaine ou péri-urbaine à partir de ces seules règles. En outre, en général, la combinatoire augmente avec la taille de la bibliothèque de modèles. Dans un autre contexte, on citera le travail d'Hassan Jibrini [1] qui utilise en entrée, en plus des images, des cartes cadastrales vectorisées fournissant l'emprise planimétrique des bâtiments (les façades). Sa méthode est basée sur la recherche, à partir d'un ensemble de plans, de la meilleure forme polyédrique continue sur un polygone cadastral donné. Cette modélisation extrêmement générique d'un bâtiment gère les sur-détections du détecteur de plans. Notre approche pour la génération d'hypothèses de bâtiments est une extension de cette méthode aux cas des plans verticaux (façades).

1.3 Objectifs

Nous avons choisi de modéliser un bâtiment par une forme polyédrale sans surplomb dont le contour externe est constitué d'un ensemble de plans verticaux. Cette définition très générique permet de générer la quasi totalité des bâtiments présents en agglomération urbaine (à l'exception des rares bâtiments aux formes arrondies). L'objectif de ces travaux est de pallier les erreurs dues aux détecteurs de primitives présentés en section 3, d'être capable de générer *à partir des primitives* un ensemble de modèles de bâtiments plausibles (section 4), d'avoir un système ouvert capable, par une modélisation simple (section 5), de gérer aisément diverses observations externes, de pouvoir intégrer une certaine forme de généralisation par le biais de

contraintes géométriques sur les primitives (orthogonalité, parallélisme...) pour choisir le bon modèle.

2 Méthodologie générale

La Figure 1 récapitule le fonctionnement général. L'algorithme fonctionne sur des régions d'intérêt ou masques de focalisation sélectionnés manuellement ou détectés par un algorithme adéquat [7] (Figure 2). Dans une zone donnée, plans et façades sont extraits. Ils constituent les primitives de base utilisées dans un arrangement qui génère un graphe 3D de facettes. Après simplification du graphe par l'utilisation de bandeau de tolérance autour du MNE issu des travaux de [8] et du masque de focalisation, l'ensemble des formes acceptables de bâtiments est recherché dans cet ensemble de facettes par un algorithme de recherche de cliques maximales. On reprend en cela les idées de [1]. A l'issue du processus, le meilleur modèle est déterminé en utilisant une modélisation bayésienne faisant intervenir des informations externes (images, masque, segments 3D...) et une modélisation de la complexité d'un modèle dans laquelle des contraintes a priori sur les primitives (orthogonalité, parallélisme, horizontalité de certaines arêtes) peuvent être intégrées. Ces contraintes ne sont pas limitatives mais assurent l'introduction d'une certaine forme de généralisation dans le processus de choix. Ces contraintes sont ensuite utilisées dans l'étape de reconstruction proprement dite qui ne sera pas détaillée dans cet article faute de place mais dont le principe est basé sur les travaux de [9].

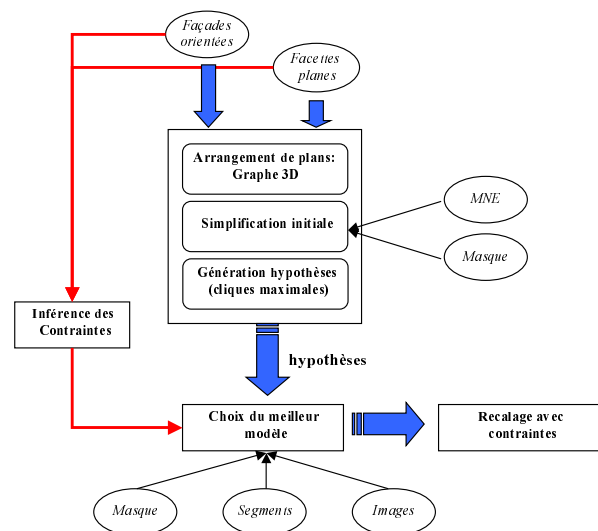


FIG. 1 – Méthodologie Générale. Les données externes apparaissent en italique dans une ellipse, les processus sont en gras.

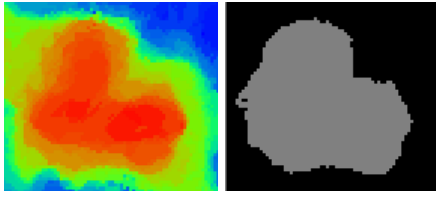


FIG. 2 – Extraits de mne et de masque de focalisation sur une zone d'exemple dont une image est présentée en Figure 3.

3 Extraction de primitives

3.1 Segments

L'algorithme d'extraction de segments 3D est décrit en détail dans [10]. Sa caractéristique principale est d'avoir une méthode d'appariement multi-images et d'utiliser la propagation d'incertitude pour valider les hypothèses d'appariement. Ces segments 3D comportent peu de fautes, on peut constater quelques manques (environ 30%) (Figure 3). Ils constituent, comme avancé dans [11], des primitives très structurantes.

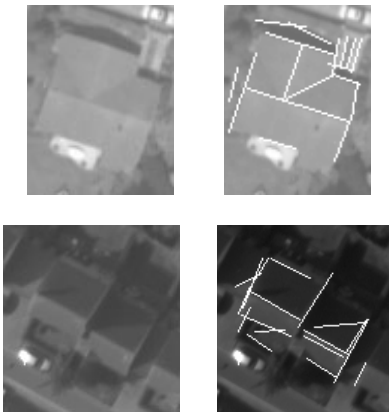


FIG. 3 – 2 images de zones test et les segments 3D détectés projetés.

3.2 Plans

L'extraction de plans se fait par un algorithme de segmentation en facettes planes de MNE décrit dans [12], dont la caractéristique principale est de pouvoir intégrer des contraintes images (gradients) et/ou externes (telles que segments 3D). Un paramètre d'échelle permet de régler le facteur de généralisation (ou caricature) désiré pour la scène. Dans notre cas, ce paramètre a été choisi très bas pour assurer une sur-détection des primitives planes (Figure 4). L'algorithme de reconstitution des formes acceptables ne peut en effet gérer les cas de sous-détection trop importants. Il est donc indispensable de s'assurer que l'ensemble des plans essentiels est présent.

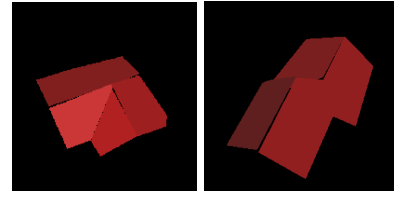


FIG. 4 – Exemple de plans obtenus sur les 2 zones données en exemple

3.3 Façades

L'extraction de façades s'appuie sur la segmentation en facettes planes précédente. Contraints par les informations image ou segments 3D, les contours de cette segmentation sont beaucoup plus réguliers que ceux d'une segmentation classique sur MNE issu de corrélation. Les arêtes qui présentent des sauts d'altitude trop importants ou qui sont en bordure du masque de focalisation sont polygonisées et on peut alors définir le plan vertical qu'elles représentent et qui constitue une façade. Dans cet algorithme, la notion d'exhaustivité est là encore essentiel (Figure 5). Il est à noter que ces façades sont *orientées* : les normales pointent vers l'extérieur du volume.

En toute rigueur, il faudrait décorrélérer les détecteurs de façon à ce qu'une erreur sur un détecteur n'influe pas sur les autres primitives. En pratique, un algorithme de détection de façades sans passer par la segmentation en facettes planes est étudié mais n'a pas encore été validé.

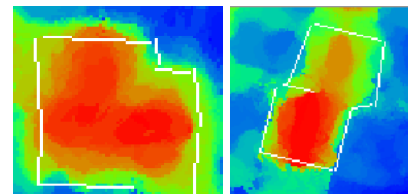


FIG. 5 – Exemple de façades obtenues sur les 2 zones données en exemple

4 Recherche des formes acceptables

4.1 Création d'un graphe 3D

Dans la suite, on appellera *polytope* tout polygone bidimensionnel plan (les points sont tous dans un même plan : le plan porteur) convexe. Soit $\mathcal{H}$ un ensemble de H plans de $\mathbb{R}^3$. On rappelle que l'intersection de ces plans définit un ensemble de polytopes d'intérieurs disjoints, d'arêtes et de points dans $\mathbb{R}^3$, qui constituent l'arrangement de plans de $\mathcal{H}$ [13, 14]. Dans la suite, par abus de langage, nous appellerons facette ou facette élémentaire tout polytope généré par l'arrangement de plans. Dans notre cas, l'ensemble des primitives planes détectées (plans et façades) est plongé dans un arrangement de plans qui permet de générer un

graphe 3D fournissant toutes les relations topologiques d'adjacence entre les points, les arêtes et les facettes. Ce graphe est confiné dans une forme prismatique convexe s'appuyant sur une caricature du masque de focalisation et délimité par une altitude minimum (le "sol") et maximum (le "ciel") (Figure 6). Le polygone délimitant cette forme prismatique au niveau du sol est intégré à l'arrangement de plans et donc découpé en facettes élémentaires appartenant au graphe 3D. Dans la suite, on dénommera "face-sol" $\mathcal{F}_s$ ce polygone située au niveau du sol, c'est à dire, dans le graphe 3D initial, l'ensemble des facettes incluses dans le plan horizontal au niveau du sol. On notera que toutes les facettes du graphe *ont une orientation* : vers le haut par défaut pour les facettes qui s'appuient sur un plan non vertical, et dans le même sens que leur plan porteur dans le cas des portions de façades.

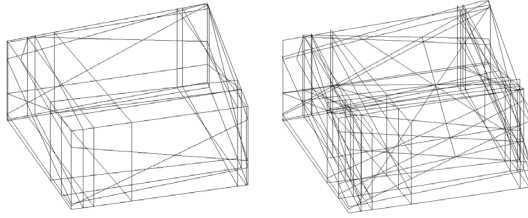


FIG. 6 – intersection de plans et Graphe 3D issu de l'arrangement (découpage en facettes)

4.2 Définitions

orientations des facettes. On rappelle dans cette section quelques définitions issues pour la plupart de [14]. La donnée de l'orientation du plan porteur d'un polytope et d'un sens de rotation arbitraire induit une orientation du polytope (ordre circulaire sur ses sommets). Dans la suite, on choisira, sans perte de généralité, une orientation arbitraire *directe* pour définir l'orientation des polytopes à partir des plans porteurs. Les orientations de deux polytopes adjacents sont *cohérentes* si elles induisent sur l'arête commune des orientations opposées (Figure 7). On peut alors montrer le lemme 1 qui permet de tester aisément pour deux polytopes adjacents leur cohérence et ainsi déterminer si une arête est localement admissible ou pas.

Lemme 1 *2 polytopes adjacents ont des orientations cohérentes autour d'un segment s si, avec les notations de la figure 7,*

$$[s, v_1, n_1] \cdot [s, v_2, n_2] < 0$$

Une surface polyédrique est l'union d'un ensemble de polytopes arrangés dans $\mathcal{R}^3$ tel que

- l'intersection de deux polytopes distincts est soit l'ensemble vide, soit un sommet commun, soit un côté commun,
- chaque côté d'un polytope quelconque est le côté d'au plus un autre polytope

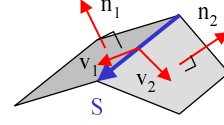
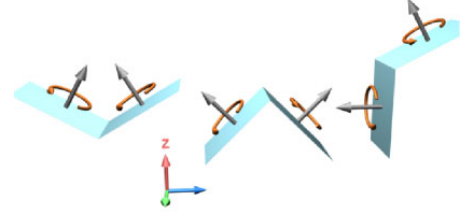
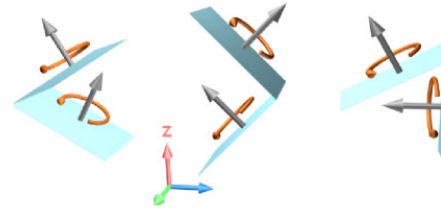


FIG. 7 – Cohérence de l'orientation de 2 polytopes



(a) facettes cohérentes



(b) facettes incohérentes

FIG. 8 – Exemple de cohérences de facettes

- tous les polytopes adjacents ont des orientations cohérentes.

Cette définition reprend la définition de triangulation orientable de [14] en l'étendant au cas des polytopes. Le *bord* d'une surface polyédrique Σ est défini comme l'union des côtés des polytopes qui n'appartiennent qu'à un seul polytope. Il est noté $\partial\Sigma$. Une surface polyédrique *fortement connexe* est une surface polyédrique Σ telle que, quels que soient deux points de Σ , il existe un arc sur Σ les reliant, en ne traversant que des arêtes de Σ (et pas de sommets, à la différence de simplement connexe). On exclut ainsi les configurations avec sommet singulier [14]. Un polyèdre est une surface polyédrique fortement connexe sans bord ($\partial\Sigma = \emptyset$).

admissibilité et compatibilité. Une *surface admissible* est une surface polyédrique fortement connexe dont le bord recouvre les arêtes de la face-sol et dont chaque polytope est une facette élémentaire de l'arrangement de plans. Une arête est localement admissible si elle appartient au bord de la face-sol ou si il existe au moins deux facettes aux orientations cohérentes ayant cette arête en commun. Une facette est localement admissible si elle est localement admissible autour de chacune de ses arêtes bordantes. Une

facette est admissible si au moins une surface admissible la contient (une facette admissible est nécessairement localement admissible). Un graphe est dit localement admissible si toutes ses facettes le sont. Un graphe est dit admissible si toutes ses facettes le sont. Deux facettes sont compatibles si elles appartiennent à au moins une surface admissible commune.

4.3 Propriétés

Dans cette section, on mentionne différentes propriétés des surfaces et graphes admissibles. Les démonstrations des deux premiers théorèmes sont des extensions de celles de [1] et ne sont pas reprises ici. La figure 9 présente les notions de surfaces enveloppes. Dans la suite, on dira qu'une facette non verticale f réalise l'altitude maximale (resp. minimale) lorsque, quel que soit un point X intérieur à f , aucune facette n'intersecte la verticale passant par X au dessus (resp. en dessous) de X .

Théorème 1 si $\mathcal{G}$ est un graphe localement admissible, son emprise planimétrique est celle de $\mathcal{F}_s$

Théorème 2 Soit $\mathcal{G}$ un graphe localement admissible. Soit $\mathcal{P}^+$ (respectivement $\mathcal{P}^-$) l'ensemble des facettes non verticales réalisant les altitudes maximales (resp. minimales) sur l'emprise planimétrique. Il existe une unique surface admissible $\mathcal{Z}^+$ (resp. $\mathcal{Z}^-$) passant par $\mathcal{P}^+$ (resp. $\mathcal{P}^-$). $\mathcal{Z}^+$ et $\mathcal{Z}^-$ sont dénotées surfaces enveloppes.

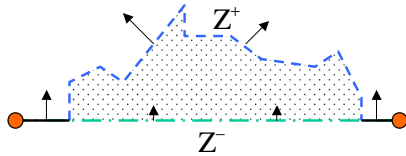


FIG. 9 – Graphe 3D issu de l'arrangement de plans. Les surfaces enveloppes sont respectivement en bleu/tiret et en vert/tiret+point. Les points rouges représentent le bord de la face-sol. Les parties attenantes aux bords de la face-sol, en noir appartiennent aux deux surfaces. Le graphe 3D est situé entre les 2 enveloppes.

Théorème 3 Dans le cas d'un arrangement de plans où tous les plans non verticaux sont orientés vers le haut, toute surface admissible Σ est sans surplomb.

Preuve de 3

Soit $\partial\Sigma$ le bord de la surface Σ qui, par définition, s'appuie sur le bord de $\mathcal{F}_s$. On considère le polyèdre construit en rajoutant des "parois latérales" et un "plancher" correctement orientés comme indiqué sur la Figure 10. La surface admissible étant une surface polyédrique fortement connexe, il est aisé de voir que la surface ainsi générée est un polyèdre. D'après [14], un polyèdre est homéomorphe à une sphère avec anses. Dans notre cas, l'absence de polytope orienté

"vers le bas" interdit la constitution d'une structure homéomorphe à une anse. Le polyèdre est donc homéomorphe à une sphère topologique et la surface Σ n'a ni anses ni surplomb (pas de facettes orientées vers le bas).

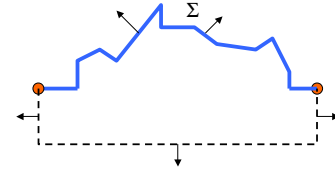


FIG. 10 – La surface Σ étudiée dans le théorème 3 est représentée en bleu trait fort, les parties en pointillés représentent les "parois latérales" et le "plancher" virtuels utilisés pour la démonstration du théorème.

4.4 Simplifications du graphe

Algorithme récursif. Considérons l'algorithme fondamental de suppression récursive de facettes à l'issue duquel il est évident que, si initialement le graphe est localement admissible, le graphe résultant est localement admissible.

Algorithme 1 (RecSup(f))

Soit $\mathcal{A}$ les facettes adjacentes à f .

- Supprimer f .
- Pour toute facette $f' \in \mathcal{A}$, si f' est localement inadmissible, RecSup(f').

suppression de facettes initiales. Dans un premier temps, en utilisant l'algorithme 1, on supprime récursivement toutes les facettes qui, dans le graphe 3D, sont localement inadmissibles. Il s'agit des facettes touchant les bords de la forme prismatique dans laquelle est confiné le graphe et qui ne touchent pas (aucune de leurs arêtes n'est incluse dans) la face-sol $\mathcal{F}_s$. Il en résulte un graphe localement admissible. On supprime ensuite récursivement les facettes non verticales attenantes au sol (Figure 11(a)). On remarquera alors que les surfaces admissibles du graphe 3D ainsi élagué sont telles que, sur toute surface admissible, tout ensemble connexe de facettes ne possédant aucune facette incluse dans $\mathcal{F}_s$ est un bâtiment. La recherche des surfaces admissibles permet, de fait, de récupérer toutes les hypothèses de bâtiment sur le masque de focalisation et, le cas échéant, de pouvoir modéliser plusieurs bâtiments sur une seule zone. On peut ainsi pallier les erreurs de focalisation.

bandeau de tolérance autour du MNE. Avant de rechercher de façon explicite les surfaces admissibles, il est indispensable de réduire le nombre des facettes du graphe. La recherche des solutions est en effet un problème NP-complet en fonction du nombre de facettes du graphe. Pour réduire ce nombre, on supprime toute facette non incluse dans $\mathcal{F}_s$ sol qui n'est pas entièrement contenue dans une "nappe" autour du MNE [15] (Figure 11(b) et 11(c)).

zone de tolérance autour du masque. De la même façon, on supprime de façon récursive toute facette attenante à la face-sol qui touche le masque érodé et toute facette verticale qui se trouve hors du masque dilaté (Figure 11(d)).

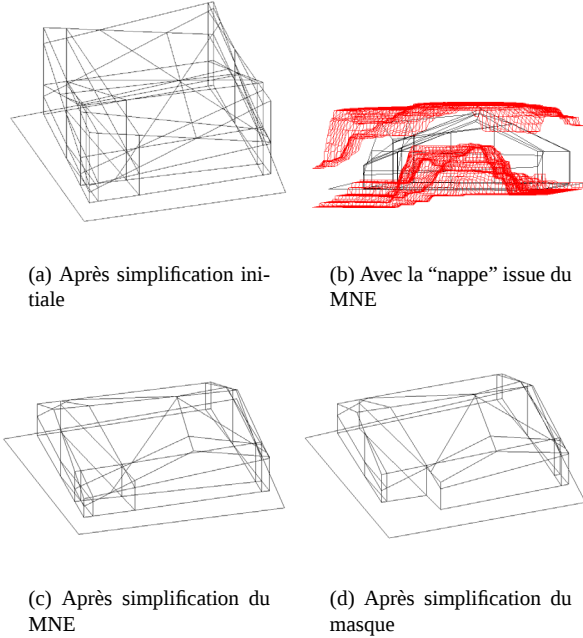


FIG. 11 – Différentes étapes de simplification

4.5 Recherche des solutions

Comme montré auparavant, pour déterminer les hypothèses de bâtiments, il faut rechercher les surfaces admissibles du graphe élagué. Nous allons maintenant détailler la technique de recherche de ces surfaces. On utilise ici le principe de recherche des cliques maximales dans un graphe de compatibilité [1].

Le calcul des facettes compatibles reprend l'idée de [1] en l'étendant aux cas des facettes verticales. Ce calcul se fait en deux temps à partir d'un graphe 3D $\mathcal{G}$ (Figure 12) :

Pour toute facette non verticale f , on supprime récursivement et virtuellement les facettes non verticales dont le projeté est d'intersection non vide avec le projeté de f . D'après le théorème 3, ces facettes ne sont pas compatibles avec f . Si le graphe virtuel $\tilde{\mathcal{G}}$ (localement admissible) résultant est non vide, toute facette *admissible* du graphe est compatible avec f qui est nécessairement admissible puisque les surfaces enveloppes de $\tilde{\mathcal{G}}$ sont des surfaces admissibles et passent par f . Dans le cas contraire, f n'est pas admissible et est supprimée récursivement du graphe $\mathcal{G}$.

Ensuite, pour toute facette verticale v , on supprime toute facette non verticale incompatible avec v déterminée d'après le calcul précédent. On peut là aussi montrer que si le graphe résultant est non vide, v est admissible

et toute facette *admissible* du graphe est compatible avec v .

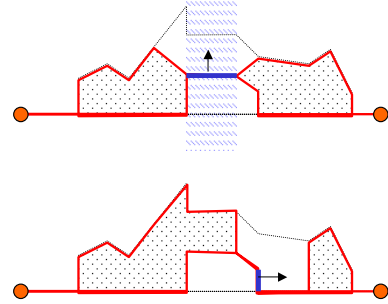


FIG. 12 – Exemples de calcul des facettes compatibles pour une facette non verticale et verticale. Toutes les facettes contenues dans les zones en pointillés sont compatibles avec la facette en bleu.

A l'issue du processus, l'ensemble des facettes restantes dans $\mathcal{G}$ est admissible et on a déterminé les couples de facettes compatibles. On peut donc construire le graphe de compatibilité $\mathcal{G}_c$ dans lequel chaque noeud est une facette admissible et chaque arête entre 2 noeuds représente la relation de compatibilité. En suivant le raisonnement de [1], on peut alors démontrer que *les surfaces admissibles du graphe sont les cliques maximales du graphe de compatibilité* ou, ce qui est équivalent, que les surfaces admissibles sont les ensembles stables maximaux du graphe d'incompatibilité $\tilde{\mathcal{G}}_c$ (Figures 13 et 14). Cette recherche de cliques maximales est un problème NP-complet. On peut montrer que par rapport à l'algorithme utilisé par [1], l'algorithme de Bron-Kerbosch optimisé [16] permet une étude beaucoup plus approfondie dans le cas de nos graphes où la densité est relativement élevée.

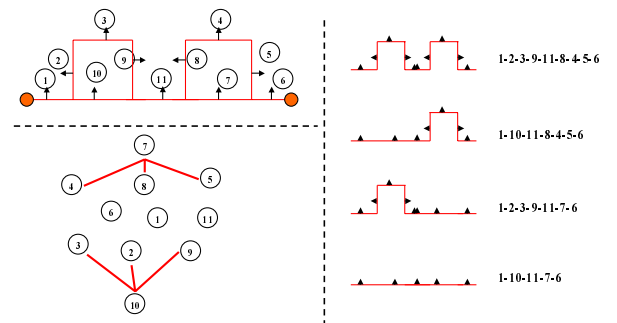


FIG. 13 – Exemple de Graphe élagué et de graphe d'incompatibilités $\tilde{\mathcal{G}}_c$ ainsi que les solutions sous forme de graphe et d'ensemble stable maximal.

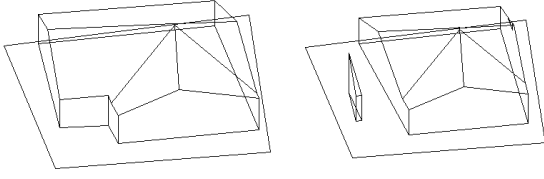


FIG. 14 – 2 solutions sur les 3196 de l'exemple présenté

5 Choix du meilleur polyèdre

5.1 Modélisation bayésienne

Le module précédent fournit un ensemble d'hypothèses de bâtiments $\mathcal{M}$. Considérons un ensemble d'observations ou données externes quelconques $\mathcal{D}$ qui vont permettre de choisir le "meilleur" modèle parmi $\mathcal{M}$. En utilisant une modélisation bayésienne de l'information, on veut trouver le modèle qui vérifie au mieux les observations, c'est à dire le modèle $\hat{M}$ tel que

$$\hat{M} = \arg \max_{M \in \mathcal{M}} P(M | \mathcal{D}) \quad (1)$$

soit en utilisant le théorème de Bayes

$$P(M | \mathcal{D}) = \frac{P(\mathcal{D} | M) \cdot P(M)}{P(\mathcal{D})} \quad (2)$$

$P(\mathcal{D})$ ne dépendant pas du modèle, on est ramené à la recherche de $\hat{M} = \arg \max_{M \in \mathcal{M}} P(\mathcal{D} | M) \cdot P(M)$. Dans la suite, nous utiliserons comme observations : les images I , les segments 3D S et le masque de focalisation M_a . En supposant les observations indépendantes, on obtient finalement :

$$\hat{M} = \arg \max_{M \in \mathcal{M}} P(S | M) \cdot P(I | M) \cdot P(M_a | M) \cdot P(M) \quad (3)$$

La probabilité du modèle dépend de la complexité du modèle de bâtiment. A priori, on cherche en effet à obtenir la "modélisation" la plus simple possible compte tenu des observations. S'inspirant de la relation de Shannon [17], on relie $P(M)$ à la longueur minimale de description par la formule 4 :

$$P(M) = C \exp \frac{-\mathcal{L}(M)}{\beta} \quad (4)$$

où C est un facteur de normalisation commun à tous les modèles que l'on omettra dans la suite. β permet de réguler le facteur de caricature désiré en donnant plus ou moins de poids au modèle ou aux données. Les sections suivantes détaillent les différents termes.

5.2 Modèle

Inférence des contraintes. Comme dit auparavant, les contraintes sur les primitives influent sur le choix du modèle, permettant ainsi une intégration d'une forme de régularisation du problème. Les contraintes sont déterminées

sur les orientations des primitives de base : plans et façades. On suit ici la logique du système à base de contraintes développé par [9] car c'est celui qui est adapté à notre problématique pour le recalage sous contraintes. L'algorithme d'inférences de contraintes utilise deux paramètres angulaires $\sigma_{\parallel}$ et $\sigma_{\perp}$. Les normales des primitives planes sont tout d'abord regroupées par un algorithme itératif de clustering qui regroupe les normales dont l'écart angulaire est faible. A l'issue du processus, pour chaque groupe de normales, on définit une *direction*, moyenne des normales de chaque groupe. L'algorithme de clustering assure que la distance angulaire minimale entre ces directions est de $\sigma_{\parallel}$ et, par le regroupement des normales permet de gérer le parallélisme. On définit ensuite sur ces directions un graphe de contraintes. Les noeuds du graphe sont les directions (on distingue 2 types de noeuds : façades et plans), les arêtes matérialisent différentes relations (Figure 15) : orthogonalité, horizontalité du produit vectoriel (donc horizontalité de l'arête intersection des 2 plans, relation très importante en milieu urbain). Chaque arête est créée si la relation est vérifiée dans la tolérance angulaire $\sigma_{\perp}$. Chaque arête q est affectée d'un poids $C(q)$ donnant le nombre de degré de libertés que la contrainte supprime (exemple : 1 pour la contrainte d'orthogonalité). On notera que l'orthogonalité entre la direction verticale et les façades n'est pas testée : elle est implicite.

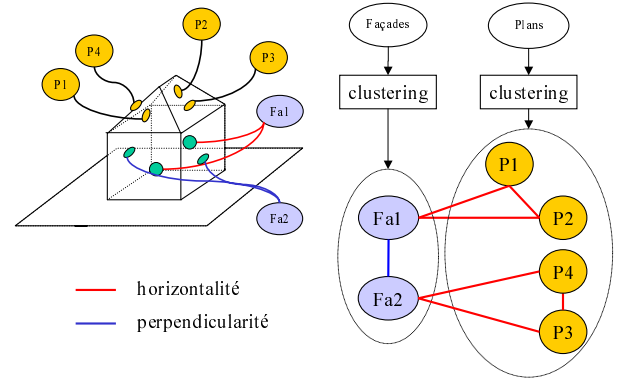


FIG. 15 – Schéma d'inférence des contraintes et représentation du graphe de contraintes. Les façades sont regroupées par le clustering en 2 directions Fa_1 et Fa_2 .

Longueur des description. On notera F_M , E_M et V_M le nombre de facettes, d'arêtes et de points d'un modèle. On notera P_M et Fa_M le nombre de plans et de façades utilisées par le modèle, $D_M = P_M + Fa_M$ le nombre de directions générales et C_M le nombre de contraintes d'un modèle. La logique utilisée ici s'inspire des travaux de [18]. On utilise le principe de transmission d'informations liées au modèle : topologique, géométrique et contraintes. A noter que la longueur de description initiale qui consiste à fournir (sur des entiers) le nombre de facettes, d'arêtes, de points, de façades, de plans et de contraintes du modèle est donnée par

$$\mathcal{L}_g(f) = \log(|V|_{\max}) + \log(|F|_{\max}) + \dots$$

Cette description a une longueur identique pour tous les modèles de $\mathcal{M}$ et représente donc un terme multiplicatif constant de $P(M)$. On peut donc le négliger.

topologie La description topologique consiste à énumérer, pour chaque facette f , l'ensemble de ses points de cardinal $|f|$, ainsi que la direction du plan porteur. Chaque point a une probabilité uniforme ce qui amène

$$\mathcal{L}_t(f) = |f| \log(V_M) + \log(D_M) \quad (5)$$

d'où en sommant sur l'ensemble des facettes et en tenant compte du fait que le nombre de points est égal au nombre d'arêtes pour un polytope fermé ou facette f de l'arrangement de plans :

$$\mathcal{L}_t(M) = 2 \cdot E_M \log(V_M) + F_M \cdot \log(D_M) \quad (6)$$

géométrie La description géométrique utilise un modèle simple d'énumération des coordonnées des points et des directions. Chaque coordonnée est codée sur un réel double précision dont la longueur de description est de 64 bits. Chaque façade équivaut à 1 inconnue, chaque plan non vertical à 2 et chaque point à 3 inconnues. Chaque facette apporte $|f| - 1$ équations de planarité du type $\mathbf{D}_M^t \cdot (\mathbf{V}_i - \mathbf{V}_j) = 0$. Chaque équation est supposée réduire d'un le nombre de degré de liberté (pas de test de non indépendance des équations pour des raisons évidentes de temps de calcul) ce qui amène :

$$\begin{aligned} \mathcal{L}_g(M) &= (2 * P_M + Fa_M + 3V_M - \sum_f (|f| - 1)) * 64 \\ &= (2 * P_M + Fa_M + 3V_M - 2E_M + F_M) * 64 \end{aligned} \quad (7)$$

contraintes Chaque modèle ou surface admissible utilise un certain nombre de directions et définit donc un sous-graphe dans le graphe complet de contraintes. Toute arête présente dans ce sous-graphe représente une contrainte du modèle. Chaque contrainte est codée par : son type (orthogonalité ou horizontalité), les directions qu'elle met en relation et chaque contrainte réduit d'un nombre $C(q)$ lié au type c de la contrainte q le nombre de degrés de liberté du problème. On notera $|c|$ le nombre de types de contraintes (2 dans notre cas) et on suppose en outre dans l'évaluation du coût du modèle que les contraintes sont indépendantes entre elles, hypothèse forte mais limitée par le fait que les normales sont déjà regroupées en directions. On obtient finalement :

$$\mathcal{L}_c(M) = \sum_q \log(|c|) + 2\log(D_M) - C(q) * 64 \quad (8)$$

modèle global Le modèle global regroupe les 3 termes et prend ainsi en compte la complexité topologique, géométrique du modèle ainsi que le gain apporté par l'application des contraintes

$$\mathcal{L}_M = \mathcal{L}_t + \mathcal{L}_g + \mathcal{L}_c \quad (9)$$

5.3 Observations

Les probabilités portant sur les observations permettent de quantifier l'adéquation du modèle aux données après avoir modélisé la complexité du modèle. On rappelle que la recherche du meilleur modèle est un compromis entre cette adéquation et cette complexité.

Images. Pour calculer le score du modèle en fonction des images, l'idéal serait de calculer le score de corrélation des orthoimages correspondant au modèle de surface de toit qui donne les meilleurs résultats. Toutefois, [1] montre également que le gain par rapport à la moyenne des scores de corrélation de l'ensemble des orthoimages horizontales régulièrement espacées et d'altitude donnée par le modèle de toit du bâtiment est marginal. En outre, cette dernière méthode présente d'indéniables avantages en termes de temps de calcul : le volume de corrélation peut-être précalculé avant la recherche des modèles (Figure 16). L'évaluation de chaque modèle devient alors beaucoup plus rapide et l'on évite ainsi les étapes de projection des images sur le modèle. Nous avons donc choisi cette solution. Dans le cas du multi-images (n images), le score de corrélation utilisé est donné par l'équation 10 dans laquelle chaque fenêtre de corrélation de taille (fu, fv) centrée en (u, v) dans une image I_i est représenté par le vecteur $\mathbf{v}_i(u, v)$ et où (u_i, v_i) est le projeté du point (x, y, z) via la matrice de projection connue correspondante de I_i :

$$C(x, y, z) = \frac{\text{Var}(\sum_i \mathbf{v}_i(u_i, v_i))}{n \cdot \sum_i (\text{Var}(\mathbf{v}_i(u_i, v_i)))} \in [0..1] \quad (10)$$

Il s'agit du score multi-images symétrique utilisé par [19] normalisé par le nombre d'images. La moyenne est calculé sur la surface planimétrique $SP(M)$ de l'ensemble $\mathcal{H}(M)$ des facettes non verticales du modèle qui n'appartiennent pas à la face-sol :

$$P(I | M) = \sum_{f \in \mathcal{H}(M)} \frac{C(x, y, M(x, y))}{|SP(M)|} \quad (11)$$

Segments 3D. Les segments 3D apportent une information de structure très forte, il semble donc logique de chercher à les appairer avec les arêtes du modèle. Pour chaque segment s , on détermine les arêtes appariées, une arête a du modèle est appariée avec s si et seulement si l'écart angulaire est inférieur à un paramètre θ_s et si l'écart en distance est inférieur à un autre paramètre d_s : on notera alors $a\mathcal{R}s$. Pour chaque arête appariée, on définit alors la longueur de recouvrement de s par a : $r(a, s)$. L'"explication" du segment s apparié par le modèle M se traduit donc par :

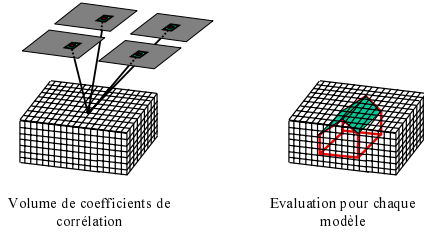


FIG. 16 – Principe du calcul de $P(I/M)$

$$P(s | M) = \sum_{a \in \mathcal{R}_s} \frac{r(a, s)}{\|s\|} \quad (12)$$

Pour prendre en compte les erreurs du détecteur de segments [20], on affecte une valeur ϵ_s par défaut à un faux appariement (ce que l'on suppose être le cas pour un segment non apparié). En pratique, nous avons choisi $\epsilon_s = 0,3$. L'équation précédente devient alors

$$P(s | M) = \max \left(\sum_{a \in \mathcal{R}_s} \frac{r(a, s)}{\|s\|}, \epsilon_s \right) \quad (13)$$

Pour l'ensemble de segments S , en supposant les segments indépendants, on obtient

$$P(S | M) = \prod_{s \in S} P(s | M) \quad (14)$$

Masque. Contrairement aux approches à partir de cadastre, il est nécessaire de prendre en compte, dans notre cas, le recouvrement de la zone de focalisation sur laquelle se trouve le bâtiment. On utilise comme mesure l'intersection de la surface planimétrique de $\mathcal{H}$ avec le masque M_a rapportée à la surface de l'union de ces 2 surfaces. On a donc :

$$P(M_a | M) = \frac{|SP(M) \cap M_a|}{|SP(M) \cup M_a|} \quad (15)$$

6 Résultats

La Figure 17 présentent les premiers résultats de l'algorithme appliqué sur 2 scènes et obtenu avec un β fixé empiriquement. De nombreuses expérimentations sont encore nécessaires pour fixer définitivement un niveau pour ce paramètre β mais ces premiers résultats témoignent de la validité et de l'intérêt de l'approche pour la reconstruction de bâtiments. On remarquera en particulier que sur la deuxième scène, l'algorithme a été capable de reconstruire les 2 bâtiments accolés.

7 Conclusion et perspectives

7.1 Conclusion

L'algorithme décrit dans cet article permet de générer à partir de primitives planes (façades et plans) un ensemble d'hypothèses de bâtiments génériques parmi lesquelles le

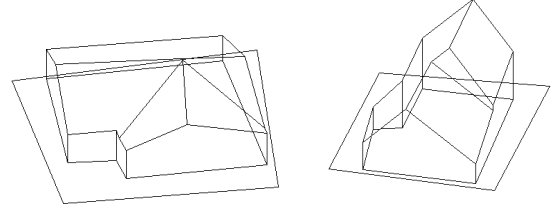


FIG. 17 – Exemples de 2 résultats.

choix est fait en utilisant le plus d'observations externes possibles ainsi que des contraintes a priori qui permettent l'introduction *souple*, dans le processus, d'une connaissance externe sur la structure des formes rencontrées.

Malgré sa généralité, l'algorithme est capable de gérer les cas d'erreurs des détecteurs. En revanche, il repose entièrement sur les caractéristiques géométriques des primitives et dépend donc beaucoup de la qualité géométrique des images et de leur calibrage ainsi que de la précision des algorithmes de détection de primitives.

Les premiers résultats sont prometteurs. Cet algorithme peut en particulier reconstruire des bâtiments accolés ou des bâtiments présentant des discontinuités altimétriques internes.

7.2 Discussion et Perspectives

Il est prévu de réutiliser en entrée de l'algorithme les polygones cadastraux qui fournissent des hypothèses de façades. Une des limitations actuelles de l'algorithme est de ne pas être capable de gérer les intersections à plus de 3 plans, problème connu comme le "vertex superstrictness" dans le cadre de la reconstitution d'objets à partir de dessins CAO. Nous travaillons actuellement à la résolution de ce problème, à l'évaluation de l'algorithme sur des zones plus étendues ainsi qu'à la détermination du rôle et de l'importance de chacun des types d'observations.

Pour traiter des zones plus complexes, il paraît également indispensable de contraindre le problème combinatoire de recherche de *toutes* les cliques maximales en privilégiant des heuristiques conduisant rapidement à des solutions acceptables [21].

Références

- [1] H. Jibrini, *Reconstruction automatique des bâtiments en modèles polyédriques 3D à partir de données cadastrales vectorisées 2D et d'un couple d'images aériennes à haute résolution*. Thèse de doctorat, Ecole Nationale Supérieure des Télécommunications, avril 2002.
- [2] C Baillard et O. Dissard, « A stereo matching algorithm for urban digital elevation models », *Photogrammetric Engineering and Remote Sensing*, vol. 66, n° 9, septembre 2000, p. 1119–1128.

- [3] S. Heuel, W. Förstner, et F. Lang, « Topological and geometrical reasoning in 3D grouping for reconstructing polyhedral surfaces », dans *ISPRS*, vol. XXIII, p. 397–404, Amsterdam, 2000.
- [4] B. Ameri et D. Fritsch, « Automatic 3D building reconstruction using plane-roof structures », dans *ASPRS*, Washington DC, 2000.
- [5] A. Fischer, T. Kolbe, F. Lang, A. Cremers, W. Förstner, L. Plümer, et V. Steinhage, « Extracting buildings from aerial images using hierarchical aggregation in 2D and 3D », *CVIU*, vol. 72, n° 2, novembre 1998, p. 163–185.
- [6] F. Fuchs, *Contribution à la reconstruction du bâti en milieu urbain, à l'aide d'images aériennes stéréoscopiques à grande échelle. Etude d'une approche structurale*. Thèse de doctorat, Université de René Descartes, avril 2001.
- [7] C. Baillard et H. Maitre, « 3-D reconstruction of urban scenes from aerial stereo imagery : A focusing strategy », *CVIU*, vol. 76, n° 3, 1999, p. 244–258.
- [8] C. Baillard, *Analyse d'images aériennes stéréoscopiques pour la restitution 3D des milieux urbains*. Thèse de doctorat, Ecole Nationale Supérieure des Télécommunications, 1997.
- [9] E. Grossmann, *Maximum Likelihood 3D Reconstruction From One or More Uncalibrated Views Under Geometric Constraints*. Thèse de doctorat, Universidade Tecnica de Lisboa, 2002.
- [10] F. Taillandier et R. Deriche, « Reconstruction de primitives linéaires 3D en multi-vues pour la modélisation de scènes urbaines », dans *RFIA 2002*, vol. I, p. 115–124, janvier 2002.
- [11] C. Baillard, C. Schmid, A. Zisserman, et A. Fitzgibbon, « Automatic line matching and 3D reconstruction of buildings from multiple views », dans *Proc of ISPRS Conference on Automatic Extraction of GIS Objects from Digital Imagery, IAPRS*, vol. 32, p. 69–80, septembre 1999.
- [12] F. Taillandier, L. Guigues, et R. Deriche, « A framework for constrained multi-scale range image segmentation », dans *ICIP*, Barcelona, 2003.
- [13] H. Edelsbrunner, *EATCS Monographs on theoretical Computer Science*, chap. Algorithms in Combinatorial Geometry. Springer-Verlag, 1987.
- [14] J-D. Boissonnat et M. Yvinec, *Géométrie Algorithmique*. Ediscience International / Collection informatique, 1995.
- [15] F. Taillandier et R. Deriche, « Reconstruction of 3D linear primitives from multiple views for urban areas modelisation », dans *PCV'02*, vol. 24, p. 267–272, septembre 2002.
- [16] C. Bron et J. Kerbosch, « Algorithm 457 : Finding all cliques of an undirected graph », *Communications of the ACM*, vol. 16, 1973, p. 575–577.
- [17] T. Kanungo, B. Dom, W. Niblack, D. Steele, et J. Sheinvald, « MDL-based multi-band image segmentation using a fast region merging scheme ». Rapport technique n° RJ 9960, IBM Research Division, Almaden Research Center, San Jose, Ca 95120-6099, mai 1995.
- [18] T-H Kolbe, *Identifikation und Rekonstruktion von Gebäuden in Luftbildern mittels unsharfer Constraints*. Thèse de doctorat, Hochschule Vechta, 1999.
- [19] N. Paparoditis, C. Thom, et H. Jibrini, « Surface reconstruction in urban areas from multiple views of aerial digital frames », dans *IAPRS*, vol. XXIII, Amsterdam, 2000.
- [20] C. Schmid, « A structured probabilistic model for recognition », dans *CVPR*, vol. 2, p. 485–490, FT Collins, CO, juin 1999.
- [21] R. Battiti et M. Protasi, « Reactive local search for the maximum clique problem », *Algorithmica*, vol. 29, n° 4, avril 2001, p. 610–637.