Modèles stochastiques pour la reconstruction tridimensionnelle d'environnements urbains

Florent Lafarge

▶ **To cite this version:**

Florent Lafarge. Modèles stochastiques pour la reconstruction tridimensionnelle d'environnements urbains. Interface homme-machine [cs.HC]. Ecole Nationale Supérieure des Mines de Paris, 2007. Français. NNT : 2007ENMP1471 . tel-00179695

HAL Id: tel-00179695
https://pastel.hal.science/tel-00179695

Submitted on 16 Oct 2007

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.

MINES PARIS

T H E S E

pour obtenir le grade de
Docteur de l'Ecole des Mines de Paris
Spécialité "Géostatistique"

présentée et soutenue publiquement par
Florent LAFARGE

le 2 octobre 2007

MODELES STOCHASTIQUES POUR LA RECONSTRUCTION TRIDIMENSIONNELLE D'ENVIRONNEMENTS URBAINS

Directeurs de thèse : Josiane Zerubia et Marc Pierrot-Deseilligny

Jury

M.	Michel SCHMITT	Président
M.	Henri MAITRE	Rapporteurs
M.	Patrick PEREZ	
M.	Xavier DESCOMBES	Examinateurs
M.	Christian LANTUEJOUL	
M.	Marc PIERROT-DESEILLIGNY	
M.	Anuj SRIVASTAVA	
Mme	Josiane ZERUBIA	
M.	Jean-Philippe CANTOU	Examinateur invité

Remerciements

J'ai eu la chance d'être bien encadré durant ces trois années. Je le dois en premier lieu à Josiane Zerubia à qui j'exprime ma reconnaissance. Je la remercie de m'avoir accueilli dans le projet Ariana. Je remercie Marc Pierrot-Deseilligny pour ses nombreux conseils avisés et sa disponibilité. Merci à Didier Boldo pour l'aide précieuse qu'il m'a apportée au quotidien. Je remercie enfin Xavier Descombes, la personne qui m'a donné goût à la recherche. Il a su m'éclairer et m'aider dans des moments importants.

Je tiens à remercier Michel Schmitt de m'avoir fait l'honneur de présider le jury de thèse. Je souhaite exprimer ma gratitude à Henri Maître et Patrick Pérez qui ont accepté avec gentillesse d'être rapporteurs. Ils ont pris le temps de lire mon travail avec attention et leurs analyses ont été précieuses. Je remercie Christian Lantuejoul et Anuj Srivastava d'avoir accepté d'être examinateurs. Les discussions avec eux ont été fructueuses. Merci à Jean-Philippe Cantou pour avoir suivi mes travaux durant ces trois années depuis Toulouse.

Je remercie le CNES et l'IGN de m'avoir accordé une bourse de financement.

Merci aux membres du projet Ariana et du laboratoire Matis pour leur convivialité. En particulier, ceux qui m'ont supporté dans leur bureau : Guillaume, Maria, Olivier, Gilles, Frédéric, Clément, Nesrine, Mathieu, Mélanie, Emilie, Arnaud. Un grand merci tout particulièrement à Guillaume pour tous les bons moments passés en B121.

Cette thèse n'aurait pas vu le jour sans l'aide d'un certain nombre de personnes . Merci à Mathias de m'avoir expliqué et confié ses codes. Merci à David, Grégroire et Mélanie pour toutes les fois où ils m'ont fourni des données et lancé des calculs. Merci à Mathieu pour les discussions techniques et sa bonne humeur. Merci à Pierre et Yann pour leur contribution à ce travail au travers de leur stage d'ingénieur. Merci à Corinne, Alain et François pour leur disponibilité et leur gentillesse sur les questions pratiques. Merci aux personnes qui m'ont aidé dans la rédaction du manuscrit : Aurélie, Emilie, Mélanie, mon père.

Enfin et surtout, je remercie toute ma famille pour son soutien.

Avant propos

Cadre de la thèse

Les travaux présentés dans cette thèse ont débuté le 1^{er} octobre 2004. Ils ont été menés sous la direction de Josiane Zerubia, Directrice de Recherche à l'INRIA, de Marc Pierrot-Deseilligny, Directeur de Recherche à l'IGN, et de Xavier Descombes, Chargé de Recherche à l'INRIA. Cette thèse a été effectuée au sein du laboratoire Matis de l'IGN, et du projet Ariana commun à l'INRIA et à l'I3S. Le laboratoire Matis est spécialisé dans les problèmes de stéréorestitution. Un de ses objectifs est notamment la mise en production de méthodes automatiques et semi-automatiques de modélisation tridimensionnelle d'environnements urbains. Le Projet Ariana aborde les problèmes inverses en télédétection par des approches variationnelles et stochastiques. Pendant la durée de sa thèse, l'auteur a bénéficié d'une bourse CNES, cofinancée par l'IGN.

Le sujet de cette thèse a été mis en place par l'INRIA, l'IGN et le CNES. L'objectif initial consistait à étudier la complémentarité des travaux de thèses menés, d'une part par Mathias Ortner au sein du projet Ariana entre 2001 et 2004 sur l'extraction automatique de bâtiments, et d'autre part, par Hassan Jibrini réalisés entre 1998 et 2002 au laboratoire Matis en collaboration avec l'ENST sur la reconstruction 3D de bâtiments. Les données utilisées (des simulations PLEIADES, fournies par le CNES) devaient également permettre d'évaluer le potentiel des futurs satellites PLEIADES, dans le cadre du programme ORFEO, en matière de reconstruction 3D d'environnements urbains. Le principal intérêt qui se dégage de cette étude réside dans la reconstruction de zones urbaines sans l'utilisation de base de données 2D. Cette perspective est particulièrement intéressante pour la modélisation des villes ne possédant pas de cadastre, situation répandue dans les pays en voie de développement.

Organisation du manuscrit

Le manuscrit présente un processus de modélisation tridimensionnelle d'environnements urbains. Il est articulé de manière linéaire. Le chapitre 1 introduit la problématique de notre étude. Le chapitre 2 expose le processus d'extraction des objets urbains. Le chapitre 3 détaille la reconstruction tridimensionnelle de ces objets. Le chapitre 4 présente et analyse les résultats obtenus.

Le déroulement des recherches n'a toutefois pas été linéaire. Des modèles intermédiaires de reconstruction 3D et des méthodes d'optimisation alternatives ont été développés

au cours de la thèse. Ceux-ci sont résumés dans les annexes du manuscrit.

Les notations utilisées dans le manuscrit sont indépendantes entre les différents chapitres.

Table des matières

Avant propos **3**

1 Introduction **7**
1.1 Problématique 8
1.1.1 Les enjeux de la modélisation urbaine en 3D 8
1.1.2 L'apport de la télédétection 9
1.2 Etat de l'art 10
1.2.1 Les données 10
1.2.2 L'automaticité 14
1.2.3 Les zones urbaines 16
1.2.4 Les modélisations 3D 17
1.2.5 Les approches 20
1.2.6 Bilan 20
1.3 Le contexte de notre étude 21
1.3.1 Données spatiales et simulations PLEIADES 21
1.3.2 Vers de l'urbain dense en automatique 25
1.4 La stratégie 26
1.4.1 L'approche structurelle 26
1.4.2 Les outils stochastiques 27
1.4.3 Démarche générale 30

2 Extraction des bâtiments **31**
2.1 Caricatures de bâtiments par agencement de rectangles : rappel et analyse de travaux antérieurs 32
2.1.1 Processus ponctuels marqués 33
2.1.2 Formulation de l'énergie 33
2.1.3 Commentaires 38
2.2 Régularisation en supports structuraux 39
2.2.1 Connexion des supports 39
2.2.2 Détection des discontinuités des hauteurs de toit 43

3 Reconstruction 3D des bâtiments **47**
3.1 Bibliothèque de modèles 3D 48
3.1.1 Typologie des toits et adéquation avec les données 48
3.1.2 La bibliothèque proposée 51
3.2 Formulation Bayésienne 53

3.2.1 Vraisemblance . . . 56
3.2.2 A priori . . . 58
3.3 Optimisation . . . 60
3.3.1 Echantillonneur RJMCMC . . . 61
3.3.2 Noyaux de propositions . . . 63
3.3.3 Recuit simulé . . . 66
3.3.4 Des méthodes alternatives . . . 69
3.4 Estimation des paramètres . . . 72
3.4.1 Paramètres de la densité . . . 72
3.4.2 Paramètres de l'optimisation . . . 73

4 Evaluation **75**
4.1 Evaluation qualitative . . . 76
4.2 Evaluation quantitative . . . 84
4.3 Influence du recuit simulé . . . 86
4.4 Influence des supports 2D . . . 91
4.5 Les limites d'utilisation . . . 94
4.6 Extension du contexte . . . 96

5 Conclusion **101**

Bibliographie **107**

6 Annexes **117**
A - Processus 3D paramétrique à partir d'emprises rectangulaires . . . 118
B - Processus 3D par squelétisation des arêtes faîtières . . . 124
C - Rappel sur les chaînes de Markov . . . 128
D - Calcul des noyaux de propositions . . . 133
E - Processus 3D après simplification séquentielle des supports 2D . . . 134
F - Résultats supplémentaires . . . 137

Chapitre 1

Introduction

1.1 Problématique

La moitié de la population mondiale vit, aujourd'hui, dans les villes. En 1950, moins de 30% des hommes habitaient en milieu urbain. D'après une étude récente des Nations Unies[1], ils seront plus de 60% en 2030. Le milieu urbain devient progressivement l'environnement naturel de l'homme. Devant ce bouleversement démographique, connaître et maîtriser l'espace urbain est un enjeu essentiel. En particulier, la reconstruction vectorielle tridimensionnelle des villes tend à devenir un outil de plus en plus important dans de nombreux secteurs, qu'ils soient publics, industriels ou militaires.

1.1.1 Les enjeux de la modélisation urbaine en 3D

La constitution de bases de données urbaines tridimensionnelles représente un enjeu important dans de nombreux domaines. Dans le secteur public, les applications se focalisent principalement autour de l'aménagement urbain. Ces bases de données sont utiles dans des études concernant l'implantation d'espaces verts, le positionnement de lampadaires (simulation d'éclairage), ou bien encore l'impact de la construction de bâtiments dans un environnement. Elles peuvent également se révéler être efficaces pour le suivi de l'évolution des villes et pour la gestion des risques tels que les inondations ou les séismes. Dans le secteur industriel, les modélisations 3D de villes intéressent beaucoup les opérateurs de téléphonie mobile. Elles leur permettent de simuler la propagation des ondes électromagnétiques et ainsi améliorer le développement de leurs réseaux de télécommunication. L'industrie cinématographique, ainsi que celles du jeu vidéo et de l'automobile, sont également demandeuses de ces modélisations 3D. Enfin, il existe des applications militaires comme l'aide à la préparation d'opérations ou au guidage d'engins autonomes tels que les drônes. Les possibilités d'utilisation sont donc très nombreuses. Le lecteur pourra trouver dans [Fritsch, 1999], une gamme d'applications possibles.

Le procédé historique d'acquisition de modélisation 3D de ville est le relevé topométrique de bâtiments. Des opérateurs saisissent manuellement des points d'intérêt sur le terrain. Ces points sont ensuite reportés à l'intérieur d'une scène virtuelle, et permettent ainsi de modéliser le bâtiment en 3D. Bien que très précise, cette méthode est lente et coûteuse. Elle est utilisée principalement pour la modélisation de bâtiments remarquables. Cependant, pour une reconstruction "en masse", c'est-à-dire de villes entières, ce procédé devient très vite pénible. Il n'est d'autant pas adapté que les modélisations urbaines ont besoin de mises à jour fréquentes. Dans ce contexte, il est nécessaire de se tourner vers des procédés où l'intervention humaine est moindre. L'imagerie terrestre est une des principales alternatives, notamment pour aborder les problèmes liés à la reconstruction et la texturation des façades. On peut citer à cet effet les travaux de [Wang et al., 2002; Zhao et Shibasaki, 2001; Frueh et al., 2005; Penard et al., 2006]. Le recours à la télédétection, c'est-à-dire aux méthodes permettant d'extraire des informations à partir de données aériennes et satellitaires, est également très prisé.

[1] "Perspectives de l'urbanisation mondiale : la révision 2001" publié en 2002 dans le volume 39(3) du magazine "Chronique de l'ONU".

1.1.2 L'apport de la télédétection

L'acquisition rapide et massive de données urbaines, la couverture importante des avions et satellites, ou bien encore les avancées récentes dans le domaine spatial concernant le coût et la qualité des images satellitaires, font de la télédétection le moyen le plus efficace aujourd'hui pour aborder la modélisation urbaine en 3D. Le temps mis pour acquérir des données aériennes ou satellitaires sur une ville est notamment beaucoup plus court que le temps mis par un géomètre pour effectuer un relevé topographique. Il existe principalement trois types de capteurs permettant d'acquérir des données aériennes et satellitaires : les capteurs optiques, LASER et RADAR. Nous reviendrons sur ces différents types de données dans la partie 1.2.1.

La modélisation urbaine en 3D par la télédétection présente de nombreuses difficultés qui doivent être prises en compte lors de la mise en place des processus. Voici les principales d'entre elles :

- **La diversité des objets :** Les scènes urbaines ne sont pas uniquement constituées de bâtiments. Arbres, pelouses, haies, voitures, routes, trottoirs, ponts ou bien encore rivières : voici un panel d'objets que l'on trouve régulièrement dans le paysage urbain et qui doivent être pris en compte lors de la mise en place du processus de modélisation 3D. Ces objets diffèrent par des caractéristiques de forme, de taille, de radiométrie ou bien encore de texture. Bien souvent une étape de classification, préliminaire au processus de reconstruction 3D, est nécessaire afin de localiser et d'identifier ces différents types d'objets dans la scène.

- **La densité des objets :** Le nombre d'objets peut varier de manière importante d'une scène urbaine à une autre. L'hétérogénéité de la répartition des objets constitue une source de difficultés. Les zones urbaines denses posent notamment des problèmes relatifs à la reconnaissance et à l'extraction des objets.

- **La complexité des bâtiments :** Afin de modéliser efficacement les zones urbaines, il est important de connaître et comprendre les différents types de bâtiments, et surtout les lois qui régissent leur construction. Or, ces dernières sont à la fois très variées et complexes. Elles dépendent de nombreux facteurs tels que le contexte historique (différentes vagues de construction au cours des époques) ou l'aspect géographique (à différentes échelles : nationale, régionale, intra-urbaine). Tout ceci contribue à créer un certain désordre au sein des zones urbaines, et rend difficile leur analyse, en particulier les études sur les typologies de bâtiments.

- **La spécificité des données :** Les différents types de données utilisés en télédétection ont chacun leurs caractéristiques propres en matière de reconstruction 3D. Tous rencontrent des difficultés pour extraire des informations 3D dans une scène. Nous détaillerons ces difficultés par la suite.

1.2 Etat de l'art

Comme cela a été souligné précédemment, la reconstruction 3D de scènes urbaines en télédétection est un problème complexe. De nombreux systèmes ont été développés dans des contextes variés. Dans cette partie, nous nous penchons sur les méthodes existantes et tentons d'en dégager les points forts et les inconvénients. Cette étude est déterminante pour le choix du système qui sera proposé dans le cadre de cette thèse. Les méthodes de reconstruction peuvent être classées en fonction de cinq critères principaux :

- **Le type des données utilisées** - optique (monoscopie, stéréoscopie, multiscopie), LASER, RADAR, produits dérivés.
- **Le niveau d'automaticité du système** - automatique, automatique avec données cadastrales, semi-automatique.
- **La complexité des zones urbaines reconstruites** - urbain dense, peri-urbain, zones d'activité et d'habitats collectifs.
- **Le degré de généralité de la modélisation 3D** - prismatique, paramétrique, structurelle, polyédrique.
- **Le type d'approche utilisé** - "Top-Down", "Bottom-Up", "Hypothesize-And-Verify".

1.2.1 Les données

Les données utilisées pour la modélisation urbaine 3D doivent pouvoir fournir, directement ou indirectement, des informations tridimensionnelles sur la scène observée. Nous évoquons brièvement ici, les principaux types de données possédant cette spécificité.

Résolution des données

La résolution des données est un point important permettant de choisir le niveau de détails de la modélisation 3D finale. Celle-ci dépend fortement du type de données utilisé. Dans ce paragraphe, nous nous efforçons d'établir un lien entre la résolution de données optiques aériennes, et le niveau de détails pouvant être exigé pour la reconstruction 3D. D'une manière générale, on peut distinguer quatre gammes de résolution :

- **>1m -** ce sont des résolutions très basses pour traiter les problèmes de reconstruction en 3D de bâtiments. On peut cependant espérer fournir une modélisation par toits plats (i.e. avec une simple estimation de la hauteur des bâtiments).
- **entre 1m et 50cm -** cette gamme permet de proposer des modélisations assez simples et généralisées, laissant apparaître les principaux plans de toits qui constituent un bâtiment.
- **entre 50cm et 20cm -** ces résolutions sont adaptées pour fournir des reconstructions détaillées des toitures, c'est-à-dire prenant en compte les décrochements de toits et les petits pans mineurs qui constituent la toiture d'un bâtiment.
- **<20cm -** cette gamme permet d'obtenir des modélisations très détaillées et complètes, en reconstruisant notamment des objets urbains très intéressants tels que les superstructures des toitures (cheminée, chien-assis,...). On peut citer à cet effet les travaux de [Brédif et al., 2007].

Ces gammes de résolution correspondent à des données optiques aériennes. Avec des données optiques satellitaires, cela n'est plus forcément valable, celles-ci ayant une moins bonne qualité d'image à résolution identique.

Monoscopie

La monoscopie consiste à extraire des informations tridimensionnelles à partir d'une seule image. [McGlone et Shufelt, 1994] et [Lin et Nevatia, 1998] proposent des approches reposant sur le "shape from shadow", c'est-à-dire en utilisant les informations inhérentes à l'ombre des objets. Ce type de procédés fournit des modélisations assez pauvres, où les bâtiments sont représentés par des parallépipèdes rectangles dont la hauteur est estimée en fonction de la projection de l'ombre dans l'image. De plus, en considérant l'ombre comme unique information tridimensionnelle, ces méthodes manquent de robustesse. Aujourd'hui, les systèmes à base de monoscopie tendent à disparaître.

Stéréoscopie

La stéréoscopie est le procédé de référence, du moins celui qui a focalisé l'attention des chercheurs durant la dernière décennie. Cette technique permet de calculer le relief d'une scène à partir de deux images prises à des points de vue différents. La vision humaine utilise ce processus et permet à l'homme d'évoluer naturellement dans un environnement en ayant cette notion de distance par rapport aux objets qui l'entourent. Ce procédé doit gérer deux difficultés principales :

- *La mise en correspondance* qui consiste à retrouver les pixels homologues dans le couple d'images (également appelée appariement). Cette opération est loin d'être évidente et fait appel à des mesures de corrélation parfois très complexes. Les algorithmes doivent faire face aux changements de luminosité, aux bruits, aux déformations des images ou bien encore aux zones homogènes.
- *Les zones d'occultation* qui rendent localement impossible la mise en correspondance puisqu'un point de la première image n'a pas d'homologue dans la seconde. Or ces points sont fréquents en stéréoscopie et obligent les algorithmes à introduire des informations a priori.

Le paramètre principal lors d'une prise de vue stéréoscopique est le rapport B/H (base sur hauteur), où B représente la distance entre les deux centres de prise de vue, et H, la hauteur entre la base des prises de vue et la scène observée (voir Figure 1.1). Ce rapport permet de déterminer la précision altimétrique théorique du relief en fonction de la précision planimétrique. La valeur de ce rapport soulève un compromis entre la difficulté d'appariement et la précision altimétrique. En théorie, un B/H faible (i.e. deux images dont les prises de vue sont très proches) permettra une mise en correspondance relativement facile en restreignant notamment les zones d'occultation, cependant la précision altimétrique sera mauvaise. Et inversement. En pratique, ce principe n'est pas toujours linéaire et l'on observe parfois des intervalles de B/H clairement avantageux.

Ce procédé, qui fournit des informations 3D fiables, a été utilisé dans de nombreux travaux. Les algorithmes proposés doivent cependant gérer les zones d'occultation. Face à ce problème, certains optent pour une modélisation 3D simplifiée. C'est le cas notamment de [Paparoditis et al., 1998] proposant uniquement des toits plats ou [Dang, 1994] reconstruisant des bâtiments parallépipédiques. D'autres, comme [Gruen, 1998; Gruen et Wang, 1998], évoluent dans un contexte semi-automatique, c'est-à-dire en permettant à un opérateur d'intervenir et de rectifier les erreurs de reconstructions. [Fuchs, 2001] et [Jibrini, 2002] ont proposé des modèles intéressants permettant des reconstructions relativement complexes.

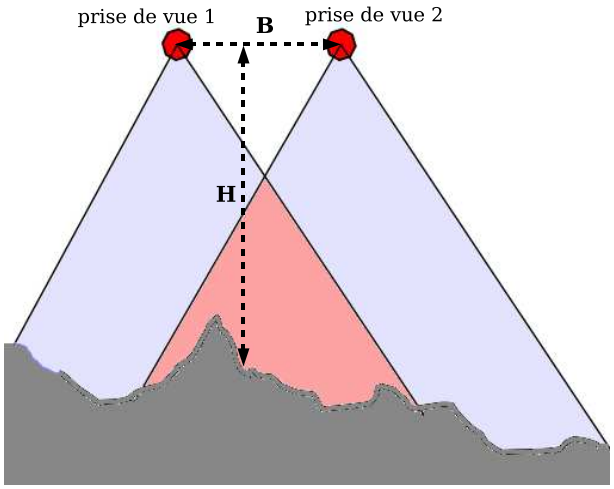

FIG. 1.1 – Principe de la stéréoscopie

Multiscopie

La multiscopie est une extension de la stéréoscopie, dans le cas où l'appariement est réalisé avec plus de deux vues. C'est aujourd'hui la méthode d'extraction d'information 3D la plus efficace. Ce procédé permet de limiter les problèmes d'occultation et d'accroître la précision du relief. La multiscopie est notamment très utile pour la détection et la texturation de façades grâce aux nombreux angles de vue qu'elle offre. La plupart des travaux réalisés proposent des modélisations 3D relativement détaillées, avec des formes de bâtiments génériques [Willuhn et Ade, 1996; Henricsson, 1998; Baillard et Zisserman, 1999; Heuel et al., 2000; Scholze et al., 2002; Taillandier, 2004; Durupt et Taillandier, 2006]. Par ailleurs, une partie importante des méthodes mises en place introduit de fortes connaissances a priori sur les structures urbaines afin de réduire les temps de calcul et proposer des modélisations 3D plus simples, comme par exemple [Jaynes et al., 1997; Collins et al., 1998; Fischer et al., 1998; Noronha et Nevatia, 2001].

LASER

Contrairement à ce qui précède, les données LASER n'utilisent pas une géométrie image. En effet, ces données ont un échantillonnage spatial irrégulier, ce qui rend les traitements plus difficiles que dans le cas de l'imagerie optique. Le second défaut majeur est le coût d'acquisition, qui reste élevé malgré le développement important du LASER ces dix dernières années. Malgré cela, ces données possèdent des caractéristiques très intéressantes. Les principales qualités du LASER résident dans la précision et la fiabilité des points relevés. Une caractéristique également profitable est le phénomène dit de double écho : le signal émis est réfléchi par la végétation, mais également par le sol après avoir traversé cette dernière, ce qui est particulièrement intéressant pour l'extraction de la végétation et la discrimination bâti / végétation. Plusieurs méthodes de modélisation urbaine 3D utilisant des données LASER ont été proposées [Haala et Brenner, 1999; Maas et Vosselman, 1999; Rottensteiner et Briese, 2002; Bretar, 2006]. La précision des points LASER permet, en

général, d'obtenir des résultats robustes.

RADAR

L'intérêt majeur de l'imagerie RADAR réside dans le caractère "tout temps" de l'acquisition. Les images RADAR peuvent, en effet, être prises aussi bien la nuit que par temps nuageux, ce qui rend leur utilisation intéressante notamment dans le domaine militaire et pour la gestion de catastrophes naturelles. Ces données occasionnent cependant des inconvénients pénalisants pour la modélisation urbaine en 3D : difficultés d'interprétation de l'image, double (voire triple) échos, bruits variés et importants qui donnent un signal très dégradé, résolution relativement faible par rapport aux autres types de données. [Soergel et al., 2005] se propose d'étudier la faisabilité des applications liées à la reconstruction urbaine. Peu de travaux ont été réalisés dans ce domaine. On peut citer les modèles mis en place par [Gamba et al., 2000; Simonetto et al., 2003; Tupin et Roux, 2003; Cellier, 2007]. Les modélisations 3D obtenues à partir de données RADAR restent très généralisées et n'ont pas encore atteint le niveau de complexité des reconstructions fournies par les données optiques et LASER.

Produits dérivés

Il existe également des produits dérivés des données évoquées précédemment. C'est le cas en particulier des Modèles Numériques d'Elévation (MNE). Un MNE décrit le relief d'une zone urbaine à travers une grille régulière de points en chacun desquels est associée une information de hauteur. Ce produit est très intéressant puisqu'il s'assimile directement à une description géométrique tridimensionnelle d'une scène.
Les MNE sont obtenus à partir de différents types de données. La plupart sont générés à partir d'images optiques stéréoscopiques, et plus généralement multiscopiques, par mise en correspondance. Parmi les méthodes proposées, on retient notamment [Collins, 1996; Baillard, 1997; Roy et Cox, 1998]. La figure 1.2 montre un MNE optique à 25cm de résolution sous deux déclinaisons : un MNE raster qui décrit l'altimétrie de la zone pixel par pixel, et un MNE ombré qui présente la scène en perspective avec un jeu d'ombre pour une visualisation plus réaliste. Cette figure montre aussi le bruit important présent dans les MNE, et particulièrement sur les plans de toits. Ce bruit souligne le problème d'incertitude en vision stéréoscopique et les difficultés à traiter les ambiguités [Senegas, 2002; De Joinville, 2003].
Les MNE obtenus à partir de données LASER sont également très intéressants. En effet, les MNE LASER sont réputés pour leur qualité, et en particulier, pour la bonne et franche localisation des façades de bâtiments, et plus généralement la bonne retranscription des discontinuités altimétriques qui est bien souvent un problème pour les MNE optiques. Enfin, il existe également des MNE générés à partir de données RADAR, comme par exemple [Tison, 2004]. Comparés aux MNE optiques et LASER, ces derniers sont malheureusement limités, principalement par les inconvénients inhérents au RADAR.

Les MNE constituent de bonnes descriptions tridimensionnelles de scènes malgré un bruit important. Beaucoup de méthodes de reconstruction de bâtiments les utilisent donc comme données initiales. C'est le cas des processus proposés par [Weidner et Forstner, 1995; Vinson et Cohen, 2002; Flamanc et al., 2003; Ortner, 2004]. Certains travaux, comme par exemple [Chehata et al., 2005], proposent de générer des MNE dits hybrides, c'est-

FIG. 1.2 – Extrait de MNE à 25 cm - version RASTER *(gauche)* et version ombrée *(droite)* - ©IGN.

à-dire dans lesquels sont introduites des contraintes planaires permettant de modéliser les facettes de toits et de façades.
Il existe un deuxième produit dérivé : l'orthoimage. L'orthoimagerie consiste à "redresser" une image de sorte à la rendre superposable à une carte. Les orthoimages sont notamment très utiles pour la détection des contours des bâtiments, comme on peut le constater dans [Vinson et Cohen, 2002]. Le calcul des orthoimages est fondé sur celui des MNE. Par conséquent, les orthoimages souffrent des mêmes problèmes que les MNE en matière de mise en correspondance.

1.2.2 L'automaticité

Un critère majeur dans la classification des méthodes est le niveau d'automaticité. Celui-ci caractérise le degré d'intervention de l'opérateur mais également l'apport éventuel de données cadastrales. En effet, les bases de données 2D fournissent une information structurelle très forte. L'utilisation du cadastre comme une donnée source change fondamentalement l'approche à mettre en œuvre. Nous présentons trois degrés d'automaticité de processus : semi-automatique, automatique avec cadastre et automatique sans cadastre.

Processus semi-automatiques

Les méthodes semi-automatiques sont, à l'heure actuelle, les seules capables de produire à grande échelle des modélisations 3D urbaines avec un taux d'erreur très faible. Il existe plusieurs processus particulièrement intéressants. C'est le cas du *CC-modeler*, développé par [Gruen et Wang, 1998], où l'opérateur saisit les points correspondant au bord du bâtiment, puis ceux correspondant à des points clé de la topologie du toit comme les extrémités d'arêtes faîtières. Le système peut alors reconstruire le bâtiment sous une forme polyédrique. Cependant, ce processus ne gère pas les points erronés (i.e. les erreurs ou inexactitudes de l'opérateur). Proposé par [Flamanc et Maillet, 2005], *BATI3D* est un système interactif permettant le partitionnement du cadastre, voire sa modification, par un opérateur. Les bâtiments sont ensuite reconstruits automatiquement par une des différentes méthodes

incorporées au système. [Li et al., 1999; Nevatia et Price, 2002] proposent un système relativement complet où l'opérateur est peu sollicité comparativement aux systèmes précédents. Une extension de ce système a été proposée, permettant un niveau de reconstruction 3D très détaillé grâce à la modélisation d'éléments structuraux urbains tels que des fenêtres, des colonnes ou bien encore des pontons [Lee et Nevatia, 2003].

Processus automatiques avec données cadastrales

La reconstruction d'une zone urbaine sans l'intervention d'un opérateur est un problème complexe. C'est pourquoi certains processus s'appuient sur des données cadastrales qui, quand elles existent, offrent une information structurelle très forte sur les bâtiments. Le cadastre fournit les emprises au sol des structures urbaines, ce qui permet de s'affranchir d'une étape de focalisation/extraction des bâtiments. Le processus proposé par [Brenner et al., 2001] est fondé sur une classification des différentes jonctions triples par intersections des plans. Le résultat est convainquant, malgré une modélisation 3D très simplifiée due notamment au fait que les éventuelles discontinuités altimétriques ne sont pas prises en compte. [Jibrini, 2002] propose une méthode générant des hypothèses de plans par des transformées de Hough 3D. L'intersection des plans forment alors un ensemble de facettes possibles parmi lesquelles on doit choisir la meilleure selon des critères de complexité sur la forme des toits, et d'attache aux données. Bien que les modélisations 3D soient très génériques, le processus souffre d'un problème combinatoire puisque le nombre de plans détectés et le nombre de facettes possibles sont extrêmement élevés. Il existe également beaucoup d'autres processus, par exemple [Suveg et Vosselman, 2001; Flamanc et al., 2003].

Processus automatiques

Cette catégorie soulève de nombreuses difficultés de mise en œuvre, puisqu'il faut à la fois localiser les bâtiments, extraire leur emprise au sol, et les reconstruire en 3D. La majorité des méthodes proposées dans la littérature se situe cependant dans ce contexte. Les processus *tout automatiques* ayant un taux faible d'erreurs constituent un des principaux objectifs de demain dans le domaine. Ils permettraient de produire à grande échelle des modélisations 3D urbaines sans coût opérateur.
Dans le contexte *tout automatique*, il faut distinguer deux types de processus : ceux qui utilisent une focalisation des bâtiments, et ceux qui utilisent une extraction des bâtiments. La focalisation et l'extraction sont des termes proches, mais l'utilisation de l'une ou l'autre a des conséquences différentes sur l'applicabilité et l'automaticité du processus.

Focalisation des emprises Les processus fondés sur une focalisation d'emprises permettent de reconstruire une à une des structures urbaines connexes. Pour ce faire, il est nécessaire de définir a priori une zone englobante dans laquelle se trouve une unique structure urbaine connexe. Ces zones englobantes de focalisation sont définies de manière automatique. En pratique, le processus fonctionne pour les bâtiments relativement bien isolés, principalement sur des zones péri-urbaines, mais montre ses limites sur des zones urbaines denses où la densité élevée des structures urbaines ne permet pas un bon fonctionnement de l'algorithme de focalisation. Les travaux utilisant cette stratégie sont nombreux, on peut citer notamment [Baillard et Zisserman, 1999; Scholze et al., 2002; Taillandier, 2004].

Extraction des emprises Les processus fondés sur une extraction des emprises ne procèdent pas "bâtiment par bâtiment" comme pour un procédé de focalisation, mais ils traitent des zones entières. Cela est réalisé généralement en deux étapes : extraction des emprises de bâtiments d'une scène (sous formes de polygones dans la plupart des cas), puis reconstruction en 3D des bâtiments à partir des emprises extraites. Les méthodes d'extraction des emprises dépassent la simple notion de bâtiments isolés et prennent en compte les interactions existantes entre bâtiments voisins, voire îlots de bâtiments. Par conséquent, ces méthodes sont bien adaptées aux zones denses. Cependant, elles sont plus complexes et plus difficiles à mettre en œuvre.
[Oriot, 2003; Bailloeul et al., 2005] utilisent des contours actifs avec des contraintes linéaires. Ces processus rencontrent cependant des problèmes relatifs à l'initialisation des contours. De plus, ce type d'outil variationnel n'est pas particulièrement bien adapté à l'extraction de formes rectilignes. [Vestri, 2000] propose un algorithme intéressant de vectorisation des contours de bâtiments préalablement extraits à partir d'une segmentation de MNE. Les travaux proposés par [Vinson et Cohen, 2002] et [Ortner et al., 2007] permettent d'extraire les emprises sous forme d'un agencement de rectangles, le premier par un modèle de rectangle déformable, le second en utilisant des processus ponctuels marqués qui sont des outils probabilistes sans contrainte d'initialisation.

1.2.3 Les zones urbaines

Le degré de complexité des zones urbaines reconstruites est un critère à prendre en compte pour classer les méthodes. Nous distinguons trois grandes familles de zones urbaines : les zones d'activité, les zones péri-urbaines et les zones urbaines denses. Les processus de reconstruction sont plus ou moins adaptés à certaines de ces familles. Il est assez rare de trouver des méthodes fournissant des résultats convaincants sur chacun de ces types de zones.

Zones urbaines denses

Les zones urbaines denses correspondent aux centres-villes, et plus généralement à des zones où la concentration de bâtiments est très élevée. La figure 1.3-(*droite*) présente une telle zone, en l'occurrence, une partie du centre-ville d'Amiens. Outre une forte densité de bâtiments, ces zones sont caractérisées par une variabilité importante des formes de toits, l'existence d'interactions entre structures voisines, et une complexité de l'environnement due aux différents types d'objets qui le composent. Pour ces raisons, ce type de zones génère des difficultés importantes dans la mise en place d'un processus de reconstruction.
D'ailleurs, quand on se penche sur la littérature, peu de méthodes automatiques ont été appliquées dans un contexte urbain dense. Avec l'utilisation de données cadastrales, les méthodes proposées par [Jibrini, 2002; Flamanc et al., 2003; Durupt et Taillandier, 2006] fournissent de bonnes modélisations 3D sur de telles zones.

Zones péri-urbaines

Les zones péri-urbaines représentent les quartiers pavillonnaires ou résidentiels. Ces zones regroupent principalement deux types d'habitations : les habitats individuels (i.e. des

maisons espacées entre elles) et les habitats groupés de style maisons mitoyennes. La densité de bâtiments est relativement faible, et bien souvent, une part importante de végétation est présente (voir figure 1.3-(*centre*)). La variabilité des bâtiments est peu prononcée, tant au niveau de la forme des toits que de la hauteur des constructions. Les difficultés générées par ce type de zones sont donc plus faciles à traiter que celles générées par les zones urbaines denses.
Beaucoup de processus automatiques sans données cadastrales ont été testés sur ce type de zones, notamment [Baillard et Zisserman, 1999; Fuchs, 2001; Scholze et al., 2002; Taillandier, 2004]. Le fait que les bâtiments soient principalement de petites structures assez simples et isolées, permet en effet d'extraire les emprises assez facilement.

Zones d'activité et d'habitats collectifs

Les zones d'activité et d'habitats collectifs sont caractérisées par un même type de constructions : des bâtiments volumineux, possédant souvent des toits plats (de type hangars, entrepôts ou grands immeubles), et généralement espacés entre eux (voir figure 1.3-(*gauche*)). Ces zones sont, en théorie, les moins difficiles à reconstruire. Elles sont souvent utilisées pour tester les modèles dits prismatiques, c'est-à-dire estimant les toits par des plans horizontaux. C'est le cas des travaux menés par [Collins et al., 1998; Lin et Nevatia, 1998; Noronha et Nevatia, 2001; Vinson et Cohen, 2002].

FIG. 1.3 – Différents types de zones urbaines - zone d'activités *(gauche)*, péri-urbaine *(centre)*, et urbaine dense *(droite)* - ©IGN.

1.2.4 Les modélisations 3D

Il existe différents types de modélisations 3D, plus ou moins complexes et détaillés. Le choix de la modélisation dépend bien souvent des données utilisées et de leurs caractéristiques (principalement la résolution). Les modélisations sont généralement des associations de facettes planes connectées entre elles. Une facette plane est, en effet, un très bon moyen de modéliser à la fois les façades et les pans de toits. Les quatre principaux types de modélisations 3D, à savoir les modélisations prismatiques, paramétriques, structurelles et polyédriques, sont présentés dans la suite.

Prismatique

Les modélisations prismatiques se caractérisent par un degré de liberté planimétrique important puisque les emprises des bâtiments sont des formes polygonales quelconques, mais cependant avec une description altimétrique très simplifiée : les formes de toits possibles sont, en effet, restreintes à la plus simple existante : la forme plate. La figure 1.4-(a) schématise cette modélisation. Une reconstruction prismatique est donc intéressante lorsque l'on cherche une bonne description planimétrique avec une simple estimation des hauteurs de bâtiments, et peut s'avérer efficace pour la description des zones d'activité ou lorsque les données sont faiblement résolues. A cet effet, on peut citer les processus mis en œuvre par [Collins et al., 1998; Lin et Nevatia, 1998; Noronha et Nevatia, 2001; Vinson et Cohen, 2002].

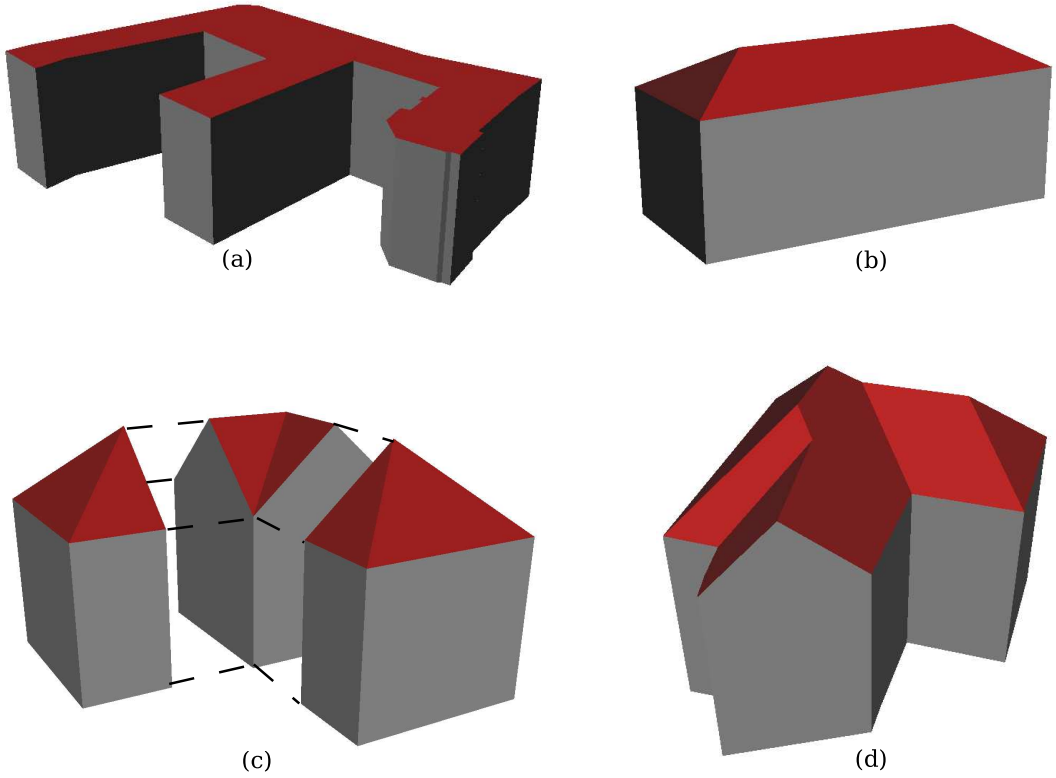

FIG. 1.4 – Les différents types de modélisations 3D - prismatique *(a)*, paramétrique *(b)*, structurelle *(c)*, et polyédrique *(d)*.

Paramétrique

Une reconstruction paramétrique fournit des formes de bâtiments assez simples, et par conséquent, une description relativement générale des zones urbaines. Elle apporte, cependant, davantage de précision que les modélisations prismatiques sur la forme de la toiture

en proposant bien souvent des toits bi-pentes, quatre-pentes et des attaques en V. Certains processus consistent à choisir parmi une collection d'hypothèses de formes simples de toits comme [Weidner et Forstner, 1995; Garcin et al., 2001; Suveg et Vosselman, 2001; Flamanc et Maillet, 2005], d'autres utilisent des algorithmes de squelétisation des emprises afin de modéliser les arêtes faîtières [Gruen et Wang, 1998]. Les modélisations paramétriques sont intéressantes pour traiter des données à faible résolution : la simplicité des modèles nécessite, en compensation, un apport de connaissances a priori sur les structures urbaines permettant de faire face au manque d'information des données. Par ailleurs, en modélisant un bâtiment par un jeu simple de paramètres, la combinatoire des processus est peu complexe, ce qui leur confère bien souvent de la robustesse.

Structurelle

Les modélisations structurelles décrivent un bâtiment comme un assemblage de modules élémentaires simples. Ces modules, qui sont souvent eux-mêmes des modèles paramétriques, représentent par exemple des terminaisons de bâtiments (terminaison droite, attaque en V,...) ou des jonctions entre modules (jonctions en "L", en "T",...). Les modélisations structurelles s'apparentent ainsi à un jeu de "LEGO". Là encore, ce type de modélisation est adapté aux données faiblement résolues. Il constitue un bon compromis entre les modélisations paramétriques et les modélisations polyédriques du point de vue de la généricité. Il nécessite néanmoins la mise en place de lois grammaticales, c'est-à-dire de contraintes sur l'assemblage des modules entre eux. [Fischer et al., 1998] propose une reconstruction structurelle relativement simple, n'utilisant qu'une seule sorte de modules (des coins) pour générer des hypothèses de bâtiments. [Fuchs, 2001] présente un processus plus abouti, avec notamment une grammaire de modèles plus complète, formalisé par une recherche d'isomorphisme de sous-graphes avec tolérance d'erreur. La complexité algorithmique est cependant assez importante.

Polyédrique

Les modélisations polyédriques sont les plus utilisées car elles sont très génériques et permettent de reconstruire une large gamme de structures urbaines sous forme de facettes planes connexes. Cette importante généricité impose des contraintes sur la qualité des données, en particulier une très bonne résolution. Beaucoup de méthodes [Henricsson, 1998; Ameri, 2000; Heuel et al., 2000; Taillandier, 2004] ne limitent pas a priori les formes à reconstruire. Cela engendre cependant une complexité algorithmique très importante. Dans les travaux de [Baillard et Zisserman, 1999], la modélisation est obtenue par intersections de demi-plans qui sont générés à partir de segments 3D (les demi-plans potentiels tournent autour des segments 3D). Celle proposée par [Scholze et al., 2002] est relativement similaire : des segments 3D sont extraits, puis groupés en facettes à travers une interprétation sémantique. La modélisation de [Heuel et al., 2000] est fondée sur la détection de coins 3D (à trois arêtes) constituant la base de la structure topologique. [Ameri, 2000] extrait des plans 3D qui sont intersectés selon un diagramme de Voronoï modélisant les interactions entre plans voisins. Ces méthodes fournissent, en général, des résultants convaincants. Cependant, quand les bâtiments ont des formes relativement simples, ce type de processus n'est pas adapté en raison des temps de calcul importants mis pour explorer l'espace des

solutions.

1.2.5 Les approches

Le dernier critère permettant de caractériser les méthodes de reconstruction est l'approche utilisée. Il existe trois grandes familles d'approches : *bottom-up*, *top-down* et *hypothesize-and-verify*.

Bottom-up

Les approches "bottom-up" partent de primitives pour proposer la reconstruction finale, sans prendre en compte d'éventuelles connaissances a priori sur les bâtiments que l'on cherche à reconstruire. Les approches "bottom-up" sont donc génériques et nécessitent des données de bonne qualité. Un inconvénient souvent inhérent à ce type d'approches est le manque de robustesse. En effet, l'extraction de primitives est sujet aux problèmes de sur-détection et de sous-détection qui influence le résultat final. De plus, l'absence (ou du moins le manque) de connaissances a priori sur les bâtiments génère parfois des résultats peu réalistes à cause d'une combinatoire importante. Dans la majorité des processus "bottom-up", les primitives extraites sont des segments 3D ou/et des plans 3D. Il existe également des primitives de nature différente comme, par exemple, des hypothèses de façades [Taillandier, 2004] ou des coins 3D [Heuel et al., 2000].

Top-Down

Dans les approches "top-down", à l'inverse, des connaissances a priori sur les bâtiments sont introduites. On part ainsi d'un ensemble de modèles connus à l'avance pour alors choisir celui qui présente la meilleure cohérence avec les données. Ces approches sont généralement plus robustes que les approches "bottom-up", mais se trouvent limitées par la librairie de modèles définie a priori. Parmi les processus "top-down", on trouve par exemple [Fuchs, 2001; Garcin et al., 2001; Suveg et Vosselman, 2001; Vinson et Cohen, 2002; Flamanc et al., 2003].

Hypothesize-And-Verify

Les approches "hypothesize-and-verify" sont des approches hybrides entre les approches "bottom-up" et "top-down". Elles consistent, dans un premier temps, à formuler des hypothèses à partir de primitives par une approche "bottom-up", puis dans un second temps, à vérifier ces hypothèses par une approche "top-down". Ces approches sont utilisées notamment dans [Willuhn et Ade, 1996; Fischer et al., 1998; Noronha et Nevatia, 2001; Scholze et al., 2002].

1.2.6 Bilan

Dans ce paragraphe, nous nous sommes penchés sur les méthodes existantes et leurs caractéristiques. Cette étude fait ressortir deux points importants :

- **Des contextes d'étude très variés.** Comme nous venons de le voir, il existe beaucoup de critères permettant de classer les méthodes existantes. En conséquence, il est souvent difficile de comparer les méthodes entre elles. Plus précisément, pour effectuer une comparaison pertinente entre deux systèmes, il paraît nécessaire que les données soient identiques (ou du moins d'une qualité équivalente, principalement la résolution), que les scènes reconstruites soient d'un même degré de complexité, et enfin que le niveau d'automaticité soit le même.

- **Deux grandes familles de processus.** Bien qu'il soit difficile de comparer les méthodes entre elles, on peut distinguer deux grandes familles de processus. La première utilise des données de basse qualité et fournit des modélisations 3D peu génériques, à partir de librairies de modèles, en utilisant des approches de type "top-down". La seconde famille, spécifique à des données de meilleure qualité, propose des reconstructions génériques fondées sur des approches "bottom-up". Ces deux familles illustrent bien le compromis existant entre la généricité des modélisations et la robustesse des processus.

1.3 Le contexte de notre étude

Dans cette partie, nous détaillons le contexte de notre étude, en particulier les données utilisées, le type de zones à reconstruire et le niveau d'automaticité exigé. Ce contexte est très spécifique : il n'existe pas, actuellement, de processus développé suivant nos contraintes. Il nous est donc malheureusement impossible de comparer pertinemment nos travaux avec des méthodes existantes. Néammoins, nous essayerons de donner des éléments de comparaison avec certaines d'entre elles.

1.3.1 Données spatiales et simulations PLEIADES

Grâce aux récents progrès réalisés dans le domaine spatial, la modélisation 3D urbaine peut, dorénavant, être abordée à partir d'images satellitaires optiques à résolution submétrique. Les intérêts de l'utilisation de ce type de données sont multiples. Tout d'abord, le coût d'acquisition au km^2 de données satellitaires est sensiblement inférieur à celui de données aériennes. La large fauchée des capteurs spatiaux permet de traiter des zones de grande superficie. Acquérir de telles zones en aérien est long et coûteux. De plus, grâce à la couverture très étendue des satellites, il est possible d'obtenir rapidement des images sur n'importe quelle partie du globe (du moins, pour les constellations de satellites majeures).
Les données satellitaires présentent néanmoins plusieurs inconvénients par rapport aux images aériennes. Le principal concerne la qualité des images. La résolution des satellites nouvelle génération, bien que sub-métrique, est encore loin d'égaler celle des images aériennes. Dans le cas des futurs satellites PLEIADES par exemple, la résolution sera de $0,7$ mètre (i.e. une densité de 2 pixels/m^2), ce qui contraste fortement avec la résolution des images aériennes comprise entre $0,25$ mètre (i.e. une densité de plus de 16 pixels/m^2) dans [Taillandier, 2004; Flamanc et Maillet, 2005; Durupt et Taillandier, 2006] et $0,08$ mètre (i.e. une densité de plus de 140 pixels/m^2) dans [Baillard et Zisserman, 1999; Scholze et al., 2002]. Outre la résolution, d'autres indicateurs soulignent la moins bonne qualité des images satellitaires, comme par exemple le rapport signal sur bruit (130 pour PLEIADES

vs 300 pour les images aériennes fournies par le capteur IGN). L'acquisition des images satellitaires optiques est également davantage dépendante des conditions météorologiques, principalement de la couverture nuageuse.

Les satellites PLEIADES

Le projet PLEIADES est une constellation de deux satellites offrant au nadir une résolution spatiale de $0,7$ mètre en panchromatique pour une fauchée de 20 km. Ils disposeront de quatre bandes spectrales (bleue, verte, rouge et proche infra rouge) avec une résolution de 2,8 mètres. Cette constellation, prévue opérationnellement pour 2010, permettra un accès journalier en tout point du globe. Par ailleurs, ces satellites offriront des capacités d'acquisition stéréoscopique particulièrement intéressantes. Des prises de vue tri-stéréoscopiques le long de la trace (i.e. réalisées quasi-simultanément) seront possibles, avec un rapport B/H pouvant varier entre $0,15$ et $0,8$.
Dans le cadre de cette thèse, des simulations PLEIADES fournies par le CNES sont utilisées. Ces simulations ont été générées à partir d'images aériennes, dégradées afin de respecter les caractéristiques PLEIADES (résolution, rapport signal sur bruit,...). Nous disposons de deux jeux de simulations : le jeu original à $0,7$ mètre de résolution, un autre à $0,5$ mètre[2]. Les images tri-stéréoscopiques simulées ont un rapport B/H valant $0,2$, ce qui facilite la mise en correspondance, mais confère une précision altimétrique moyenne.

Le contenu informatif des images PLEIADES

La résolution sub-métrique des satellites PLEIADES permet d'envisager de nombreuses applications en télédétection, en particulier la reconstruction en 3D de bâtiments. Il est important d'analyser le contenu informatif des images PLEIADES afin d'en connaître le potentiel en matière de modélisation urbaine 3D.
De manière générale, la résolution de ces images ne permet pas l'extraction des superstructures de toit (cheminée ou chien-assis), des décrochements de toit ou des petits pans de toit. Elle permet cependant d'extraire la forme générale des toits, i.e. les principaux plans qui composent une toiture. La figure 1.5 montre deux bâtiments présentés à travers trois jeux de données. Le premier jeu correspond à des images aériennes à $0,25m$, images à partir desquelles ont été générées les simulations PLEIADES. Les deux autres jeux représentent respectivement les simulations PLEIADES à $0,5m$ et $0,7m$. On peut noter une nette dégradation de la qualité image entre le jeu aérien et les deux jeux de simulations PLEIADES. Cette dégradation est notamment symbolisée par la différence de qualité des superstructures de toits. Les cheminées sont facilement identifiables sur le jeu aérien (une quarantaine de pixels pour une cheminée), mais pas sur les jeux satellitaires (quelques pixels) où elles apparaissent plutôt comme du bruit au sein des principaux pans de toits. La différence de qualité se manifeste également lorsque l'on cherche à extraire des primitives de ces images, telles que des segments 3D par l'algorithme de [Taillandier et Deriche, 2002]. Très bruitées, les données satellitaires ne permettent pas d'obtenir des segments 3D d'intérêt de manière robuste.

[2]Ces deux jeux de simulations ont été fournis en 2004. A ce moment, la résolution finale des satellites PLEIADES n'avait pas encore été définie et une incertitude entre $0,5$ et $0,7$ mètre demeurait.

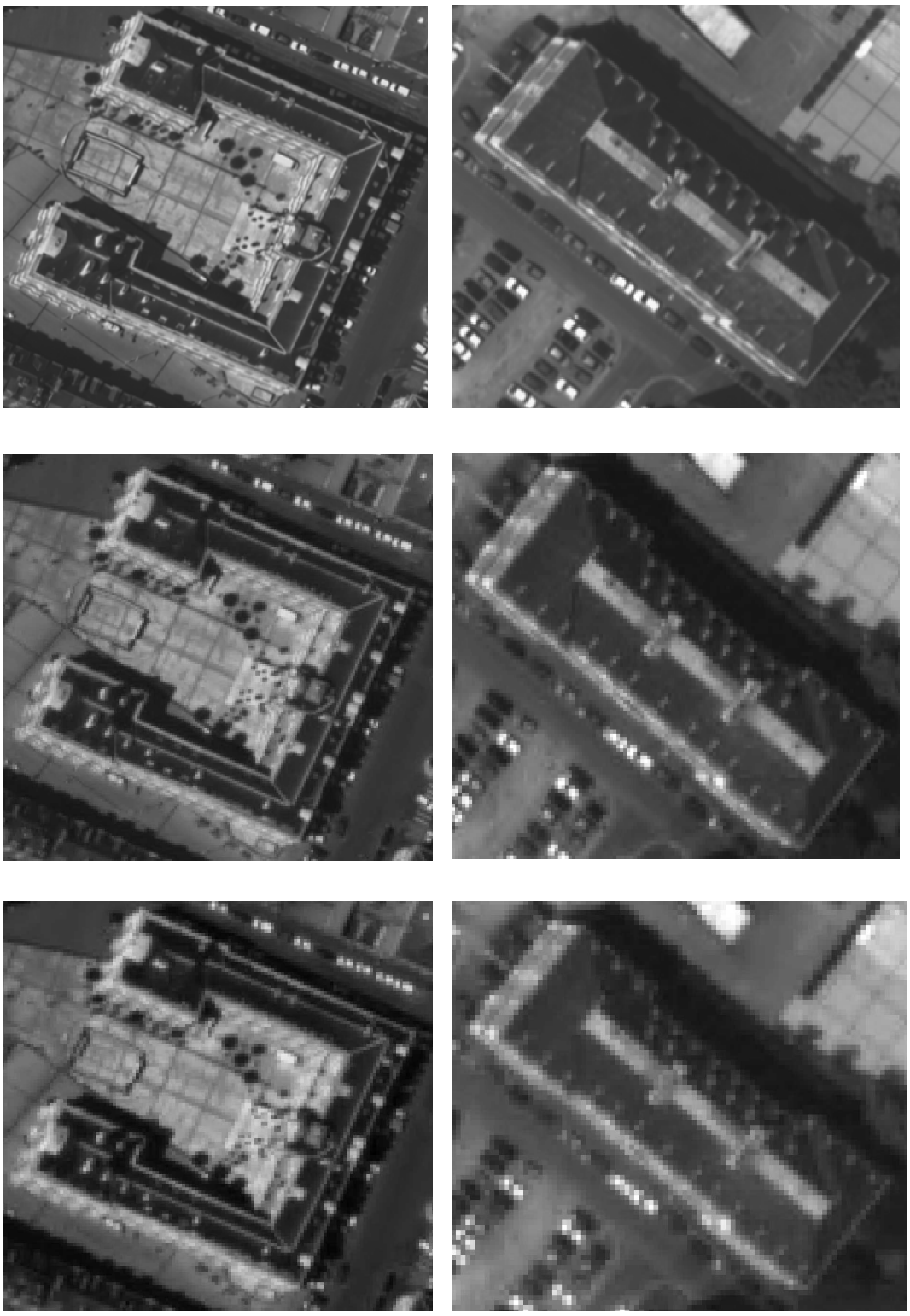

FIG. 1.5 – Différents jeux de données optiques - images aériennes à $25cm$ de résolution *(haut)*, simulations PLEIADES à $0,5m$ *(centre)* et $0,7m$ *(bas)* - ©IGN/CNES.

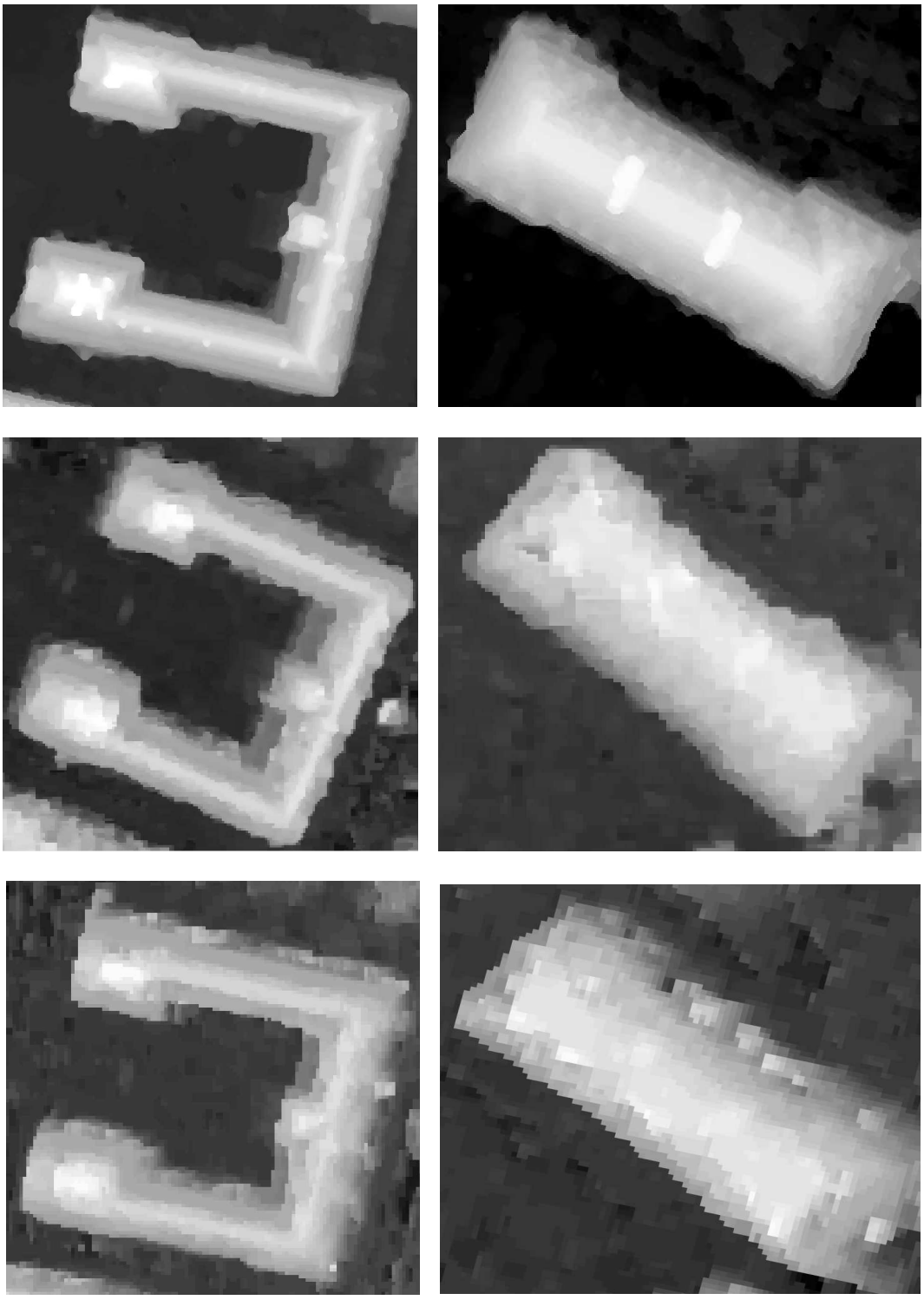

FIG. 1.6 – MNE associés aux données de la figure 1.5, générés par l'algorithme MICMAC - ©IGN.

Les MNE PLEIADES

Les MNE générés à partir des images PLEIADES sont intéressants à analyser. Ils permettent de mettre en avant les informations 3D pouvant être extraites, et ainsi d'estimer le niveau de détails de la modélisation finale. De nombreux algorithmes de génération de MNE ont été développés. L'Institut Géographique National a principalement mis au point deux processus d'appariement, [Baillard, 1997] et [Pierrot-Deseilligny et Paparoditis, 2006]. Le plus récent, et également le plus performant, est l'algorithme MICMAC (Multi Images Correspondances par Méthodes Automatiques de Corrélation). L'approche de MICMAC est fondée sur une analyse multi-résolution formulée par un problème de minimisation énergétique. L'énergie combine un terme d'attache aux données avec un terme a priori permettant d'introduire des contraintes de régularisation. Elle est optimisée par une technique de "graph-cut", fondée sur les travaux de [Roy et Cox, 1998]. La figure 1.6 présente les MNE obtenus par MICMAC en tri-stéréoscopie, correspondant aux données présentées sur la figure 1.5. La qualité des MNE obtenus confirme ce qui a été dit dans le paragraphe précédent : les données satellitaires permettent, au mieux, d'extraire les plans de toits principaux des bâtiments. Dans le chapitre 3, nous reviendrons de manière plus détaillée sur le potentiel des MNE "satellitaires" concernant la restitution des formes de toits.

Un MNE comme unique donnée d'entrée ?

Nous venons de voir que les MNE de type PLEIADES transposaient bien les informations contenues dans les images, à savoir les plans principaux des toits. La question de l'utilisation d'un MNE comme unique donnée d'entrée se pose. L'étude s'apparenterait, dans ce cas, à un problème de régularisation d'une nappe 3D.
Comme nous l'avons vu précédemment, les simulations PLEIADES ne permettent pas d'extraire, de manière robuste, des primitives 3D tels que des plans ou des segments 3D. Nous pourrions envisager de mettre en place une méthode qui, via un terme d'attache aux données, calculerait la vraisemblance d'une solution directement à partir des images par un procédé d'appariement. Cependant, une telle méthode serait extrêmement complexe puisqu'il faudrait également qu'elle gère le manque d'information des données (et donc un apport de connaissances a priori). Les travaux menés par [Garcin et al., 2001] témoignent de la difficulté à concilier ces deux aspects. Avec une qualité moyenne d'image, l'utilisation d'un MNE comme unique donnée d'entrée est donc une option tout à fait intéressante. D'ailleurs, de nombreuses méthodes comme [Vinson et Cohen, 2002; Ortner, 2004; Durupt et Taillandier, 2006] ont choisi cette solution.

1.3.2 Vers de l'urbain dense en automatique

L'intérêt du tout-automatique

Notre étude se situe dans un contexte automatique, c'est-à-dire sans l'intervention d'un opérateur, et sans données cadastrales. Dans le cadre satellitaire, ces processus automatiques sont très intéressants, en particulier pour les pays en voie de développement. Tout d'abord, les bases de données cadastrales de ces pays sont souvent incomplètes, parfois obsolètes, voire inexistantes. Par ailleurs, l'acquisition de données aériennes ou terrestres y

est généralement difficile et onéreuse à cause du manque de moyens locaux et du besoin de sous-traitance à des pays plus développés.

De l'urbain dense

Parmi les différents types de zones urbaines, celles qui suscitent le plus d'intérêt de la part des urbanistes, industriels et militaires, sont les zones urbaines denses. L'urbain dense correspond également aux zones les plus difficiles à modéliser. Sans données cadastrales, les processus fondés sur une focalisation des bâtiments sont mal adaptés. Il est donc nécessaire de proposer des processus d'extraction des emprises dans ce contexte.
Il est intéressant que les processus soient applicables sur d'autres types de zones urbaines (i.e. péri-urbaines et d'activité). Cependant, développer un algorithme qui fonctionne pour l'urbain dense ne garantit pas forcement de bons résultats sur d'autres types de zones. En effet, la densité importante des bâtiments et l'existence d'interactions fortes entre les structures voisines dans le cadre urbain dense, engendrent des processus avec de fortes connaissances a priori, difficilement transposables à d'autres contextes urbains.

1.4 La stratégie

Dans cette partie, nous définissons et justifions la stratégie choisie, et évoquons brièvement les différentes étapes permettant la mise en place de la méthode proposée.

1.4.1 L'approche structurelle

Le contexte imposé par l'étude oriente fondamentalement le choix de l'approche. Celle qui apparaît la plus légitime est l'approche structurelle. Son choix se justifie par les raisons suivantes :

- Nous avons vu que la qualité des données permettait, au mieux, de reconstruire les pans de toits principaux, c'est-à-dire la forme générale des toitures. Ce premier point signifie que les modélisations polyédriques de bâtiments ne sont pas adaptées à notre contexte, celles-ci exigeant des données d'une meilleure qualité. D'autre part, des modélisations prismatiques, voire paramétriques, seraient relativement grossières, et ne permettraient pas d'exploiter tout le potentiel des images. Une modélisation structurelle apparaît donc comme un compromis intéressant.
- Le premier point est renforcé par l'impossibilité d'utiliser les méthodes classiques d'extraction de primitives 3D, due principalement à un important bruit dans les données. Cela favorise l'utilisation d'approches "top-down" face aux approches "bottom-up". Ces premières sont spécialement adaptées aux modélisations fondées sur l'utilisation d'une bibliothèque de modèles.
- Le contexte de notre étude (traiter des zones urbaines denses en automatique à partir de données satellitaires) est particulièrement difficile et exigeant. Face à cela, il est nécessaire d'introduire de fortes connaissances a priori sur les structures urbaines. L'approche structurelle est alors idéale, puisque par définition, elle permet de modéliser des bâtiments en assemblant des modules urbains élémentaires à travers des connaissances a priori sur les modèles et leur assemblage.

Par ailleurs, la modélisation structurelle est très intéressante en elle-même à travers son aspect évolutif (on choisit les types de modèles constituant la bibliothèque, on peut notamment ajouter ou enlever des types de modèles en fonction de l'application). Elle a également été peu exploitée dans la littérature, ce qui rend intéressant son approfondissement.

1.4.2 Les outils stochastiques

Les modèles probabilistes sont des outils très utiles pour traiter les problèmes de traitement d'images. Utilisés initialement au niveau pixélique pour aborder les problèmes de segmentation et de restauration d'images [Geman et Geman, 1984; Besag, 1986], ils se sont ensuite étendus à la notion d'objet [Stoyan et al., 1987; Grenander et al., 1991; Pennec, 1996; Descombes, 2004], c'est-à-dire en associant à des variables aléatoires, non plus des pixels, mais des objets et des formes paramétriques relativement complexes. Les modèles probabilistes sont particulièrement intéressants car ils permettent :

- d'introduire des contraintes géométriques sur la nature et la forme des objets recherchés (dans notre cas, divers modèles paramétriques 3D constituant la bibliothèque de modules urbains),
- d'utiliser des informations statistiques sur l'agencement géométrique des objets dans la scène (dans notre cas, modéliser efficacement les lois d'assemblage, points fondamentaux d'une approche structurelle permettant de représenter un bâtiment comme une association de modules urbains),
- de traiter des espaces d'état de très grande dimension (traiter des scènes urbaines entières composées de plusieurs centaines de bâtiments par des objets paramétriquement complexes).

Pour ces raisons, les modèles stochastiques sont particulièrement adaptés à la mise en place d'une approche structurelle. Dans la suite, nous présentons brièvement ce type d'outils, ainsi que certaines de leurs propriétés. Pour un point de vue plus complet sur ces techniques et leurs évolutions, on pourra consulter [Pérez, 2003; Descombes, 2004].

Modèles probabilistes pour les problèmes inverses

Les problèmes inverses sont très répandus en traitement d'images. Ce type de problème consiste à retrouver une grandeur x à travers une donnée observée y. C'est le cas de notre étude : on cherche une configuration de formes paramétriques 3D à partir d'une image, plus précisément d'un MNE. Pour aborder ce type de problème, on fait appel à deux processus aléatoires :

- Y, correspond aux données ou observations (généralement une ou plusieurs images et/ou primitives extraites d'images).
- X, correspond aux configurations d'objets ou étiquettes (i.e. une variable aléatoire définie dans l'ensemble des configurations d'objets possibles).

Dans de nombreux cas, la loi a posteriori $P(X = x|Y = y)$ est utilisée pour modéliser le problème inverse. Par cette loi, on mesure la probabilité d'une réalisation x connaissant les données observées y.
Un estimateur de X particulièrement intéressant est celui qui permet de trouver la réalisation qui maximise la probabilité a posteriori : c'est l'estimateur du Maximum A Posteriori

(X_{MAP}) :

$$X_{MAP} = \arg\max_x P(X = x|Y = y) \tag{1.1}$$

Formulation Bayésienne

Pour mettre en place la loi a posteriori, on utilise généralement une formulation Bayésienne. La formule de Bayes permet d'exprimer la loi a posteriori à travers la loi marginale des observations $P(Y = y)$, la loi des observations conditionnellement à X $P(Y = y|X = x)$ et la loi a priori $P(X = x)$:

$$P(X = x|Y = y) = \frac{P(Y = y|X = x)P(X = x)}{P(Y = y)} \tag{1.2}$$

Les observations étant connues, la loi marginale des observations peut être omise pour l'optimisation :

$$P(X = x|Y = y) \propto P(Y = y|X = x)P(X = x) \tag{1.3}$$

Pour résumer, la loi a posteriori est directement proportionnelle au produit de deux lois : la loi conditionnelle des observations, ou vraisemblance, qui mesure la cohérence des données avec une réalisation x, et la loi a priori, qui impose des contraintes de régularisation sur x. Dans de nombreux cas, l'hypothèse d'indépendance conditionnelle des données est réalisée afin de simplifier le calcul de la vraisemblance. En indiçant par i les différents composants de X, on a alors :

$$P(Y = y|X = x) = \prod_i P(Y_i = y_i|X_i = x_i) \tag{1.4}$$

Le cadre Markovien

Rechercher l'estimateur du MAP est souvent un problème d'optimisation très difficile, principalement dû à la complexité du calcul de la loi a priori. Le cadre Markovien permet de simplifier cela en restreignant le problème d'optimisation globale à des problèmes d'optimisation locaux [Winkler, 1995; Li, 1995]. Concrètement, on contraint chaque élément d'une réalisation x à une dépendance locale, c'est-à-dire à considérer qu'il est en interaction uniquement avec les éléments qui lui sont proches.
Cette notion de proximité se caractérise à travers la mise en place d'un système de voisinage, noté V. La propriété de Markovianité de X se formule alors par l'expression suivante :

$$P(X_i = x_i|X_j = x_j, i \neq j) = P(X_i = x_i|X_j = x_j, j \in V(i)) \tag{1.5}$$

où i et j indicent les éléments du processus X.
Un des principaux intérêts du cadre Markovien est la possibilité d'exprimer la loi a priori à travers une énergie U, appelée énergie de Gibbs. La loi a priori peut ainsi s'écrire :

$$P(X = x) = \frac{1}{Z} \exp -U(x) \tag{1.6}$$

où Z est une constante de normalisation définie par $Z = \sum_x \exp -U(x)$, et U, une énergie pouvant se décomposer en une somme de potentiels, c'est-à-dire d'énergies d'interaction entre éléments voisins. D'une manière générale, la recherche de l'estimateur du MAP peut donc s'apparenter à un problème de minimisation énergétique.

Recherche du MAP et techniques d'optimisation

La recherche de l'estimateur du MAP cache bien souvent un problème d'optimisation difficile, non convexe, dans un espace d'état très vaste. Plusieurs techniques d'optimisation sont envisageables. Celles-ci peuvent être classées en deux catégories : les algorithmes déterministes et les algorithmes probabilistes.

Algorithmes déterministes Ce sont, pour la plupart, des algorithmes sous-optimaux, c'est-à-dire ne permettant pas de trouver le minimum global d'une fonction, mais seulement des minima locaux. Par conséquent, ils dépendent fortement des conditions initiales. Nous abordons brièvement dans ce paragraphe quelques algorithmes utilisés en traitement d'images. Les techniques de descente de gradient permettent rapidement d'atteindre un minimum local, la dérivée de la fonction doit cependant être facilement calculable ou du moins estimable en tout point. Les techniques fondées sur la théorie des graphes, particulièrement les méthodes de coupure de graphe ("graph-cut"), sont particulièrement prisées [Greig et al., 1989; Boykov et al., 1999]. L'ICM ("Iterated Conditional Modes"), présenté dans [Besag, 1986], est une méthode d'optimisation itérative fondée sur un parcours déterministe des éléments. Son efficacité est cependant assez limitée puisqu'elle ne permet pas l'optimisation de problèmes non convexes et est dépendante de la configuration initiale. Cette méthode est souvent utilisée pour traiter les problèmes de régularisation d'images. Le GNC ("Graduated Non Convexity"), proposé par [Blake et Zisserman, 1987], permet l'optimisation d'une classe particulière de fonctions non convexes, par une approximation itérative en fonctions convexes. Cette classe est malheureusement assez restreinte. On peut également citer des algorithmes pour les modèles causaux, tels que [Viterbi, 1967; Forney, 1973] sur des chaînes ou [Laferté, 1996] sur des quadrabres, qui permettent d'optimiser globalement une séquence d'états.

Algorithmes probabilistes Ce type d'algorithmes se place dans une autre optique : trouver le minimum global d'une fonction non convexe, sans tenir compte des contraintes de temps et sans dépendance aux conditions initiales. La plupart sont fondés sur une relaxation stochastique de type recuit simulé [Metropolis et al., 1953]. Une famille importante d'algorithmes est celle utilisant les échantillonneurs de type Monte Carlo par Chaînes de Markov [Hastings, 1970; Green, 1995; Robert, 1996]. Cela consiste à construire une chaîne de Markov à temps discret sur l'espace d'état, qui converge vers la solution désirée. Les transitions correspondent à des perturbations simples de la configuration courante, ce qui permet de les simuler facilement, surtout lorsque la fonction (ou énergie) a été mise en place dans un cadre Markovien. En pratique, on atteint rarement le minimum global, mais une solution proche en utilisant des relaxations "rapides". Une autre famille intéressante est celle des processus de diffusion, fondés notamment sur les équations de Langevin [Geman et Huang, 1986; Grenander et Miller, 1994]. Ils s'apparentent à des techniques de descente de gradient dans lesquelles une partie aléatoire, de type mouvement Brownien, a été ajoutée afin de permettre à la trajectoire de sortir des puits énergétiques.

Enfin, d'autres algorithmes, comme ceux fondés sur le filtrage particulaire [Doucet et al., 2000; Pérez, 2003] ou les algorithmes génétiques [Mitchell, 1996], consistent à suivre les évolutions de particules ou individus correspondant à des états. Un des avantages de ces

algorithmes réside dans le fait qu'ils fournissent non pas une solution, mais un ensemble d'états intéressants.

1.4.3 Démarche générale

Pour résumer, la méthode que nous souhaitons mettre en place, utilise comme unique donnée d'entrée un MNE. Ce dernier a été préalablement généré à partir d'images optiques satellitaires, en l'occurrence des simulations PLEIADES. L'objectif est de reconstruire automatiquement des zones urbaines denses. Pour ce faire, nous avons opté pour une modélisation structurelle formulée dans un cadre stochastique. La démarche de ce travail se décompose en trois parties résumées ci-dessous.

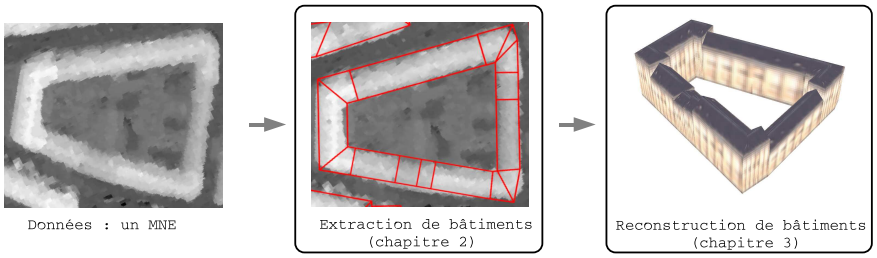

FIG. 1.7 – Illustration de la démarche générale

Extraction des bâtiments

La première étape, abordée dans le chapitre 2, consiste à extraire les bâtiments à partir d'un MNE. Premièrement, nous utilisons des travaux antérieurs, développés dans le projet Ariana, fondés sur les processus ponctuels marqués afin d'extraire des caricatures des emprises de bâtiments à travers un agencement de rectangles. Ensuite, nous régularisons cet agencement afin d'obtenir des supports 2D adaptés à une approche structurelle, en l'occurrence un agencement de quadrilatères quelconques qui sont connectés entre voisins et qui correspondent à des parties spécifiques d'un bâtiment.

Reconstruction des bâtiments

L'étape de reconstruction des bâtiments est traitée dans le chapitre 3. Après avoir défini une bibliothèque de modèles 3D, nous reconstruisons les bâtiments en recherchant la configuration optimale de modèles 3D se fixant sur les supports précédemment extraits. Cette configuration correspond à la réalisation qui maximise une densité mesurant la cohérence entre la réalisation et le MNE, mais également prenant en compte des connaissances a priori telles que les lois d'assemblage des modules.

Expérimentations

Dans le chapitre 4, nous discutons de la pertinence de cette approche en analysant les résultats obtenus. Nous abordons également l'influence de la technique d'optimisation et des supports 2D sur les résultats et évoquons les limites de notre méthode.

Chapitre 2

Extraction des bâtiments

Dans ce chapitre, nous présentons le processus mis en place pour extraire les emprises de bâtiments à partir d'un MNE. Dans le cadre de notre approche structurelle, l'objectif de cette première étape est de fournir les supports 2D (ou emprises) des modules urbains composant un bâtiment. Cela signifie que le résultat attendu devra satisfaire certaines contraintes relatives à :

- la nature des supports 2D - Ce doit être des objets géométriques facilement paramétrables, et de plus, chacun doit correspondre à une partie spécifique d'un bâtiment afin de modéliser au mieux les jonctions de toitures ou les discontinuités de hauteurs de toits par exemple.

- la connexion des supports entre eux - Les supports voisins doivent pouvoir se connecter parfaitement entre eux, c'est-à-dire sans recouvrement surfacique. C'est un point important qui permet d'éviter la présence d'artefacts, et également de faciliter la mise en place des lois d'assemblage.

Ce problème est donc plus complexe qu'une simple extraction surfacique de bâtiments. Le processus proposé se décompose en deux étapes. Premièrement, nous extrayons des caricatures des emprises de bâtiments à travers un agencement de rectangles. Pour cela, nous utilisons des travaux antérieurs fondés sur les processus ponctuels marqués. Une fois la forme générale des emprises obtenue, un processus de régularisation est appliqué afin d'avoir des emprises plus précises et mieux adaptées à une approche structurelle, c'est-à-dire des **supports structuraux**. Ainsi, cette seconde étape permet de transformer un agencement de rectangles en objets plus complexes (en l'occurrence des quadrilatères et triangles quelconques) qui sont connectés entre voisins, puis de les diviser en supports possédant des hauteurs de toit différentes.

2.1 Caricatures de bâtiments par agencement de rectangles : rappel et analyse de travaux antérieurs

La première étape consiste à extraire une description générale des emprises de bâtiments. La méthode doit être automatique et sans focalisation, avec pour unique donnée d'entrée un MNE satellitaire, c'est-à-dire d'une qualité moyenne. Bien que nous ne cherchons qu'une forme générale d'emprises, ce problème est difficile comme nous avons pu le souligner dans la partie 1.2.2.

Pour ce faire, nous utilisons les travaux d'Ortner [Ortner, 2004; Ortner et al., 2007]. Ils consistent à extraire les emprises de bâtiments sous la forme d'un agencement de rectangles. Ces objets sont probablement les formes géométriques les plus simples permettant une description générale des emprises de bâtiments. Le cadre général de ces travaux est fondé sur les processus ponctuels marqués : il s'agit de variables aléatoires dont les réalisations sont des configurations d'objets géométriques. Face à la complexité et la diversité des formes de bâtiments, une telle approche est particulièrement intéressante car elle fournit une modélisation des emprises à travers des objets géométriques simples. De plus, ces travaux prennent en compte des connaissances a priori sur la disposition spatiale des objets entre eux. Le

problème est défini à travers la minimisation d'une fonction d'énergie. La configuration de rectangles qui minimise l'énergie est trouvée au moyen d'un recuit simulé. Dans la suite de cette partie, nous présentons ces travaux, sans toutefois rentrer dans les détails.

2.1.1 Processus ponctuels marqués

Considérons un processus ponctuel X défini sur le compact $K = [0, X_{max}] \times [0, Y_{max}]$ qui représente une image. X est une application mesurable d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ dans l'ensemble des configurations de points de K :

$$\forall \omega \in \Omega, x_i \in K, n \in \mathbb{N}, X(\omega) = \{x_1, ..., x_n\} \tag{2.1}$$

Un processus ponctuel marqué est un processus ponctuel où à chaque point est associé une marque, c'est-à-dire des paramètres permettant de définir un objet. Un processus ponctuel marqué est défini sur $U = K \times M$ où M est l'espace des marques (dans notre cas, l'espace des paramètres du rectangle duquel on exclut le centre) :

$$M = [-\frac{\pi}{2}, \frac{\pi}{2}] \times [L_{min}, L_{max}] \times [l_{min}, l_{max}] \tag{2.2}$$

On note $\mathcal{C}$, l'espace des configurations finies d'objets de U.

$$\mathcal{C} = \bigcup_{n \in \mathbb{N}} (K \times M)^n \tag{2.3}$$

La figure 2.1 représente une réalisation d'un processus ponctuel dans K, une réalisation d'un processus ponctuel marqué dans U, et un élément de U. Pour davantage de détails sur les processus ponctuels et leurs applications, le lecteur pourra consulter [Van Lieshout, 2000; Descombes, 2004].

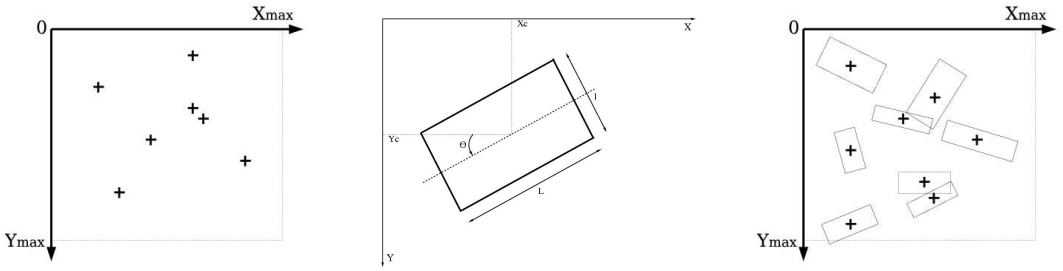

FIG. 2.1 – Illustration d'un processus ponctuel marqué - réalisation d'un processus ponctuel dans K *(gauche)*, élément de U (avec $(X_c, Y_c) \in K$, le centre du rectangle, $(\theta, L, l) \in M$, l'orientation, la longueur et la largeur du rectangle) *(centre)*, réalisation d'un processus ponctuel marqué dans U *(droite)*.

2.1.2 Formulation de l'énergie

Considérons un processus ponctuel marqué X spécifié par une densité non normalisée $h(.)$ définie dans $\mathcal{C}$. La densité $h(.)$ se décompose en différents termes qui peuvent être

définis à travers des énergies. Ainsi, pour une configuration $x \in \mathcal{C}$, la densité est donnée par :

$$h(x) = \exp -(U_{ext}(x) + U_{int}(x) + U_{excl}(x)) \tag{2.4}$$

Ces trois termes énergétiques sont présentés brièvement dans la suite. Pour plus de détails, le lecteur pourra consulter [Ortner, 2004; Ortner et al., 2007].

Terme d'attache aux données U_{ext}

Ce terme permet de mesurer la qualité d'une configuration de rectangles connaissant les données, en l'occurrence un MNE. Il peut être décomposé en une somme d'énergies $U_d(.)$ associées à chaque objet de la configuration, c'est-à-dire tel que $U_{ext}(x) = \sum_{u \in x} U_d(u)$. Ainsi, pour chaque rectangle $u \in x$, une énergie locale d'attache aux données est calculée grace à un algorithme fondé sur la comparaison entre les discontinuités présentes dans le MNE et les contours du rectangle.

La figure 2.2-(a) présente un extrait de MNE, ainsi qu'un rectangle sur lequel des coupes ont été disposées. Afin de mesurer la qualité du positionnement du rectangle par rapport au MNE, on extrait, sur chaque coupe, des points d'intérêt correspondant à des discontinuités du MNE (voir figure 2.2-(b)). Ces points d'intérêt sont détectés par un algorithme de simplification de profil détaillé dans [Ortner et al., 2007]. On vérifie ensuite la cohérence entre ces points et le contour rectangulaire de l'objet en utilisant trois critères. Le premier critère permet de calculer un taux de volume, c'est-à-dire de vérifier si la surface définie par les points d'intérêt a un recouvrement compatible avec celle du rectangle (voir la figure 2.2-(c) - les lignes grises représentent des segments utilisés pour calculer le taux de volume). Le second critère correspond à un taux de moment, et permet de mesurer l'orientation du rectangle relativement aux points d'intérêt (voir la figure 2.2-(d) - les lignes grises représentent des segments utilisés pour calculer le taux de moment). Enfin, le troisième critère vérifie si les points d'intérêt sont bien localisés relativement aux contours du rectangle (voir la figure 2.2-(e) - les points d'intérêt suffisamment proches des contours du rectangle sont entourés).

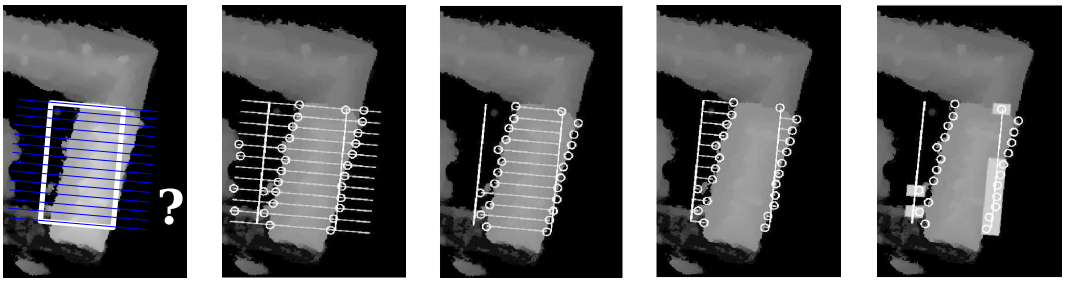

FIG. 2.2 – Energie d'attache aux données : un extrait de MNE avec un rectangle proposé *(a)*, points d'intérêt détectés *(b)*, taux de volume *(c)*, taux de moment *(d)*, taux de localisation *(e)*.

Energie interne U_{int}

L'énergie interne $U_{int}(x)$ permet de donner une structure spatiale particulière à la configuration x. C'est un terme de régularisation qui prend en compte trois types d'interactions entre objets voisins. Le premier permet de favoriser l'alignement des rectangles. Cela est particulièrement efficace pour extraire les longues structures urbaines, souvent parallèles aux réseaux routiers, sans pour autant perdre la notion de connexité entre les objets. Le deuxième type d'interactions, dit de pavage, agit sur les configurations parallèles de rectangles afin de modéliser les configurations de bâtiments juxtaposés. Enfin, les interactions de complétion favorisent le raccordement des rectangles très proches par l'ajout de rectangles complémentaires. Cela permet d'obtenir des configurations dont les rectangles sont emboîtés les uns aux autres, ce qui est particulièrement utile dans le cas des zones urbaines denses. La figure 2.3 illustre ces trois types d'interactions.

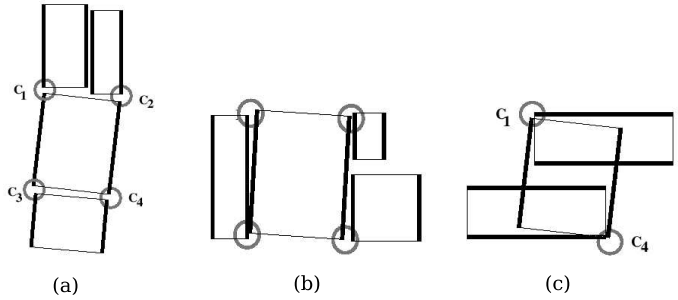

FIG. 2.3 – Energie interne : interactions d'alignement *(a)*, interactions de pavage *(b)*, interactions de complétion *(c)*

Energie d'exclusion U_{excl}

Ce terme permet d'éviter la présence d'objets redondants. Il empêche, principalement, les interactions attractives de provoquer une accumulation inutile d'objets à certains endroits. Ce terme pénalise ainsi les objets ayant un taux de recouvrement important avec d'autres objets.

Optimisation

Une fois la densité h définie, l'objectif est de trouver la configuration de rectangles $\widehat{x}$ qui maximise h :

$$\widehat{x} = \arg\max_{x} h(x) \tag{2.5}$$

Un échantillonneur de Monte Carlo par Chaînes de Markov à Sauts Réversibles (RJMCMC) [Green, 1995], couplé à un recuit simulé est particulièrement adapté pour aborder les problèmes fondés sur une modélisation par processus ponctuels marqués [Van Lieshout, 2000]. Cet échantillonneur consiste à construire une chaîne de Markov à temps discret sur l'espace d'état. Un des principaux avantages réside dans le fait que la chaîne converge asymptotiquement vers la densité désirée quelque soit la configuration initiale. Les transitions de cette

chaîne correspondent à des transformations locales de la configuration d'objets, appelées mouvements ou sauts. Ces mouvements sont simulés par des noyaux de propositions. Nous ne rentrerons pas dans les détails de cet algorithme (le lecteur pourra se référer à [Ortner, 2004] ou à la partie 3.3.1 du manuscrit), cependant il est intéressant de décrire les différents noyaux de propositions qui génèrent les transformations d'objets.

La figure 2.4 présente les noyaux utilisés. Les deux premiers types de noyaux sont les naissances et morts uniformes, qui permettent respectivement d'ajouter et de supprimer aléatoirement un rectangle dans la configuration d'objets. Ces deux transformations, qui correspondent à des sauts dans des espaces de dimension supérieure (naissance) et inférieure (mort), sont les deux noyaux de base, suffisant pour garantir que la chaîne de Markov puisse atteindre n'importe quelle configuration de l'espace d'état. Il est cependant important de définir d'autres noyaux, plus pertinents, afin d'accélérer la convergence du processus. Ainsi, on introduit des noyaux de naissance et mort dans un voisinage [Van Lieshout, 2000; Lacoste et al., 2005], permettant d'ajouter/supprimer des objets dans un voisinage de rectangles de la configuration courante. Trois autres noyaux sont également introduits, à savoir la rotation, la dilatation et la translation : il s'agit de perturbations visant à modifier les paramètres des objets.

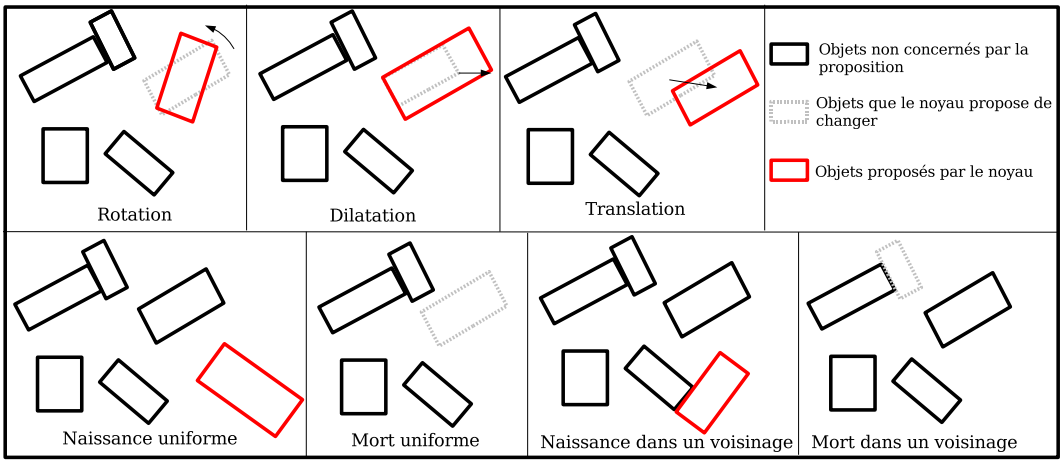

FIG. 2.4 – Les différents noyaux de propositions qui régissent l'échantillonneur RJMCMC

L'échantillonneur RJMCMC nécessite une relaxation stochastique pour assurer la convergence du processus. Il est ainsi couplé à un recuit simulé [Metropolis et al., 1953]. Ainsi, la densité $h(.)$ est remplacée par $h(.)^{\frac{1}{T_t}}$ où T_t est une séquence de température qui tend vers 0 quand t tend vers l'infini. En pratique, une décroissance géométrique, détaillée dans [Van Laarhoven et Aarts, 1987], est utilisée. Ce schéma de décroissance est particulièrement efficace pour obtenir une solution approchée de manière relativement rapide. Au début de l'algorithme (i.e. quand la température est élevée), le processus n'est pas très sélectif, ce qui permet d'explorer les différents modes de la densité. Quand la température décroît, les configurations d'objets ayant une densité élevée sont favorisées. A basse température, le processus est proche de la solution optimale. La figure 2.5 présente un exemple de l'évolution d'une configuration d'objets au cours de la décroissance de température.

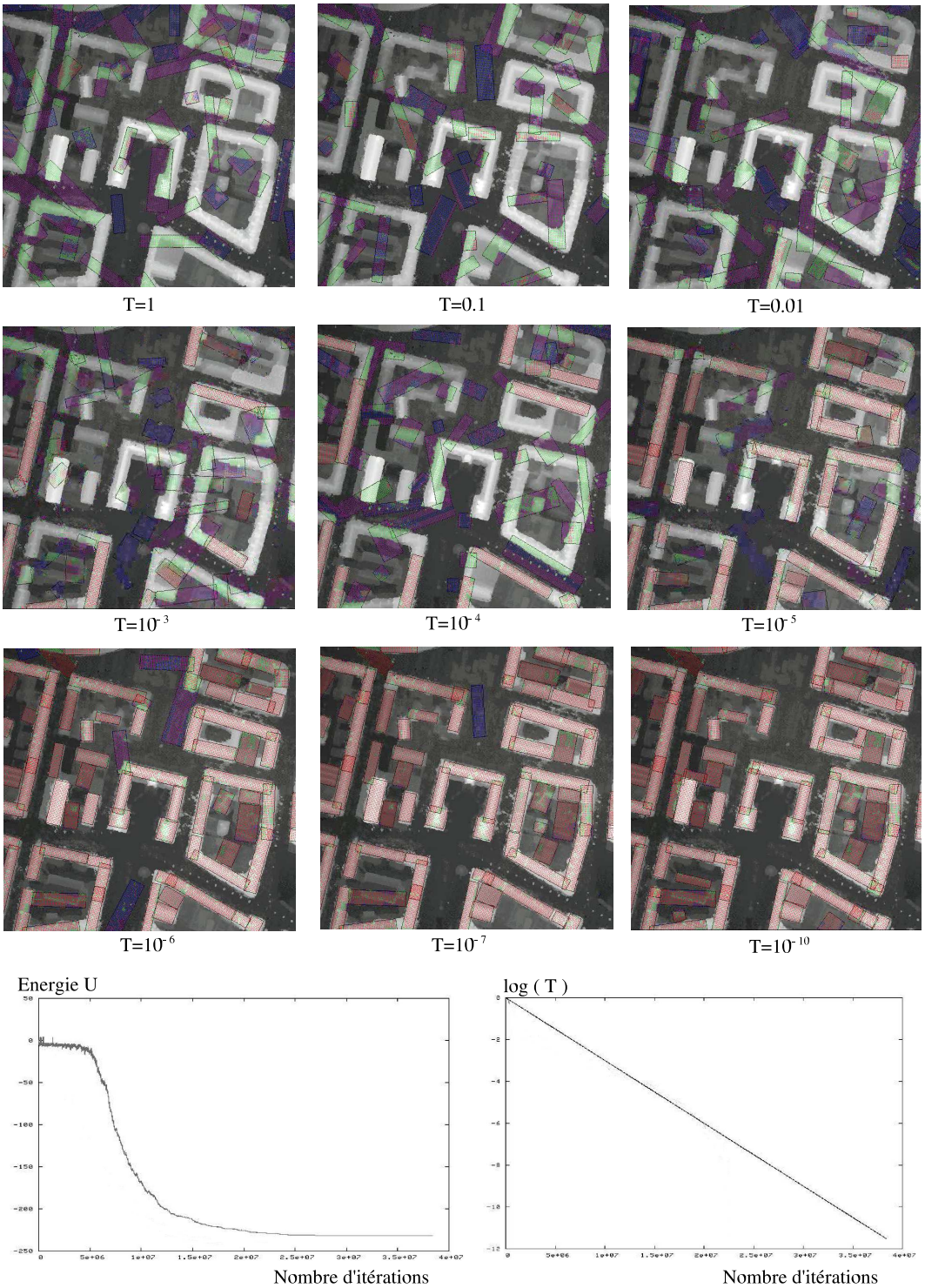

FIG. 2.5 – Processus d'optimisation - Exemple de l'évolution d'une configurations d'objets au cours de la décroissance de température *(haut)*, décroissance énergétique associée, et décroissance de température *(bas)*.

2.1.3 Commentaires

La figure 2.6 présente des résultats obtenus à partir de MNE de type PLEIADES, correspondant à deux bâtiments de la ville d'Amiens (France). De manière générale, ces résultats sont satisfaisants aussi bien du point de vue du recouvrement surfacique que de la localisation des façades. La figure 6.14, présentée en annexe F, montre les résultats obtenus sur trois zones urbaines denses de la ville d'Amiens. Nous reviendrons plus en détail sur la qualité de ces résultats dans le chapitre 4.

Cette méthode fournit la forme générale des emprises de bâtiments à travers un agencement de rectangles. Il est possible d'envisager directement l'utilisation de ces rectangles comme supports de modèles paramétriques 3D à base rectangulaire. Cela a été testé à travers deux processus de reconstruction, présentés dans l'annexe A. Cependant, ces processus fournissent des résultats peu convaincants, dus notamment à la présence de nombreux artefacts dans les modélisations 3D de bâtiments. L'utilisation de rectangles comme supports est, en effet, limitée : ces formes géométriques sont trop simples pour permettre une bonne connexion des supports entre eux, comme on peut le constater sur la figure 2.7. Outre l'aspect visuel des résultats, ces processus nécessitent des formulations énergétiques complexes, c'est-à-dire intégrant des interactions nombreuses et variées, afin d'assembler entre eux les modèles 3D à base rectangulaire. Cela a l'inconvénient d'introduire beaucoup de paramètres dans l'algorithme, et par conséquent, de le rendre peu robuste.

Pour développer une méthode efficace, il est nécessaire de régulariser a posteriori les agencements de rectangles. L'objectif est d'améliorer la connexion des rectangles voisins (voir figure 2.7). Pour ce faire, les rectangles doivent être transformés en objets géométriquement plus complexes. Ce problème est traité dans la seconde partie de ce chapitre.

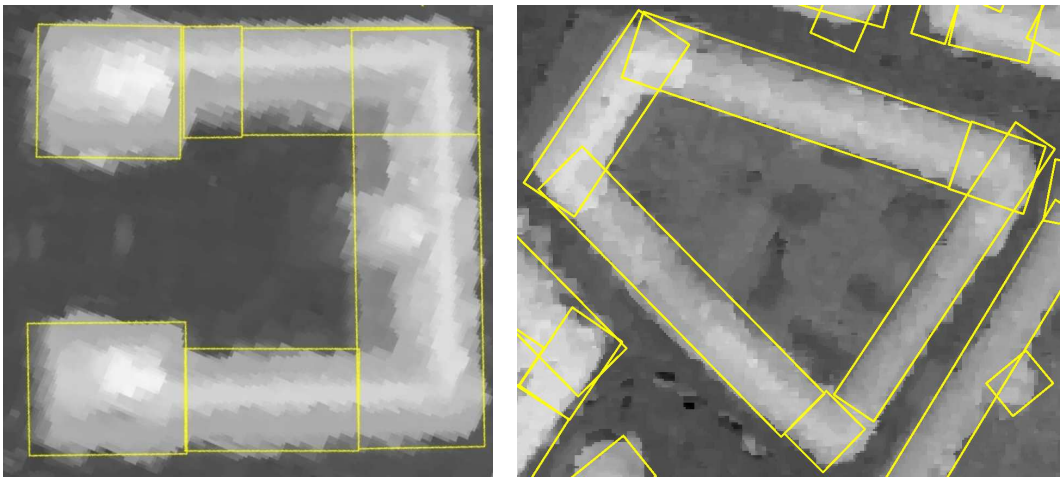

FIG. 2.6 – Résultats d'extraction de bâtiments par agencement de rectangles.

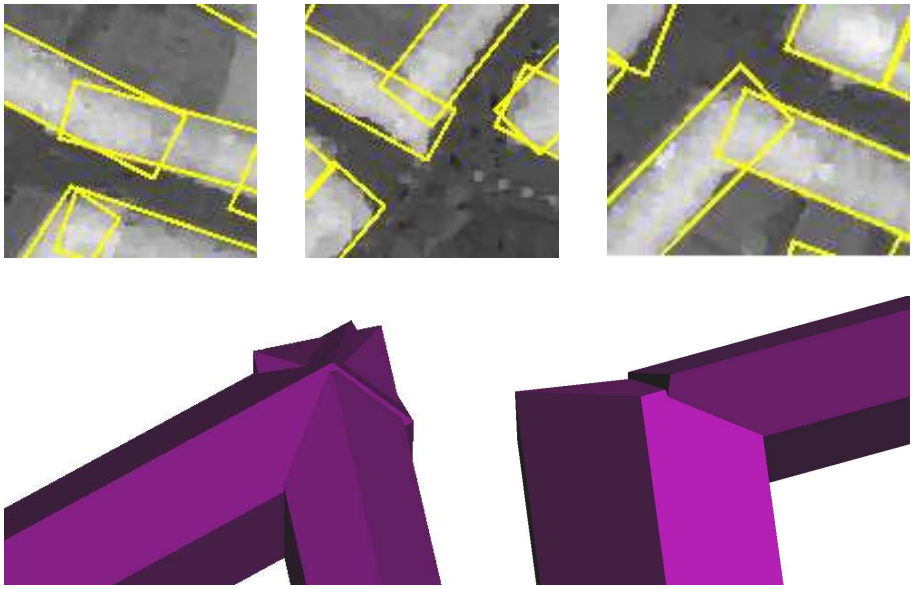

FIG. 2.7 – Illustrations de mauvaises connexions entre rectangles voisins *(haut)*, engendrant la présence d'artefacts dans les modélisations 3D *(bas)*

2.2 Régularisation en supports structuraux

La partie précédente nous a permis d'obtenir la forme générale des emprises de bâtiments à travers des agencements de rectangles. L'objectif est désormais de régulariser ces configurations de rectangles en supports structuraux, c'est-à-dire en emprises plus précises et adaptées à une approche structurelle.

Pour ce faire, nous procédons en deux étapes. D'abord, nous transformons les rectangles en objets géométriquement plus complexes, pouvant se connecter entre voisins sans recouvrement de surface et sans que la forme générale des emprises soit changée. Ensuite, nous effectuons une division des objets possédant des hauteurs de toit différentes.

2.2.1 Connexion des supports

Le but est de transformer un agencement de rectangles en un agencement d'objets géométriques dont les éléments voisins peuvent se connecter sans recouvrement de surface. En excluant les formes courbées, les polygones sont les objets naturellement adaptés à ce problème. Nous introduisons dès lors la notion de connexion entre deux polygones :

Définition 1 *soient P_1 et P_2, deux polygones, et S_1 et S_2, leur surface respective. P_1 et P_2 seront dits **connectés** si les deux conditions suivantes sont vérifiées :*

- $\overset{\circ}{S_1} \bigcap \overset{\circ}{S_2} = \emptyset$
- *P_1 et P_2 ont au moins une arête commune*

Les polygones sont toutefois des objets pouvant être très complexes. Or, il est important qu'ils puissent être décrits à travers un nombre restreint de paramètres, afin de ne pas com-

plexifier la combinatoire du processus de reconstruction. La solution consiste à limiter les formes polygonales considérées à des quadrilatères et triangles quelconques. La figure 2.8 illustre une transformation de deux rectangles en une association de quadrilatères et triangles connectés.

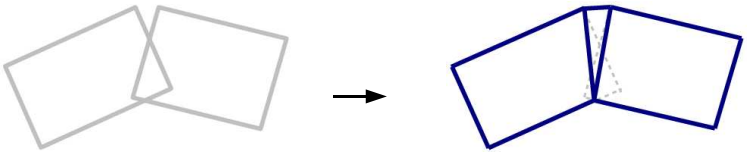

FIG. 2.8 – Illustration d'une transformation de deux rectangles voisins en une association de quadrilatères quelconques connectés

La méthode retenue consiste à transformer localement les couples de rectangles voisins en association de quadrilatères quelconques connectés (ces quadrilatères pouvant être des triangles). La relation de voisinage, permettant de définir si deux rectangles sont voisins, consiste à vérifier si l'intersection de leur "sur-rectangle" est non vide (cette relation est détaillée en annexe A). La figure 2.9 permet de schématiser le problème : les points A et B doivent être mis en relation dans le secteur extérieur, de même que les points C et D dans le secteur intérieur[1]. Ce problème n'est pas trivial notamment à cause des nombreuses dispositions possibles de paires de rectangles voisins. Ces dernières dépendent de la largeur des rectangles, de leur orientation relative, et de la distance les séparant. La figure 2.9 présente quelques exemples de dispositions possibles.

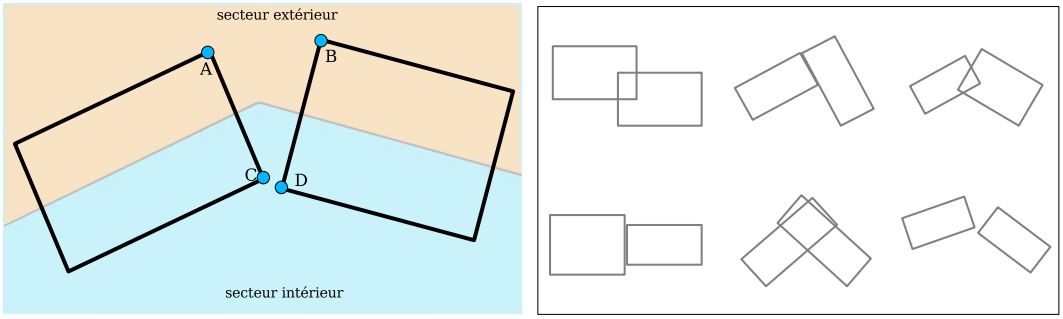

FIG. 2.9 – Deux rectangles voisins *(gauche)*, différents exemples de dispositions de deux rectangles voisins *(droite)*.

Près d'une centaine de dispositions de couples de rectangles sont possibles : il n'est pas envisageable de les traiter au cas par cas. Il faut définir un processus général de transforma-

[1]La frontière séparant le secteur intérieur du secteur extérieur est définie par les deux médianes des rectangles. Dans certaines situations (médianes parallèles disjointes), ces deux secteurs ne sont pas représentés. Les points A, B, C et D sont alors définis comme les deux paires de points de chaque rectangle ayant la distance deux à deux la plus courte.

tion des couples de rectangles voisins. Nous proposons seize configurations différentes de transformations permettant de traiter les dispositions possibles. La figure 2.10 schématise ces seize configurations. Dans le secteur extérieur, quatre possibilités de mise en relation des points A et B sont permises. Celles-ci sont fondées sur des intersections issues de droites prolongeant les arêtes des rectangles. Concernant le secteur intérieur, les quatre mêmes possibilités sont autorisées. Finalement, quatre $\times$ quatre configurations sont proposées.

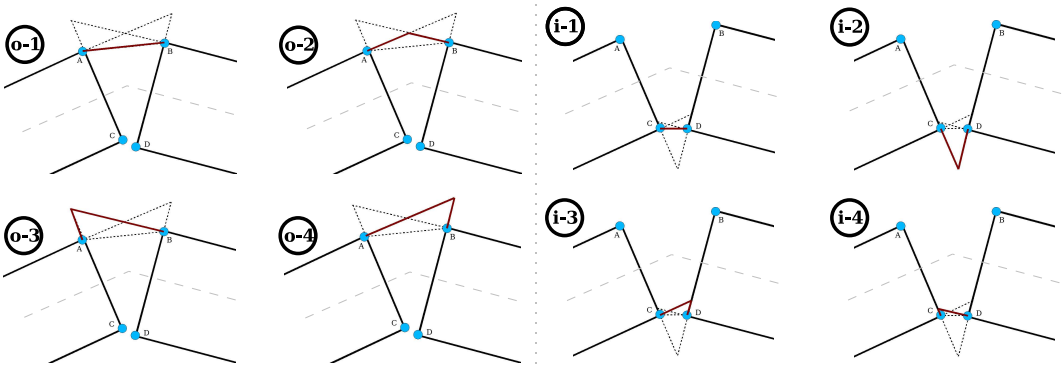

FIG. 2.10 – Les quatre possibilités de mise en relation du secteur extérieur *(gauche)*, les quatre possibilités de mise en relation du secteur intérieur *(droite)*.

Afin de choisir la configuration la mieux adaptée parmi les seize pour une disposition de rectangles donnée, nous définissons une fonction de coût C qui prend en compte la disposition relative des rectangles, mais également le MNE. Il peut arriver que la configuration choisie ne soit pas pertinente. Ainsi, la configuration choisie doit avoir un coût inférieur à un coût de référence C_{ref}, sans quoi la transformation n'est pas appliquée et les deux rectangles restent inchangés. La fonction de coût C se décompose en trois termes :

- un terme d'attache au MNE, qui évalue la qualité de la configuration par rapport au MNE.
- un terme de recouvrement surfacique, qui favorise les taux de recouvrement élevés entre les deux rectangles et les quadrilatères proposés.
- un terme de correspondance des contours, qui permet de comparer la pertinence des contours des quadrilatères relativement à ceux des deux rectangles initiaux.

Ces trois critères sont pondérés entre eux par deux poids ω_1 et ω_2. La fonction de coût s'exprime donc par :

$$C = C_{MNE} + \omega_1 C_{recouvrement} + \omega_2 C_{contour} \tag{2.6}$$

ω_1 and ω_2 ont été fixés respectivement à $0,5$ et $0,8$. Le coût de référence C_{ref}, qui permet d'autoriser ou non la transformation, est fixé à $0,2$.

Critère d'attache au MNE

Ce terme permet de mesurer la qualité de la configuration proposée par rapport au MNE. Pour cela, nous calculons le taux de pixels inclus dans les quadrilatères proposés, dont la valeur est supérieure à l'altitude d'un bâtiment d'un demi étage (en pratique, $1,5$ mètre). Plus ce taux est élevé, plus la configuration est favorisée et donc plus le coût est faible. En notant N, le nombre de ces pixels, et N_t, le nombre total de pixels contenu dans les quadrilatères, le critère d'attache au MNE s'exprime par :

$$C_{MNE} = 1 - \frac{N}{N_t} \tag{2.7}$$

Critère de recouvrement surfacique

Ce critère prend en compte le recouvrement surfacique entre les deux rectangles initiaux, et les quadrilatères proposés. Nous voulons que la surface avant et après transformation soit sensiblement inchangée. Soit S_R, la surface de l'union des deux rectangles, et S_Q, la surface de l'union des quadrilatères proposés. Le critère de recouvrement est donné par :

$$C_{recouvrement} = \frac{|\, S_Q - S_R \,|}{S_R} \tag{2.8}$$

Critère de cohérence des contours

Ce terme mesure la cohérence des contours entre les deux rectangles initiaux, et les quadrilatères proposés. Il est indispensable au bon fonctionnement du processus puisqu'il permet d'éviter les configurations complexes et peu réalistes, comme celle schématisée sur la figure 2.11. Notons L_{in} (respectivement L_{out}), la longueur totale intérieure (respectivement extérieure) des quadrilatères proposés. Notons également L_1 et L_2, les longueurs des deux rectangles. Le coût de ce critère est :

$$C_{contour} = \frac{|\, L_{out} + L_{in} - 2(L_1 + L_2) \,|}{2(L_1 + L_2)} \tag{2.9}$$

Pour un agencement de rectangles complet, nous traitons localement tous les couples de rectangles voisins par ce procédé. Le résultat est une association de quadrilatères connectés, comme le montre la figure 2.13 (*centre* - emprises bleu). La forme générale des emprises a été conservée. De plus, grâce au terme d'attache au MNE utilisé dans la fonction de coût, nous pouvons remarquer que les petits artefacts générés par le recouvrement de rectangles ont été éliminés, ce qui permet d'améliorer nettement la qualité des contours des emprises.

Le processus mis en place est relativement simple. Cependant, il constitue une étape nécessaire à la mise en place de notre approche structurelle. En théorie, le résultat de cette étape aurait pu être obtenu directement en utilisant une approche par processus ponctuels marqués, c'est-à-dire une approche similaire à celle développée dans la partie 2.1, où les objets auraient été des quadrilatères quelconques. En pratique, cela n'est pas envisageable : les objets, bien que simples, sont déjà paramétriquement trop complexes pour espérer des temps de calculs abordables.

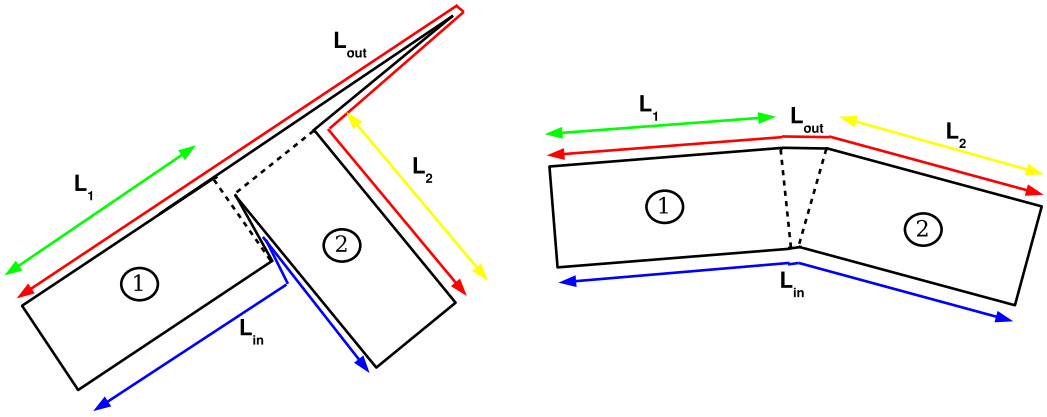

FIG. 2.11 – Critère de cohérence des contours : configuration pénalisée *(gauche)*, configuration favorisée *(droite)*.

2.2.2 Détection des discontinuités des hauteurs de toit

Dans ce paragraphe, l'objectif est de détecter, à l'intérieur des emprises, les éventuelles discontinuités des hauteurs de toits. Les emprises quadrilatérales obtenues précédemment peuvent, en effet, contenir différentes structures urbaines au sein d'un même quadrilatère. Ces structures urbaines sont différentiables, dans une grande majorité de cas, grâce à leur hauteur de toit. La figure 2.13 illustre ce problème. Il s'agit donc de localiser les éventuelles discontinuités des hauteurs de toit, et le cas échéant, de diviser le quadrilatère en plusieurs quadrilatères correspondant chacun à une structure urbaine. Nous restreignons ce problème aux emprises de tailles importantes, c'est-à-dire susceptibles d'inclure plusieurs structures urbaines. Pour cette raison, les quadrilatères dont la surface est inférieur à $100m^2$, ainsi que les triangles ne seront pas traités.

Le partitionnement des emprises est effectué de manière à détecter les discontinuités de hauteurs de toit perpendiculairement au grand axe de l'emprise (c'est-à-dire l'axe de la longueur de l'emprise). Ce procédé fournit un partitionnement certes très simple, mais en pratique, cette hypothèse est souvent vérifiée notamment pour les longues structures le long des réseaux routiers.

Pour ce faire, nous calculons, pour chaque quadrilatère, un profil d'arêtes faîtières dans le sens de la longueur. Ce profil est obtenu à partir du MNE, en estimant la hauteur de toit moyenne dans une fenêtre glissante. Ensuite, nous extrayons les forts gradients par un seuillage sur la dérivée du profil. Le seuil utilisé est un paramètre qui a été fixé à 3 : cela signifie que la variation altimétrique du profil doit être au moins trois fois plus importante que sa variation planimétrique. Il est fréquent, au niveau d'une discontinuité, de détecter plusieurs forts gradients à proximité les uns des autres : il faut, dans ce cas, réaliser une accumulation de gradients afin de localiser l'endroit précis où se situe la discontinuité de hauteur de toit. Très simplement, nous choisissons de prendre le barycentre de ces forts gradients. Ces différentes étapes sont illustrées sur la figure 2.12. La figure 2.13 présente

deux exemples de résultats (*bas* - emprises rouges). Les discontinuités y sont correctement détectées. Sur l'exemple de droite, nous détectons également des discontinuités au niveau d'un décrochement de toiture. Dans notre contexte, cela n'est pas très préjudiciable : il est préférable de sur-détecter les discontinuités plutôt que d'en omettre. Dans la suite, nous nommerons ces emprises des supports structuraux.

Les supports structuraux fournissent des informations planimétriques, proches de données cadastrales. Ils apportent également des connaissances sur la structuration des bâtiments. Ces supports structuraux ont été utilisés dans une méthode paramétrique simple de reconstruction 3D, fondée sur un processus de squelétisation des arêtes faîtières. Cette méthode, simple mais peu précise, présente un intérêt mineur. Elle est présentée en annexe B.

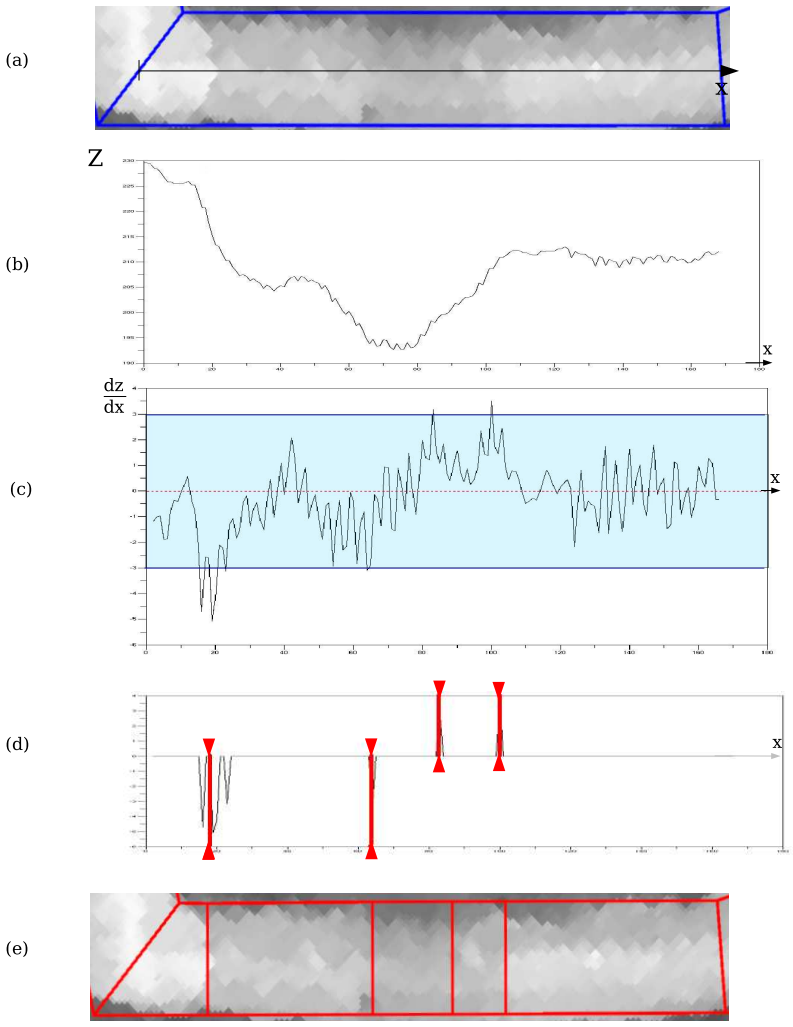

FIG. 2.12 – Détection des discontinuités des hauteurs de toit - emprise initiale *(a)*, estimation du profil du toit *(b)*, seuillage de la dérivée du profil *(c)*, accumulation des gradients *(d)*, emprises finales avec les discontinuités en rouge *(e)*.

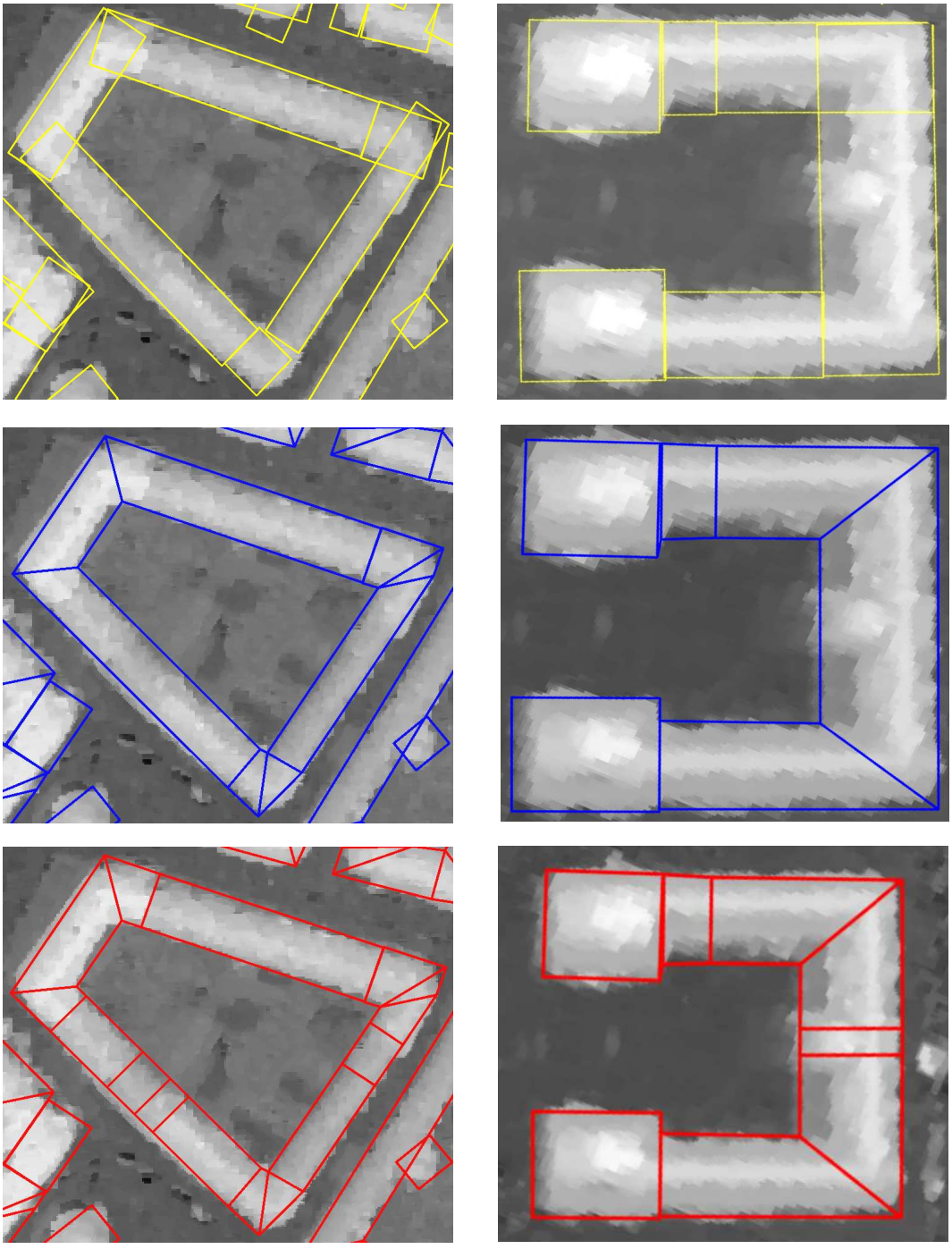

FIG. 2.13 – Deux exemples de régularisation en supports structuraux - emprises rectangulaires initiales *(haut)*, emprises quadrilatérales après connexion des rectangles *(centre)*, emprises quadrilatérales après extraction des discontinuités de hauteurs de toit *(bas)*.

Chapitre 3

Reconstruction 3D des bâtiments

Dans ce chapitre, nous présentons l'étape de reconstruction en 3D des bâtiments. Les données d'entrée sont constituées par un MNE et les supports structuraux 2D extraits dans le chapitre précédent. Tout d'abord, nous mettons en place une bibliothèque de modèles 3D représentant des modules élémentaires de bâtiments. Nous reconstruisons alors chaque bâtiment en recherchant la configuration optimale d'éléments de cette bibliothèque se fixant sur les supports structuraux. Dans un cadre stochastique, cette configuration correspond à la réalisation qui maximise une densité. Deux problèmes majeurs se posent alors : Comment définir cette densité ? Comment trouver la réalisation qui la maximise ? Les réponses à ces deux questions induisent la mise en place de paramètres. La dernière partie de ce chapitre est consacrée à l'estimation de ces paramètres.

3.1 Bibliothèque de modèles 3D

La première étape consiste à définir une bibliothèque de modèles 3D. La définition de la bibliothèque est un point fondamental puisque celle-ci permet de fixer le niveau de généricité de la modélisation 3D. Elle doit être représentative du paysage urbain et être également en adéquation avec les données que nous utilisons. Par ailleurs, les modèles 3D doivent être des objets simples définis à travers un nombre restreint de paramètres afin de ne pas exploser la combinatoire de notre méthode.

3.1.1 Typologie des toits et adéquation avec les données

Afin de proposer une bibliothèque pertinente, nous procédons en deux étapes. D'abord, nous étudions la typologie des structures urbaines et, plus spécifiquement, la morphologie des toitures. L'objectif est de déterminer les formes de toits dont l'occurrence n'est pas négligeable dans un paysage urbain donné. Ensuite, nous vérifions l'adéquation entre ces formes et les données utilisées : la qualité des MNE doit être suffisante pour permettre la reconnaissance de ces différentes formes.

Morphologie des toitures dans le paysage urbain européen

Il existe une très grande variété de toitures à travers le monde. Il n'est pas envisageable de les rassembler dans une même bibliothèque : cela serait à la fois très fastidieux à mettre en place et très pénalisant au niveau des temps de calcul. Nous devons donc nous focaliser sur un type de zones et nous limiter aux formes de toits les plus représentatives. Nous décidons de construire notre bibliothèque par rapport aux centres-villes européens. Les raisons de ce choix résident dans l'aspect général des bâtiments européens et les données tests utilisées (villes d'Amiens et de Toulouse).

Pour déterminer les formes de toits les plus communes, nous nous sommes appuyés sur les travaux de [Ching, 1996; Gaminet, 1999; Allain, 2004]. Nous avons ainsi mis en avant trois grandes familles de toits, illustrées sur la figure 3.1.

- **Les mono-plans** - Ils sont souvent utilisés pour construire de grandes structures, à l'image des centres-villes américains. Ils sont présents dans les zones urbaines denses mais également dans les zones d'activité. Nous pouvons distinguer trois déclinaisons

possibles : les très classiques toits plats, les toits terrasse (qui sont des dérivés des toits plats auxquels on ajoute un module souvent de forme parallélépipédique permettant soit l'accès à la terrasse, soit l'implantation de machineries d'ascenseurs), et enfin les toits en appentis (qui sont des mono-plans inclinés, correspondant souvent à des modules annexes de la structure principale d'un bâtiment).

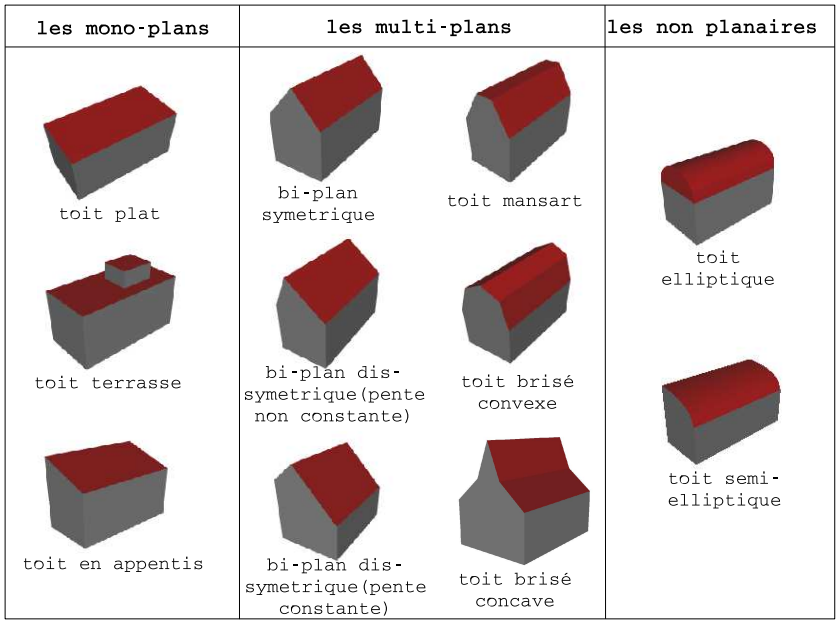

FIG. 3.1 – Formes de toit les plus communes

- **Les multi-plans** - Ils constituent la famille de toits la plus répandue dans le paysage urbain européen. Nous avons distingué six déclinaisons. La première est le bi-plan symétrique, très courant. Sur le même principe, mais avec l'arête faîtière située de manière dissymétrique par rapport au support, nous avons deux autres déclinaisons : le toit dissymétrique à pente constante et le toit dissymétrique à pente variable. La quatrième, du nom de son créateur, est le toit Mansart, toit tri-plan intermédiaire entre le toit plat et le bi-plan symétrique. Il est relativement répandu sur les structures anciennes (historiquement, ce toit au volume habitable important a été imaginé au 17^{e} siècle pour contrer la taxe d'habitation fondée sur le nombre d'étages). Enfin, les deux dernières déclinaisons sont les toits brisés convexes et concaves. Ce sont des quadri-plans composés de deux plans extérieurs de même pente et de deux plans intérieurs de même pente. Le toit convexe, qui possède un volume habitable important, est fréquent en centre-ville ; son homologue concave est lui plutôt répandu au sein des zones péri-urbaines.

- **Les non planaires** - Ce sont des toitures courbées. Elles sont principalement utilisées pour la construction d'édifices remarquables dans les centres-villes, et aussi de grands bâtiments type entrepôts dans les zones d'activité. Nous distinguons deux

déclinaisons : les toits elliptiques (i.e. courbés suivant une demi-ellipse) et les toits semi-elliptiques (i.e. courbés suivant un quart d'ellipse).

Adéquation entre la morphologie de toits et la qualité des MNE PLEIADES

La qualité des MNE PLEIADES doit être suffisante pour permettre la reconnaissance des formes de toits précédemment citées. Ce point n'est pas trivial car ces MNE sont des données particulièrement bruitées. Pour vérifier cela, nous comparons les profils de MNE avec les profils réels associés.

La figure 3.2 présente quatre couples de profil qui sont représentatifs du potentiel des MNE PLEIADES. Ces profils ont été extraits à partir du centre-ville d'Amiens. Le premier couple illustre le profil d'un toit bi-plan symétrique. Nous constatons que le positionnement de l'arête faîtière peut être restitué sans grande ambiguïté. De plus, les discontinuités des façades sont relativement franches. La reconnaissance des bi-plans, symétriques ou dissymétriques, ne doit donc pas poser de difficultés majeures. Le deuxième point positif réside dans la bonne restitution des plans horizontaux, comme le montrent les deuxième et troisième couples de profil de la figure 3.2. Par contre, les arêtes des brisures de plans, spécifiques aux toits brisés, ne sont pas distinguables. De plus, les profils des MNE ont une tendance à être concaves comme on peut le constater sur le premier couple : la présence des toits brisés concaves dans la bibliothèque porterait à confusion. Il n'est donc pas envisageable d'inclure les toitures brisées dans la bibliothèque. Le dernier couple de profil illustre la restitution d'un toit courbé. Celle-ci est bonne comme nous pouvons le constater, à condition que la structure soit d'une taille importante (i.e. avec une largeur de support d'au moins une vingtaine de mètres).

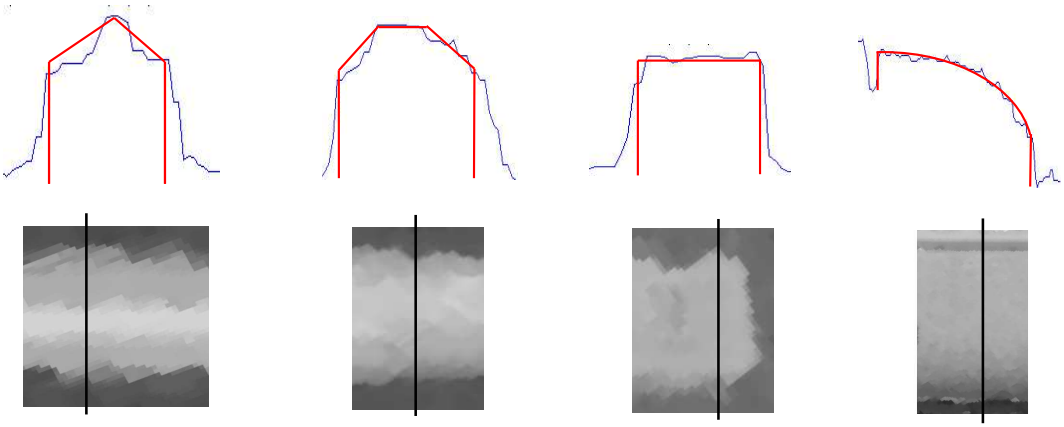

FIG. 3.2 – Potentiel des MNE PLEIADES en matière de restitution des formes de toits - profil réel en rouge, profil du MNE en bleu.

Pour résumer, à l'exception des toitures brisées, les MNE PLEIADES montrent des caractéristiques intéressantes pour permettre une reconnaissance efficace des formes de toits regroupées sur la figure 3.1. Une simple étude des profils des MNE ne garantit pas, cependant, une distinction correcte des différentes formes de toit retenues. Celle-ci devra être

confirmer dans le chapitre 4 lors de l'évaluation des résultats.

3.1.2 La bibliothèque proposée

La bibliothèque que nous utilisons, notée $\mathcal{M}$, est composée de différents modèles auxquels on associe une forme de toit $\mathcal{F}$ (détaillée dans la partie précédente) et une variante $\mathcal{V}$ (spécifique aux terminaisons ou jonctions de modules). Chaque élément de cette bibliothèque est ainsi caractérisé par un modèle spécifié par le couple $(\mathcal{F}, \mathcal{V})$ auquel est associé un jeu de paramètres $\theta = (F, V)$.

Les différentes formes de toit $\mathcal{F}$ et leurs paramètres F

On dispose de neuf formes de toit, proposées dans la partie précédente (plat $\mathcal{F}_{11}$, terrasse $\mathcal{F}_{12}$, en appentis $\mathcal{F}_{13}$, bi-plan symétrique $\mathcal{F}_{21}$, bi-plan dissymétrique à pente constante $\mathcal{F}_{22}$, bi-plan dissymétrique à pente variable $\mathcal{F}_{23}$, mansart $\mathcal{F}_{24}$, elliptique $\mathcal{F}_{31}$ et semi-elliptique $\mathcal{F}_{32}$). La figure 3.3 présente ces différentes formes de toit. Chacune des neuf formes possède son propre jeu de paramètres, que nous appelons les paramètres de formes F. Ils sont communs à toutes les variantes du modèle. Ils correspondent à des paramètres altimétriques (hauteur de gouttière et hauteur de faîtière) et également planimétriques (dissymétrie de l'arête faîtière, largeur du plateau d'un toit mansart,...). Le nombre de paramètres varie entre un (pour le toit plat) et six (pour le toit terrasse). Ils sont à valeurs continues dans des compacts de $\mathbb{R}$. Pour chaque type de toit, ces paramètres sont détaillés dans la table 3.2.

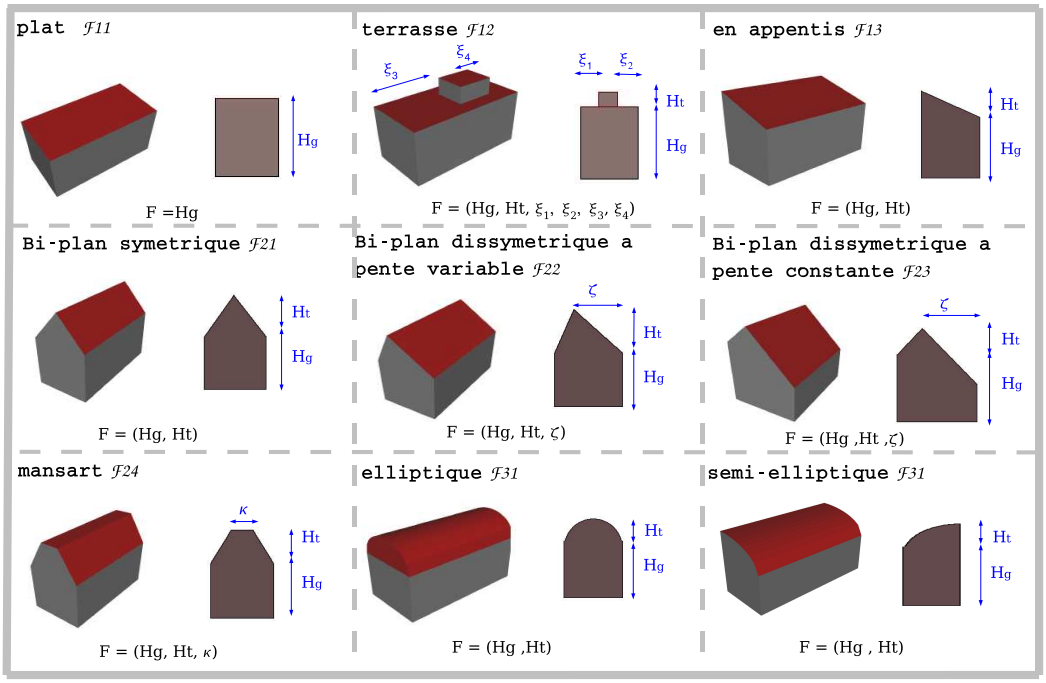

FIG. 3.3 – Les différentes formes de toits $\mathcal{F}$ de la bibliothèque

Les différentes variantes $\mathcal{V}$ et leurs paramètres V

Chaque type de toit possède un certain nombre de variantes possibles $\mathcal{V}$ (au plus six) permettant de décrire différentes situations de terminaisons et jonctions de toits. Il en existe six différentes (la forme de base (notée $\mathcal{V}_{-}$), les variantes avec attaque en V simple ($\mathcal{V}_V$) et double ($\mathcal{V}_{2V}$), et les variantes de jonction en L ($\mathcal{V}_L$), en T ($\mathcal{V}_T$) et en croix ($\mathcal{V}_+$)). Elles sont illustrées sur la figure 3.4 dans le cas d'un toit bi-plan symétrique, seul modèle de la bibliothèque à posséder ces six variantes. Ces variantes sont plus ou moins spécifiques à différentes situations de voisinage :

- les modules isolés, qui correspondent à des bâtiments très simples dont le support est un quadrilatère sans voisin,
- les modules de terminaison, qui ont exactement un voisin,
- les modules de jonction, qui ont au moins deux voisins.

La figure 3.4 montre les variantes qui sont préférentielles selon la situation de voisinage du module et qui, par la suite, devront être favorisées dans notre modèle. Chaque variante possède un jeu de paramètres spécifiques noté V. On trouve des paramètres d'orientation du modèle, à valeur discrète. Le support ayant quatre côtés, ϕ_2 permet d'orienter le modèle suivant quatre directions. Si le modèle a deux symétries axiales, il n'aura plus que deux orientations possibles : dans ce cas, on utilise le paramètre ϕ_1 à valeur dans $\{1,2\}$. On trouve également un paramètre η définissant la profondeur d'une attaque en V. Ces paramètres sont illustrés sur la figure 3.5.

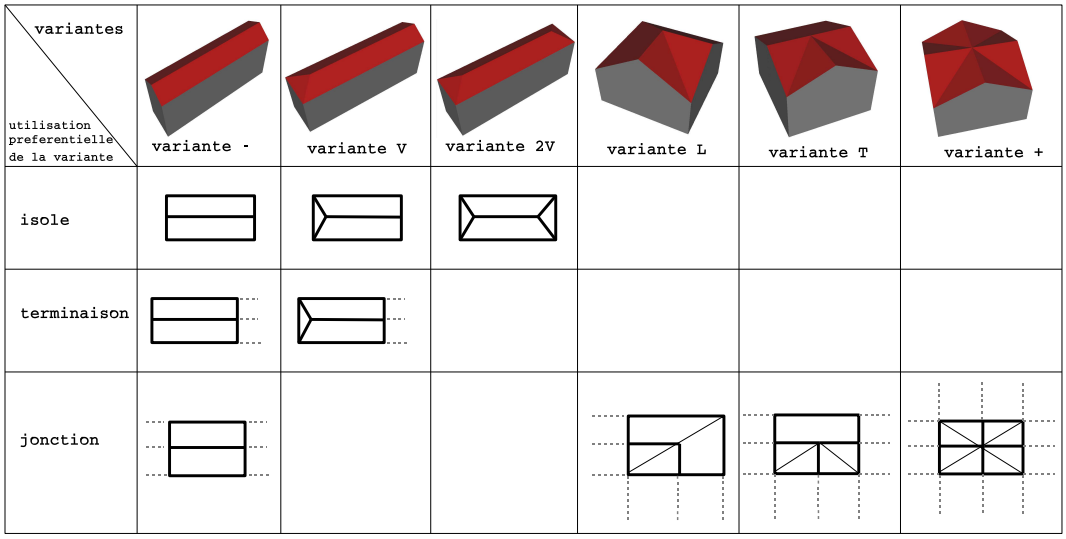

FIG. 3.4 – Illustration des variantes dans le cas du modèle bi-plan symétrique

Les paramètres de chaque modèle sont résumés dans la table 3.1. Il faut noter que, dans le cas de supports 2D triangulaires, certains modèles peuvent être dégénérés en fonction de la valeur du paramètre d'orientation. Ces modèles sont logiquement interdits. Du point de vu de la généricité, les modèles qui composent la bibliothèque peuvent paraitrent assez

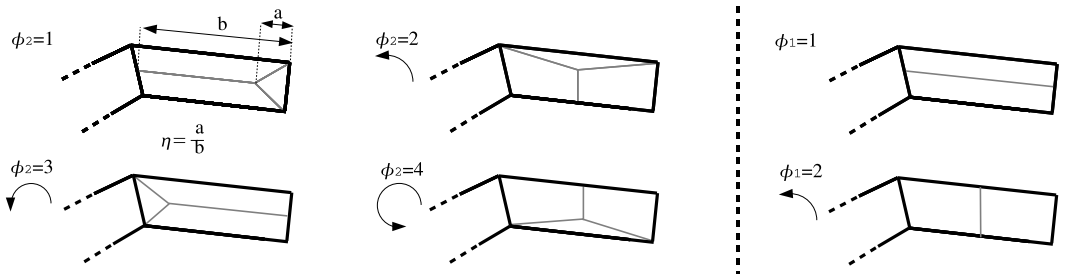

FIG. 3.5 – Paramètres ϕ et η - variante ayant deux symétries axiales *(droite)*, variante n'en possédant qu'une *(gauche)*.

restreints au premier abord. Cependant, en assemblant ces modèles les un aux autres, cette bibliothèque permet de modéliser une gamme très importante de structures urbaines. Par exemple, bien que nous ne disposons pas de modèles spécifiques aux toits "d'usine" dans la bibliothèque, on peut facilement les restituer en assemblant des modèles de toits en appentis de manière parallèle.

Il est important de noter que cette bibliothèque est évolutive et peut être étendue à d'autres modèles afin de gagner en généricité. Inversement, il est facile de la restreindre à des modèles très représentatifs d'un environnement urbain, afin de gagner en robustesse et en temps de calcul. Le contenu de la bibliothèque joue donc un rôle important dans le processus global de reconstruction.

3.2 Formulation Bayésienne

Dans cette partie, nous proposons une densité de probabilité permettant de caractériser la qualité d'une configuration de modèles 3D étant donné un MNE. Cette densité est définie dans un cadre bayésien : elle permet de mesurer la cohérence entre la configuration d'objets 3D et le MNE, et également de prendre en compte des connaissances a priori sur l'agencement des objets entre eux. Ce dernier point est fondamental pour la mise en place de la loi d'assemblage des modèles 3D.

La formulation de la densité requiert quelques notations, développées ci-dessous.

- Soient S, un ensemble de sites, et $\Lambda = \{\Lambda(s)/s \in S\}$, un MNE où $\Lambda(s)$ représente l'élévation au site s.

- Soit $\mathcal{Q}$, la configuration de quadrilatères et triangles représentant les supports 2D structuraux associée au MNE Λ. On note N, le nombre de quadrilatères inclus dans $\mathcal{Q}$.

- Soit S_i, le sous-ensemble de S dont les sites sont à l'intérieur du quadrilatère $i \in \mathcal{Q}$ et

Formes de toit	Paramètres de formes F	Variantes	
		Variantes possibles $\mathcal{V}$	Paramètres des variantes V
$\mathcal{F}_{11}$	H_g	$\mathcal{V}_-$	
$\mathcal{F}_{12}$	$H_g, H_t, \xi_1, \xi_2, \xi_3, \xi_4$	$\mathcal{V}_-$	
$\mathcal{F}_{13}$	H_g, H_t	$\mathcal{V}_-$	ϕ_2
		$\mathcal{V}_L$	ϕ_2
$\mathcal{F}_{21}$	H_g, H_t	$\mathcal{V}_-$	ϕ_1
		$\mathcal{V}_V$	ϕ_2, η
		$\mathcal{V}_{2V}$	ϕ_1, η
		$\mathcal{V}_L$	ϕ_2
		$\mathcal{V}_T$	ϕ_2
		$\mathcal{V}_+$	
$\mathcal{F}_{22}$	H_g, H_t, ζ	$\mathcal{V}_-$	ϕ_1
		$\mathcal{V}_V$	ϕ_2, η
		$\mathcal{V}_{2V}$	ϕ_1, η
		$\mathcal{V}_L$	ϕ_2
$\mathcal{F}_{23}$	H_g, H_t, ζ	$\mathcal{V}_-$	ϕ_2
		$\mathcal{V}_L$	ϕ_2
$\mathcal{F}_{24}$	H_g, H_t, κ	$\mathcal{V}_-$	ϕ_1
		$\mathcal{V}_L$	ϕ_2
		$\mathcal{V}_T$	ϕ_2
		$\mathcal{V}_+$	
$\mathcal{F}_{31}$	H_g, H_t	$\mathcal{V}_-$	ϕ_1
		$\mathcal{V}_L$	ϕ_2
		$\mathcal{V}_T$	ϕ_2
		$\mathcal{V}_+$	
$\mathcal{F}_{32}$	H_g, H_t	$\mathcal{V}_-$	ϕ_2
		$\mathcal{V}_L$	ϕ_2

TAB. 3.1 – Table récapitulative des paramètres de chaque modèle

Paramètres	Signification	Domaine de définition
H_g	Hauteur de gouttière d'un bâtiment	$[H_{gmin}, H_{gmax}]$
H_t	Hauteur de faîtière d'un bâtiment	$[H_{tmin}, H_{tmax}]$
$\xi_1, \xi_2, \xi_3, \xi_4$	Position du support rectangulaire du toit terrasse	$[0,1]^4$
ζ	Dissymétrie de l'arête faîtière	$[0,1]$
κ	Largeur du plateau d'un toit mansart	$[0,1]$
ϕ_1	Orientation d'un modèle axialement symétrique	$\{1,2\}$
ϕ_2	Orientation d'un modèle non symétrique	$\{1,2,3,4\}$
η	Profondeur de l'attaque en V	$[0,1]$

TAB. 3.2 – Les paramètres et leur signification

dont l'élévation est supérieure à la hauteur d'un demi-étage[1] H_e. On a donc $S_i = \{s \in int(i)/\Lambda(s) > H_e\}$.

- Soit $\mathcal{Y} = (\mathcal{Y}_i)_{i\in\mathcal{Q}}$, les données de notre problème inverse, où $\mathcal{Y}_i = \{\Lambda(s)/s \in S_i\}$.

- Soit x, une configuration de l'espace d'état $\mathcal{C}$, qui correspond à une configuration d'objets 3D paramétriques spécifiés par la bibliothèque $\mathcal{M}$ et les supports structuraux $\mathcal{Q}$. On écrit cette configuration $x = (x_i)_{i\in\mathcal{Q}} = (m_i, \theta_i)_{i\in\mathcal{Q}}$ où chaque objet x_i est caractérisé à la fois par un modèle de toit[2] $\mathcal{M}_{m_i} = (\mathcal{F}_{f_i}, \mathcal{V}_{v_i}) \in \mathcal{M}$, et par un ensemble de paramètres $\theta_i = (F_i, V_i)$ associé à $\mathcal{M}_{m_i}$. Dans la suite, $x_i = (m_i, \theta_i)$ et $\mathcal{M}_{m_i}$ seront appelés respectivement un objet et un modèle.

- Soit d_m, le nombre de paramètres à valeurs continues décrivant le modèle $\mathcal{M}_m$.

- Soit enfin $\mathcal{S}_{x_i}$, la fonction de S_i à valeur dans $\mathbb{R}$ qui associe l'altitude du toit de l'objet x_i pour chaque site de S_i.

En notant la fonction caractéristique $\mathbb{1}_{\{.\}}$, la mesure associée à l'ensemble des modèles de $\mathcal{M}$ est donnée par :

$$\mu(u) = \sum_{k\in\mathbb{N}^\star} \mathbb{1}_{\{m=k\}} \nu_m(\theta) \tag{3.1}$$

où $u = (m, \theta)$ est un objet de la bibliothèque et $\nu_m(.)$, la mesure associée au modèle m. La majorité des modèles ont à la fois des paramètres continus et discrets. Dans ce cas, $\nu_m(\theta)$ est le produit de deux mesures $\nu_m^{(c)}(\theta^{(c)}) \times \nu_m^{(d)}(\theta^{(d)})$ où $\nu_m^{(c)}(.)$ correspond à la mesure de Lebesgue sur $\mathbb{R}^{d_m}$ relativement aux paramètres continus $\theta^{(c)}$ du modèle , et $\nu_m^{(d)}(.)$ correspond à la mesure de comptage sur $\mathbb{N}$ relativement au paramètre discret $\theta^{(d)}$ du modèle (c'est-à-dire, le paramètre d'orientation du modèle ϕ). Pour les modèles ne possédant pas de paramètres discrets, $\nu_m(.)$ correspond simplement à la mesure de Lebesgue sur $\mathbb{R}^{d_m}$.

On peut maintenant considérer l'espace mesurable $(\mathcal{C}, \mathcal{B}(\mathcal{C}), \mu^N(.))$ associé à l'ensemble des configurations d'objets $\mathcal{C}$. On peut désormais définir la variable aléatoire X, distribuée dans l'espace des configurations d'objets $\mathcal{C}$ suivant une densité non normalisée $h(.)$ selon la mesure $\mu^N(.)$.

Dans notre problème inverse, $h(.)$ représente la densité a posteriori d'une configuration d'objets x, étant donné $\mathcal{Y}$. Dans un cadre bayésien, cette densité peut s'exprimer sous la forme suivante :

$$h(x) = h(x/\mathcal{Y}) \propto h_p(x)\mathcal{L}(\mathcal{Y}/x) \tag{3.2}$$

$\mathcal{L}(\mathcal{Y}/x)$ est la vraisemblance, c'est-à-dire la loi des observations. Elle permet de mesurer la cohérence entre les données $\mathcal{Y}$ et une configuration d'objets x. Le deuxième terme, $h_p(x)$,

[1] Lors de la phase d'extraction des emprises de bâtiments, certains supports ne sont pas correctement placés comme on peut le constater sur la figure 2.13. Il est important d'exclure les pixels correspondant au sol qui sont à l'intérieur des supports, sinon le calcul de l'attache aux MNE se trouverait fortement biaisé. Ces pixels sont donc retirés de l'ensemble des données en seuillant à un demi-étage du sol.

[2] L'indice m_i d'un modèle est défini par le couple d'indices (f_i, v_i) où $f_i \in \{11, 12, 13, 21, 22, 23, 24, 31, 32\}$ et $v_i \in \{-, V, 2V, L, T, +\}$.

correspond à la loi a priori. Cette loi permet d'introduire des connaissances sur l'agencement des objets. Le but est de mettre en place une vraisemblance et un a priori qui modélisent de manière pertinente notre problème.

Densité non normalisée et énergie de Gibbs

Comme cela a été souligné dans la partie 1.4.2, il est possible d'exprimer une densité normalisée h_{norm} par une énergie U à travers la loi de Gibbs. Cette relation peut s'avérer être très utile lors de la mise en place des différents termes qui composent une densité, mais également lors de la phase d'optimisation afin de poser le problème sous une formulation énergétique. Si $x \in \mathcal{C}$, cette loi s'exprime par :

$$h_{norm}(x) = \frac{1}{Z} \exp -U(x) \tag{3.3}$$

où Z est une constante de normalisation qui s'écrit :

$$Z = \int_{x \in \mathcal{C}} \exp -U(x) \tag{3.4}$$

Cependant, dans de nombreux problèmes où l'espace des configurations est très grand, cette constante de normalisation est très difficilement calculable ou estimable. C'est notre cas : l'ensemble des configurations d'objets est très important à cause notamment de la taille de notre bibliothèque, et des nombreux paramètres à valeurs continues inhérents aux différents modèles de toit. Le nombre d'objets contenus dans une scène peut également être élevé (plusieurs centaines en général). Il est toutefois possible de décomposer la scène en plusieurs problèmes indépendants correspondant à des bâtiments ou pâtés de maisons, afin de réduire le nombre d'objets (à quelques dizaines en général).
Pour cette raison, nous utilisons des densités non normalisées. Cela n'est pas préjudiciable dans la mesure où les méthodes d'optimisation que nous utilisons sont adaptées à ce type de densités. Ainsi, une densité non normalisée h peut s'exprimer à travers une énergie U par :

$$h(x) = \exp -U(x) \tag{3.5}$$

3.2.1 Vraisemblance

La vraisemblance est une densité qui mesure la cohérence entre les données observées $\mathcal{Y}$ et une configuration d'objets x. Afin de simplifier sa mise en place, nous considérons l'hypothèse d'indépendance conditionnelle des données. Cette hypothèse est plausible dans la mesure où les données locales $\mathcal{Y}_i$ ne se recouvrent pas entre elles. La vraisemblance peut ainsi s'exprimer sous la forme du produit des vraisemblances locales $\mathcal{L}(\mathcal{Y}_i/x_i)$ des objets :

$$\mathcal{L}(\mathcal{Y}/x) = \prod_{i \in Q} \mathcal{L}(\mathcal{Y}_i/x_i) \tag{3.6}$$

Pour estimer la pertinence d'un objet x_i par rapport à une partie de MNE, la solution généralement utilisée consiste à introduire une distance qui compare pixel à pixel l'altitude du toit de l'objet avec celle du MNE. Plusieurs distances peuvent correspondre à notre problème, parmi lesquelles la distance de Cauchy ou de Tukey [Xu et Zhang, 1996]. Cette dernière est

d'ailleurs particulièrement efficace pour traiter les problèmes dont les données sont bruitées puisque les valeurs aberrantes ont une pénalisation constante. L'utilisation de cette distance n'est cependant pas nécessaire puisque nous excluons les valeurs aberrantes dans la définition des données $\mathcal{Y}_i$. Nous utilisons plus classiquement une distance en norme L_α. Nous exprimons dès lors la vraisemblance par :

$$\mathcal{L}(\mathcal{Y}/x) = \prod_{i \in Q} \frac{1}{Z(x_i)} \exp -\Gamma^{\alpha}_{(i)}(\mathcal{S}_{x_i}, \mathcal{Y}_i) \tag{3.7}$$

où $Z(x_i)$ est la constante de normalisation de la vraisemblance locale, et $\Gamma^{\alpha}_{(i)}(.,.)$ correspond à la distance en norme L_α, définie de $\mathbb{R}^{card(S_i)} \times \mathbb{R}^{card(S_i)}$ dans $\mathbb{R}$.

Cette distance s'exprime par :

$$\Gamma^{\alpha}_{(i)}(\mathcal{S}_{x_i}, \mathcal{Y}_i) = \left(\sum_{s \in S_i} |\mathcal{S}_{x_i}(s) - \Lambda(s)|^{\alpha} \right)^{\frac{1}{\alpha}} \tag{3.8}$$

La distance Γ utilisée est fondée sur une comparaison entre les pixels du MNE et l'altitude du toit de l'objet x_i notée $\mathcal{S}_{x_i}$. Bien que relativement simple, elle permet de modéliser efficacement l'attache aux données de notre problème. La manière dont est définie cette distance, c'est-à-dire à travers une différence "pixel à pixel" sans autre intervention de paramètres de l'objet x_i dans l'expression, permet par ailleurs de rendre la constante de normalisation $Z(x_i)$ indépendante[3] de l'objet x_i. Cela signifie que le calcul de cette constante de normalisation n'est pas nécessaire : on résonne donc avec des vraisemblances locales non normalisées.

Le choix de α se tourne généralement vers la norme L_1, voire la norme L_2. La norme L_2 est particulièrement sensible aux fortes variations. Cette norme n'est donc pas spécialement adaptée à notre problème puisque les MNE que nous utilisons sont fortement bruités[4]. La norme L_1 est plus robuste au bruit. Cependant, nous avons remarqué expérimentalement que les profils des toits sur les MNE utilisés sont généralement concaves pour décrire, par exemple, des bi-plans symétriques (voir la figure 3.2). La norme L_1 réagit mal à cette concavité en ayant tendance à sous-estimer la pente du toit. La norme $L_{1,5}$ constitue, au regard des expérimentations réalisées, un bon compromis entre les normes L_1 et L_2. D'un point de vue algorithmique, la norme $L_{1,5}$ est légèrement plus lente à calculer que la norme L_2 ou L_1, ce qui représente son principal défaut.

En définitive, la vraisemblance est définie de manière relativement simple, sans introduction de paramètres qui pourraient faire perdre de la robustesse au processus global.

[3] Pour vérifier cela, il suffit d'écrire l'expression de la constante de normalisation $Z(x_i)$ et de poser le changement de variable $u(s) = \mathcal{S}_{x_i}(s) - \Lambda(s)$ quand on intègre sur les données, l'expression obtenue ne dépend alors plus de x_i. Cela suppose juste que les données soient intégrées sur un espace non borné.

[4] En plus du bruit caractérisant les erreurs de mise en correspondance, nous devons également faire face aux petits éléments de toiture comme les cheminées ou chien-assis qui sont perçus, à notre échelle de reconstruction, comme du bruit au sein des principaux plans de toit.

3.2.2 A priori

La densité a priori constitue un point fondamental de notre approche structurelle. Ce terme permet de régulariser la configuration d'objets afin de combler le manque d'information contenu dans les données et, également, d'avoir un rendu visuel de bonne qualité. Pour cela, les différents modules élémentaires constituant un bâtiment doivent être correctement assemblés entre eux.

Ce terme doit être défini simplement et faire intervenir peu de paramètres. La méthode présentée en annexe A, dont l'a priori est composé de trois types différents d'interactions pondérés entre eux par des paramètres, présente un manque de robustesse et souligne la nécessité de minimiser le nombre de types d'interactions. Nous nous imposons ainsi une contrainte sur le nombre de paramètres à introduire dans l'a priori. Plus précisément, ce terme ne devra comporter qu'un seul paramètre, c'est-à-dire le nombre minimal permettant la pondération avec le terme de vraisemblance. Nous cherchons donc à modéliser l'a priori à travers un seul type d'interactions devant regrouper les différentes contraintes sur l'assemblage des modules entres eux.

Relation de voisinage sur les supports des objets

La première étape consiste à définir une relation de voisinage sur $\mathcal{Q}$, c'est-à-dire sur la configuration des supports des objets. De manière naturelle, nous utilisons la relation de connexion entre quadrilatères, définie dans la partie 2.2.1. Ainsi, deux quadrilatères i et $j \in \mathcal{Q}$ sont dits voisins s'ils sont connectés. Cette relation de voisinage est notée V. L'ensemble des voisins du quadrilatère i sont ainsi notés $V(i)$. $i \bowtie j$ représente l'ensemble des paires de supports voisins dans $\mathcal{Q}$. Par abus de langage, nous nommerons par la suite "objets voisins", les objets dont les supports 2D sont voisins.

Loi d'assemblage entre deux objets voisins

Afin d'avoir une densité a priori simple et efficace, nous définissons une loi d'assemblage permettant de tester si deux objets voisins peuvent se raccorder entre eux. Deux objets $x_i = (m_i, \theta_i)$ et $x_j = (m_j, \theta_j)$ seront dits **assemblables** (noté $x_i \sim_a x_j$) s'ils vérifient les trois points suivants :

- **Compatibilité des formes de toit** - les deux objets doivent avoir le même type de toit. Cela signifie que deux objets voisins, l'un ayant par exemple un toit mansart, l'autre un toit elliptique, ne seront pas assemblables. Une exception est faite pour les toits plats avec les toits terrasse qui peuvent parfaitement se raccorder entre eux. Pour résumer, la compatibilité des formes de toit est satisfaite si $\mathcal{F}_{f_i} = \mathcal{F}_{f_j}$ ou si $\{\mathcal{F}_{f_i}, \mathcal{F}_{f_j}\} = \{\mathcal{F}_{11}, \mathcal{F}_{12}\}$.

- **Compatibilité des arêtes faîtières** - Les arêtes faîtières des deux objets doivent être disposées dans le prolongement de l'une par rapport à l'autre. En d'autres termes, nous voulons que la continuité planimétrique des arêtes faîtières d'un objet à l'autre soit possible. Cette contrainte est vérifiée par un test sur les paramètres spécifiques

des variantes des modèles, prenant en compte les paramètres d'orientation et les possibilités de jonction des variantes données par la figure 3.4.

- **Continuité des hauteurs de toit** - Les hauteurs des deux modules doivent être proches afin de permettre une continuité altimétrique des arêtes faîtières. Cela revient à tester si l'arête commune aux deux supports est une discontinuité de hauteur de toits, procédé établi dans la partie 2.2.2.

La première contrainte assure que les deux objets aient la même forme de toit. Les deux autres contraintes garantissent la possibilité de continuité des arêtes faîtières d'un objet à l'autre.

Formulation de l'a priori

La densité a priori que nous proposons consiste à favoriser les paires d'objets voisins qui sont assemblables. Cette interaction ne doit cependant pas être binaire. En effet, la loi d'assemblage n'assure pas l'assemblage parfait de deux objets, mais garantit juste que certaines conditions indispensables à leur assemblage sont vérifiées. Ainsi, nous introduisons une fonction g qui mesure la distance entre les paramètres de forme de deux objets assemblables.

Nous exprimons alors la densité a priori h_p à travers une énergie de Gibbs U_p (i.e. $h_p(x) = \exp -U_p(x)$) par :

$$\forall x \in \mathcal{C}\,, U_p(x) = \beta \sum_{i \bowtie j} \mathbb{1}_{\{x_i \sim_a x_j\}} g(x_i, x_j) \tag{3.9}$$

où $\beta \in \mathbb{R}^+$ représente le poids qui pondère l'importance de la densité a priori par rapport à la vraisemblance. La fonction g est à valeurs dans $[-1,0]$. Elle permet d'uniformiser les paramètres de deux objets voisins en les attirant vers un assemblage homogène, sans la présence d'artefacts. Plus précisément, la fonction g est formulée de la manière suivante :

$$g(x_i, x_j) = \frac{D(x_i, x_j)}{D_{max}} - 1 = \frac{\sum_k \omega_k |F_{i,(k)} - F_{j,(k)}|}{D_{max}} - 1 \tag{3.10}$$

où $F_{i,(k)}$ et $F_{j,(k)}$ correspondent au k^{ieme} élément de l'ensemble des paramètres de forme des objets x_i et x_j respectivement. $D_{max} = \max_{x_i, x_j} D(x_i, x_j)$ représente la valeur maximale de la distance. Les ω_k sont des poids qui sont introduits dans cette distance afin de normaliser les valeurs des paramètres selon le système métrique. Ces poids sont connus et sont calculés en tenant compte de la résolution en XY, de la résolution en Z et de la configuration de quadrilatères $\mathcal{Q}$.

La figure 3.6 illustre le principe de cette interaction. Si les deux objets ont des formes de toit différentes (configuration en haut à droite) ou si la continuité des arêtes faîtières n'est pas assurée (configuration en bas à droite), les deux objets ne sont pas assemblables : l'énergie associée à ces configurations est nulle. A l'inverse, si les deux objets sont assemblables (les autres configurations de la figure), l'énergie est négative : ces configurations sont favorisées. Grâce à la fonction g, elles seront d'autant plus favorisées que les objets seront

paramétriquement proches. La configuration de gauche est ainsi la meilleure : l'assemblage entre les deux objets est parfaitement homogène.

En définitive, la densité a priori permet de favoriser l'assemblage d'objets de manière proportionnelle à la qualité de l'assemblage. Elle n'interdit toutefois pas les configurations ne satisfaisant pas la loi d'assemblage : celles-ci ne sont juste pas favorisées. Les interactions proposées sont de cardinal deux. Nous aurions pu imaginer introduire des interactions de cardinal trois et quatre, favorisant spécifiquement les cas de jonctions en "T" et en "+". Cela aurait, cependant, complexifié la densité a priori et introduit de nouveaux paramètres pondérateurs.

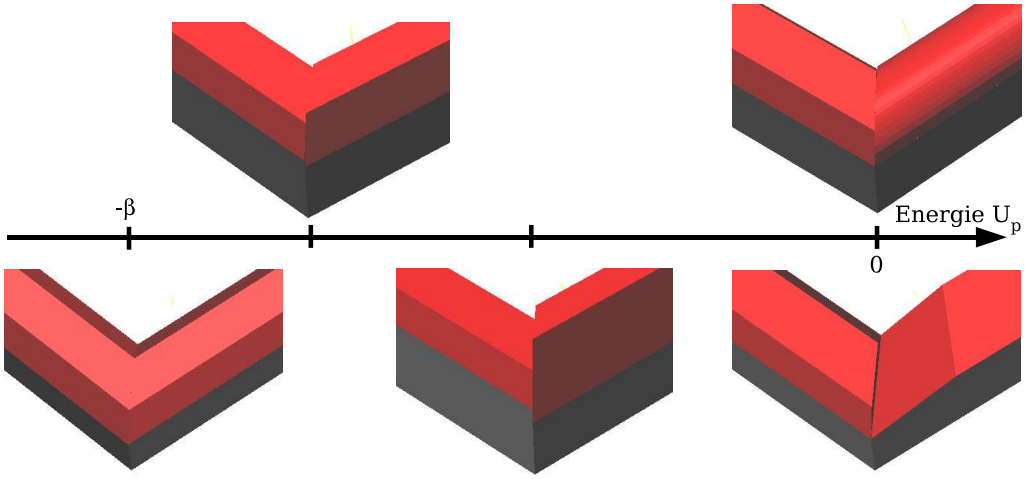

FIG. 3.6 – Principe de l'énergie a priori - exemples de divers cas d'interaction entre deux objets.

Finalement, la densité $h(.)$ que nous proposons s'exprime sous la forme suivante :

$$h(x) = \exp - \left(\sum_{i \in Q} \Gamma^{\frac{3}{2}}_{(i)}(S_{x_i}, \mathcal{Y}_i) + \beta \sum_{i \lhd j} \mathbb{1}_{\{x_i \sim_a x_j\}} g(x_i, x_j) \right) \tag{3.11}$$

Un seul paramètre pondérateur intervient dans cette densité : il s'agit de β. Cela accorde de la robustesse au processus et permet de limiter les problèmes engendrés par le réglage des paramètres. Dans la suite, l'énergie de Gibbs associée à la densité non normalisée h sera notée U ($U = -\ln h$).

3.3 Optimisation

L'objectif consiste à trouver la configuration d'objets qui maximise la densité $h(.)$. Cela revient à déterminer l'estimateur du Maximum A Posteriori (MAP) x_{MAP} :

$$x_{MAP} = \arg\max_{x \in \mathcal{C}} h(x) \tag{3.12}$$

C'est un problème d'optimisation non convexe dans un espace $\mathcal{C}$ très grand, dont les variables sont, pour la plupart, à valeurs continues dans des compacts de $\mathbb{R}$. De plus, comme les modèles de la bibliothèque $\mathcal{M}$ sont définis par un nombre différent de paramètres, $\mathcal{C}$ est une union d'espaces de dimension variable.
Très peu d'algorithmes peuvent être utilisés efficacement sous ces contraintes. Dès lors, deux options sont possibles :

- Soit utiliser un algorithme stochastique adapté aux caractéristiques de notre problème, mais relativement coûteux en temps de calcul - Nous pensons en particulier aux méthodes de Monte Carlo par Chaînes de Markov (MCMC) [Robert et Casella, 1999] .

- Soit simplifier notre problème d'optimisation afin d'utiliser des techniques moins coûteuses - Nous envisageons notamment de restreindre l'espace d'état (discrétisation des paramètres) ou de simplifier la densité proposée (modifier la relation de voisinage en séquence d'objets par exemple, ou poser des hypothèses sur l'énergie U afin de la rendre dérivable).

Nous optons pour la première solution. L'algorithme d'optimisation stochastique par MCMC que nous proposons est détaillé dans la suite de cette partie. Pour comprendre son fonctionnement, le lecteurs pourra consulter l'annexe C dans laquelle la définition des chaînes de Markov, leurs propriétés et les conditions de convergence sont rappelées. Cependant, nous évoquons également des méthodes d'optimisation alternatives fondées sur une simplification de notre problème. Celles-ci sont abordées dans le dernier paragraphe de cette partie.

3.3.1 Echantillonneur RJMCMC

Les techniques d'échantillonnage de Monte Carlo par Chaînes de Markov (MCMC) consistent à simuler une chaîne de Markov discrète $(X_t)_{t\in\mathbb{N}}$ sur l'espace des configurations qui converge vers une mesure cible. Les transitions de cette chaîne correspondent à des perturbations locales de la configuration courante. Elles sont locales, ce qui signifie qu'en général, seul un composant (ou objet) de la configuration courante est concerné par une perturbation. Cette chaîne est conçue de manière à être ergodique. Cela permet, sous certaines conditions de relaxation probabiliste que nous détaillerons dans la suite, d'assurer la convergence de la chaîne vers la mesure cible quelle que soit la configuration initiale. Ces algorithmes construisent itérativement les différents états de la chaîne à travers la mise en place des transitions. Celles-ci sont définies en deux étapes : une phase de proposition de perturbation de l'état courant, suivie d'une phase de décision où l'on choisit d'accepter ou de refuser la perturbation.

Ce type d'algorithmes a été introduit initialement par Metropolis [Metropolis et al., 1953]. Hastings [Hastings, 1970] a ensuite proposé un formalisme plus général dans lequel les probabilités de transitions ne sont plus forcement symétriques. Enfin, la dernière extension, celle de Green [Green, 1995], permet à ces échantillonneurs de traiter les cas où l'espace des configurations est de dimension variable. Cette dernière, appelée échantillonneur de Monte Carlo par Chaînes de Markov à sauts réversibles (RJMCMC pour "Reversible

Jump Markov Chain Monte Carlo" en anglais), est particulièrement efficace pour aborder les problèmes de reconnaissance d'objets multiples. Plusieurs travaux ont d'ailleurs démontré l'efficacité de l'algorithme RJMCMC dans ce type de problèmes. On peut citer notamment [Dick et al., 2004] qui reconstruit des bâtiments remarquables à partir d'imagerie terrestre en utilisant des objets 3D paramétriques multiples tels que des colonnes, des fenêtres ou bien encore des pontons. Il est aussi utilisé par [Brenner et Ripperda, 2006] pour extraire des façades de bâtiments. Cet échantillonneur est également bien adapté dans le cadre des processus ponctuels marqués comme cela a été souligné dans la partie 2.1, entre autres avec les travaux de [Lacoste et al., 2005; Perrin et al., 2005; Ortner et al., 2007].

Afin de construire une chaîne de Markov pour notre problème, plusieurs outils doivent être introduits :

- La mesure cible π de la chaîne définie sur $\mathcal{C}$ et spécifiée par la densité a posteriori h (elle est, dans notre cas, non normalisée, c'est-à-dire connue à une constante de normalisation près).

- Les noyaux de proposition de perturbations $Q_k(.,.)$ définis sur $\mathcal{C} \times \mathcal{B}(\mathcal{C})$. Ils permettent de proposer des perturbations de différents types k que nous détaillons dans la partie 3.3.2.

- Une mesure symétrique $\varphi_k(.,.)$ associée au noyau $Q_k(.,.)$ et définie sur $\mathcal{C} \times \mathcal{C}$, telle que $\pi(.)Q_k(.,.)$ soit absolument continue par rapport à $\varphi_k(.,.)$. Cette condition permet d'assurer l'existence et l'unicité de la dérivée de Radon-Nikodym $f_k(.,.)$ définie par :

$$f_k(x,y) = \frac{\pi(dx)Q_k(x,dy)}{\varphi_k(dx,dy)} \tag{3.13}$$

Dès lors, on définit la probabilité d'acceptation d'une proposition $x \rightarrow y$ par $\min(1, R_k(x,y))$ où R_k représente le taux de Green :

$$R_k(x,y) = \frac{f_k(y,x)}{f_k(x,y)} = \frac{\pi(dy)Q_k(y,dx)}{\pi(dx)Q_k(x,dy)} \tag{3.14}$$

Cette probabilité d'acceptation est définie de manière à vérifier la condition de réversibilité de la chaîne (impliquant la stationnarité de π), condition nécessaire pour obtenir l'ergodicité de la chaîne.

Pour résumer, l'échantillonneur RJMCMC peut se formuler de la manière suivante. A l'itération t, si $X_t = x$:

- 1- Choisir un noyau de proposition $Q_k(x,.)$ avec la probabilité q_k.

- 2- Proposer selon Q_k une nouvelle configuration y.

- 3- Accepter $x^{(t+1)} = y$ avec la probabilité $\min(1, R_k(x,y))$, et conserver $x^{(t+1)} = x$ sinon.

3.3.2 Noyaux de propositions

La mise en place des noyaux de propositions constitue un point fondamental dans l'efficacité de l'échantillonneur. Il est possible d'utiliser les noyaux dits "classiques", qui permettent de proposer des perturbations uniformément sur l'espace d'état. Ces noyaux sont suffisants pour garantir l'irréductibilité de la chaîne puisqu'ils permettent d'atteindre n'importe quelle configuration de l'espace d'état. Cependant, il est préférable d'utiliser également d'autres noyaux plus pertinents et mieux adaptés à notre problème. L'objectif est d'accélérer la convergence du processus en proposant plus régulièrement des configurations d'intérêt. La mise en place de ces noyaux est, cela dit, un problème délicat. Il ne suffit pas, en effet, de proposer en quantité importante des configurations d'intérêt pour accélérer l'algorithme, car ce n'est pas pour autant qu'elles seront acceptées. En effet, le taux de Green prend en compte la perturbation réciproque et, si celle-ci est peu probable, le rapport $\frac{Q_k(y,dx)}{Q_k(x,dy)}$ sera faible. Et la proposition, bien que pertinente, sera facilement rejetée.

Les noyaux de propositions se calculent de manière relativement simple pour l'échantillonneur d'Hastings, les propositions étant fondées sur des modifications de paramètres. Pour l'échantillonneur de Green, la mise en place des noyaux est plus délicate puisqu'il faut gérer le changement de dimension des objets. Considérons deux modèles $\mathcal{M}_m$ et $\mathcal{M}_n$ et une perturbation (appelée également saut) d'un objet $x_i = (m, \theta_i)$ vers $\widehat{x_i} = (n, \widehat{\theta_i})$ de telle sorte que la configuration courante $x = (x_p)_{p \in \mathcal{Q}}$ soit perturbée en la configuration $y = (x_p)_{p \in \mathcal{Q} - \{i\}} \cup \widehat{x_i}$. L'idée introduite par Green consiste à créer une bijection entre l'espace des paramètres des modèles $\mathcal{M}_m$ et $\mathcal{M}_n$. Pour cela, θ_i est complété en simulant $u_{mn} \sim \varphi_{mn}(.)$ à travers (θ_i, u_{mn}), et $\widehat{\theta_i}$ en simulant $v_{nm} \sim \varphi_{nm}(.)$ à travers $(\widehat{\theta_i}, v_{mn})$ de manière à ce que l'application Ψ_{mn} entre (θ_i, u_{mn}) et $(\widehat{\theta_i}, v_{mn})$ soit une bijection :

$$(\widehat{\theta_i}, v_{mn}) = \Psi_{mn}(\theta_i, u_{mn}) \tag{3.15}$$

Le rapport des noyaux présent dans le taux de Green peut alors s'exprimer par :

$$\frac{Q_k(y,dx)}{Q_k(x,dy)} = \frac{J_{nm}^{(k)} \varphi_{nm}^{(k)}(v_{nm})}{J_{mn}^{(k)} \varphi_{mn}^{(k)}(u_{mn})} \left| \frac{\partial \Psi_{mn}(\theta_i, u_{mn})}{\partial(\theta_i, u_{mn})} \right| \tag{3.16}$$

où J_{mn} représente la probabilité de faire un saut du modèle $\mathcal{M}_m$ vers le modèle $\mathcal{M}_n$. Calculer la bijection Ψ_{mn} est un problème difficile quand les modèles $\mathcal{M}_m$ et $\mathcal{M}_n$ diffèrent par de nombreux paramètres. Dans notre cas, les modèles possèdent entre un et six paramètres seulement, ce qui facilite le calcul des bijections, détaillé en annexe D.

La mise en place des noyaux de proposition s'effectue en spécifiant à la fois les probabilités J_{mn} de saut entre modèles et les distributions φ_{mn} des paramètres de complétion du modèle $\mathcal{M}_n$ par rapport au modèle $\mathcal{M}_m$. Nous proposons trois types de noyaux, c'est-à-dire trois jeux de distributions (J_{mn}, φ_{mn}).

- **Noyau Q_1 : perturbation uniforme**

  Il constitue le noyau de base, du moins le plus simple que l'on puisse mettre en place tout en garantissant l'apériodicité de la chaîne. Le nouvel état est proposé de manière

uniforme sur l'espace des configurations. En d'autres termes, nous tirons aléatoirement une proposition, sans a priori. Ce noyau est intéressant dans la mesure où, même s'il n'est pas fréquemment proposé ($0 < q_1 << 1$), il permet de garantir l'apériodicité de la chaîne sous certaines conditions de relaxation. Cependant, la convergence est très lente comme nous avons pu le constater à travers la méthode de reconstruction présentée en annexe A. Le jeu de distributions $(J_{mn}^{(1)}, \varphi_{mn}^{(1)})_{m,n}$ de ce noyau est donné par :

$$J_{mn}^{(1)} = \frac{1}{card(\mathcal{M})} \tag{3.17}$$

$$\varphi_{mn}^{(1)} = \mathcal{U}_{K_{mn}} \tag{3.18}$$

où $\mathcal{U}_{K_{mn}}$ est une distribution uniforme sur le compact K_{mn} qui représente le domaine de définition des paramètres de complétion du modèle $\mathcal{M}_n$ par rapport au modèle $\mathcal{M}_m$.

- **Noyau Q_2 : perturbation dirigée par les données**

Ce noyau permet d'explorer, de manière plus pertinente que le noyau Q_1, l'espace des configurations en introduisant des connaissances provenant des données, comme l'illustrent notamment les travaux de [Tu et Zhu, 2002]. Comme nous l'avons souligné précédemment, il est difficile d'extraire dans le MNE des informations fiables sur les objets. Les seuls paramètres pouvant être estimés avec robustesse sont les paramètres altimétriques, c'est-à-dire H_g et H_t, respectivement la hauteur de gouttière et la hauteur de faîtière d'un objet. Il est très intéressant d'avoir une estimation de ces paramètres car ce sont les plus utilisés dans la bibliothèque (voir la table 3.1). De plus, ces deux paramètres sont importants dans la mesure où ils constituent la principale source d'erreur lors des reconstructions de bâtiments. La méthode estimant ces deux paramètres à partir du MNE est présentée en annexe B-2. On note $\widehat{H_g}$ et $\widehat{H_t}$, les estimations de H_g et H_t fournies par cette méthode.

Pour ces deux paramètres, il faut utiliser des distributions plus pertinentes que la distribution uniforme, de préférence centrées en $\widehat{H_g}$ et $\widehat{H_t}$. Nous choisissons la distribution gaussienne. Ainsi, les distributions φ_{nm} des paramètres de complétion suivent les mêmes lois (uniformes) que pour le noyau Q_1, à l'exception des paramètres H_g et H_t qui sont tirés suivant des lois normales $\mathcal{N}(\widehat{H_g}, \sigma_1)$ et $\mathcal{N}(\widehat{H_t}, \sigma_2)$ respectivement. Les variances σ_1 et σ_2 sont des paramètres qui sont fixés respectivement à 1 et $0,5$ mètre. Ces valeurs ont été choisies en fonction de la précision fournie par la méthode d'estimation de $\widehat{H_g}$ et $\widehat{H_t}$.

Concernant le choix du modèle, il est envisageable de mettre en place une méthode estimant la probabilité d'un modèle de toits à partir d'un MNE. Nous n'utiliserons pas le MNE pour définir les probabilités $J_{nm}^{(2)}$, mais simplement des informations a priori sur l'occurrence des différents modèles en milieu urbain, fournies dans les tables 3.3 et 3.4 à travers un système de poids (plus le poids est important, plus l'occurrence est forte). Les J_{mn} sont calculées indépendamment du modèle de départ

$\mathcal{M}_m$. Si $\mathcal{M}_n = (\mathcal{F}_p, \mathcal{V}_q)$, on définit J_{mn} par :

$$J_{mn} = \frac{P_p^{(f)}}{\sum\limits_{k \in \mathcal{F}} P_k^{(f)}} \times \frac{P_q^{(v)}}{\sum\limits_{k' \in \mathcal{V}_{\mathcal{F}_p}} P_{k'}^{(v)}} \tag{3.19}$$

Forme	$\mathcal{F}_{11}$	$\mathcal{F}_{12}$	$\mathcal{F}_{13}$	$\mathcal{F}_{21}$	$\mathcal{F}_{22}$	$\mathcal{F}_{23}$	$\mathcal{F}_{24}$	$\mathcal{F}_{31}$	$\mathcal{F}_{32}$
Poids $P^{(f)}$	2	1	1	5	2	1	1	1	1

TAB. 3.3 – Occurrence des formes de toits

Variante	$\mathcal{V}_-$	$\mathcal{V}_V$	$\mathcal{V}_{2V}$	$\mathcal{V}_L$	$\mathcal{V}_T$	$\mathcal{V}_+$
Poids $P^{(v)}$	5	2	2	2	1	1

TAB. 3.4 – Occurrence des variantes

- **Noyau Q_3 : perturbation de régularisation**

  Dans notre application, le rendu visuel du résultat est un aspect très important. En particulier, les différents modules constituant un bâtiment doivent être correctement assemblés entre eux afin de minimiser le nombre d'artefacts. En pratique, l'algorithme d'optimisation que nous utilisons procède en deux phases : une première dans laquelle il explore l'espace des configurations et favorise les modes de forte densité et une seconde dans laquelle il cherche à ajuster minutieusement les objets dans une configuration courante proche de l'optimale. Cette seconde phase, qui est capitale pour le rendu visuel du résultat, est longue et fastidieuse puisque la grande majorité des propositions sont mauvaises et donc rejetées.

  L'idée de ce noyau consiste alors à proposer des objets paramétriquement proches de leurs voisins afin d'accélérer la régularisation de la configuration d'objets. Contrairement au noyau Q_2 qui prend en compte des informations provenant des données pour orienter les propositions, ce noyau utilise des informations provenant des objets voisins de l'objet concerné par une proposition. Le modèle de l'objet est proposé parmi ceux des objets voisins (on y inclut le modèle courant de l'objet en question, afin d'avoir la réversibilité du saut). Les probabilités J_{mn} de saut entre modèles sont ainsi données par :

  $$J_{mn}^{(3)} = \frac{N_{\mathcal{F}_n}}{N_t} \tag{3.20}$$

  où $N_{\mathcal{F}_n}$ représente le nombre de voisins de l'objet concerné dont la forme de toit est $\mathcal{F}_n$, et N_t, le nombre total de voisins. Chaque paramètre du modèle est proposé suivant une loi gaussienne dont l'espérance correspond à la moyenne des paramètres voisins.

  $$\varphi_{mn}^{(3)} = (\mathcal{N}(\widehat{F}_k, \sigma_2))_k \tag{3.21}$$

  où $\widehat{F}_k$ est la moyenne des k_{ieme} paramètres des objets voisins dont la forme est $\mathcal{F}_n$.
  Ce noyau agit dans un espace de configurations réduit en proposant de "petites" perturbations. Cela permet par ailleurs d'avoir des perturbations réciproques tout-à-fait

probables et un taux de Green non nul. L'utilisation de ce noyau est dangereuse car il va avoir tendance à geler la configuration courante dans un optimum local. C'est pour cette raison qu'il doit être employé uniquement en fin de processus, c'est-à-dire quand la configuration courante est proche de la solution optimale.

3.3.3 Recuit simulé

L'échantillonneur RJMCMC est couplé avec un algorithme de relaxation stochastique, en l'occurrence un recuit simulé, afin de permettre la convergence vers la solution désirée. Pour ce faire, la densité $h(.)$ est remplacée par la densité $h(.)^{\frac{1}{T_t}}$ dans le taux de Green. T_t est une suite de températures qui tend vers 0 quand t tend vers l'infini. Le recuit simulé permet d'assurer la convergence du processus vers l'optimum global, quelle que soit la configuration initiale d'objets, en utilisant une décroissance de température logarithmique. En pratique, on préfère utiliser des schémas de décroissance plus rapides, convergeant vers une solution approchée de l'optimum global. La décroissance géométrique est, dans ce cas, le schéma de référence :

$$T_t = T_o.\alpha^t \tag{3.22}$$

où α et T_0 sont respectivement le coefficient de décroissance et la température initiale. Davantage de détails sur ce schéma de décroissance pourront être trouvés dans [Van Laarhoven et Aarts, 1987]. Il existe également d'autres schémas intéressants qui consistent à adapter la décroissance de la température en fonction de la variation de l'énergie [Haario et Saksman, 1991; Varanelli, 1996; Perrin et al., 2005]. De manière générale, ces schémas ralentissent la décroissance de la température lorsqu'on est proche de la température critique[5].

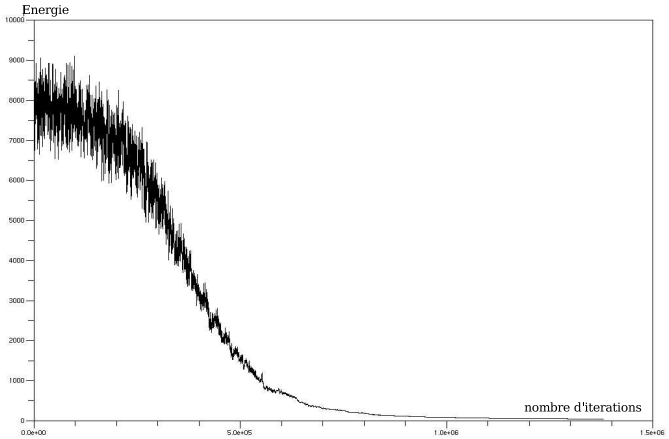

FIG. 3.7 – Illustration d'une décroissance d'énergie par un schéma géométrique

[5]La température critique est une température théorique durant laquelle le processus doit choisir entre plusieurs puits énergétiques. Elle correspond généralement à la période pendant laquelle la variation d'énergie est forte. Le processus doit donc faire le bon choix à cette température, c'est pourquoi plusieurs schémas adaptatifs consistent à ralentir la décroissance à cet endroit.

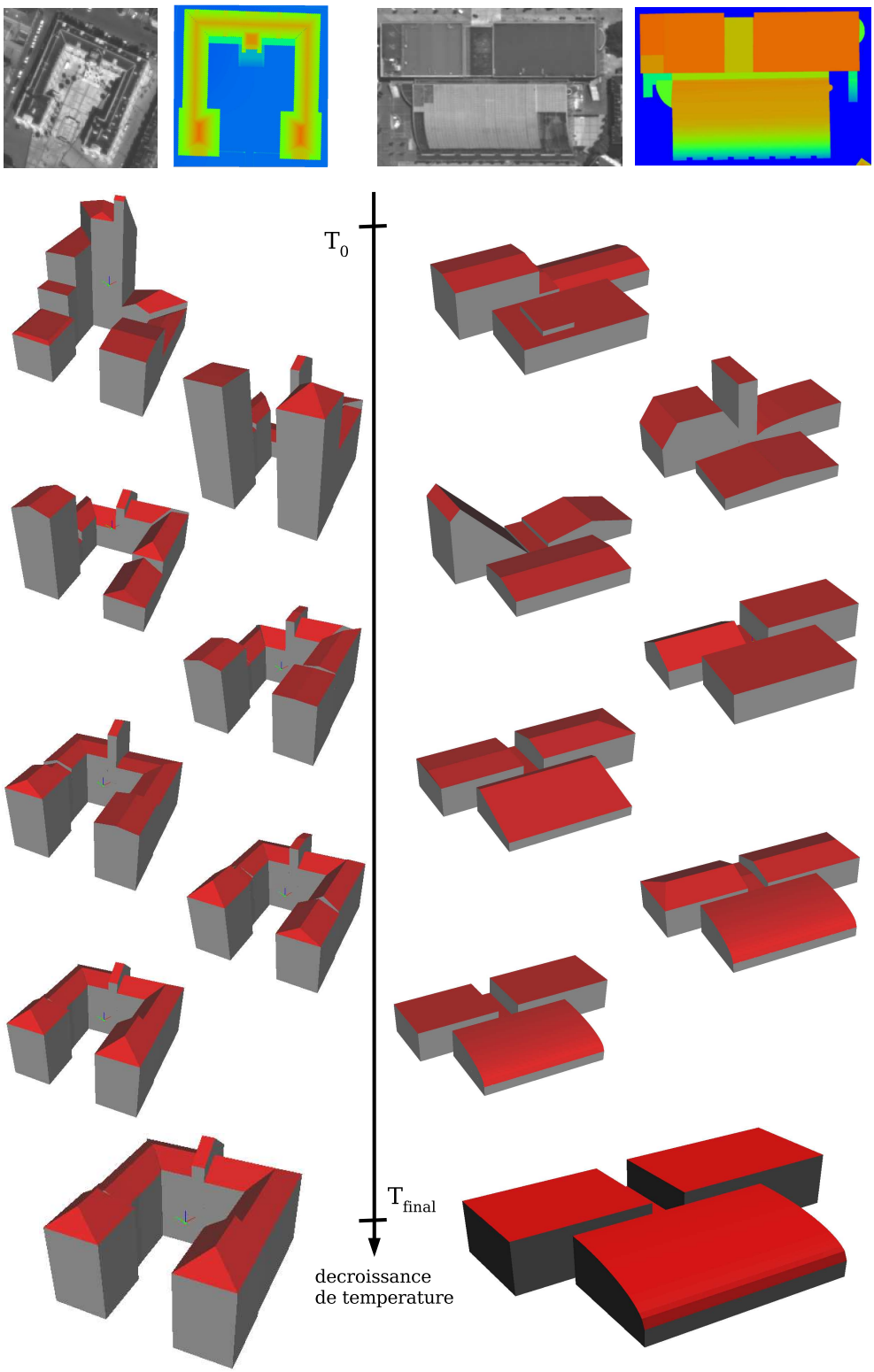

FIG. 3.8 – Deux exemples simples d'évolution de configuration d'objets au cours de la décroissance de température *(bas)*, simulations PLEIADES et vérités terrain associées *(haut)*.

Le processus de décroissance se décompose en deux phases[6]. La figure 3.7 représente une évolution d'énergie typique pour une décroissance de température géométrique relativement lente. Au début de l'algorithme, c'est-à-dire quand la température est élevée, l'échantillonneur explore les modes de la densité. C'est la phase d'exploration, durant laquelle les configurations d'objets ayant une densité élevée sont favorisées au fur et à mesure que la température décroît. Quand la configuration commence à se stabiliser (à faible température), on entre dans la seconde phase qui correspond à un ajustement minutieux des objets de la configuration[7]. Le taux de rejet de propositions est alors très important. La figure 3.8 illustre deux exemples simples d'évolution de configuration d'objets au cours de la décroissance de température.

Les noyaux Q_2 et Q_3 permettent de proposer des perturbations adaptées à ces deux phases. Q_2 est un noyau d'exploration, utile pour la première phase. Q_3 doit être choisi principalement durant la seconde phase, afin de régulariser la configuration d'objets.

Le piège des minima locaux

Le principal danger lors de l'utilisation du recuit simulé consiste à utiliser un schéma de décroissance trop rapide. Le processus risque alors de ne pas sélectionner le bon mode de densité. Cela l'amène à rester bloqué dans un minimum local de l'énergie.

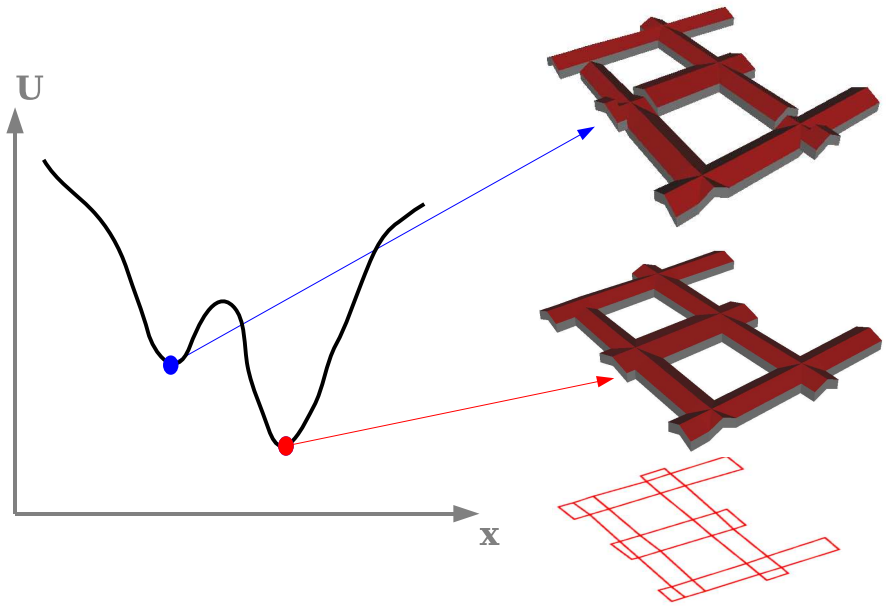

FIG. 3.9 – Illustration du piège des minima locaux.

[6]Parfois, on considère qu'il existe une phase supplémentaire, préliminaire aux deux autres. C'est, comme on peut le voir sur la figure 3.7, un plateau énergétique où une large majorité de propositions sont acceptées et qui permet de s'assurer que la température initiale est assez élevée.

[7]Durant cette phase de régularisation, on se trouve régulièrement piégé dans de petits puits énergétiques (voir le prochain paragraphe). C'est pour cette raison qu'il n'est pas possible d'utiliser en fin de processus un algorithme déterministe de type descente de gradient pour régulariser la configuration.

La figure 3.9 illustre ce problème. La configuration correspondante au point bleu représente un minimum local de notre énergie : deux groupes d'objets, relativement bien attachés aux données, se sont régularisés entre eux. Or la configuration optimale (point rouge) correspond à un seul groupe d'objets, dont tous les éléments voisins sont assemblés entre eux. Ainsi, pour passer du point bleu au point rouge, il faut casser au moins la régularisation d'un des deux groupes d'objets pour que les éléments concernés s'assemblent progressivement avec les objets de l'autre groupe. Pour réaliser cela, il faut que le processus visite des configurations intermédiaires dont le niveau énergétique est mauvais (ou du moins supérieur à l'énergie du point bleu). La décroissance de température doit être assez lente pour assurer la visite de ces configurations intermédiaires.

3.3.4 Des méthodes alternatives

La méthode d'optimisation stochastique par MCMC qui vient d'être proposée est adaptée aux caractéristiques de notre problème, mais est relativement coûteuse en temps de calcul. Cette partie constitue un aparté dans lequel nous évoquons plusieurs méthodes alternatives permettant de simplifier notre problème d'optimisation, et qui soulèvent des pistes intéressantes pour le futur.

Modélisation par chaînes de Markov cachées après simplification des supports 2D

La première méthode consiste à transformer notre modèle en processus causaux dans lesquels un bâtiment est considéré comme une association de séquences d'objets (voir la figure 3.10), tout en réduisant fortement l'espace des configurations par une discrétisation des paramètres. Nous simplifions ainsi les supports 2D de manière à obtenir des séquences de quadrilatères connectés (ces quadrilatères ont donc au plus deux voisins).

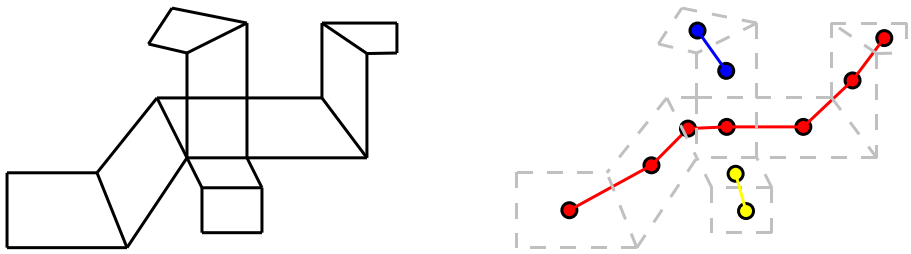

FIG. 3.10 – Simplification des emprises en séquences de supports 2D.

Pour chaque séquence de N supports, nous utilisons une chaîne de Markov cachée homogène $(X_t, Y_t)_{t=1}^{N}$ dont le graphe de dépendance est illustré sur la figure 3.11. $X = (X_t)_t$ représente les états cachés (i.e. la séquence x d'objets paramétriques 3D) et $Y = (Y_t)_t$, la séquence des observations correspondantes. La probabilité de transition $P(X_{t+1} = x_{t+1} | X_t = x_t)$ et la vraisemblance locale $P(Y_t = y_t | X_t = x_t)$ sont directement calculées à partir de l'expression de la densité h définie dans le chapitre 3.2. L'espace des configurations est discrétisé afin d'utiliser un processus d'optimisation causal, en l'occurrence l'algorithme de Viterbi [Viterbi, 1967; Forney, 1973]. Le pas de discrétisation des paramètres joue un rôle

important dans le processus : il détermine le degré de précision de la modélisation, mais également les temps de calcul.

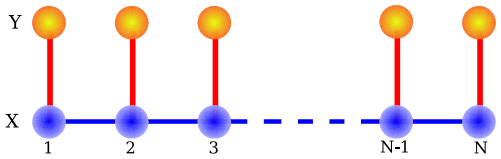

FIG. 3.11 – Graphe de dépendance de la chaîne de Markov cachée $(X_t, Y_t)_{t=1}^{N}$

Cette simplification en séquences d'objets est pénalisante dans la mesure où les jonctions complexes (jonctions en T ou en +) ne pourront plus être restituées correctement puisque les chaînes de Markov cachées sont indépendantes entre elles. Une des solutions consisterait à utiliser une modélisation par graphe hiérarchique à deux niveaux : le niveau inférieur correspondant aux séquences d'objets, le niveau supérieur permettant d'introduire des lois de dépendance entre les séquences. Une autre possibilité pourrait être d'utiliser les techniques de propagation des messages. Les méthodes de "belief propagation" semblent être bien adaptées, d'autant que les graphes de notre problème possèdent souvent peu de cycles. Les approches utilisant un formalisme fondée sur les modèles graphiques ouvrent, en tout cas, des perspectives particulièrement intéressantes.

En définitive, cette méthode est intéressante pour obtenir rapidement une modélisation approchée (i.e. avec un pas de discrétisation de l'ordre de deux mètres, une bonne régularisation, mais sans restitution des jonctions complexes) avec des temps de calcul très faibles (de l'ordre d'une seconde pour un bâtiment). Si l'on souhaite un degré de précision supérieur, le processus par optimisation stochastique est alors mieux adapté. Cette méthode alternative est présentée de manière plus détaillée en annexe E.

Utilisation des méthodes de diffusion ?

Parmi les techniques d'optimisation relativement rapides, les méthodes utilisant les gradients énergétiques pour se déplacer dans l'espace des configurations sont très efficaces. Si l'énergie est convexe, une simple descente de gradient est suffisante pour atteindre la solution optimale. Dans le cas contraire, il existe des techniques de diffusion stochastique notamment les méthodes fondées sur les équations différentielles stochastiques [Kloeden et Platen, 1992; Salomon et al., 2002; Descombes et Zhizhina, 2004] ou celles couplées à des processus de sauts ("Jump-Diffusion") [Grenander et Miller, 1994; Srivastava et al., 2002; Han et al., 2004].

Dans ce paragraphe, nous regardons dans quelles mesures nous pourrions utiliser ce type de techniques d'optimisation. La principale contrainte est la dérivabilité de l'énergie. Du moins, la dérivée de l'énergie doit être facilement calculable ou estimable en tout point.

L'énergie de notre problème est donnée par l'expression suivante :

$$U(x) = \sum_{i \in Q} \Gamma^{\frac{3}{2}}_{(i)}(\mathcal{S}_{x_i}, \mathcal{Y}_i) + \beta \sum_{i \bowtie j} \mathbb{1}_{\{x_i \sim_a x_j\}} g(x_i, x_j) \tag{3.23}$$

Or $U(x)$ n'est pas dérivable en tout point de son espace d'état. Le terme a priori contient une fonction caractéristique et une fonction g non dérivables en tout point. C'est également le cas du terme d'attache aux données, où les fonctions d'équations des toits $\mathcal{S}_{x_i}$ sont C_1 par morceaux (les toits étant des associations connexes de plans). Une solution simple permettant de contourner le problème consiste à restreindre l'espace d'état et/ou modifier la formulation de U. Ainsi, nous pourrions simplifier le problème en considérant l'hypothèse suivante : les configurations dans lesquelles les objets voisins ne sont pas assemblables sont improbables. Cela signifie que l'on réduit l'espace aux configurations dans lesquelles les objets voisins sont assemblables. De plus, nous pourrions modifier certaines fonctions de manière à redéfinir notre énergie sous la forme :

$$\widehat{U}(x) = \sum_{i \in Q} \Gamma^{\frac{3}{2}}_{(i)}(\mathcal{S}_{x_i}, \mathcal{Y}_i) + \beta \sum_{i \bowtie j} g'(x_i, x_j) \tag{3.24}$$

où g' est une fonction de distance analogue à g mais définie en norme L_2. $\widehat{U}$ diffère de U par les normes utilisées dans l'a priori. Dès lors, si l'on dérive $\widehat{U}$ par rapport au k^{ieme} paramètre de l'objet X_i, on obtient l'expression suivante :

$$\frac{d\widehat{U}(x)}{dx_i^{(k)}} = \frac{d}{dx_i^{(k)}} \Gamma^{\frac{3}{2}}_{(i)}(\mathcal{S}_{x_i}, \mathcal{Y}_i) + \beta \sum_{j \in V(i)} \frac{dg'(x_i, x_j)}{dx_i^{(k)}} \tag{3.25}$$

La dernière difficulté concerne les fonctions d'équation des toits $\mathcal{S}_{x_i}$ qui sont C_1 par morceaux. Cela rend, par conséquent, le terme d'attache aux données non dérivable sur les bords des différentes parties des supports de toit. Une solution possible consiste à négliger la variation des différentes parties des supports dans l'expression de la dérivée. Dans ce cas, nous avons :

$$\begin{aligned} \frac{d\Gamma^{\frac{3}{2}}_{(i)}(\mathcal{S}_{x_i}, \mathcal{Y}_i)}{dx_i^{(k)}} &= \frac{d}{dx_i^{(k)}} \left(\int_{S_i} |z - \mathcal{Y}_i|^{\frac{3}{2}} dm \right)^{\frac{2}{3}} \\ &\approx \left(\int_{S_i} |z - \mathcal{Y}_i|^{\frac{3}{2}} dm \right)^{-\frac{1}{3}} \sum_p \int_{S_i^{(p)}} \left(sign(z^{(p)} - \mathcal{Y}_i) \sqrt{|z^{(p)} - \mathcal{Y}_i|} \frac{dz^{(p)}}{dx_i^{(k)}} \right) dm \end{aligned} \tag{3.26}$$

où les $S_i^{(p)}$ correspondent aux différentes parties du support S_i, chacune associée à un plan du toit, C^1 sur $S_i^{(p)}$. z est l'équation du toit et m représente la mesure de Lebesgue dans $\mathbb{R}^2$.

Notre problème pourrait donc être abordé par des processus de diffusion, notamment par des techniques de "jump-diffusion" adaptées aux problèmes de reconnaissance d'objets multiples. Il faudrait alors voir dans quelles mesures les hypothèses et approximations réalisées pour calculer le gradient de l'énergie pénalisent les résultats. Il serait particulièrement intéressant d'approfondir les recherches dans cette direction.

3.4 Estimation des paramètres

L'étape de reconstruction 3D fait intervenir différents types de paramètres au niveau de la formulation de la densité et, surtout, au niveau de la technique d'optimisation. Dans cette partie, nous présentons les méthodes utilisées pour estimer ces paramètres.

3.4.1 Paramètres de la densité

La densité proposée a été mise en place de manière à minimiser le nombre de paramètres. On distingue deux types de paramètres : les paramètres physiques et les paramètres pondérateurs. Les paramètres physiques, peu nombreux, ne posent pas de problème de réglage puisqu'ils ont un sens physique dans l'application. Par exemple, le paramètre H_e, introduit afin d'exclure les pixels aberrants des données $\mathcal{Y}$, représente la hauteur d'un demi-étage, soit $1,5$ mètre. Il existe également H_{gmin},H_{gmax}, H_{tmin} et H_{tmax} permettant de définir les domaines de définition de H_g et H_t. Ces quatre paramètres physiques ont été fixés respectivement à 2, 30, 2 et 6 mètres. En réglant de la sorte, on réalise l'hypothèse que les bâtiments ne dépassent pas une trentaine de mètres de hauteur. Cela est le cas sur nos zones test.

L'estimation du paramètre pondérateur β de la densité est plus délicate. Il existe différentes méthodes permettant d'estimer ce genre de paramètres, comme les algorithmes EM [Dempster et al., 1977] ou SEM [Celeux et Diebolt, 1986]. Ces algorithmes sont cependant très généraux et ne sont pas spécifiquement adaptés à notre problème d'estimation.
Si l'on cherche à régler β manuellement, la démarche intuitive va consister à rechercher par dichotomie la valeur de β qui maximise la vraisemblance tout en garantissant une qualité de régularisation suffisante pour avoir un rendu visuel correct. Formellement, cela peut s'écrire sous le problème d'optimisation suivant :

$$\begin{aligned} \widehat{\beta} = \quad & \arg\max_{\beta} \mathcal{L}\left(\mathcal{Y}^{(e)} / \widehat{x}_{(\beta)}\right) \\ & sous\, U_p(\widehat{x}_{(\beta)}) + \beta \times N_I \leq \frac{\varepsilon}{N_I} \end{aligned} \tag{3.27}$$

où $\widehat{x}_{(\beta)} = \arg\max_{x} h(x,\beta)$.
$\mathcal{Y}^{(e)}$ est un échantillon représentatif de données dont on connaît N_I, le nombre de paires d'objets voisins qui doivent être assemblés. ε est un paramètre fixé qui permet de définir la qualité minimale de la régularisation exigée. Cette estimation de β revient à introduire un nouveau paramètre ε qui a lui un sens physique puisqu'il correspond, en mètres, à l'écart moyen maximal toléré entre les paramètres des modules assemblables.

En considérant que, sur l'échantillon $\mathcal{Y}^{(e)}$, on connaît parfaitement les N_I cas réels d'assemblage, c'est-à-dire les paires d'objets voisins qui doivent être assemblés, le terme a priori se simplifie et s'écrit sous une somme de fonctions g. Sous cette hypothèse, la vraisemblance et l'a priori de $\widehat{x}_{(\beta)}$ évoluent de manière opposée en fonction de β. Si l'on diminue β, l'a priori a moins de poids dans h. La régularisation est alors de moins bonne qualité. A l'inverse, l'attache aux données est meilleure. L'estimation de β peut être faite itérativement par un principe dichotomique de la manière suivante :

- Prendre un β_0 relativement grand, et qui vérifie la contrainte $U_p(\widehat{x}_{(\beta_0)}) + \beta_0 \times N_I \leq \frac{\varepsilon}{N_I}$.
- A t,
  - tant que $U_p(\widehat{x}_{(\beta_t)}) + \beta_0 \times N_I \leq \frac{\varepsilon}{N_I}$, prendre $\beta_t = \beta_{t-1} - P_t$ avec $P_t = P_0$.
  - ensuite, prendre $\beta_t = \beta_{t-1} - P_t$ si $U_p(\widehat{x}_{(\beta_t)}) + \beta_0 \times N_I \leq \frac{\varepsilon}{N_I}$, et prendre $\beta_t = \beta_{t-1} + P_t$ sinon, avec $P_t = \frac{P_{t-1}}{2}$.
- S'arrêter quand $P_t < P_{limite}$: $\widehat{\beta} = \beta_t$.

P_t correspond au pas d'exploration à l'itération t. Ce processus est assez long puisqu'il faut, à chaque itération, calculer $\widehat{x}_{(\beta_t)}$ par un recuit simulé. Il est donc très important de choisir un échantillon de données $\mathcal{Y}^{(e)}$ de taille relativement petite (quelques bâtiments), mais représentatif de $\mathcal{Y}$. En définitive, cet algorithme automatique d'estimation de β procède d'une manière similaire à un réglage manuel par un opérateur. Il est à noter que cette estimation est effectuée une seule fois pour une gamme donnée de MNE.

3.4.2 Paramètres de l'optimisation

Le processus d'optimisation fait intervenir plusieurs paramètres, certains relatifs au recuit simulé comme la température initiale, les autres à l'échantillonneur MCMC comme par exemple les probabilités q_i de proposition des noyaux Q_i.

D'après [Salamon et al., 2002], le schéma de décroissance du recuit simulé et les paramètres qui en découlent doivent prendre en compte plusieurs éléments : l'énergie à optimiser, son échelle, son nombre de minima locaux ainsi que leur taille. En pratique, les informations relatives aux minima locaux sont difficilement accessibles. On utilise surtout des informations liées à la variation de l'énergie.

- **Coefficient α de décroissance de la température -** Le coefficient de décroissance peut, au cours du recuit, varier et être adapté aux variations d'énergie [Varanelli, 1996; Perrin et al., 2005]. Le gain de temps est cependant relativement faible dans notre cas. Nous préférons utiliser plus classiquement un coefficient de décroissance constant qui dépend du nombre d'objets N dans la scène. En pratique, on prend $\alpha = 0.9999^{\frac{1}{N}}$.

- **Température initiale T_0 -** Elle est estimée en fonction des variations de l'énergie sur un échantillon de configurations aléatoires, à partir des travaux de [White, 1984]. On choisit, en effet, T_0 comme deux fois l'écart-type de U calculé à température infinie sur l'échantillon des configurations[8] :

$$T_0 = 2.\sigma(U_{T=\infty}) = 2.\sqrt{\langle U^2_{T=\infty}\rangle - \langle U_{T=\infty}\rangle^2} \tag{3.28}$$

  où $\langle U \rangle$ est la moyenne empirique de l'énergie calculée sur l'échantillon d'exemples. En pratique, on constate, au début de la décroissance d'énergie, un petit plateau durant lequel les propositions sont très majoritairement acceptées : cela permet de confirmer que cette estimation de la température initiale est efficace.

[8] En pratique, plusieurs centaines de configurations sont nécessaires pour une estimation robuste, ce qui est, en temps de calcul, négligeable par rapport au processus d'optimisation global.

- **Température finale T_{final} -** Il existe principalement deux moyens de déterminer la température finale, c'est-à-dire la fin du processus. Le premier, proposé par [White, 1984], consiste à choisir T_{final} de l'ordre de la plus petite échelle d'énergie (ce qui sous-entend avoir un problème discret), c'est-à-dire à prendre la plus petite variation d'énergie d'un mouvement simple. La seconde solution, plus usuelle, est celle que nous retenons : l'algorithme est arrêté quand l'énergie n'a pas évolué durant un certain nombre d'itérations. En pratique, on teste si le taux d'acceptation des propositions calculé sur 1000 itérations est nul.

La majorité des paramètres intervenant dans l'échantillonneur MCMC, notamment les distributions des différents paramètres, ont été explicités dans la partie 3.3.2. Il reste cependant à déterminer les probabilités q_i de proposition des noyaux Q_i qui doivent être en corrélation avec le schéma de décroissance de l'énergie. En effet, durant la phase d'exploration du processus, le noyau dirigé par les données Q_2 doit être majoritairement proposé. Les trois noyaux sont alors choisis selon les probabilités $q_2 = 3q_1 = 3q_3 = 0,6$. Lorsque l'on entre dans la phase d'ajustement minutieux des objets[9], c'est le noyau de régularisation Q_3 qui doit être proposé en priorité : on choisit alors $q_3 = 3q_1 = 3q_2$.

[9]En pratique, le changement de phase est détecté quand le taux d'acceptation des propositions calculé sur 1000 itérations devient inférieur à $0,05$.

Chapitre 4

Evaluation

Dans ce chapitre, nous discutons de la pertinence de l'approche proposée en analysant les résultats obtenus. Nous abordons également l'influence des techniques d'optimisation sur les résultats et évoquons les limites de notre méthode. Les données utilisées sont des MNE de type PLEIADES dont les caractéristiques sont détaillées dans le chapitre 1. La vérité terrain planimétrique est fournie par l'Institut Géographique National (IGN), et correspond aux plans cadastraux nationaux. La vérité terrain 3D, présentée sous forme d'images RASTER, est également fournie par l'IGN. Elle a été réalisée par stéréorestitution de points 3D par des opérateurs. Sur certains résultats présentés, des textures génériques ont été appliquées afin d'améliorer le rendu visuel.

4.1 Evaluation qualitative

Dans cette partie, nous évaluons qualitativement les résultats du processus d'extraction des supports 2D (Chapitre 2) et du processus de reconstruction 3D (Chapitre 3). Plusieurs jeux de résultats sont illustrés. Les figures 4.1 et 4.2 présentent l'extraction et la reconstruction de divers bâtiments (possédant des formes de toit variées, des discontinuités de hauteur de toiture, des structures ouvertes et fermées, et des cas de jonctions complexes). Les figures 4.3, 4.4 et 4.5 montrent les résultats obtenus sur trois zones urbaines denses de la ville d'Amiens.

Extraction des supports 2D

Les résultats fournis par le processus d'extraction des supports 2D sont, de manière générale, satisfaisants. Plusieurs points intéressants peuvent être dégagés :

- **Précision des contours -** Le processus restitue les contours des bâtiments avec précision. Cela est principalement dû à la méthode d'extraction de rectangles de [Ortner, 2004] dont le critère d'attache aux données est fondé sur les discontinuités de façades, critère permettant d'obtenir de meilleures performances que par des considérations surfaciques avec les MNE utilisés. Certains polygones sont toutefois mal positionnés, ce qui est le cas notamment du support de gauche du bâtiment B3 de la figure 4.2. Cela engendre des décalages de 1 à 2 mètres dans le positionnement des façades, qui s'expliquent par la technique d'optimisation : en pratique, celle-ci permet rarement d'atteindre en des temps raisonnables l'optimum global, mais une solution approchée.

- **Détection des discontinuités des hauteurs de toit -** La détection des discontinuités des hauteurs de toit est également convaincante. Les discontinuités présentes sont quasiment toutes correctement situées, avec très peu de sous-détection. Leur localisation est relativement précise malgré la simplicité du processus mis en place comme nous pouvons le voir sur la figure 4.2 - bâtiment B3. Un processus de localisation fondé sur une méthode plus complexe qu'un simple barycentre, pourrait permettre néanmoins de gagner un peu plus en précision.

- **Connexion des supports 2D -** Le processus de transformation des rectangles en quadrilatères connectés donne satisfaction comme on peut le constater sur les figures 4.3, 4.4 et 4.5. La fonction de coût utilisée (équation 2.6) permet d'obtenir une bonne régularisation des supports. Cela est notamment dû au coût de référence C_{ref} qui permet d'éviter la présence de mauvaises configurations de transformation. La robustesse apportée par ce coût a une contre partie : il n'autorise pas la transformation de certains rectangles générant des jonctions complexes dans certains cas, comme notamment pour le bâtiment B4 de la figure 4.2 où la double jonction en T n'a pas été restituée.

Le processus d'extraction des supports 2D présente néanmoins des défauts relativement pénalisant pour des applications liées à l'aménagement urbain par exemple. Il s'agit de la sur-détection de végétation et la sous-détection de petites structures urbaines.

- **Sur-détection de végétation -** Les arbres présents dans les zones urbaines sont détectés comme étant des bâtiments par le processus d'extraction de rectangles. On peut le constater sur la figure 4.4 (en haut de l'image) où deux rangées d'arbres sont extraites. Cette sur-détection est gênante pour de nombreuses applications. Cependant, il est envisageable de coupler le processus d'extraction d'emprises avec un masque de végétation urbain. Etant donné le masque, ce procédé serait simple à mettre en œuvre puisqu'il suffirait de soustraire les zones de végétation au MNE pour que le processus d'extraction n'ait plus ces sur-détections.

- **Sous-détection de petites structures -** La majorité des petites structures urbaines situées dans les cours intérieures des îlots de bâtiments ne sont pas détectées comme on peut le voir sur les figures 4.3, 4.4 et 4.5. Ces petites structures correspondent à des bâtiments d'un étage, souvent des garages. Le processus d'extraction de rectangles, fondé sur la détection des discontinuités de façades, rencontre des difficultés pour faire la distinction entre le sol et ces structures de faible hauteur. Une solution consisterait à détecter ces petites structures dans les cours intérieures a posteriori, en utilisant par exemple un algorithme de segmentation par ligne de partage des eaux ("watershed").

- **Emprises non rectilignes -** Le processus proposé a été mis en place pour traiter des formes rectilignes, formes "naturelles" dans le milieu urbain. Cependant, certaines structures, souvent des ouvrages remarquables à l'image du bâtiment B6 de la figure 4.2, possèdent des emprises courbées. Le résultat est, dans ce cas, particulièrement mauvais comme on pouvait s'y attendre. Il n'y a pas de réelle solution permettant de faire face à ce défaut.

Reconstruction 3D

Les reconstructions 3D obtenues dépendent de la qualité de l'extraction des supports 2D. Les défauts de cette première étape, soulignés précédemment, sont donc visibles sur les résultats 3D. Cela n'empêche pas d'obtenir des modélisations 3D de bâtiments globalement satisfaisantes.

- **Niveau de généralisation des bâtiments -** Les plans principaux des bâtiments sont correctement restitués, ce qui constituait notre objectif initial. Cela est suffisant pour une majorité des bâtiments. Certaines structures, comme l'hôtel de ville d'Amiens (voir le bâtiment B1 de la figure 4.1), possèdent cependant des décrochements de toits et des avancées de façades. Ces "détails" ne sont pas restitués par notre processus. La bibliothèque de modèles est pourtant assez riche pour modéliser la plupart de ces éléments : un partitionnement plus pertinent des supports 2D pourrait être une solution pour obtenir une modélisation plus détaillée. Le niveau de généralisation des bâtiments reste tout de même acceptable compte tenu de la qualité des données.

- **Restitution des différentes formes de toit -** De manière générale, les différentes formes de toit sont correctement identifiées comme on peut le constater sur la figure 4.1. Bien que relativement simple, la distance introduite dans le terme d'attache aux données permet une bonne distinction des différentes formes contenues dans la bibliothèque. Par ailleurs, quand l'identification de la forme de toit d'un module n'est pas évidente, le terme a priori apporte une aide à la décision en prenant en compte la forme de toit des modules voisins : c'est un des principaux avantages de la formulation bayésienne. Il existe cependant des confusions entre les formes de toit proches, particulièrement avec le toit mansart. Lorsque le paramètre de largeur de plateau κ de ce dernier est proche de 0 (respectivement de 1), la distinction avec le bi-plan symétrique (respectivement avec le toit plat) sera délicate. Ce type de confusions n'est pas pénalisant du point de vue de la fidélité altimétrique des modélisations. Du point de vue de la qualité visuelle, cela peut être plus gênant pour certaines applications. Une solution simple consisterait à restreindre le domaine de définition de certains paramètres de formes de toit. Cela nécessiterait néanmoins de définir des paramètres supplémentaires (des marges dans les domaines de définition) et complexifierait la résolution de la méthode d'optimisation détaillée dans la partie 3.3.

- **Qualité de l'assemblage des objets -** Dans les problèmes de reconstruction 3D, le rendu visuel est un des critères majeurs permettant d'évaluer les résultats. Cet aspect visuel dépend majoritairement, dans notre cas, de la qualité d'assemblage des objets, et plus précisément de la pertinence du terme a priori développé dans la partie 3.2.2. De ce point de vue, les bâtiments reconstruits sur la figure 4.1 sont très satisfaisants. Ce point est souligné sur la figure 4.2 - bâtiment B2. Comparativement au modèle présenté dans l'annexe A, il y a très peu d'artefacts, et les assemblages des modules sont dans leur configuration optimale quand cela a lieu d'être. Quand des ruptures de toits sont présentes (comme sur le bâtiment B2 ou le bâtiment B3), les modules ne sont pas, à juste titre, assemblés. L'a priori proposé est donc efficace.

- **Jonctions complexes -** Les cas de jonctions complexes sont souvent mal restitués, comme on peut le voir sur la figure 4.2 à travers la double jonction en T du bâtiment B4. Ce défaut est inhérent à la mauvaise connexion des quadrilatères dans certaines situations, lors de la phase d'extraction des emprises. La modélisation des bâtiments concernés reste cependant acceptable, le principal inconvénient provenant de l'absence de petites parties d'arêtes faîtières au niveau des jonctions concernées.

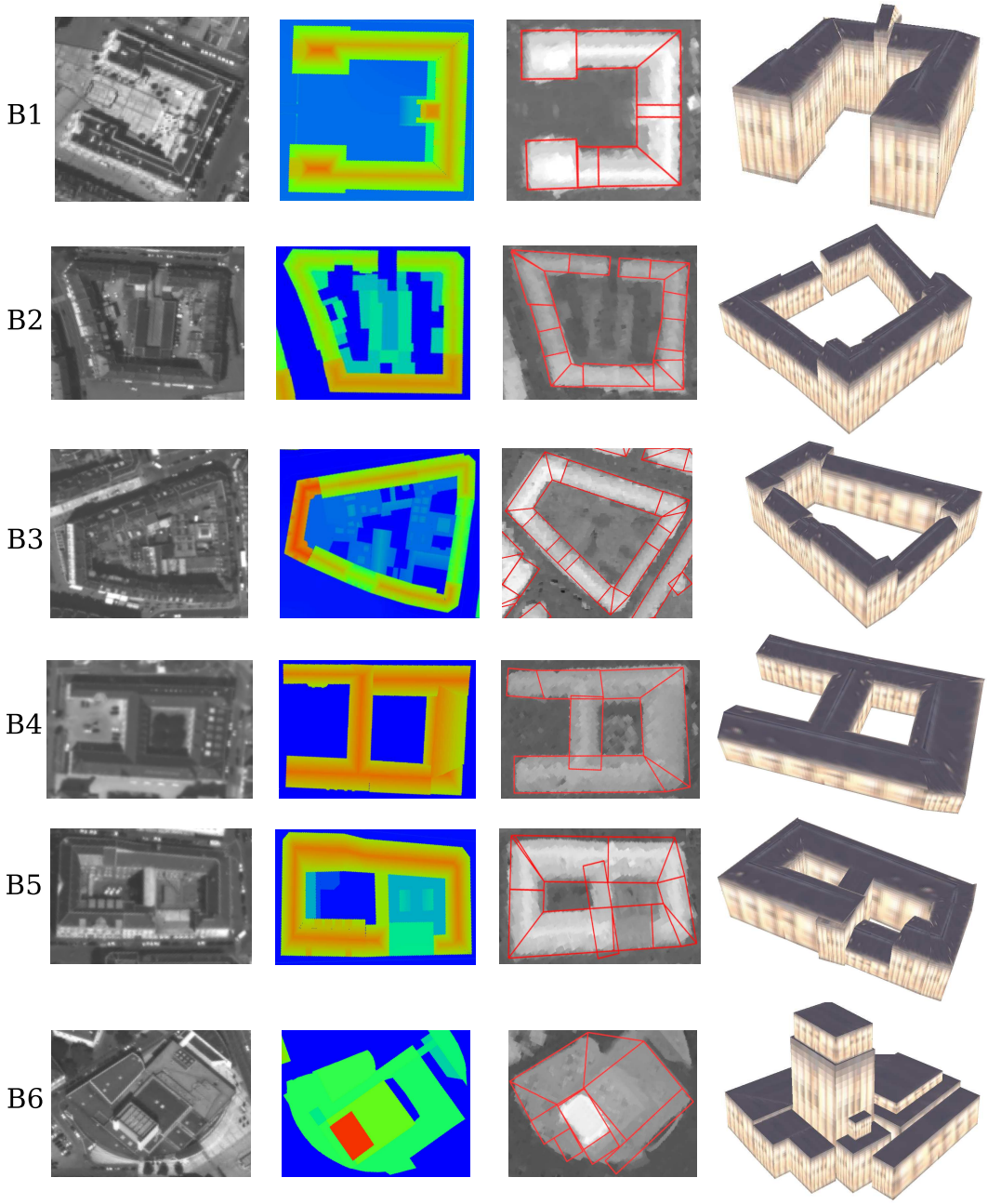

FIG. 4.1 – Différents exemples de bâtiments reconstruits *(4^{ieme} colonne)* associés avec les simulations PLEIADES *(1^{ere} colonne)*, la vérité terrain *(2^{ieme} colonne)* et les supports 2D extraits *(3^{ieme} colonne)*.

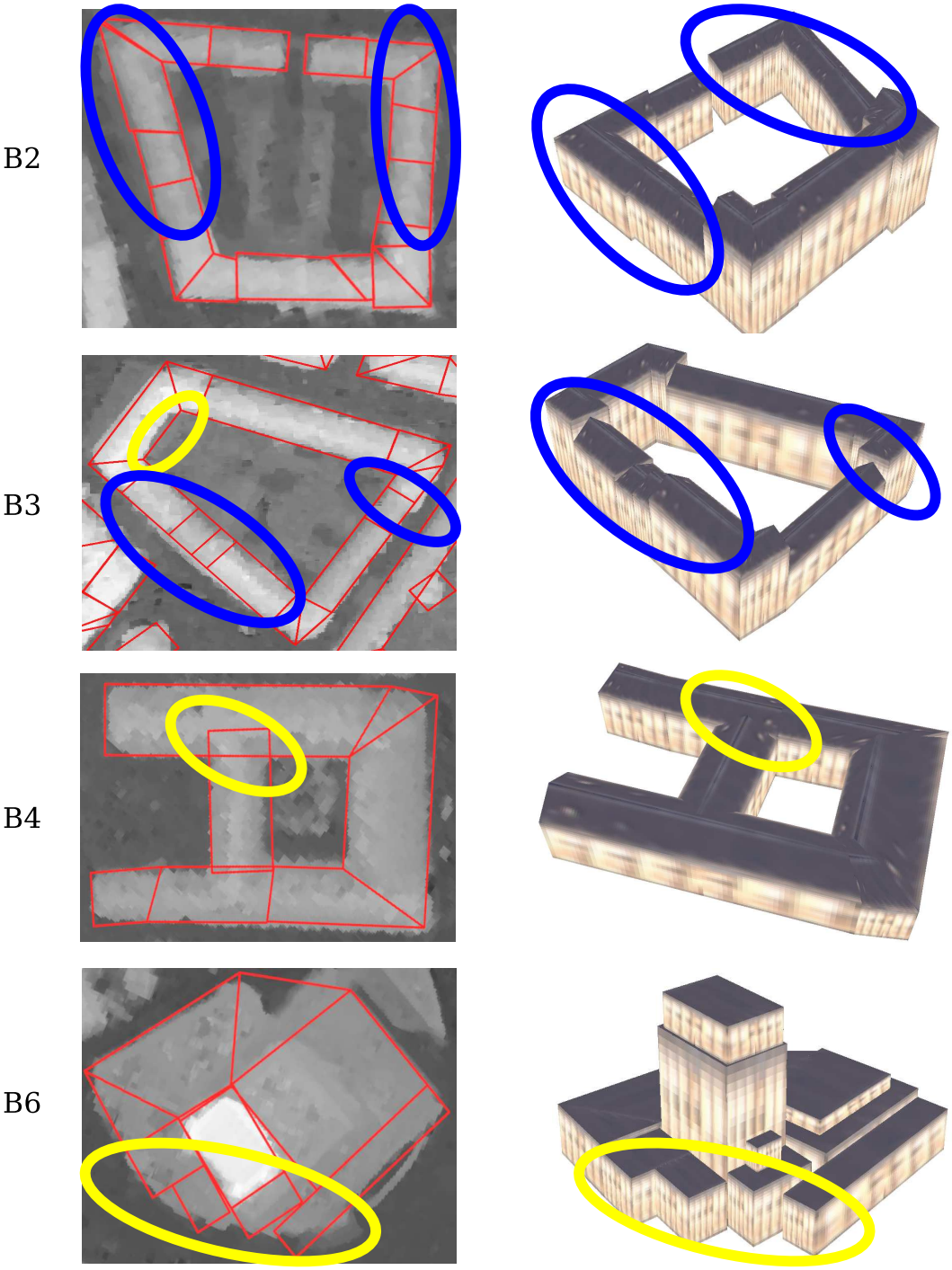

FIG. 4.2 – Extraits des bâtiments présentés sur la figure 4.1 soulignant certains défauts (en jaune) et points positifs (en bleu) - bâtiments reconstruits *(2^{ieme} colonne)*, supports 2D extraits *(1^{iere} colonne)*.

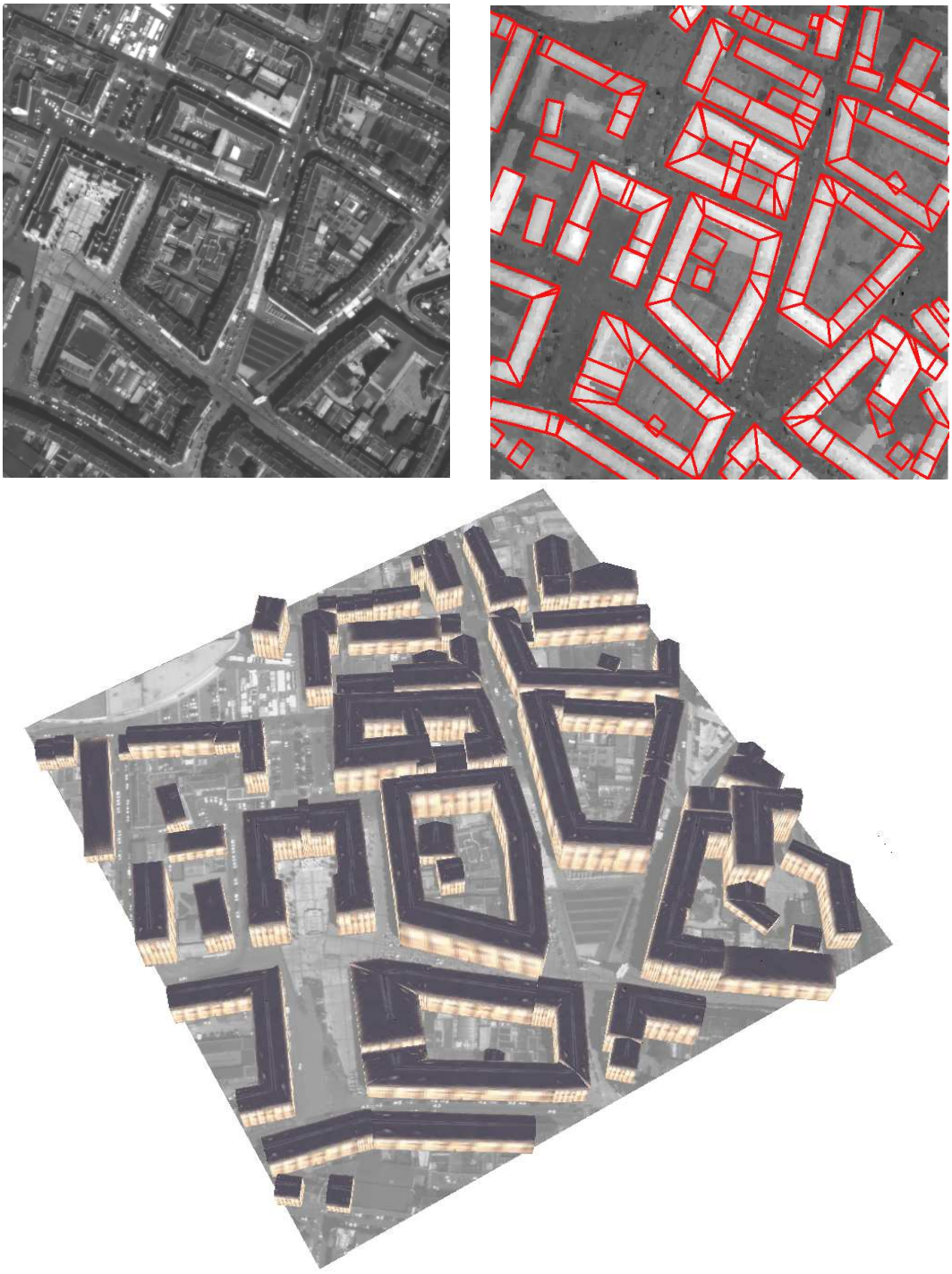

FIG. 4.3 – Reconstruction 3D d'un extrait d'Amiens - zone A - *(bas)* associée avec une simulation PLEIADES et l'extraction des supports 2D *(haut)*.

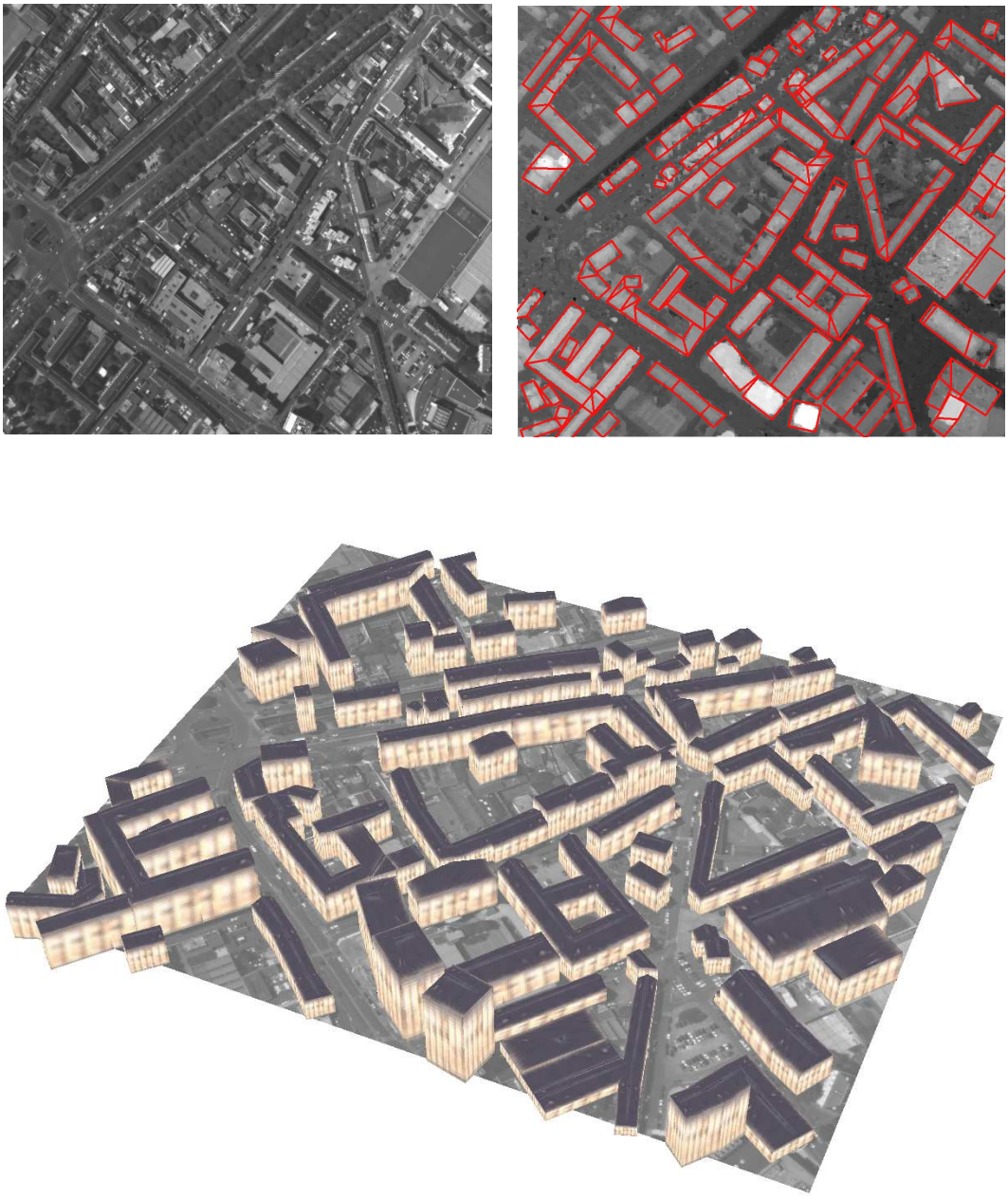

FIG. 4.4 – Reconstruction 3D d'un extrait d'Amiens - zone B - *(bas)* associée avec une simulation PLEIADES et l'extraction des supports 2D *(haut)*.

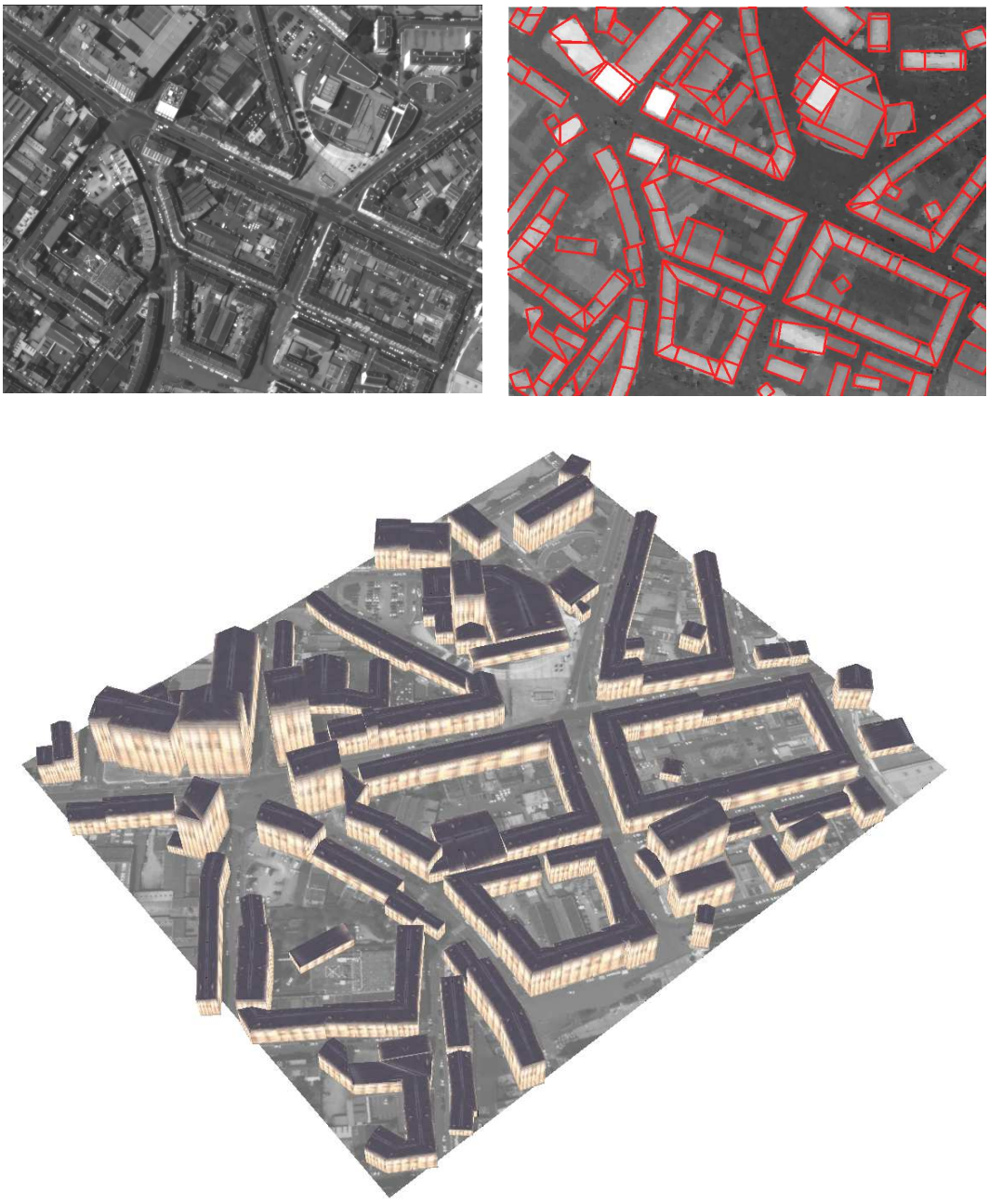

FIG. 4.5 – Reconstruction 3D d'un extrait d'Amiens - zone C - *(bas)* associée avec une simulation PLEIADES et l'extraction des supports 2D *(haut)*.

4.2 Evaluation quantitative

La fidélité à la réalité des résultats est un point tout aussi important que la qualité du rendu visuel. Dans les applications techniques, cela constitue même le critère majeur permettant d'évaluer un processus de reconstruction 3D. Les résultats présentés dans la partie précédente sont analysés quantitativement en planimétrie et en altimétrie.

Qualité planimétrique

Les erreurs planimétriques des zones A, B et C (figures 4.3, 4.4 et 4.5) sont détaillées dans le tableau 4.1 et illustrées sur la figure 4.6. Le taux général de sur-détection (en terme de surface) est de 9.7%. Cette valeur est principalement due au problème de végétation souligné dans la partie précédente. On constate également à certains endroits des sur-détections à proximité des contours des emprises. Elles correspondent à des décalages de 1 à 2 mètres du positionnement des façades dont les causes ont été expliquées précédemment. Le taux de sous-détection est relativement élevé : 15.3%. Cependant, si l'on exclut les sous-détections des petites structures d'un étage situées dans les cours intérieures, ce taux chute à 4.5%, ce qui constitue une valeur tout à fait satisfaisante.

D'une manière générale, l'évaluation planimétrique fournit des résultats corrects qui pourraient être très nettement améliorés en résolvant les problèmes de sous-détection des petites structures et de sur-détection de la végétation. Les temps de calcul mis pour extraire les supports 2D sont de l'ordre de 3 heures pour une zone d'environ $0,5\ km^2$ (soit environ 500 objets) avec un processeur 3 GHz. Cette durée correspond, très majoritairement, aux temps de calcul du processus d'extraction des rectangles.

	sur-détection	sous-détection	
		petites structures d'un étage	le reste
Zone A	7.9%	14.3%	2.9%
Zone B	11.6%	8.1%	5.4%
Zone C	9%	11.3%	4.6%

TAB. 4.1 – Erreurs planimétriques des zones A, B et C.

Qualité altimétrique

L'évaluation altimétrique fournit des résultats globalement satisfaisants puisque l'erreur quadratique moyenne en Z entre les reconstructions et la vérité terrain (calculée sur les emprises communes des bâtiments) est de $2,3$ mètres, c'est-à-dire une valeur légèrement inférieure à la hauteur d'un étage. Cette erreur a été nettement améliorée par rapport à celle obtenue à partir du processus proposé en annexe A (c'est-à-dire $3,2$ mètres).

Cette erreur quadratique moyenne en Z reste toutefois élevée. Cela s'explique par des erreurs locales importantes (i.e. supérieures à 3 mètres) à certains endroits qui sont dues

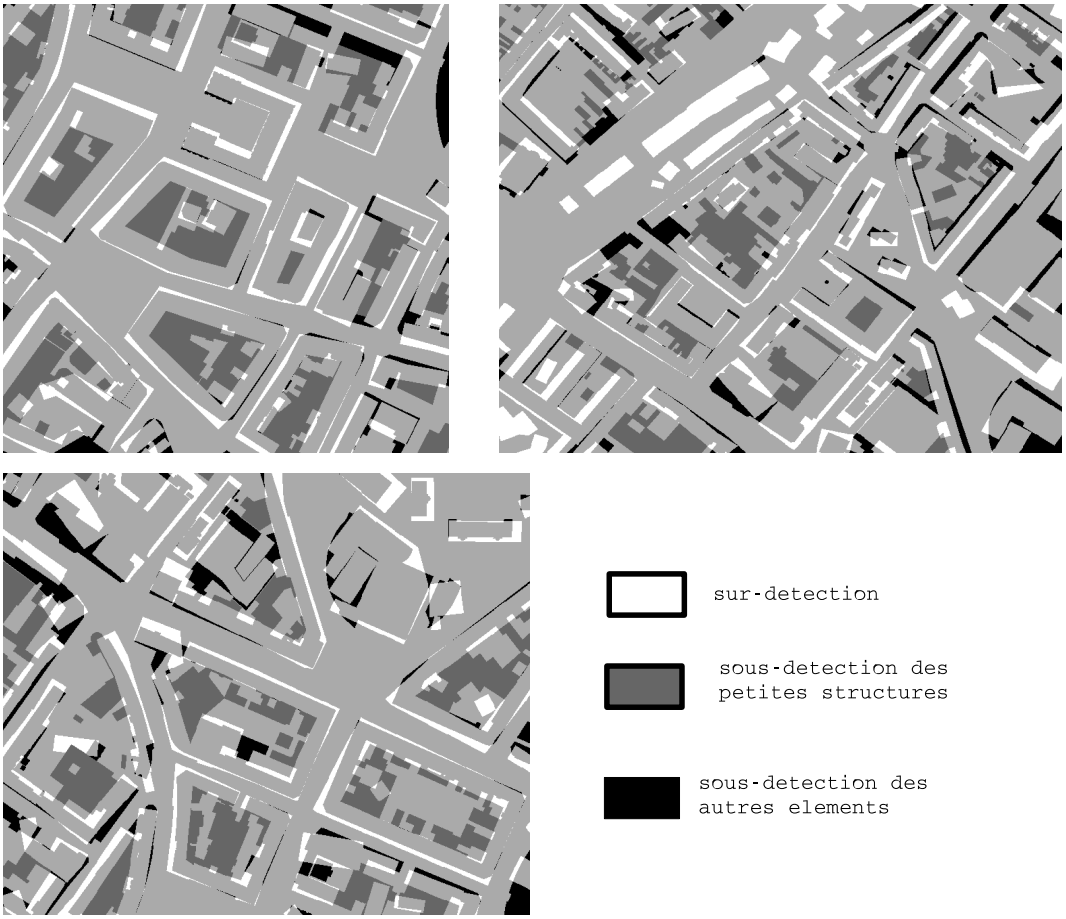

FIG. 4.6 – Carte d'erreur planimétrique des zones A, B et C.

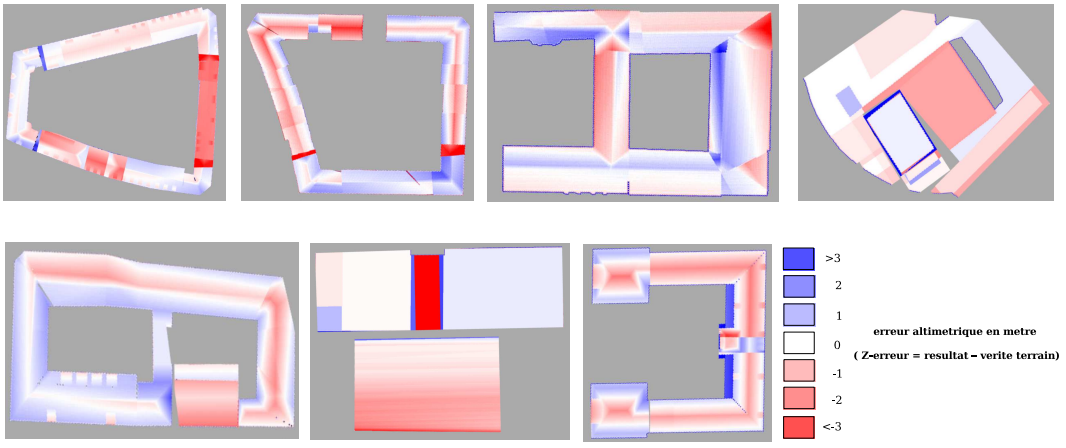

FIG. 4.7 – Carte d'erreur altimétrique des bâtiments reconstruits sur la figure 4.1.

à plusieurs raisons : un mauvais positionnement des façades entraînant des erreurs altimétriques de plusieurs étages, des inexactitudes dans le MNE principalement dues à des corrélations de surfaces non lambertiennes (transparence des verrières, spécularité sur les ardoises,...), et des détails de toitures non restitués comme des avancées de façades, des décrochements de toit ou des superstructures.

La figure 4.7 présente, de manière détaillée, la carte d'erreur en Z des bâtiments reconstruits sur la figure 4.1. Plusieurs points attirent l'attention. Tout d'abord, il est difficile d'obtenir une estimation précise des pentes de toits comme on peut le constater sur certaines parties de toitures composées à la fois de pixels bleu et rouges. Ce manque de précision est dû au bruit présent dans le MNE, mais également aux imprécisions générées par le positionnement des supports 2D. Ensuite, on peut voir que certaines discontinuités de hauteurs de toit ne sont pas très bien localisées (voir les petites zones rouges et bleu foncé). Cela pourrait être amélioré en reconsidérant le processus d'accumulation des gradients fondé sur une opération barycentrique dans la partie 2.2.2. Enfin, de nombreuses petites erreurs peuvent être localisées sur les plans de toits qui pénalisent localement l'erreur quadratique moyenne : il s'agit de superstructures, principalement des cheminées et chien-assis.

Les temps de calcul mis par le processus de reconstruction 3D sont acceptables malgré l'utilisation de techniques d'optimisation stochastiques. Par exemple, moins d'une minute est nécessaire pour obtenir les bâtiments présentés sur la figure 4.1. Les zones reconstruites sur les figures 4.3, 4.4 et 4.5 sont, elles, réalisées en une vingtaine de minutes environ.

4.3 Influence du recuit simulé

L'algorithme d'optimisation mis en place dans la partie 3.3 joue un rôle fondamental sur la qualité des résultats et les temps de calcul. Nous analysons, dans cette partie, la pertinence des trois noyaux de propositions établis dans la partie 3.3.2, et particulièrement le noyau dirigé par les données Q_2 et le noyau de régularisation Q_3. Nous discutons également de l'apport d'une discrétisation de l'espace d'état.

Efficacité des noyaux de propositions

Les trois noyaux de propositions mis en place dans la partie 3.3.2 ont chacun un rôle précis. Nous analysons ici leur pertinence et leur efficacité au sein du processus d'optimisation.

Le noyau uniforme Q_1 est le noyau de référence. Il permet d'atteindre la configuration optimale sous des temps de calculs importants. La figure 4.10 illustre l'évolution des configurations d'objets durant le recuit simulé en utilisant le noyau uniforme. Les variations altimétriques des objets sont importantes, ce qui prouve que ce noyau n'est pas particulièrement pertinent pour notre problème.
Les courbes illustrées sur la figure 4.8-(*haut*) correspondent à deux décroissances énergétiques, l'une utilisant uniquement le noyau uniforme, l'autre utilisant la combinaison des trois noyaux proposée dans la partie 3.3.3. Ce graphe permet de voir la pertinence de notre schéma de décroissance : malgré une décroissance plus rapide ($\alpha = 0,9997$), la configuration finale de notre schéma de décroissance a une énergie meilleure ($U = 971$) que le schéma

uniforme ($\alpha = 0,9999$ et $U = 986$). Le gain de temps est, par conséquent, très intéressant. Il est difficile de le définir avec exactitude puisque, en pratique, aucune configuration finale n'est identique à une autre avec des schémas de décroissance différents. On peut cependant affirmer que le gain de temps est au moins d'un facteur 3.

Le graphe sur la figure 4.8-(*bas*) montre la décroissance énergétique d'un schéma utilisant le noyau Q_2 durant la phase d'exploration du recuit et le noyau Q_3 durant la phase de régularisation. La température initiale est très basse, car estimée avec le noyau Q_2 (voir le principe de l'estimation de la température initiale dans la partie 3.4.2). Durant la phase d'exploration, l'énergie de la configuration courante varie peu, celle-ci étant déjà initialement bonne. Quand on entre dans la phase de régularisation, on constate une légère descente énergétique qui amène à la convergence. Cette descente est provoquée par l'utilisation du noyau Q_3 qui ajuste et régularise les paramètres des objets de la configuration. En pratique, l'utilisation d'un tel schéma n'est pas recommandée : on se trouve assez régulièrement piégé dans des minima locaux. Il est important d'inclure dans le schéma de décroissance un noyau uniforme qui explore aléatoirement l'espace d'état, et de ne pas utiliser que des noyaux focalisant leurs propositions sur des configurations a priori qui peuvent parfois être mauvaises. L'utilisation du noyau uniforme est donc indispensable pour faire face aux éventuelles erreurs d'estimation des paramètres intervenant dans les noyaux Q_2 et Q_3.

La figure 4.9 présente deux graphes, le premier illustrant un schéma de décroissance fondé uniquement sur l'utilisation du noyau Q_2, et le second, sur l'utilisation du noyau de régularisation Q_3. Comme nous l'avons souligné dans le paragraphe précédent, l'énergie évolue peu en utilisant le noyau d'exploration Q_2. La convergence est relativement lente (même si elle est plus rapide que pour un noyau uniforme). La solution obtenue, bien que souvent moyennement régularisée, a cependant l'avantage d'être fidèle aux données, c'est-à-dire au MNE. Le noyau Q_3 (voir figure 4.9-(*bas*)) fonctionne de manière inverse. La solution est bien régularisée, mais peut être altimétriquement éloignée de la réalité. Le schéma de décroissance illustre ces propos : la convergence rapide et brutale caractérise une bonne qualité de régularisation. Par contre, l'énergie de la solution peut être très mauvaise : le résultat est quantitativement médiocre. L'utilisation de ce noyau est donc particulièrement dangereuse car elle force la convergence vers une configuration régularisée de la configuration courante. Il doit être utilisé en fin de processus, à proximité de la solution globale.

Discrétisation de l'espace d'état ?

La méthode de reconstruction en 3D a été présentée en utilisant un espace d'état dont les variables sont, pour la plupart, à valeurs continues sur des compacts de $\mathbb{R}$. Il n'est cependant pas exclu de discrétiser cet espace d'état afin notamment de réduire les temps de calcul. Concrètement, cela revient à introduire un pas de discrétisation (en mètre). Plus le pas sera important, moins la précision de la modélisation sera grande, et plus les temps de calcul seront faibles. Fixer un pas de discrétisation revient donc à jouer sur le compromis temps de calcul / précision de la modélisation.

La discrétisation de l'espace d'état présente également un autre avantage : la qualité de la régularisation. Celle-ci est un critère important dans notre approche structurelle puis-

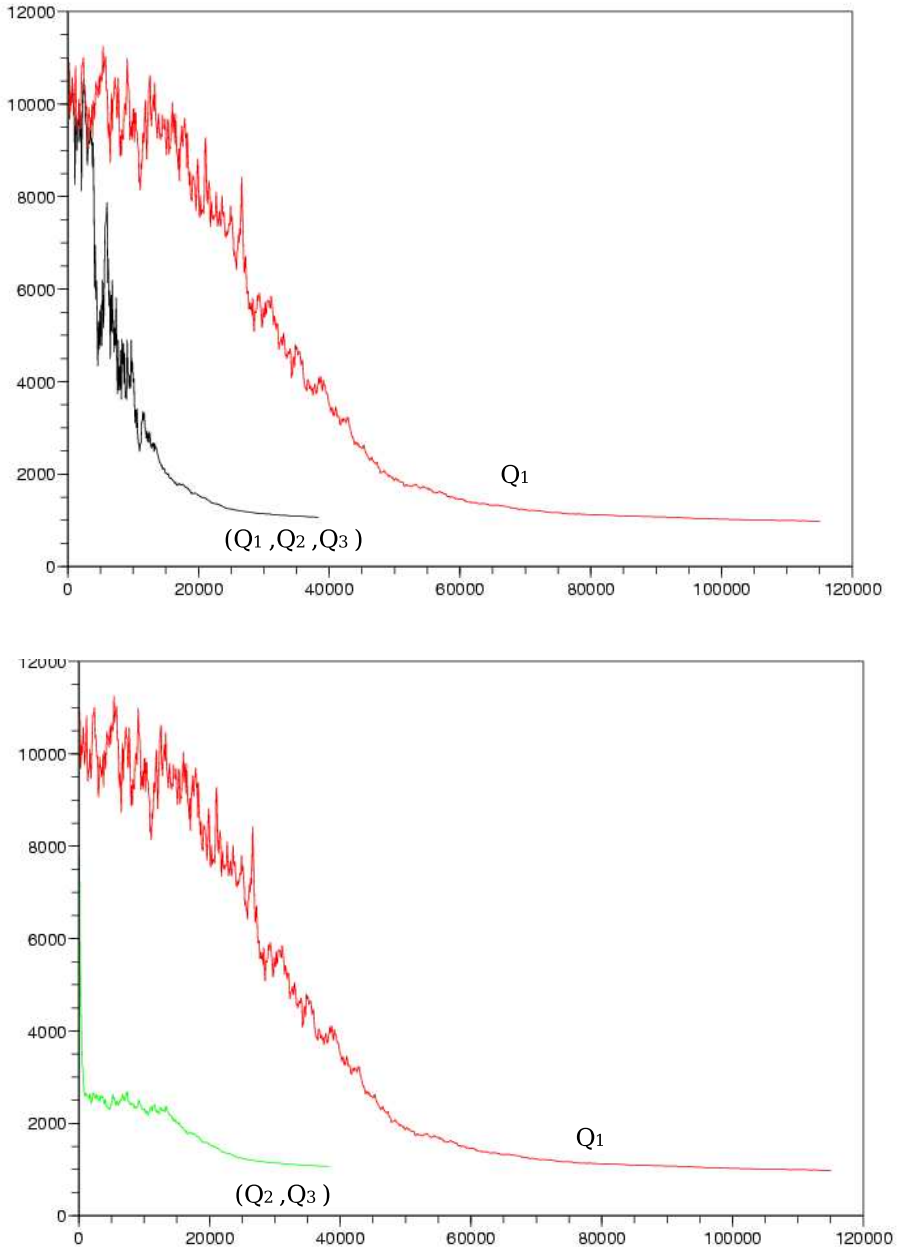

FIG. 4.8 – Décroissance énergétique $x = t$, $y = U$ - avec le noyau Q_1 en rouge, avec la combinaison (Q_1, Q_2, Q_3) en noir *(haut)*, avec la combinaison (Q_2, Q_3) en vert *(bas)*.

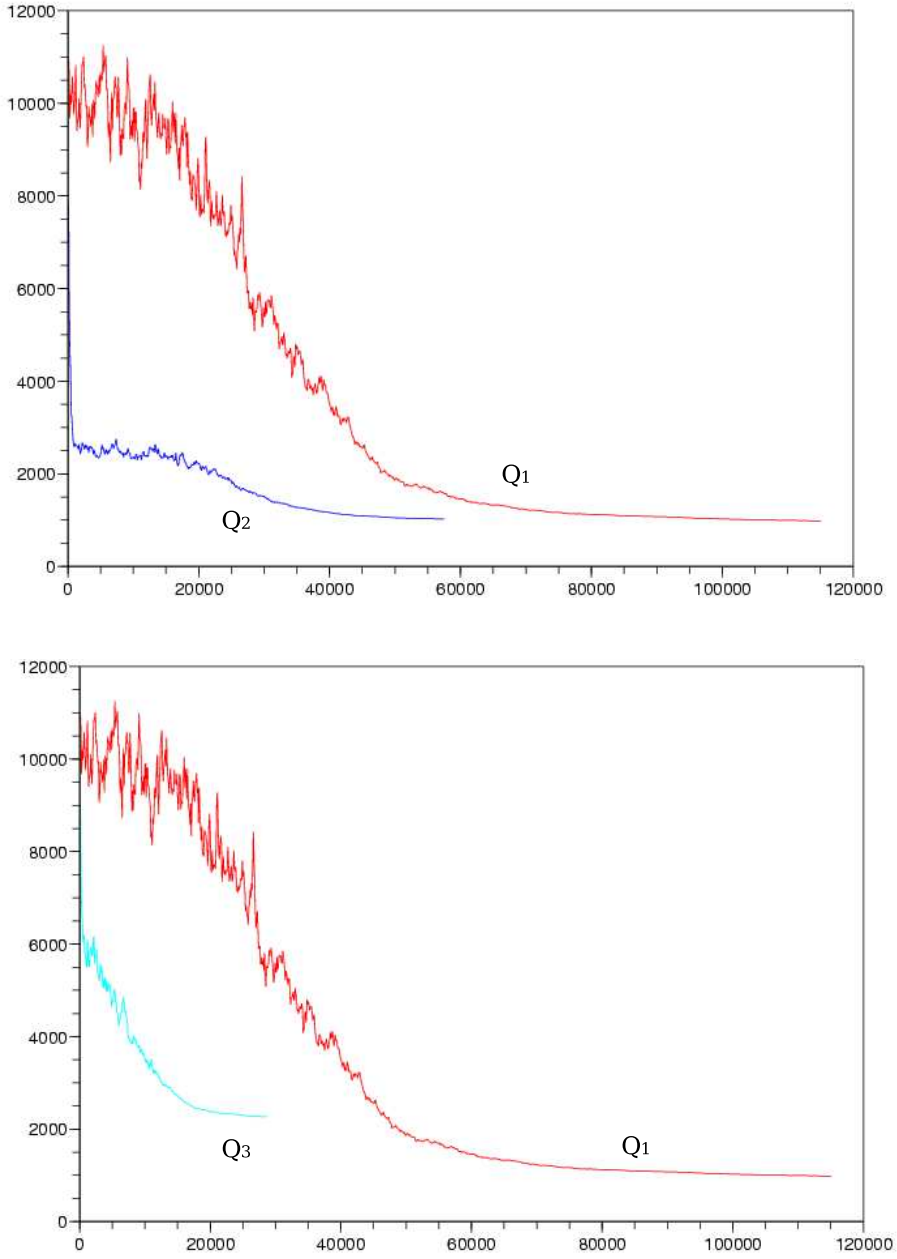

FIG. 4.9 – Décroissance énergétique $x = t$, $y = U$ - avec le noyau Q_1 en rouge, avec le noyau Q_2 en bleu *(haut)*, avec le noyau Q_3 en mauve *(bas)*.

FIG. 4.10 – Evolution de la configuration d'objets au cours d'un schéma de décroissance utilisant le noyau uniforme (à lire de gauche à droite, et de haut en bas).

qu'elle permet le bon assemblage des modules 3D entre eux. Avec un espace discrétisé, les paramètres d'objets voisins peuvent prendre la même valeur, et l'assemblage peut ainsi être optimal. Quand l'espace d'état n'est pas discrétisé, il est très difficile d'obtenir, sous des temps de calcul raisonnables, un assemblage parfait sans la présence de petits décalages d'ajustement des paramètres. Cet inconvénient a toutefois des conséquences mineures sur la qualité du rendu visuel, surtout quand les facettes de toits et de façades sont texturées.

Au vu de la résolution planimétrique et altimétrique de nos données, la discrétisation de l'espace d'état est une opération tout à fait intéressante. En effet, devant la précision décimétrique des MNE, il n'est pas particulièrement pertinent d'exiger une précision absolue pour la modélisation 3D. Le choix du pas de discrétisation dépend des contraintes d'utilisations. Un pas pris à $20cm$ apparaît comme un bon compromis entre la précision et les temps de calcul.

4.4 Influence des supports 2D

Les modélisations 3D obtenues par l'approche proposée dépendent fortement de la qualité de l'extraction des supports 2D. Il est intéressant de comprendre l'influence des supports 2D sur la qualité des reconstructions. Les résultats présentés jusqu'alors montrent des reconstructions globales de bâtiments, c'est-à-dire en modélisant uniquement les plans principaux des toitures. Cependant, notre bibliothèque de modèles est assez riche pour permettre de modéliser certains détails comme par exemple des décrochements de toits ou des avancées de façades. En d'autres termes, en considérant des supports 2D plus précis, est-il possible d'obtenir des modélisations 3D plus détaillées ?

La figure 4.11 permet d'apporter une réponse. Sur celle-ci, on compare le résultat 3D obtenu par le processus de modélisation 3D intégral avec celui obtenu uniquement par la méthode de reconstruction 3D où les supports 2D ont été extraits par un opérateur (et qui sont, par conséquent, très détaillés). La différence de qualité est très nette. En ayant pris soin de partitionner correctement les supports 2D, on arrive à obtenir un niveau de détails comparable à celui proposé par la plupart des méthodes dites génériques. Sur l'hôtel de ville d'Amiens, on restitue notamment les deux avancées de façades présentent dans la cour intérieure ainsi que le décrochement de toit central. La figure 4.12 montre trois autres résultats convaincants obtenus avec une extraction manuelle des supports 2D. On remarque notamment qu'il est possible d'obtenir de bons résultats sur les bâtiments dont les emprises sont courbées. Il suffit pour cela de partitionner assez finement les supports 2D en approximant la courbe par une ligne brisée.

En définitive, la méthode de reconstruction 3D ne permet pas uniquement d'obtenir des modélisations globales. Elle peut fournir des modélisations avec un très bon niveau de détails, comparable aux modélisations génériques. Cela est conditionné par la qualité des supports 2D. Dans l'état actuel de nos recherches, la méthode d'extraction des supports ne permet pas d'atteindre un niveau de description détaillé. C'est là une des principales limites de notre approche.

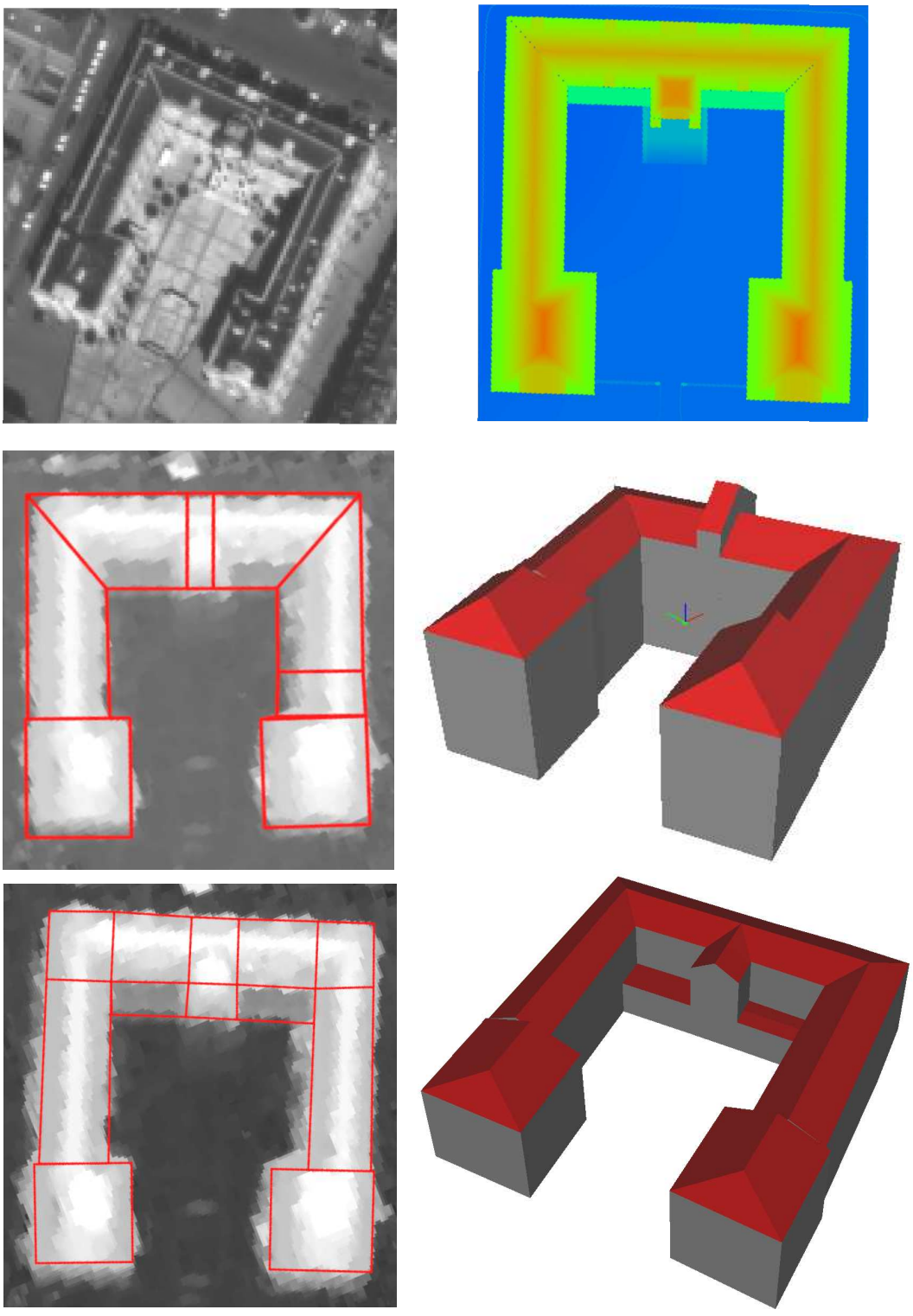

FIG. 4.11 – Influence des supports 2D sur la reconstruction - exemple avec l'hôtel de ville d'Amiens - Simulation PLEIADES et vérité terrain 3D *(haut)*, supports 2D automatiques et reconstruction associée *(centre)*, supports 2D manuels et reconstruction associée *(bas)*

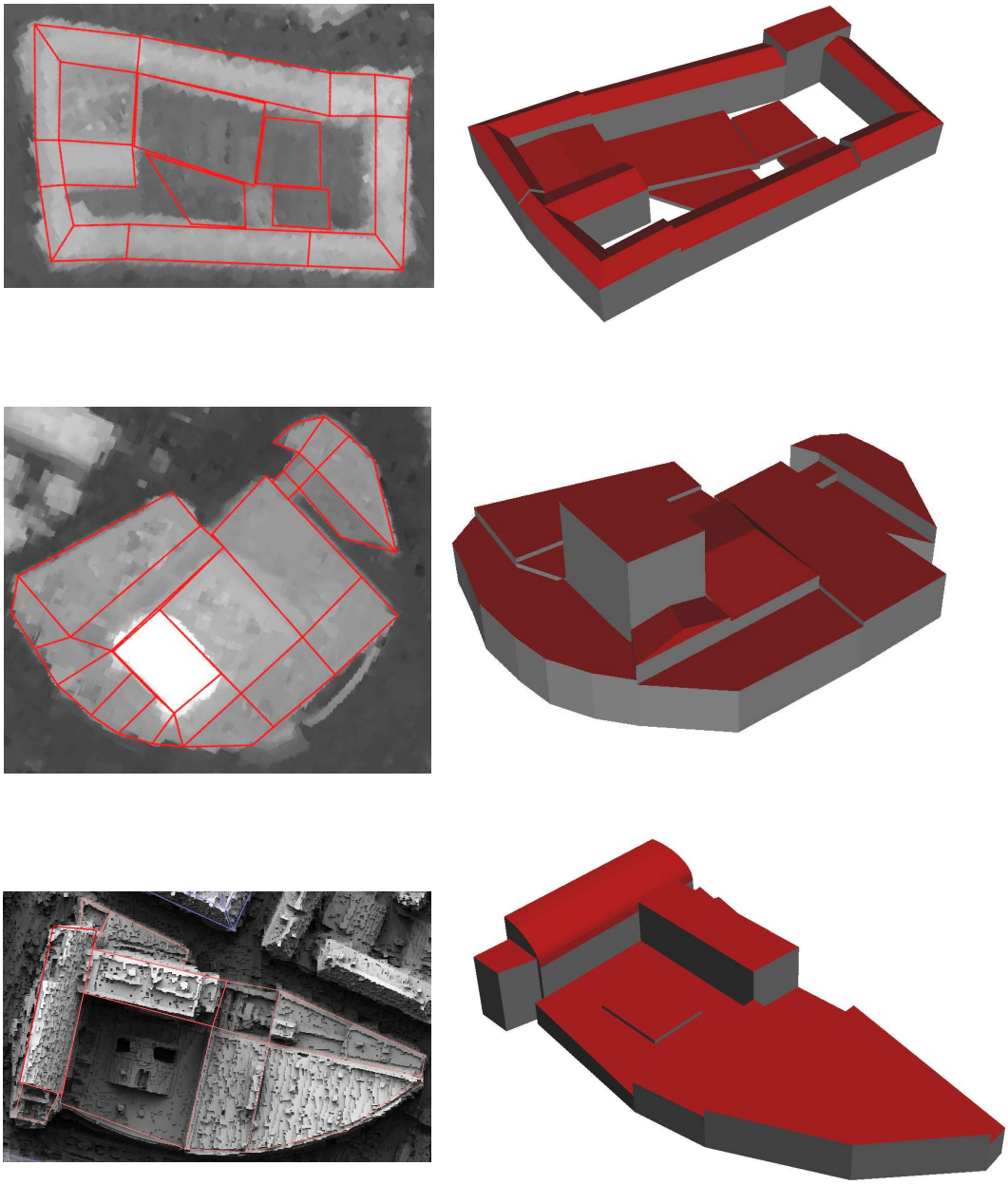

FIG. 4.12 – Trois exemples de reconstructions obtenues avec une saisie manuelle des supports 2D

4.5 Les limites d'utilisation

Dans cette partie, nous détaillons les principales limites d'utilisation de notre approche en spécifiant les situations dans lesquelles notre méthode ne permet pas d'obtenir des résultats optimaux. C'est le cas des zones urbaines très denses pour lesquelles les discontinuités de façades sont mal restituées dans le MNE, et également des zones péri-urbaines constituées de petites structures isolées.

Mauvaises restitutions des discontinuités de façades dans le MNE

La première limite est directement donnée par la qualité du MNE, et plus précisément son potentiel à restituer les discontinuités de façades. En effet, la première étape de notre méthode consiste à extraire la forme globale des emprises de bâtiments à travers un agencement de rectangles en utilisant les travaux d'Ortner [Ortner, 2004]. Ces derniers sont fondés sur une détection des discontinuités de façades dans le MNE : si celles-ci ne sont pas distinguables, l'algorithme fournira inexorablement des résultats non exploitables. La netteté des discontinuités dans le MNE est donc une condition nécessaire. La figure 4.13 illustre ces propos : les résultats obtenus sur la zone présentée sont inexploitables car le processus d'extraction n'arrive pas à localiser les emprises de bâtiments. Même un utilisateur éprouve visuellement beaucoup de difficultés pour identifier sur le MNE les différents éléments de la scène, comme par exemple la route située verticalement au centre de l'image. Ce problème apparaît lorsque l'on traite des zones urbaines très denses où, à l'image de la figure 4.13, les bâtiments ont été construits les uns sur les autres au fil du temps (souvent dans les quartiers anciens), les teintes des toitures sont homogènes, et les routes séparant les îlots de bâtiments sont très étroites. Tous ces facteurs contribuent à rendre la corrélation des images stéréoscopiques délicate.

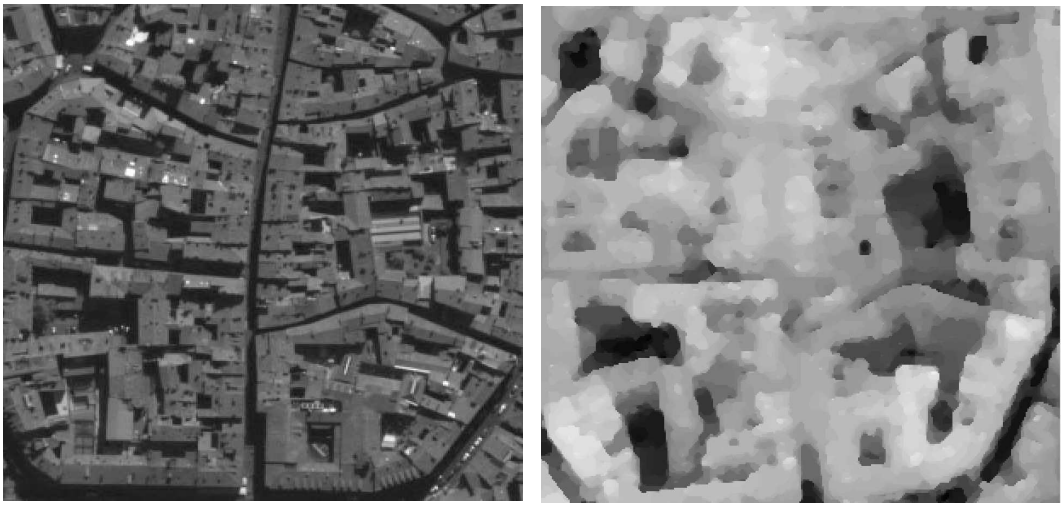

FIG. 4.13 – Extrait du centre-ville de Toulouse - Simulation PLEIADES *(gauche)*, MNE associé *(droite)*

Petites structures isolées

Le processus automatique de modélisation 3D a été proposé afin de traiter les zones urbaines denses. La phase d'extraction à travers l'utilisation de processus ponctuels marqués, comme la phase de reconstruction utilisant des relations d'assemblage de modules urbains, sont spécifiquement adaptés à une forte densité d'objets complexes et variés. Dans le cadre des zones péri-urbaines, les objets sont des bâtiments relativement simples, très petits, possédant souvent des emprises rectangulaires et des toits bi-plans symétriques. Bien que la modélisation 3D de ces zones présente beaucoup moins de difficultés, notre processus n'est cependant pas réellement approprié.

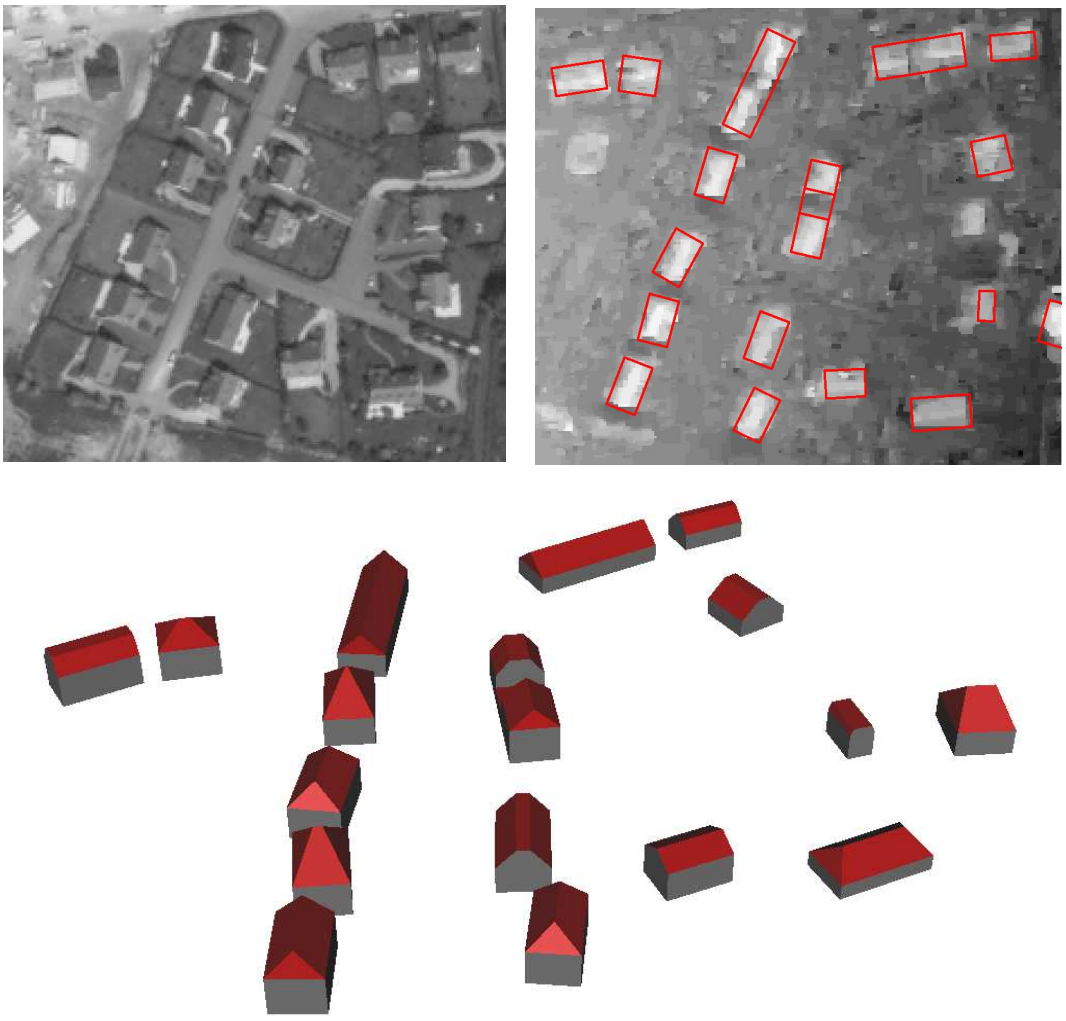

FIG. 4.14 – Reconstruction 3D automatique d'un extrait d'un quartier pavillonnaire d'Amiens *(bas)*, orthoimage PLEIADES et supports 2D extraits sur le MNE associés *(haut)*.

Nous avons testé le processus sur une zone pavillonnaire présentée sur la figure 4.14. Pour ce faire, les termes permettant d'introduire des interactions entre les objets ont été négligés (seuls les termes d'attache aux données ont été pris en compte). Nos a priori sont, en effet, inutiles dans le cadre péri-urbain et leur utilisation peut même provoquer des ambiguïtés et erreurs à travers une "sur-connexion" d'objets. Les maisons sont en général correctement localisées comme on peut le constater sur la figure 4.14. Cependant, lorsque deux maisons sont très proches l'une de l'autre ou lorsque des arbres sont disposés entre elles, le processus d'extraction distingue une seule structure. Le principal point négatif provient de la mauvaise reconnaissance des formes de toit. Les bâtiments sont très petits d'un point de vue surfacique sur le MNE à $0,7$ mètre de résolution : il est très difficile de restituer la forme correcte du toit. De plus, les structures sont isolées et n'ont pas d'objets voisins : les interactions de connexion entre objets ne peuvent pas être utilisées pour aider dans la décision. Ces raisons expliquent pourquoi un nombre non négligeable de maisons ont une forme de toit erronée. En définitive, notre processus fournit des résultats exploitables sur ce type de zones, mais n'est toutefois pas adapté.

4.6 Extension du contexte

Dans cette partie, nous quittons le contexte tout-automatique de notre étude et analysons le comportement de notre processus de reconstruction 3D. Celui-ci a été intégré à la plateforme de modélisation urbaine de l'IGN. Cette plateforme permet d'introduire de nombreux éléments dans les scènes 3D (Modèle Numérique de Terrain (MNT), textures à partir des orthoimages,...) soignant le rendu visuel. Cette plateforme permet de gérer la saisie manuelle des supports 2D par un opérateur. Les figures 6.15, 6.16, et 6.17 présentées en annexe F montrent les résultats obtenus sur cette plateforme par notre processus de reconstruction 3D à partir de MNE PLEIADES sur Toulouse (avec une saisie opérateur des emprises).

Comparaison avec les outils existants sur la plateforme de reconstruction IGN

Un des principaux intérêts de cette intégration réside dans la possibilité de comparer, dans un contexte identique, notre processus de reconstruction 3D à celui déjà présent sur la plateforme, fondé sur les travaux de [Taillandier, 2004; Durupt et Taillandier, 2006]. La figure 4.15 présente trois bâtiments reconstruits par les deux processus (l'hôtel de ville d'Amiens, la prison de Toulouse et la piscine d'Amiens).
D'une manière générale, les deux types de reconstructions sont très proches. Les temps de calculs mis par le processus original sont beaucoup plus courts (quelques secondes pour un bâtiments contre près d'une minute pour notre processus). Notre processus présente néanmoins deux avantages. Le premier, typiquement associé aux approches structurelles, est la qualité de la régularisation des toits et de leurs jonctions. Cela joue, par conséquent, sur la présence d'artefacts comme on peut le constater avec les reconstructions de la prison de Toulouse. Le second avantage réside, comme on peut le voir sur les reconstructions de la piscine d'Amiens, dans la possibilité de modéliser des toits courbés. Le processus original doit approximer la surface courbée par une succession de monoplans. Cela donne un résultat visuel médiocre et génère des artefacts au niveau des connexions des monoplans.

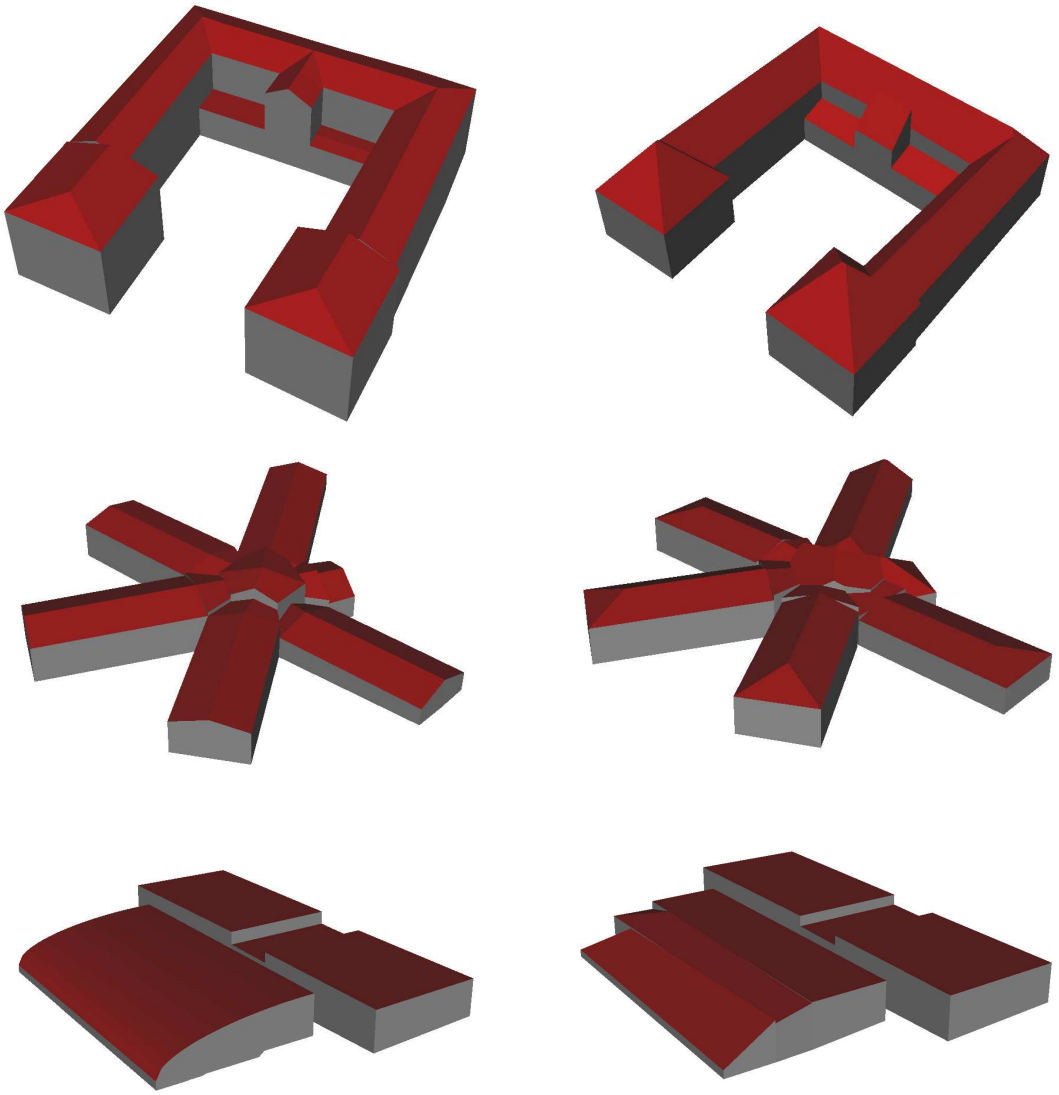

FIG. 4.15 – Comparaison entre deux processus de reconstruction 3D à partir d'une saisie manuelle des emprises - celui présenté dans le chapitre 3 *(gauche)*, celui fondé sur les travaux de [Taillandier, 2004; Durupt et Taillandier, 2006] *(droite)*.

On peut, par ailleurs, noter que les deux processus proposent des terminaisons de bâtiments souvent erronées. C'est le cas de l'hôtel de ville d'Amiens sur lequel les deux tourelles au premier plan sont des toits avec doubles attaques en V. Notre processus propose une attaque en V simple tandis que le processus original propose lui un toit pyramidal. On retrouve également ce problème sur la prison de Toulouse : plusieurs terminaisons sont, dans les deux cas, mal restituées. Ce point souligne la difficulté à reconnaître les terminaisons des bâtiments, qui sont généralement de très petites structures, à partir des MNE PLEIADES.

Résultats à partir d'images aériennes haute résolution

Nous avons constaté, dans la partie 4.4, que notre processus de reconstruction 3D ne se limitait pas à une utilisation "satellitaire" à travers des modélisations globales de bâtiments (principaux pans de toit). Le processus peut également fournir des modélisations plus détaillées (avec des avancées de façades, des décrochements de toit,...) sous réserve d'avoir un partitionnement précis des supports 2D. Nous présentons ici des résultats obtenus dans un contexte aérien, à partir de données haute résolution ($0,25$ cm). Des résultats complémentaires sont illustrés en annexe F.

La figure 4.16 présente le résultat obtenu sur l'Université des Sciences d'Amiens à partir d'un MNE aérien à $0,25$cm. La qualité de la modélisation est très satisfaisante. De nombreux détails sont restitués comme par exemple des balcons ou de petites cours intérieures. La précision altimétrique est très bonne et les formes de toit sont correctement restituées comme on peut le constater sur la figure 6.18. Le processus s'adapte très bien à des MNE aériens haute résolution.

Reconstruction de superstructures à partir d'images aériennes très haute résolution

Le processus 3D permet également de reconstruire les superstructures des bâtiments à partir d'images aériennes très haute résolution (à $0,1$ mètre). Les superstructures sont des éléments contenus sur les toits des bâtiments, le plus souvent correspondant à des cheminées, des chien-assis et des verrières.

La figure 4.17 présente un résultat obtenu sur un extrait de toit contenant des différentes sortes de superstructures. La qualité de la modélisation est globalement satisfaisante. Les résultats obtenus sont encourageants. Les superstructures sont majoritairement bien reconstruites. Les cheminées sont modélisées par des toits plats. Les chien-assis et verrières sont eux reconstruits respectivement par des attaques en V et par des bi-plans et toits en appentis. Des erreurs fréquentes sont toutefois présentes, principalement sur les cheminées. Cela s'explique par le fait que notre bibliothèque de modèles n'est pas spécifiquement adaptée pour la reconnaissances de ces petites formes. Pour améliorer les résultats, il faudrait inclure des modèles de superstructures dans la bibliothèque et introduire, dans la formulation de la densité, des connaissances a priori sur la taille des objets et des interactions d'inclusion superstructures/toits. Des résultats complémentaires sont présentés en annexe F.

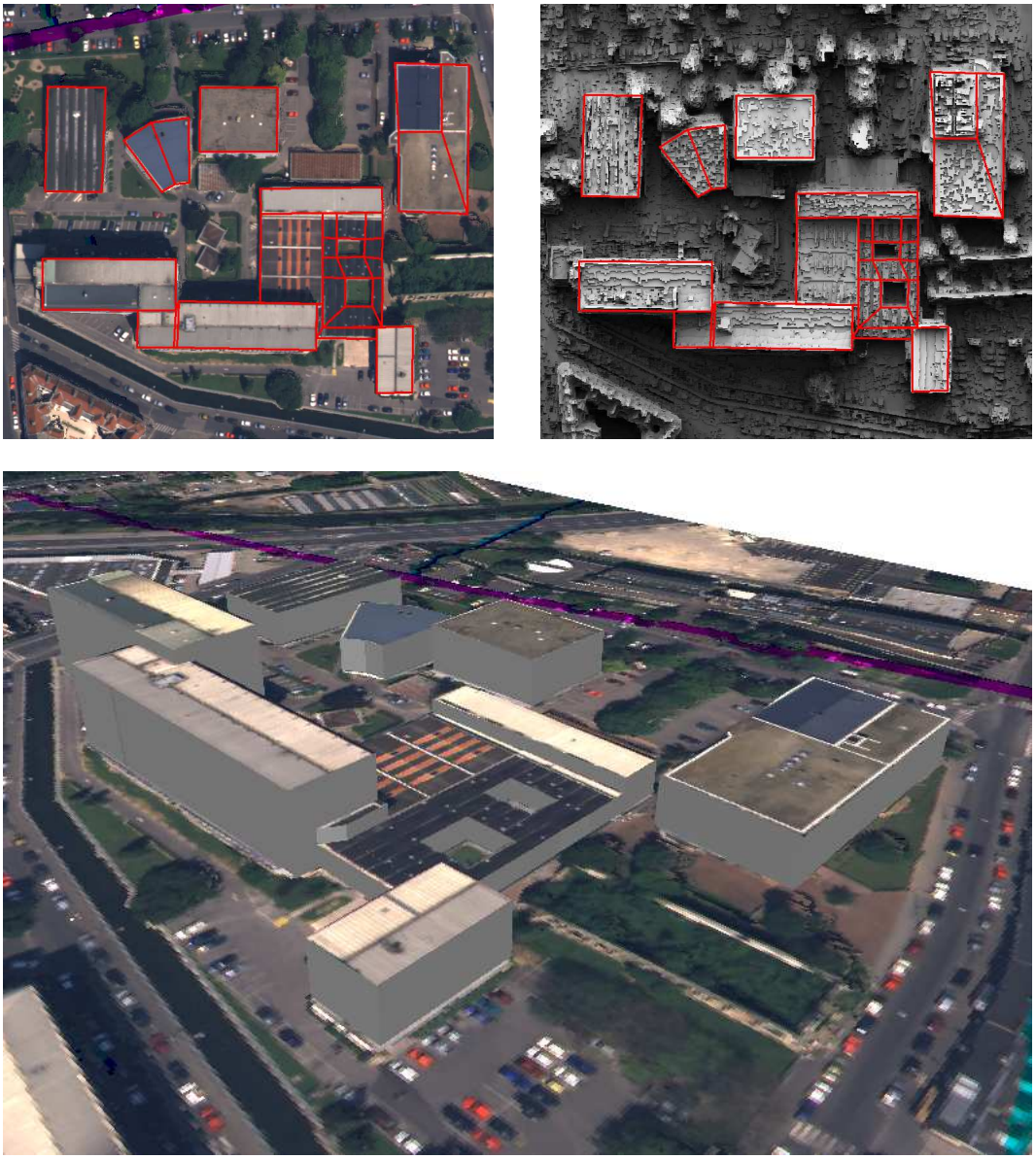

FIG. 4.16 – Reconstruction 3D de l'Université des Sciences d'Amiens à partir d'images aériennes ($0,25$ mètre) et d'une extraction manuelle des supports 2D.

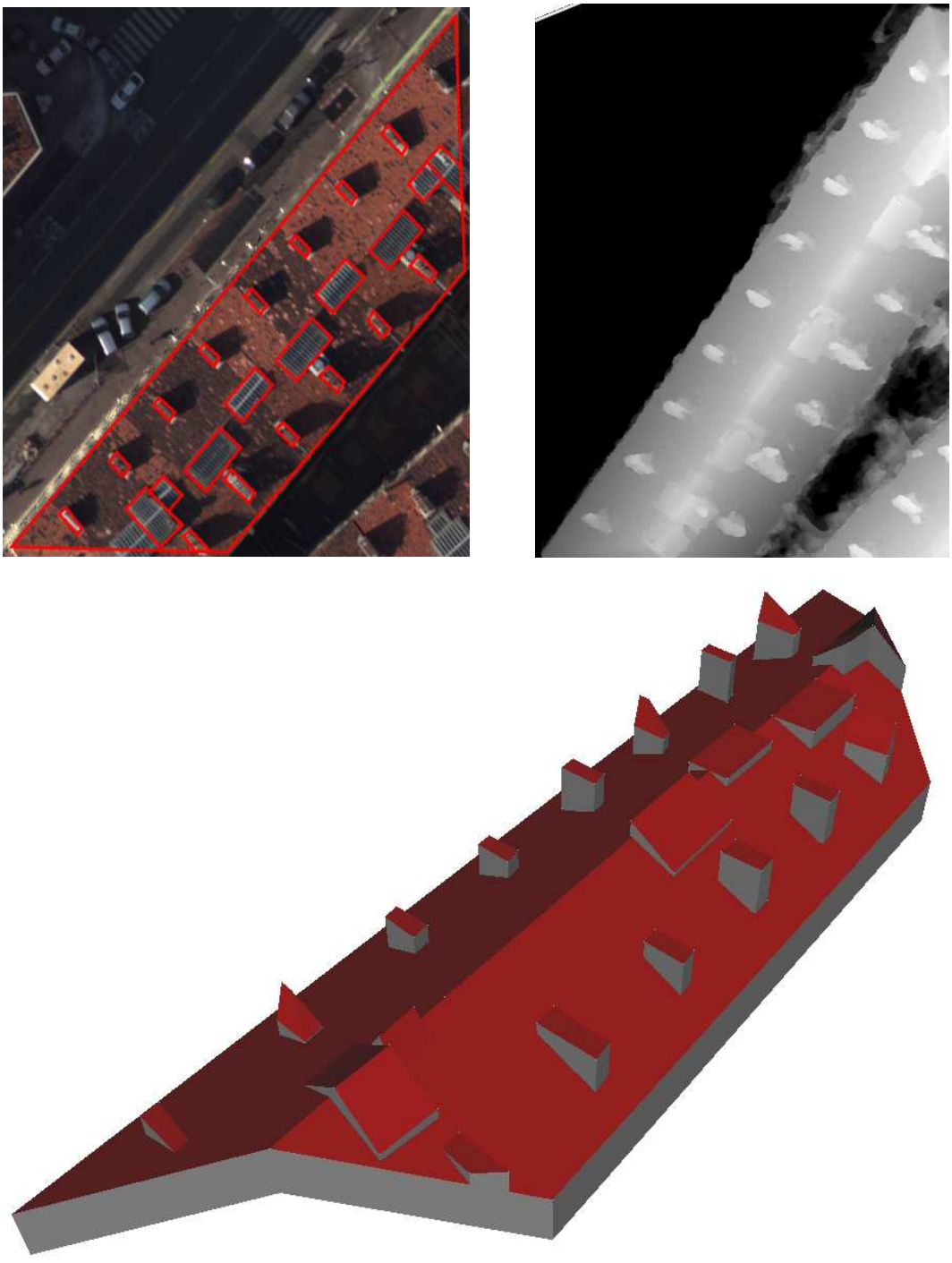

FIG. 4.17 – Reconstruction 3D d'un extrait de toit contenant des superstructures à partir d'images aériennes (0,1 mètre) et d'une extraction manuelle des supports 2D.

Chapitre 5

Conclusion

Synthèse des travaux effectués

Les travaux présentés dans cette thèse ont permis d'évaluer un processus automatique de modélisation tridimensionnelle de zones urbaines fondé sur une approche structurelle. Le contexte satellitaire imposait de travailler avec des données au contenu informatif faible par rapport aux images aériennes. L'utilisation d'outils stochastiques a permis de traiter cette problématique.

Contributions

La contribution majeure de ces travaux a consisté à proposer un processus original de modélisation 3D fondé sur une approche structurelle. Une telle approche n'est pas nouvelle dans le domaine. Cependant, les méthodes développées sur ce principe [Fischer et al., 1998; Fuchs, 2001] ont été améliorées, aussi bien du point de vue de l'automaticité du processus (sans cadastre, sur des zones urbaines denses), que de la généricité des modélisations (bibliothèque composée de neuf types de toits, de divers modules de jonctions et terminaisons, et surtout évolutive). Cette méthode a été élaborée en deux phases.

La première étape a consisté à reprendre les travaux d'Ortner [Ortner, 2004] afin d'obtenir des caricatures de bâtiments sous formes d'agencements de rectangles. Nous avons alors proposé un processus permettant de régulariser ces agencements de rectangles en supports 2D adaptés à une approche structurelle. Le résultat correspond à des agencements de quadrilatères quelconques connectés entre voisins et associés à des parties spécifiques d'un bâtiment.

La deuxième étape a constitué le coeur de la thèse : la reconstruction en 3D des bâtiments à partir des supports 2D extraits dans l'étape précédente. Nous avons proposé une bibliothèque de toits adaptée au paysage européen, relativement complète, variée et évolutive. La seconde contribution de cette partie a été de formuler le problème d'assemblage des modules à travers une densité définie dans un cadre bayésien, tout en prenant soin de faire intervenir un minimum de paramètres pour ne pas perdre en robustesse. La dernière contribution réside dans la mise en place d'un schéma d'optimisation original fondé sur les techniques MCMC. Nous avons notamment proposé deux noyaux de propositions adaptés aux problèmes où rendu visuel (à travers la régularisation des modules 3D) et fidélité à la réalité (à travers l'attache aux données) sont d'égale importance.

D'autres processus de reconstruction ont été brièvement présentés en annexe. Certains constituent des modèles intermédiaires (annexe A) dont les résultats sont peu convaincants. Ils ont toutefois permis de comprendre la manière d'aborder certaines difficultés de notre problème, et ont ainsi amené à proposer le processus structurel décrit dans le manuscrit. D'autres modèles (annexes B et E) sont des algorithmes proposant une modélisation simplifiée, pouvant être efficaces sous des contraintes d'utilisation spécifiques comme des temps de calcul faibles.

Ce travail a également permis d'évaluer le potentiel des futurs satellites PLEIADES en matière de modélisation 3D d'environnement urbain. Nous avons démontré que les MNE générés à partir de simulations PLEIADES ont une qualité suffisante pour permettre une

modélisation globale automatique des bâtiments en utilisant des approches robustes de type objets. Cependant, la densité des zones urbaines ne doit pas être trop élevée. A l'image du centre-ville de Toulouse, les environnements très denses semblent difficilement modélisables par un processus automatique.

Intérêts du processus proposé

Le processus de modélisation tridimensionnelle mis en place présente plusieurs points de satisfaction qui justifient l'utilisation d'outils stochastiques au sein d'une approche objet structurelle.

- La méthode permet d'obtenir de manière relativement robuste des résultats satisfaisants, aussi bien du point de vue quantitatif que qualitatif, dans un cadre difficile (images satellitaires de qualité moyenne pour la problématique de modélisation urbaine 3D, méthode automatique sans utilisation de données cadastrales et sans focalisation sur des bâtiments, zones urbaines denses relativement complexes et variées).

- Le processus est évolutif dans la mesure où on peut ajouter ou supprimer avec facilité des modèles 3D dans la bibliothèque. On peut donc espérer adapter notre méthode sur des paysages urbains différents des zones européennes en créant de nouveaux modèles de formes de toit à inclure dans la bibliothèque.

- La méthode ne fonctionne pas uniquement à partir de données satellitaires. Elle peut, en effet, être utilisée à des résolutions plus fines, typiquement à partir de données aériennes. Avec un partitionnement précis des supports 2D, elle permet d'obtenir, à partir de MNE aériens haute et très haute résolution, des modélisations détaillées avec des avancées de façades, des décrochements de toits et même des superstructures.

- L'approche permet d'utiliser de manière séparée soit le processus d'extraction des supports 2D (chapitre 2), soit le processus de reconstruction 3D (chapitre 3). Intégré à la plateforme de modélisation 3D de l'IGN, ce dernier constitue un processus semi-automatique intéressant permettant de fournir des modélisations très détaillées à partir d'un partitionnement manuel des supports 2D.

Les limites

Le processus automatique de modélisation 3D possède néanmoins plusieurs défauts relativement pénalisant en vue d'une utilisation à grande échelle.

- Le processus rencontre des difficultés dans certaines situations. Ainsi, le MNE utilisé doit impérativement avoir des discontinuités de façades franches, sans quoi le processus d'extraction des supports fournit des résultats inexploitables. Par ailleurs, la méthode ne permet pas, à partir de MNE à $0,7$ mètre de résolution, une bonne restitution des toits pour les petites structures isolées style maisons mitoyennes. Enfin, le processus connaît des problèmes de sur et sous-détections. La végétation n'est pas gérée, ce qui engendre une confusion entre arbres et bâtiments. De même, les petites

structures dans les cours intérieures sont difficilement, voire pas du tout, détectées.

- L'étape d'extraction des supports 2D, comme l'étape de reconstruction 3D font appel à des méthodes d'optimisation stochastiques qui sont assez lentes. Les temps de calculs du processus restent très importants malgré les techniques d'accélération mises en place.

- Le processus est particulièrement dépendant de la méthode d'extraction des supports 2D. La méthode de reconstruction 3D ne permet pas, en effet, de corriger les éventuelles erreurs de positionnement des supports 2D. De ce fait, le processus peut être sujet à des accumulations d'erreurs au cours des deux phases successives de la modélisation. Il paraît important, afin de gagner en robustesse, d'introduire, entre les deux phases du processus, une méthode permettant l'évaluation et la correction éventuelle des supports 2D.

Perspectives

Les perspectives d'amélioration du processus s'articulent essentiellement autour de trois axes de recherche : l'extraction des supports 2D, l'optimisation de la méthode de reconstruction 3D, et la gestion des détails dans les modélisations.

L'extraction des supports 2D

Comme nous avons pu le voir précédemment, la qualité des reconstructions dépend fortement de la qualité des supports 2D qui ont été extraits. Le processus automatique d'extraction des supports permet essentiellement de reconstruire la forme générale des toits. Or, avec des supports plus précis et mieux partitionnés, on obtiendrait des modélisations plus détaillées. L'enjeu est donc important. Deux pistes paraissent particulièrement intéressantes.

La première consiste à reprendre le processus d'extraction proposé et à chercher à affiner le partitionnement des supports. La méthode prendrait ainsi en compte différents niveaux de détails et permettrait d'extraire notamment les avancées de toits et les décrochements de façades. Dans le même temps, on pourrait mettre en place une méthode permettant de détecter les petites structures à l'intérieur des cours intérieures.

La seconde solution consiste à proposer un nouveau processus d'extraction. La piste la plus intéressante est probablement apportée par les champs polygonaux de Markov [Arak et Surgailis, 1989]. Les processus ponctuels marqués sont des outils stochastiques efficaces pour extraire des objets dans les images. Ces objets doivent, cependant, être paramétriquement simples pour avoir des temps de calcul raisonnables (au plus cinq à six paramètres par objet). Or, les emprises de bâtiments sont des formes complexes pouvant difficilement être décrites par des objets simples sans l'utilisation a posteriori d'un processus de transformation en objets plus complexes. Les champs polygonaux de Markov peuvent potentiellement fournir des descriptions beaucoup plus fines à travers des polygones [Kluszczynski et al., 2004], tout en gardant les aspects intéressants des modélisations stochastiques (introduction de contraintes géométriques sur les polygones, mises en place d'interactions, possibilité

d'utiliser des espaces d'état continus,...).

L'optimisation

Dans la partie 3.3.4, nous avons évoqué et discuté de l'utilisation de techniques d'optimisation alternatives à celle proposée dans le manuscrit. Il apparaît important d'approfondir cette direction. Nous avons montré que les processus de diffusion à sauts pouvaient être utilisés pour optimiser notre énergie. Ce type d'algorithmes est particulièrement intéressant dans la mesure où les temps de calcul seraient significativement réduits. Une autre possibilité pourrait être d'utiliser les techniques fondées sur les modèles graphiques. Un procédé intéressant consisterait à utiliser une modélisation par graphe hiérarchique à deux niveaux : le niveau inférieur correspondant à des séquences d'objets (dont l'optimisation locale est relativement simple et rapide), le niveau supérieur permettant d'introduire des lois de dépendance sur le raccordement des séquences entre elles. On pourrait également se pencher sur les techniques de propagation des messages. Les méthodes de "belief propagation" semblent être particulièrement bien adaptées, d'autant que les graphes de notre problème possèdent souvent peu de cycles.

Nous pourrions également imaginer utiliser des modèles de reconstruction 3D peu précis mais rapides (à l'image de ceux proposés en annexe) comme une initialisation pour un modèle déterministe plus précis. Ainsi, les résultats fournis par le processus présenté en annexe E permettraient de donner le type de toit des objets ainsi qu'une estimée des valeurs des paramètres. Nous pourrions alors utiliser ces informations pour fournir une bonne initialisation à un processus déterministe rapide et précis.

La gestion des détails

Il semble particulièrement intéressant d'adapter plus spécifiquement la méthode de reconstruction 3D à des objets représentant des niveaux de détails différents. On pourrait ainsi alimenter la bibliothèque d'objets permettant la restitution de nombreux détails situés sur les toits ou à coté des façades. Les résultats encourageants obtenus sur la reconstruction de superstructures permettent d'envisager la mise en place d'une densité a posteriori plus complexe, permettant de gérer des interactions d'inclusion entre les objets (notamment entre les superstructures et les toits).

Bilan

L'approche structurelle est, d'une manière générale, une solution très intéressante pour aborder la modélisation 3D d'environnements urbains. D'un intérêt limité dans les milieux péri-urbains, cette approche est particulièrement bien adaptée pour traiter les zones urbaines denses. A mi-chemin entre une modélisation purement paramétrique et une modélisation polyédrique, elle apparaît comme un compromis très intéressant permettant notamment de concilier les deux aspects majeurs de l'évaluation d'une scène 3D : la fidélité à la réalité et la qualité du rendu visuel.

L'approche structurelle idéale serait un processus monolithique, c'est-à-dire qu'elle permettrait de réaliser simultanément la phase d'extraction des supports 2D et la phase de

reconstruction 3D. Mais au vu du problème combinatoire extrêmement complexe qu'elle engendrerait, une telle approche semble difficilement concevable aujourd'hui pour traiter des environnements urbains denses.

La modélisation 3D d'environnement urbain par télédétection reste un thème de recherche très ouvert. La mise en production d'un processus complètement automatique est, au vu de la problématique, un objectif qui semble lointain. Pour l'atteindre, des progrès concernant la physique des capteurs et la qualité des données en général seront probablement nécessaires.

Bibliographie

Allain, R., 2004. Morphologie Urbaine : Géographie, aménagement et architecture de la ville. A. Colin.

Ameri, B., 2000. Feature based model verification : A new concept for hypothesis validation in building reconstruction. Dans : Proc. of the XIXth ISPRS Congress. Vol. 33, B3. Amsterdam, Pays-Bas, pp. 24–35.

Arak, T., Surgailis, D., 1989. Markov fields with polygonal realisations. Probability Theory and Related Fields 80, 543–579.

Baillard, C., 1997. Analyse d'images aériennes stéréoscopique pour la restitution 3D des milieux urbains. Détection et caractéristique du sursol. Thèse de doctorat, Ecole Nationale Supérieure des Télécommunications de Paris.

Baillard, C., Zisserman, A., 1999. Automatic reconstruction of piecewise planar models from multiple views. Dans : Proc. of the IEEE Conference on Computer Vision and Pattern Recognition. Corfu, Grèce.

Bailloeul, T., Prinet, V., Serra, B., Marthon, P., 2005. Spatio-temporal prior shape constraint for level set segmentation. Dans : Proc. Energy Minimization Methods in Computer Vision and Pattern Recognition. St. Augustine, Etats-Unis.

Besag, J., 1986. On statistical analysis of dirty pictures. Journal of the Royal Statistical Society 48 (3), 259–302.

Blake, A., Zisserman, A., 1987. Visual reconstruction. Dans : MIT Press. M.A., Cambridge.

Boykov, Y., Veksler, O., Zabih, R., 1999. Fast approximate energy minimization via graph cuts. Dans : IEEE proc. of International Conference on Computer Vision. Corfu, Grèce.

Brédif, M., Boldo, D., Pierrot-Deseilligny, M., Maître, H., 2007. 3D building reconstruction with parametric roof superstructures. Dans : Proc. of the IEEE International Conference on Image Processing. San Antonio, Etats-Unis.

Brenner, C., Haala, N., Fritsch, D., 2001. Towards fully automated 3D city model generation. Dans : Proc. of the Workshop on Automatic Extraction of Man-Made Objects from Aerial and Space Images. Ascona, Suisse.

Brenner, C., Ripperda, N., 2006. Extraction of facades using RJMCMC and constraint equations. Dans : Proc. of the ISPRS Commission III Symposium on Photogrammetric and Computer Vision. Bonn, Allemagne.

Bretar, F., 2006. Couplage de données laser aéroporté et photogrammétriques pour l'analyse de scènes tridimensionnelles. Thèse de doctorat, Ecole Nationale Supérieure des Télécommunications de Paris.

Celeux, G., Diebolt, J., 1986. L'algorithme SEM : un algorithme d'apprentissage probabiliste pour la reconnaissance de mélange de densités. Revue de Statistique Appliquée 34(2), 35–52.

Cellier, F., 2007. Reconstruction 3d de bâtiments en interférométrie rso haute résolution, approche par gestion d'hypothèses. Thèse de doctorat, Ecole Nationale Superieure des Telecommunications de Paris.

Chehata, N., Pierrot-Deseilligny, M., Stamon, G., 2005. Hybrid digital elevation models computation constrained by 3d-primitives : A global optimization algorithm using graph cuts. Dans : Proc. of the IEEE International Conference on Image Processing. Genova, Italie.

Ching, F., 1996. Building Construction Illustrated : A Visual Dictionary of Architecture, Form, Space and Order. John Wiley & Sons.

Cocquerez, J., Philipp, S. (Eds.), 1995. Analyse d'images : filtrage et segmentation. MASSON.

Collins, R., 1996. A space-weep approach to true multi-image matching. Dans : Proc. of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, San Francisco, Etats-Unis, pp. 358–365.

Collins, R., Jaynes, C., Cheng, Y.-Q., Wang, X., Stolle, F., Riseman, E., Hanson, A., 1998. The ascender system : Automated site modeling from multiple aerial images. Computer Vision and Image Understanding 72 (2), 143–162.

Dang, T., 1994. Interpretation et restitution automatique des batiments isoles a partir d'un couple stereoscopique d'images aeriennes. These de doctorat, Ecole Nationale Superieure des Telecommunications de Paris.

De Joinville, O., 2003. Evaluation de la qualite d'une cartographie urbaine a l'aide d'images aeriennes a haute resolution. Thèse de doctorat, Ecole Nationale Superieure des Telecommunications de Paris.

Dempster, A., Laird, N., Rubin, D., 1977. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society B 39, 1–38.

Descombes, X., 2004. Méthodes stochastiques en analyse d'image : des champsde markov aux processus ponctuels marqués. Habilitation à diriger des recherches, Universite de Nice Sophia Antipolis.

Descombes, X., Zhizhina, E., 2004. Applications of gibbs fields methods to image processing problems. Problems of Information Transmission 40 (3), 108–125.

Dick, A., Torr, P., Cipolla, R., Nov 2004. Modelling and interpretation of architecture from several images. International Journal of Computer Vision 60 (2), 111–134.

Doucet, A., Godsill, S., Andrieu, C., 2000. On Sequential Monte Carlo Sampling Methods for Bayesian Filtering. Statistics and Computing 10 (3), 197–208.

Durupt, M., Taillandier, F., 2006. Automatic building reconstruction from a digital elevation model and cadastral data : an operational approach. Dans : Proc. of the ISPRS Commission III Symposium on Photogrammetric and Computer Vision. ISPRS, Bonn, Allemagne.

Fischer, A., Kolbe, T., Lang, F., Cremers, A., Forstner, W., Plumer, L., Steinhage, V., 1998. Extracting buildings from aerial images using hierarchical aggregation in 2D and 3D. Computer Vision and Image Understanding 72 :2, 163–185.

Flamanc, D., Maillet, G., 2005. Evaluation of 3d city model production from pleiades-hr images and 2d ground maps. Dans : URBAN, GRSS/ISPRS Joint Workshop on Data Fusion and Remote Sensing over Urban Areas. Tempe, Etats-Unis.

Flamanc, D., Maillet, G., Jibrini, H., Sep 2003. 3-d city models : an operational approach using aerial images and cadastral maps. Dans : ISPRS Conference Photogrammetric Image Analysis (PIA). Munich, Allemagne, pp. 53–58.

Forney, J., 1973. The Viterbi algorithm. Proceedings of the IEEE 13, 268–278.

Fritsch, D., 1999. Virtual cities and landscape models - what has photogrammetry to offer ? Dans : Proceedings of the Photogrammetric Week. pp. 3–14.

Frueh, C., Jain, S., Zakhor, A., 2005. Data processing algorithms for generating textured 3D building facade meshes from laser scans and camera images. International Journal of Computer Vision 61 (2), 159–184.

Fuchs, F., 2001. Contribution à la reconstruction du bâti en milieu urbain, à l'aide d'images aériennes stéréoscopiques à grande échelle. Etude d'une approche structurelle. Thèse de doctorat, Universite René Descartes - Paris 5.

Gamba, P., Houshmand, B., Saccani, M., 2000. Detection and extraction of buildings from interferometric sar data. IEEE Transactions on Geoscience and Remote Sensing 38 (1), 611–617.

Gaminet, O., 1999. Architecture et Urbanisme : l'espace moderne et la ville. Editions de l'Ecole Polytechnique.

Garcin, L., Descombes, X., Zerubia, J., Le Men, H., 2001. Buiding extraction using a Markov point process. Dans : Proc. of the IEEE International Conference on Image Processing. Thessalonique, Grèce.

Geman, S., Geman, D., 1984. Stochastic relaxation, gibbs distribution ans the Bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence 6, 721–741.

Geman, S., Huang, C., 1986. Diffusion for global optimization. Journal of Control and Optimization 24 (5), 1031–1043.

Green, P., 1995. Reversible Jump Markov Chains Monte Carlo computation and Bayesian model determination. Biometrika 57, 97–109.

Greig, D., Porteous, T., Seheult, A., 1989. Exact maximum a posteriori estimation for binar images. Journal of the Royal Statistical Society 51 (B), 271–279.

Grenander, U., Chow, Y., Keenan, D. M., 1991. Hands : a pattern theoretic study of biological shapes. Springer-Verlag New York, Inc., New York, Etats-Unis.

Grenander, U., Miller, M., 1994. Representations of Knowledge in Complex Systems. Journal of the Royal Statistical Society 56 (4).

Gruen, A., 1998. Tobago - a semi-automated approach for the generation of 3D building models. ISPRS Journal of Photogrammetry and Remote Sensing 53 (2), 108–118.

Gruen, A., Wang, X., 1998. CC-modeler : a topology generator for 3D city models. ISPRS Journal of Photogrammetry and Remote Sensing 53 (5), 286–295.

Haala, N., Brenner, C., 1999. Extraction of buildings and trees in urban environments. ISPRS Journal of Photogrammetry and Remote Sensing 54 (2-3), 130–137.

Haario, H., Saksman, E., 1991. Simulated annealing process in general state space. Advances in Applied Probability (23), 866–893.

Han, F., Tu, Z., Zhu, S., 2004. Range Image Segmentation by an Effective Jump-Diffusion Method. IEEE Trans. on Pattern Analysis and Machine Intelligence 26 (9), 1138–1153.

Hastings, W., 1970. Monte Carlo sampling using Markov chains and their applications. Biometrica 57(1), 97–109.

Henricsson, O., 1998. The role of color attributes and similarity grouping in 3D building reconstruction. Computer Vision and Image Understanding 72 (2), 163–184.

Heuel, S., Lang, F., Forstner, W., 2000. Topological and geometrical reasoning in 3D grouping for reconstructing polyhedral surfaces. Dans : Proc. of the XIXth ISPRS Congress. Amsterdam, Pays-Bas.

Jaynes, C., Hanson, A., Riseman, E., Schultz, H., 1997. Building reconstruction from optical and range images. Dans : Proc. of the IEEE Conference on Computer Vision and Pattern Recognition. San Juan, Porto Rico.

Jibrini, H., 2002. Reconstruction automatique de bâtiments en modèles polyhédriques 3D à partir de données cadastrales vectorisées 2D et un couple d'images aériennes à haute résolution. Thèse de doctorat, Ecole Nationale Supérieure des Télécommunications de Paris.

Kada, M., 2006. 3D building generalization based on half-space modeling. Dans : workshop ISPRS, commission II. Hannovre, Allemagne.

Kloeden, P., Platen, E., 1992. Numerical solution of stochastic differential equations. Springer-Verlag, New York.

Kluszczynski, R., VanLieshout, M. N. M., Schreiber, T., 2004. Image segmentation by polygonal markov fields. Research Report 0409, PNA.

Lacoste, C., Descombes, X., Zerubia, J., October 2005. Point processes for unsupervised line network extraction in remote sensing. IEEE Trans. Pattern Analysis and Machine Intelligence 27 (10), 1568–1579.

Laferté, J., 1996. Contribution à l'analyse d'images par modèles markoviens sur des graphes hiérarchiques. applications à la fusion de données multirésolutions. Thèse de doctorat, Université de Rennes 1.

Lee, S., Nevatia, R., 2003. Interactive 3D building modeling using a hierarchical representation. Dans : IEEE Workshop on Higher-Level Knowledge in 3D Modeling and Motion (HLK). Nice, France.

Li, J., Nevatia, R., Noronha, S., 1999. User assisted modeling of buildings from aerial images. Dans : Proc. of the IEEE Conference on Computer Vision and Pattern Recognition. Fort Collins, Etats-Unis.

Li, S., 1995. Markov Random Field modeling in computer vision. Springer-Verlag.

Lin, C., Nevatia, R., 1998. Building detection and description from a single intensity image. Computer Vision and Image Understanding 72 (2), 101–121.

Maas, H., Vosselman, G., 1999. Two algorithms for extracting building models from raw laser altimetry data. ISPRS Journal of Photogrammetry and Remote Sensing 54, 153–163.

McGlone, J., Shufelt, J., 1994. Projective and object space geometry for monocular building extraction. Rapport technique, Carnegie Mellon University.

Metropolis, M., Rosenbluth, A., Teller, A., Teller, E., 1953. Equation of state calculations by fast computing machines. Journal of Chemical Physics 21, 1087–1092.

Mitchell, M., 1996. Introduction to genetic algorithms. Dans : MIT Press. M.A., Cambridge.

Nevatia, R., Price, K., 2002. Automatic and interactive modeling of buildings in urban environments from aerial images. Dans : Proc. of the IEEE International Conference on Image Processing. Rochester, Etats-Unis.

Noronha, S., Nevatia, R., 2001. Detection and modeling of buildings from multiple aerial images. IEEE Transactions on Pattern Analysis and Machine Intelligence 23 (5), 501–518.

Oriot, H., 2003. Statistical snakes for building extraction from stereoscopic aerial images. Dans : ISPRS Conference on Photographic Image Analyses. Hannovre, Allemagne.

Ortner, M., 2004. Processus ponctuels marqués pour l'extraction automatique de caricatures de bâtiments à partir de Modèles Numériques d'Elévation. Thèse de doctorat, Université de Nice Sophia Antipolis.

Ortner, M., Descombes, X., Zerubia, J., 2007. Building outline extraction from digital elevation models using marked point processes. International Journal of Computer Vision 72 (2), 107–132.

Palubinskas, G., Descombes, X., Kruggel, F., August 1998. An unsupervised clustering method using the entropy minimization. Dans : Proc. International Conference on Pattern Recognition. Brisbane, Australie.

Paparoditis, N., Cord, M., Jordan, M., Cocquerez, J.-P., 1998. Building detection and reconstruction from mid-and high-resolution aerial imagery. Computer Vision and Image Understanding 72 (2), 122–142.

Penard, L., Paparoditis, N., Pierrot-Deseilligny, M., 2006. Reconstruction 3D automatique de façades de bâtiments en multi-vues. Dans : Reconnaissance des Formes et Intelligence Artificielle. Tours, France.

Pennec, X., 1996. L'incertitude dans les problèmes de reconnaissance et de recalage – applications en imagerie médicale et biologie moléculaire. Thèse de doctorat, Ecole Polytechnique, Palaiseau, France.

Perrin, G., Descombes, X., Zerubia, J., November 2005. Adaptive simulated annealing for energy minimization problem in a marked point process application. Dans : Proc. Energy Minimization Methods in Computer Vision and Pattern Recognition. St Augustine, Etats-Unis.

Pierrot-Deseilligny, M., Paparoditis, N., 2006. A multiresolution and optimization-based image matching approach : an application to surface reconstruction from SPOT5-HRS stereo imagery. Dans : Workshop ISPRS commission I. Ankara, Turquie.

Pérez, P., 2003. Modèles et algorithmes pour l'analyse probabilistiques des images. Habilitation à diriger des recherches, Universite de Rennes 1.

Robert, C., 1996. Méthodes de Monte Carlo par Chaînes de Markov. Statistique mathématique et probabilité. Economica.

Robert, C., Casella, G., 1999. Monte Carlo Statistical Methods. Springer-Verlag, New York.

Rottensteiner, F., Briese, C., 2002. A new method for building extraction in urban areas from high-resolution lidar data. Dans : Proc. of the ISPRS Commission III Symposium on Photogrammetric and Computer Vision. Vol. 34. Graz, Autriche.

Roy, S., Cox, I., 1998. A maximum-flow formulation of the n-camera stereo correspondence problem. Dans : Proc. of the IEEE International Conference on Computer Vision. Bombay, Inde.

Salamon, P., Sibani, P., Frost, R., 2002. Facts, conjectures, and improvements for simulated annealing. SIAM Monographs on Mathematical Modeling and Computation. Society for Industrial and Applied Mathematics, Philadelphia.

Salomon, M., Perrin, G., Heitz, F., 2002. Parallel sampling with stochastic differential equations for 3D deformable matching of medical images. Dans : Proc. International Conference on Parallel and Distributed Processing Techniques and Applications. Las Vegas, Etats-Unis.

Scholze, S., Moons, T., Van Gool, L., 2002. A probabilistic approach to building roof reconstruction using semantic labelling. Dans : Proc. of the DAGM Conference. Vol. 2449 de Lecture Notes in Computer Science. Zurich, Suisse.

Senegas, J., 2002. Méthodes de Monte Carlo en vision stéréoscopique -Application à l'étude de Modèles Numériques de Terrain. Thèse de doctorat, Ecole des Mines de Paris.

Simonetto, E., Oriot, H., Garello, R., Le Caillec, J., 2003. Radargrammetric processing for 3D building extraction from high-resolution airborne SAR data. Dans : IEEE International Geoscience and Remote Sensing Symposium. Toulouse, France.

Soergel, U., Michaelsen, E., Thoennessen, U., Stilla, U., 2005. Potential of high-resolution sar images for urban analysis. Dans : URBAN, GRSS/ISPRS Joint Workshop on Data Fusion and Remote Sensing over Urban Areas. Tempe, Etats-Unis.

Srivastava, A., Grenander, U., Jensen, G., Miller, M., 2002. Jump-Diffusion Markov processes on orthogonal groups for object recognition. Journal of Statistical Planning and Inference 103(1/2), 17–37.

Stoyan, D., Kendall, W., Mecke, J., 1987. Stochastic Geometry and its Applications. John Wiley and Sons.

Suveg, I., Vosselman, G., 2001. 3D building reconstruction by map based generation and evaluation of hypotheses. Dans : British Machine Vision Conference. Manchester, Angleterre.

Taillandier, F., 2004. Reconstruction de bâti en milieu urbain : une approche multi-vues. Thèse de doctorat, Ecole Polytechnique.

Taillandier, F., Deriche, R., 2002. Reconstruction of 3D linear primitives from multiple views for urban areas modelisation. Dans : Proc. of the ISPRS Commission III Symposium on Photogrammetric and Computer Vision. Graz, Autriche.

Tison, C., 2004. Interférométrie RSO à haute résolution en milieu urbain : application au calcul de MNS urbain. Thèse de doctorat, Ecole Nationale Supérieure des Télécommunications de Paris.

Tu, Z., Zhu, S., 2002. Image Segmentation by Data-Driven Markov Chain Monte Carlo. IEEE Trans. on Pattern Analysis and Machine Intelligence 24 (5), 657–673.

Tupin, F., Roux, M., 2003. Detection of building outlines based on the fusion of sar and optical features. ISPRS Journal of Photogrammetry and Remote Sensing 58 (1), 71–82.

Van Laarhoven, P., Aarts, E., 1987. Simulated Annealing : Theory and Applications. D. Reidel, Boston.

Van Lieshout, M., 2000. Markov point processes and their applications. Dans : Imperial College Press. Londre.

Varanelli, J., 1996. On the acceleration of the simulated annealing. Thèse de doctorat, University of Virginia, Charlottesville, Etats-Unis.

Vestri, C., 2000. Outils pour la reconstruction automatique de bâtiments à partir d'imagerie aérienne. Thèse de doctorat, Université de Nice-Sophia Antipolis.

Vinson, S., Cohen, L., 2002. Multiple rectangle model for buildings segmentation and 3D scene reconstruction. Dans : Proc. of the International Conference on Pattern Recognition. Quebec, Canada.

Viterbi, A., 1967. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. IEEE Transactions on Information Theory 13, 260–269.

Wang, X., Totaro, S., Taillandier, F., Hanson, A., Teller, S., 2002. Recovering facade texture and microstructure from real-world images. Dans : Proc. of ISPRS Commission III Symposium on Photogrammetric Computer Vision. Graz, Autriche.

Weidner, U., Forstner, W., 1995. Towards automatic building reconstruction from high-resolution digital elevation models. ISPRS Journal of Photogrammetry and Remote Sensing 50 (4), 38–49.

White, S., 1984. Concepts of scale in simulated annealing. Dans : IEEE proc. of International Conference on Computer Design. pp. 646–651.

Willuhn, W., Ade, F., 1996. A rule-based system for house reconstruction from aerial images. Dans : Proc. of the International Conference on Pattern Recognition. Vienne, Autriche.

Winkler, G., 1995. Image analysis, random fields and Markov chain Monte Carlo methods a mathematical introduction. Springer.

Xu, G., Zhang, Z., 1996. Epipolar Geometry in stereo, motion and object recognition. Kluwer Academic Publishers.

Zhao, H., Shibasaki, R., 2001. Reconstructing textured CAD model of urban environment using vehicle-borne laser range scanners and line cameras. Lecture Notes in Computer Science 2095, 284–297.

Publications de l'auteur

Journaux

- F. Lafarge, X. Descombes, J. Zerubia et M. Pierrot-Deseilligny. Automatic building extraction from DEMs using an object approach. Journal of Photogrammetry and Remote Sensing, à paraître, 2007.
- F. Lafarge, X. Descombes, J. Zerubia et M. Pierrot-Deseilligny. Automatic 3D Building Reconstruction from DEMs. Revue Française de Photogrammétrie et de Télédétection (SFPT), à paraître, 2007.
- F. Lafarge, X. Descombes, J. Zerubia et S. Mathieu. Détection de feux de forêt par analyse statistique d'événements rares à partir d'images infrarouges thermiques. Traitement du Signal , Vol.24(1), 2007.
- F. Lafarge, X. Descombes, J. Zerubia et M. Pierrot-Deseilligny. Modèle Paramétrique pour la Reconstruction Automatique en 3D de Zones Urbaines Denses à partir d'Images Satellitaires Haute Résolution. Revue Française de Photogrammétrie et de Télédétection (SFPT) 180 :pages 4-12, 2005.

Conférences

- F. Lafarge, X. Descombes, J. Zerubia et M. Pierrot-Deseilligny. 3D city modeling based on Hidden Markov Model. IEEE ICIP'07, San Antonio, Etats-Unis, Sept. 2007.
- F. Lafarge, X. Descombes et J. Zerubia. Forest Fire Detection based on Gaussian Field Analysis. EURASIP EUSIPCO'07, Poznan, Pologne, Sept. 2007.
- O. Tournaire, N. Paparoditis et F. Lafarge. Rectangular road marking detection with marked point processes. ISPRS PIA'07, Munich, Allemagne, Sept. 2007.
- F. Lafarge, X. Descombes, J. Zerubia et M. Pierrot-Deseilligny. An Automatic Building Reconstruction Method : A Structural Approach Using HR Images. IEEE ICIP'06, Atlanta, Etats-Unis, Oct. 2006.
- F. Lafarge, X. Descombes, J. Zerubia et M. Pierrot-Deseilligny. Automatic 3D Building Reconstruction from DEMs : an Application to PLEIADES Simulations. In Proc. ISPRS Commission I Symposium, Marne La Vallee, France, Juil. 2006.
- F. Lafarge, X. Descombes, J. Zerubia et M. Pierrot-Deseilligny. An automatic 3D city model : A Bayesian approach using satellite images. IEEE ICASSP'06, Toulouse, France, Mai 2006.
- F. Lafarge, X. Descombes et J. Zerubia. Textural kernel for SVM classification in remote sensing : application to forest fire detection and urban area extraction. IEEE ICIP'05, Gêne, Italie, Sept. 2005.
- F. Lafarge, X. Descombes, J. Zerubia et S. Mathieu-Marni. Détection de feux de forêt à partir d'images satellitaires IRT par analyse statistique d'événements rares. GRETSI'05, Louvain la neuve, Belgique, Sept. 2005.

Rapports

- F. Lafarge, X. Descombes, J. Zerubia et M. Pierrot-Deseilligny. A structural approach for 3D building reconstruction. Rapport de Recherche 6048, INRIA, Nov. 2006.

- F. Lafarge, P. Trontin, X. Descombes, J. Zerubia et M. Pierrot-Deseilligny. An automatic building extraction method : Application to the 3D-city modeling. Rapport de Recherche 5925, INRIA, Mai 2006.
- F. Lafarge, X. Descombes, J. Zerubia et M. Pierrot-Deseilligny. A parametric model for automatic 3D building reconstruction from HR satellite images. Rapport de Recherche 5687, INRIA, Sept. 2005.
- F. Lafarge, X. Descombes et J. Zerubia. Détection de feux de forêt par analyse statistique de la radiométrie d'images satellitaires. Rapport de Recherche 5369, INRIA, Nov. 2004.
- F. Lafarge, X. Descombes et J. Zerubia. Noyaux texturaux pour les problèmes de classification par SVM en télédétection. Rapport de Recherche 5370, INRIA, Nov. 2004.
- F. Lafarge. Détection de feux de forêt par analyses radiométriques et texturales. Rapport de DEA, SUPAERO, Sep. 2004.

Workshops

- Une approche structurelle pour la modélisation urbaine en 3D à partir d'images satellitaires THR et de données cadastrales. Atelier PNTS "La très haute résolution spatiale en télédétection", Nantes, France, Sep. 2007.
- 3D city modelling using RJMCMC sampler. EURANDOM workshop on image analysis and inverse problems, Eindhoven, Pays-Bas, Dec. 2006.
- A structural approach for automatic 3D city modeling. Mathematical methods for multi-channel image processing workshop (MULTIM), Pékin, Chine, Juil. 2006.
- Urban 3D reconstruction from PLEIADES simulations. 6th WORKSHOP CNES-DLR, DLR, Oberpfaffenoffen, Allemagne, Nov. 2005.
- Textural kernel for SVM classification applied to remote sensing. 5th WORKSHOP CNES-DLR, DLR, Oberpfaffenoffen, Allemagne, Nov. 2004.

Séminaires et communications diverses

- Reconstruction en 3D de bâtiments par approche stochastique. DGA, Arcueil, France, Janv. 2007.
- Reconstruction en 3D de bâtiments à partir de simulations PLEIADES. Journées CNES jeunes chercheurs (JC^2), Toulouse, France, Oct. 2006.
- Modèles paramétriques pour la reconstruction en 3D de zones urbaines à partir de données satellitaires. Journées de la recherche IGN, Saint-Mandé, France, Mars 2006.
- Parametric models for building 3D reconstruction using satellite data. MTA-SZTAKI, Académie des Sciences, Budapest, Hongrie, Dec. 2005.
- Parametric 3D city models using satellite data. University of science, Szeged, Hongrie, Nov. 2005.

Chapitre 6

Annexes

A - Processus 3D paramétriques à partir d'emprises rectangulaires

Deux modèles de reconstruction 3D, à partir d'un MNE et des emprises rectangulaires associées, sont présentés dans cette annexe. Ces deux modèles sont les premiers processus de reconstruction 3D de bâtiments que nous avons proposés. Les résultats obtenus sont limités, surtout au niveau du rendu visuel. Cependant, ces deux modèles sont à l'origine de réflexions qui nous ont amenés à proposer le processus présenté dans le corps de la thèse.

1- Modèle simple

Dans le premier modèle, nous cherchons à fixer sur chaque emprise rectangulaire, un modèle de bâtiments à 6 paramètres. La principale difficulté consiste à gérer les interactions entre les modèles voisins. Pour ce faire, nous utilisons une formulation énergétique bayésienne.

Modèle paramétrique de bâtiments

Le modèle de bâtiments retenu est schématisé sur la figure 6.1. Il est défini à travers 2 paramètres altimétriques (la hauteur de gouttière et la hauteur de l'arête faîtière) et 4 paramètres morphologiques (avec la contrainte que l'aire du plan supérieur soit nulle : $a+b=l$ ou $c+d=L$, sans quoi, la configuration est peu réaliste). Ce modèle est intéressant dans la mesure où il permet de générer différents types de toits, notamment des toits plats ($H_g = H_c$), des bi-plans symétriques et dissymétriques ($a=b=0$ ou $c=d=0$), ou des attaques en V ($a=b=\frac{l}{2}$ et $(c,d)\neq(0,0)$ par exemple).

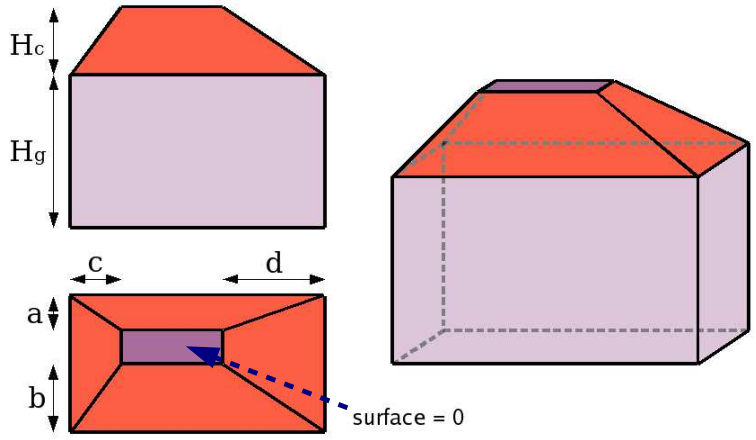

FIG. 6.1 – Modèle 3D paramétrique à base rectangulaire (1^{er} modèle)

Formulation énergétique

L'énergie est mise en place dans un cadre bayésien. Soient y, un MNE marqué de N rectangles, et x, une configuration de N modèles paramétriques dont les supports corres-

pondent aux rectangles de y. Notons $\mathcal{C}$, l'ensemble des configurations x possibles. Alors, nous définissons l'énergie U d'une configuration $x \in \mathcal{C}$, connaissant y, par :

$$U(x) = U_d(x) + U_p(x) \tag{6.1}$$

U_d correspond à l'énergie d'attache aux données. En considérant l'hypothèse d'indépendance conditionnelle des données (c'est-à-dire notamment que nous négligeons le recouvrement des rectangles, hypothèse quelque peu discutable), cette énergie peut être décomposée en une somme d'énergies locales associées à chacun des N modèles paramétriques de la configuration x :

$$U_d(x) = \sum_{x_i \in x} U_d^l(x_i) \tag{6.2}$$

où $U_d^l(x_i)$ correspond à la distance en norme L_1 entre les pixels inclus dans le rectangle i et la côte des points correspondant du modèle x_i (distance au sens de la somme "pixels par pixels").

$U_p(x)$ représente l'énergie a priori de la configuration x. Elle se décompose en 3 termes :

$$U_p(x) = U_p^{(1)}(x) + U_p^{(2)}(x) + U_p^{(3)}(x) \tag{6.3}$$

Les deux derniers termes sont des interactions sur les paires de rectangles voisins. Deux rectangles sont considérés comme voisins si l'intersection de leur sur-rectangle (i.e. rectangle de longueur $L+\varepsilon$ et de largeur $l+\varepsilon$ - en pratique, $\varepsilon = 1m$) est nulle. La figure 6.2 illustre la relation de voisinage que l'on note $\Diamond$.

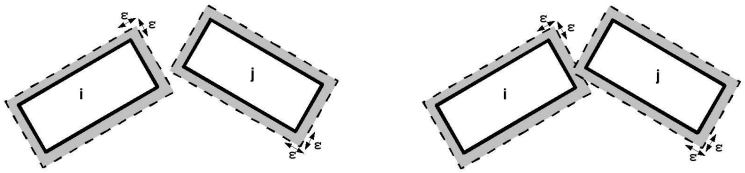

FIG. 6.2 – Illustration de la relation de voisinage - paire de rectangles non voisins *(gauche)*, paire de rectangles voisins *(droite)*.

- $U_p^{(1)}(x)$ est une énergie qui mesure la complexité des N modèles de x. Pour ce faire, les modèles, étant axialement symétriques par rapport à la largeur et/ou la longueur de leur emprise rectangulaire, sont favorisés. C'est le cas des toits les plus courants tels que les toits plats ou les bi-plans. Ainsi, en notant $n_f(x)$ et $n_F(x)$, le nombre de modèles de x ayant respectivement une et deux symétries, et (ω_f, ω_F), un couple de poids à valeurs réelles négatives, nous exprimons $U_p^{(1)}(x)$ par :

$$U_p^{(1)}(x) = \omega_f n_f(x) + \omega_F n_F(x) \tag{6.4}$$

- $U_p^{(2)}(x)$ correspond à des interactions d'ordre 2 permettant l'ajustement des hauteurs de gouttières entre les modèles voisins, tout en autorisant les sauts supérieurs à la hauteur d'un étage (noté H_f - en pratique, 3 mètres). Nous avons :

$$U_p^{(2)}(x) = \sum_{i \Diamond j} f_h(|H_{g_i} - H_{g_j}|) \tag{6.5}$$

où f_h est la fonction réelle représentée sur la figure 6.3, et $i \Diamond j$ représente les paires de modèles voisins.

- $U_p^{(3)}(x)$ favorise la continuité des arêtes faîtières sur un bâtiment, toujours par des interactions d'ordre 2. Plus précisément, ce terme permet l'attraction des extrémités des arêtes faîtières des modèles voisins. En notant e_i et e_j, les extrémités des arêtes faîtières des modèles x_i et x_j (dans $\mathbb{R}^3$), nous avons :

$$U_p^{(3)}(x) = \sum_{i \Diamond j} \omega_r \, d_c(e_i, e_j) \tag{6.6}$$

$d_c(.,.)$ est une fonction correspondant à la distance en norme L_2 dans $\mathbb{R}^3$ si $|(H_{g_i} + H_{c_i}) - (H_{g_j} + H_{c_j})| < H_f$ et valant 0 sinon. ω_r est un poids à valeur réelle positive.

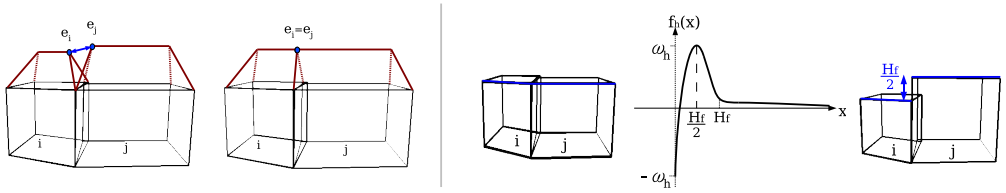

FIG. 6.3 – Illustration des interactions d'ordre 2 - attraction des extrémités des arêtes faîtières *(gauche)*, ajustement des hauteurs de gouttières *(droite)*

Optimisation par la dynamique de Metropolis

L'énergie $U(x)$ est une fonction non convexe qui est définie sur un espace d'état de grande dimension. Afin de trouver la configuration qui minimise U, nous utilisons l'algorithme de Metropolis [Metropolis et al., 1953]. Cette technique procède itérativement selon une relaxation stochastique gérée par une température T_t qui décroît vers 0 quand t tend vers l'infini. Si à l'itération t, la configuration courante est $x^{(t)} = x$, alors l'algorithme consiste à :

- choisir aléatoirement le rectangle i (parmi les N)
- tirer un nouveau jeu de paramètres $\widehat{x_i}$, permettant de définir une nouvelle configuration d'objets $\widehat{x} = (x_1, ..., x_{i-1}, \widehat{x_i}, x_{i+1}, ..., x_N)$
- accepter $x^{(t+1)} = \widehat{x}$
  - si $U(\widehat{x}) - U(x) \leq 0$
  - ou si $U(\widehat{x}) - U(x) > 0$ et $p < \exp - \frac{U(\widehat{x}) - U(x)}{T_t}$ où p est tiré uniformément entre $[0,1]$
- sinon accepter $x^{(t+1)} = x$
- mettre à jour la température T_t

Nous choisissons une décroissance de température géométrique, détaillée dans [Van Laarhoven et Aarts, 1987]. En pratique, nous n'avons aucune garantie d'atteindre la configuration optimale, mais une configuration relativement proche. Les temps de calcul sont, de plus, très importants. Le deuxième modèle, présenté ci-dessous, permet de réduire efficacement les temps de calcul.

2- Modèle à objets multiples

Cette deuxième méthode est une version améliorée de la première, utilisant la même approche énergétique. Elle diffère par le modèle 3D paramétrique de bâtiments mis en place, et également par la technique d'optimisation utilisée. La fonction énergétique est, elle, inchangée.

Vers une collection de modèles paramétriques

Le principal inconvénient du modèle 3D paramétrique présenté sur la figure 6.1 est qu'il est trop générique. La plupart des jeux de paramètres correspondent à des modèles 3D peu réalistes. Les jeux de paramètres intéressants représentant des toits plats, des bi-plans et des attaques en V, sont tirés de manière peu fréquente par l'algorithme de Metropolis.

La solution consiste à se limiter à une collection de modèles 3D spécifiques tels que les toits plats, les bi-plans et les attaques en V. De cette manière, nous restreignons fortement la taille de l'espace d'états. La figure 6.4 représente une collection de modèles correspondant à des types de toit de forte occurrence (6 au total). Ils possèdent de 1 à 5 paramètres. H_g et H_c représentent toujours la hauteur de gouttière et la hauteur de l'arête faîtière du modèle. α, β et γ sont des paramètres de formes contenus dans $[0,1]$. ϕ oriente le modèle par rapport à son support 2D (il prend 4 valeurs possibles). Cette collection de modèles permet de réduire fortement l'espace des configurations.

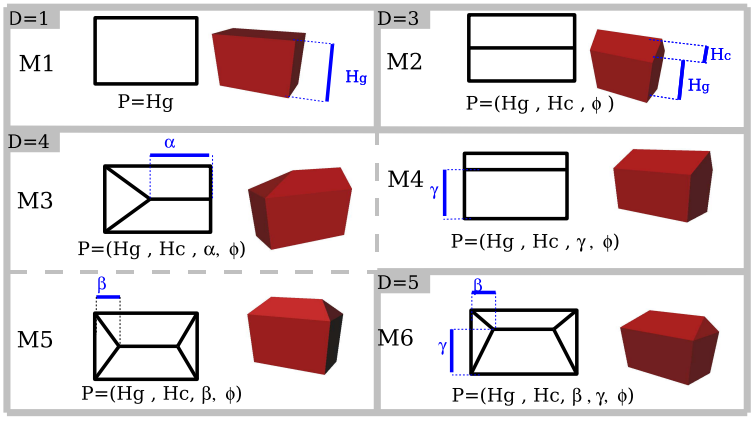

FIG. 6.4 – Collection de Modèles 3D "spécifiques" à base rectangulaire

Optimisation par RJMCMC

Nous cherchons toujours à optimiser l'énergie U. Cependant, l'espace de recherche de la solution est désormais une union de sous-espaces de dimension variée. En effet, les modèles 3D de la collection ont un nombre de paramètres variant entre 1 et 5.
Nous échantillonnons notre espace d'états par l'algorithme de Green [Green, 1995] que nous couplons à un recuit simulé. Cette technique d'optimisation est utilisée à plusieurs

reprise dans le manuscrit, c'est pourquoi nous ne nous attarderons pas à la détailler ici. Nous précisons juste que les sauts sont proposés de manière uniforme sur l'espace des configurations, et que le schéma de décroissance de température est identique à celui utilisé dans le premier modèle.

3- Résultats

Les résultats obtenus par ces deux modèles sont visuellement identiques, le deuxième modèle étant une version améliorée du premier dans lequel nous avons restreint l'espace des configurations. Le deuxième modèle a des temps de calcul réduit d'un facteur 3 environ par rapport au premier modèle. Pour un bâtiment composé d'une dizaine d'emprises rectangulaires, environ 5 minutes seront nécessaires pour le premier modèle, tandis que le second mettra un peu moins de 2 minutes, sur un PC standard (2 Ghz).

La figure 6.5 présente les reconstructions de deux bâtiments. Ces résultats donnent plusieurs points de satisfaction. D'abord, la forme générale des bâtiments est globalement respectée comparativement aux vérités terrain associées. Ensuite, le terme de régularisation permet un ajustement efficace des hauteurs de gouttière et une bonne connexion des extrémités des arêtes faîtières.

Cependant, ces modélisations possèdent des défauts très pénalisants. Les discontinuités des hauteurs de toits ne sont pas retranscrites. L'inconvénient majeur est la présence de nombreux artefacts, dus à une mauvaise connexion des supports rectangulaires voisins. Des algorithmes de simplifications de formes 3D, comme celui proposé par [Kada, 2006], ne sont pas adaptés à ce type d'artefacts, et ne sauraient résoudre le problème. Ce défaut ne peut pas être traité dans le processus de reconstruction 3D : la solution qui s'impose consiste à repenser les supports 2D. Nous abordons ce problème dans la partie 2.2. Les résultats présentés sur la figure 6.6 soulignent la présence importante des artefacts, pénalisant fortement le rendu visuel des reconstructions. Les évaluations altimétriques sont toutefois correctes. L'erreur quadratique moyenne en Z est de 3,2 mètres.

L'énergie inhérente à ces deux modèles est régie par un jeu de paramètres pondérateurs. Ce jeu comporte 4 potentiels ($\omega_f, \omega_F, \omega_h$ et ω_r), particulièrement difficiles à régler. Cela représente un inconvénient majeur qui rend le processus de reconstruction peu robuste. La solution consiste à proposer une énergie plus pertinente, et dépendant de moins de paramètres.

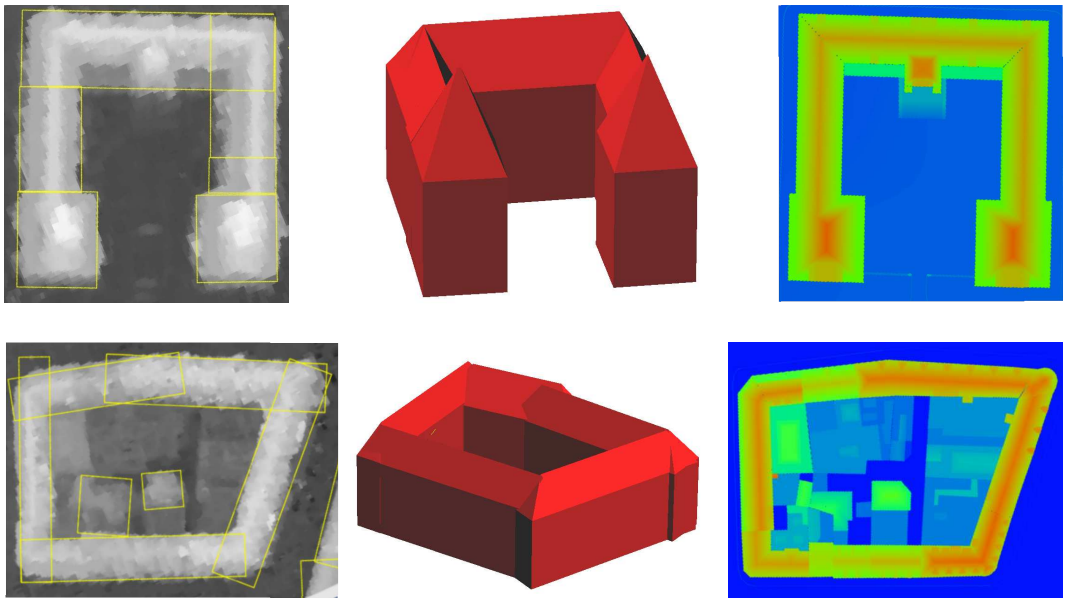

FIG. 6.5 – Résultat sur deux bâtiments - emprises rectangulaires *(gauche)*, résultats 3D *(centre)*, vérités terrain *(droite)*.

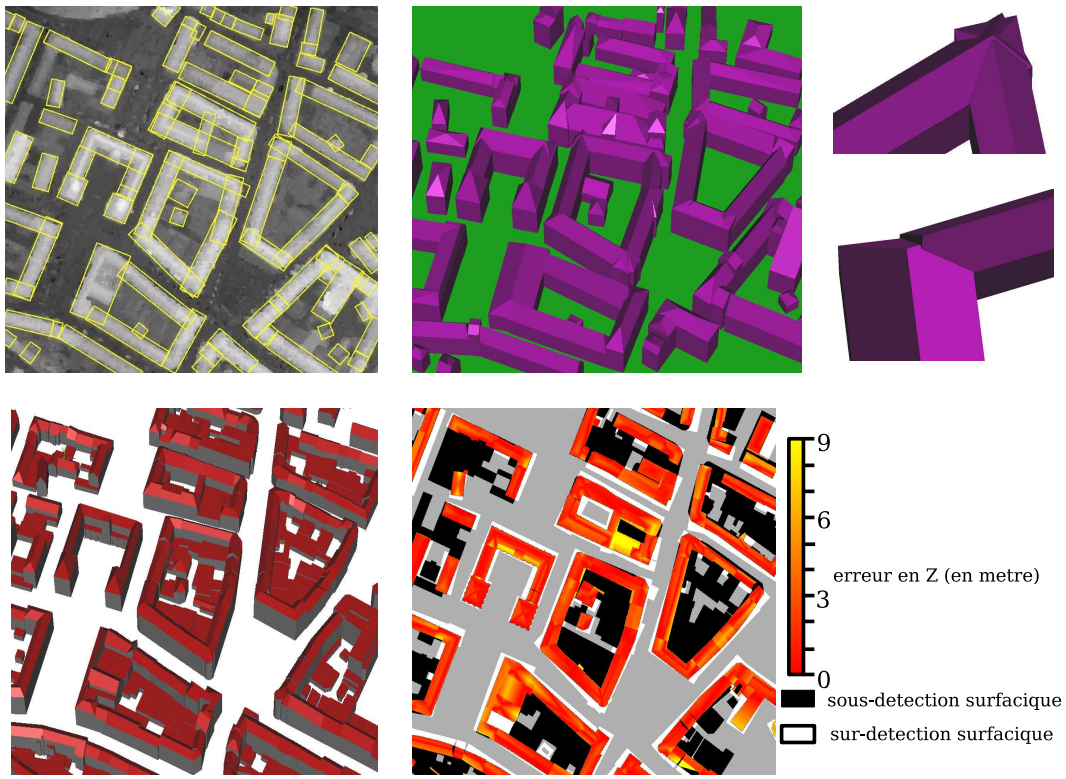

FIG. 6.6 – Résultat sur une zone dense - emprises rectangulaires, résultat 3D et exemples d'artefacts *(haut)*, vérité terrain et carte d'évaluation *(bas)*.

B - Processus 3D par squelétisation des arêtes faîtières

La méthode de reconstruction 3D présentée dans cette annexe est fondée sur un processus de squelétisation des arêtes faîtières à partir de supports structuraux (présentés dans le chapitre 2). Cette approche est relativement simple, rapide et robuste. Cependant, la complexité de la modélisation 3D est de bas niveau. Les toits sont, en effet, limités à des bi-plans symétriques et des toits plats. Les seules inconnues du modèle sont la hauteur de gouttière et la hauteur de faîtière à déterminer pour chaque support 2D.

1-Squelétisation des supports 2D

La forme des toits, et plus précisément la disposition des arêtes faîtières, est entièrement définie par un algorithme de squelétisation des supports 2D. Pour un bâtiment, nous schématisons les relations de connexion entre supports par un graphe dont chaque noeud correspond à un support. La méthode consiste à rechercher le plus long chemin simple, lequel va permettre de définir l'arête faîtière principale du bâtiment (voir la figure 6.7 - chaîne rouge). Ensuite, les noeuds non parcourus sont traités indépendemment, en squelétisant les supports associés (voir la figure 6.7 - chaînes bleues).

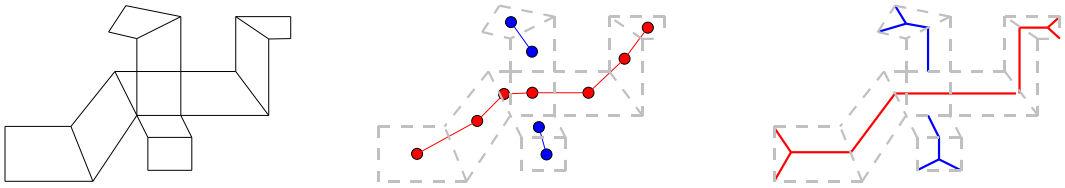

FIG. 6.7 – Illustration de la squeletisation des supports - la chaîne principale en rouge, chaînes secondaires en bleu

2-Estimation des différentes hauteurs

Pour chaque support 2D, nous estimons la hauteur de gouttière ainsi que la hauteur d'arête faîtière en nous focalisant sur des régions d'intérêt à l'intérieur du support. Ces masques sont illustrés sur la figure 6.8, où $b = \frac{a}{10}$ et $c = \frac{a}{4}$. Les valeurs des pixels du MNE contenus dans ces deux masques sont ensuite classées de manière décroissante dans deux tables. L'estimation de la hauteur de gouttière est donnée par la moyenne des valeurs rangées entre 20% et 80% de toutes les valeurs de la table des gouttières. L'estimation de la hauteur d'arête faîtière est, quant-à elle, effectuée en calculant la moyenne des valeurs rangées entre 10% et 20% de toutes les valeurs de la table des arêtes faîtières.

3-Régularisation des hauteurs

Pour chaque support d'un bâtiment, nous avons estimé les deux inconnues du modèle, à savoir la hauteur de gouttière ainsi que la hauteur de faîtière. Or, chaque support a un couple unique d'inconnues. Il est nécessaire de régulariser ces valeurs entre les supports voisins, c'est-à-dire faire en sorte que la continuité du toit d'un support à un autre soit préservée s'il

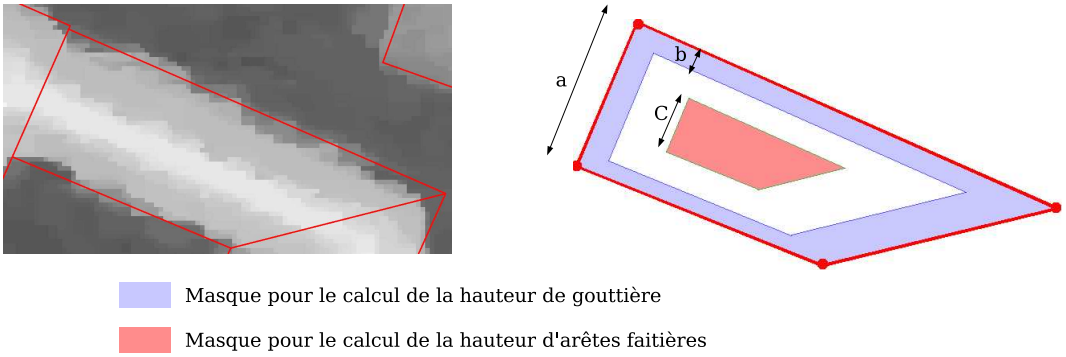

FIG. 6.8 – Illustration des masques sur lesquels les différentes hauteurs sont estimées

y a lieu d'être. La figure 6.9 présente un bâtiment non régularisé, ainsi que son homologue régularisé par la méthode que nous proposons.

Nous utilisons, pour chaque groupe de supports connectés, l'algorithme du "K-means", présenté dans [Cocquerez et Philipp, 1995], incluant un terme permettant de pénaliser les nombres de classes élevés. Cet algorithme s'apparente à celui proposé par [Palubinskas et al., 1998] utilisant un terme entropique pour minimiser le nombre de classes. L'avantage d'un tel algorithme est que le nombre de classes n'est pas connu a priori.

Considérons un ensemble de supports connectés composé de N éléments. Notons H_{g_i} l'estimation de la hauteur de gouttière du i^{eme} élément de cet ensemble. L'algorithme du "K-means" avec terme de régularisation consiste à rechercher les c_j, centroïde de la j^{eme} classe, qui maximisent le critère J définie par :

$$J = \sum_{j=1}^{C} \sum_{i=1}^{N} d^2(Hg_i, c_j) - \alpha C \tag{6.7}$$

où C représente le nombre de classes (non connu a priori). En pratique, on calcule ce terme pour chaque possibilité de nombre de classes (c'est-à-dire pour chaque $C \in [1,N]$). On conserve alors le nombre de classes et les centroides c_j associés qui minimisent le critère J.

Le critère J est composé de deux termes. Le premier correspond au critère classique de l'algorithme du "K-means". Le second, qui est pondéré par le paramètre $\alpha \in \mathbb{R}^{+\star}$, représente un a priori sur les classes. Ce terme favorise linéairement les nombres de classes faibles. Finalement, le centroïde c_j du i^{eme} support correspond à la valeur régularisée de la hauteur de gouttière pour ce support. Le paramètre α a été fixé à 90. Une valeur trop faible de α rend la modélisation 3D mal régularisée. A l'inverse, un α trop élevé uniformisera à l'excès les toits et fera disparaître les discontinuités des hauteurs de toit détectées dans le chapitre 2 du manuscrit. Le même procédé est appliqué pour régulariser les hauteurs d'arête faîtière.

4-Quelques résultats

Les figures 6.10 et 6.11 présentent les résultats de cette méthode. Les modélisations obtenues, bien que très simples, fournissent une description générale des bâtiments qui peut être intéressante pour des applications exigeant de la robustesse et un faible temps de calcul.

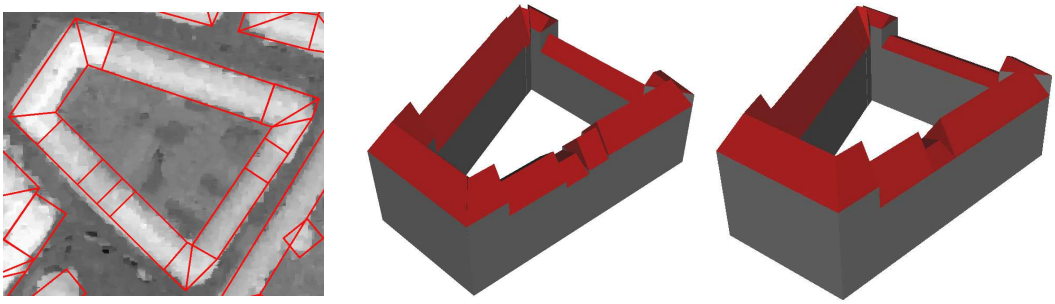

FIG. 6.9 – Illustration de la régularisation des hauteurs - supports 2D et MNE *(gauche)*, modélisation non régularisée *(centre)*, modélisation régularisée *(droite)*.

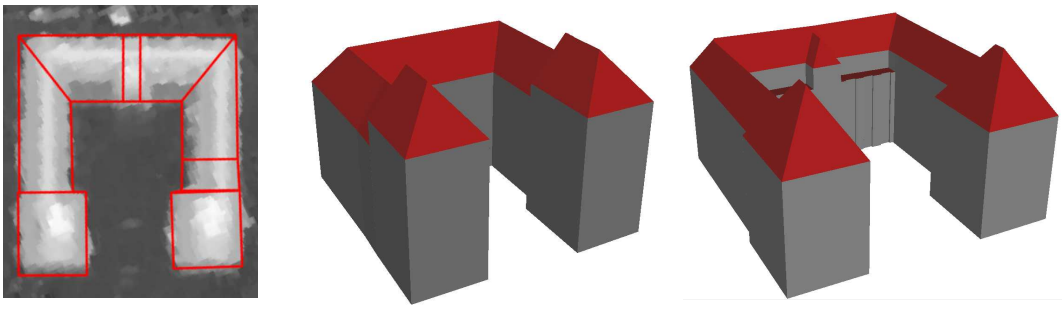

FIG. 6.10 – Résultat sur l'hôtel de ville d'Amiens - supports 2D et MNE *(gauche)*, modélisation obtenue *(centre)*, vérité terrain 3D *(droite)*.

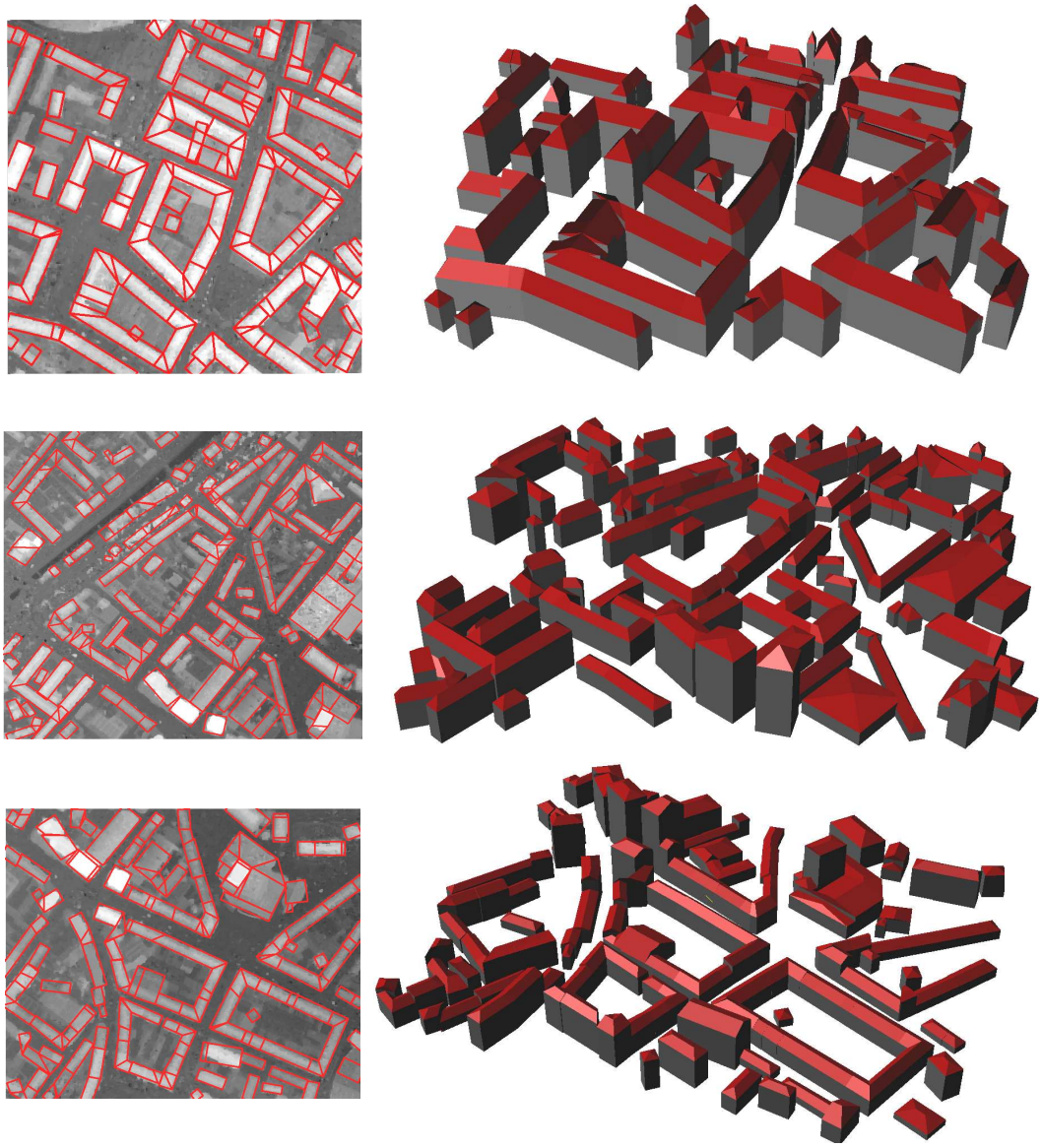

FIG. 6.11 – Résultats sur trois zones urbaines - supports 2D et MNE *(gauche)*, modélisations 3D *(droite)*.

C - Rappel sur les chaînes de Markov

La mise en place des algorithmes d'échantillonnage présentés dans la partie 3.3.1 du manuscrit reposent sur la construction d'une chaîne de Markov convergeant vers la mesure à échantillonner. Cette annexe fournit des rappels sur les chaînes de Markov et leur convergence, notions utiles pour la compréhension du fonctionnement de ces échantillonneurs.

1- Définitions

Définition 2 *Une suite de variables aléatoires (X_n) à valeurs dans un espace E muni de sa tribu $\mathcal{B}$ est une* ***chaîne de Markov*** *si :*

$$p(X_{t+1} \in A | X_0 = x_0, ..., X_t = x_t) = p(X_{t+1} \in A | X_t = x_t) \quad \forall A \in \mathcal{B} \tag{6.8}$$

L'évolution d'une chaîne de Markov (c'est-à-dire le passage de X_t à X_{t+1}) a ainsi la propriété de ne pas dépendre du passé de la chaîne, mais uniquement de son état courant $(X_t = x_t)$. Cette propriété est particulièrement intéressante lors de la mise en œuvre informatique puisqu'il n'est pas nécessaire de garder en mémoire les états passés du processus pour effectuer les calculs.
Une chaîne de Markov est homogène si son évolution est indépendante de l'instant t. Dans la suite, nous considérons uniquement ce type de chaînes de Markov.

Définition 3 *Un* ***noyau de transition*** *est une fonction P définie sur $E \times \mathcal{B}$ telle que :*
- *$\forall x \in E$, $P(x,.)$ est une mesure de probabilité*
- *$\forall A \in \mathcal{B}$, $P(.,A)$ est mesurable.*

Le noyau de transition P associé à une chaîne de Makov homogène est donné par :

$$P(x,A) = p(X_{t+1} \in A | X_t = x) \tag{6.9}$$

2- Stationnarité

La propriété de stationnarité est essentielle pour un échantillonneur de type Monte Carlo par chaîne de Markov. La chaîne à construire (X_t) doit en effet avoir une mesure stationnaire π. Cette propriété est équivalente à l'invariance de π par rapport à (X_t).

Propriété 1 *Une mesure π est* ***stationnaire*** *ou* ***invariante*** *pour la chaîne de Markov de noyau de transition P si :*

$$\pi(A) = \int P(x,A) d\pi(x) \quad \forall A \in \mathcal{B} \tag{6.10}$$

Afin d'établir la stationnarité de π pour un noyau P, on vérifie la réversibilité de la chaîne de Markov. En effet, la réversibilité étant plus facile à vérifier que la stationnarité, la plupart des algorithmes d'échantillonnage vérifient la condition de réversibilité qui implique l'invariance pour π.

Propriété 2 *Une chaîne de Markov est* ***réversible*** *si son noyau de transition P vérifie l'équation :*

$$\int_A P(x,B) d\pi(x) = \int_B P(y,A) d\pi(y) \quad \forall A, B \in \mathcal{B} \tag{6.11}$$

En d'autres termes, la probabilité de passer de A à B est, sous π, est la même que la probabilité de passer de B à A.

3- Convergence

Dans cette partie, nous présentons les notions d'*irréductibilité*, d'*apériodicité*, de *récurrence* et d'*ergodicité* qui vont permettre de garantir la convergence de la chaîne de Markov vers sa mesure stationnaire. La convergence de la chaîne de Markov vers sa mesure invariante π s'exprime par :

$$P^t(x,A) \underset{t\to\infty}{\longrightarrow} \pi(A) \qquad \forall x \in E, A \in \mathcal{B} \tag{6.12}$$

où $P^t(x,A) = p(X_t \in A | X_0 = x)$.

Irréductibilité

Propriété 3 *Une chaîne de Markov (X_t) sur un espace mesurable $(E,\mathcal{B})$ est ϕ-**irréductible** s'il existe une mesure non nulle ϕ sur $\mathcal{B}$ telle que $\forall x \in E, A \in \mathcal{B}$,*

$$\phi(A) > 0 \Rightarrow \exists k \in \mathbb{N} : P^k(x,A) > 0 \tag{6.13}$$

La convergence d'une chaîne de Markov nécessite sa π-irréductibilité (également appelée plus simplement irréductibilité). L'irréductibilité signifie que la chaîne a une probabilité non nulle d'atteindre en temps fini tout ensemble π-*probable* quelque soit la condition initiale.
Si une chaîne de Makov (X_t) est ϕ-irréductible et si elle possède une mesure invariante π, alors :

- π est l'unique mesure invariante pour le noyau P
- (X_t) est π-irréductible
- $\pi(A) = 0 \Rightarrow \phi(A) = 0$

La π-irréductibilité d'une chaîne peut donc être démontrée à travers une étude sur une autre mesure.

Apériodicité

Propriété 4 *Soit (X_t), une chaîne de Markov π-irréductible. Les ensembles $A_1,...,A_m$ de $\mathcal{B}$ forment un m-cycle si :*

$$\begin{array}{lll} x \in A_1 & \Longrightarrow & P(x,A_2) = 1 \\ & \vdots & \\ x \in A_{m-1} & \Longrightarrow & P(x,A_m) = 1 \\ x \in A_m & \Longrightarrow & P(x,A_1) = 1 \\ et & \pi(\cup_{i=1}^m A_i) = 1 & \end{array}$$

*Le plus grand entier d pour lequel un d-cycle est formé est appelé la période de la chaîne. Si $d = 1$, la chaîne est dite **apériodique**.*

L'apériodicité de la chaîne est nécessaire pour avoir sa convergence. Pour avoir en pratique l'apériodicité d'une chaîne, on vérifie la condition suffisante suivante :

$$\exists x \in E : P(x,\{x\}) > 0 \tag{6.14}$$

Si cette condition est vérifiée, la chaîne est dite **fortement apériodique**. L'apériodicité d'une chaîne est donc toujours vérifiée dès que la probabilité de rester dans l'état courant est positive (c'est-à-dire avec une probabilité de rejet non nulle dans les algorithmes de type Metropolis-Hastings).

Lorsque la chaîne est π-irréductible et apériodique, on obtient la convergence pour presque tout $x \in E$:

$$\|P^t(x,.) - \pi\| \underset{t\to\infty}{\longrightarrow} 0 \quad \forall x \in E \backslash N \tag{6.15}$$

où $N \in \mathcal{B}$ tel que $\pi(N) = 0$, et $\|.\|$ représente la norme en variation totale (c'est-à-dire $\|\mu_1 - \mu_2\| = \sup_A |\mu_1(A) - \mu_2(A)|$).

Récurrence au sens de Harris

Afin d'introduire la convergence indépendamment de toute condition initiale, la notion de récurrence au sens de Harris est introduite.

Propriété 5 *une chaîne de Markov est* ***récurrente au sens de Harris*** *si :*

$$p(\{\exists t / X_t \in A\} | X_0 = x) = 1 \quad \forall x \in E\,, \forall A \in \mathcal{B} / \pi(A) > 0 \tag{6.16}$$

La récurrence au sens de Harris d'une chaîne implique son irréductibilité (c'est-à-dire sa π-irréductibilité où π est la mesure stationnaire de la chaîne). Si l'espace d'état est fini, l'irréductibilité et la récurrence au sens de Harris sont équivalentes.

Ergodicité

Propriété 6 *Une chaîne de Markov de mesure stationnaire* π ***converge ergodiquement*** *vers* π *si elle est apériodique et récurrente au sens de Harris. L'ergodicité d'une chaîne est équivalente à :*

$$\|P^t(x,.) - \pi\| \underset{t\to\infty}{\longrightarrow} 0 \quad \forall x \in E \tag{6.17}$$

La convergence ergodique d'une chaîne permet d'assurer la convergence d'une mesure vers une autre mesure indépendamment de la condition initiale.

Pour vérifier la récurrence au sens de Harris dans le cas des espaces d'état non finis, on utilise le concept des ensembles petits.

Définition 4 *Un ensemble* $C \in \mathcal{B}$ *est dit* ***petit*** *s'il existe un entier m, un réel* $\varepsilon > 0$ *et une mesure de probabilité* κ *sur* $\mathcal{B}$ *tels que :*

$$P^m(x,A) \geq \varepsilon \kappa(A) \quad \forall x \in C\,, \forall A \in \mathcal{B} \tag{6.18}$$

Propriété 7 *Soit (X_t), une chaîne de Markov irréductible et apériodique. Soient C, un ensemble petit de $\mathcal{B}$, et une fonction $V : E \to \mathbb{R}$ telle que l'ensemble $\{x/V(x) \leq n\}$ est petit pour n'importe quel $n \in \mathbb{N}$ et que :*

$$\mathbb{E}[V(X_1)|X_0 = x] \leq V(x) \quad \forall x \in E \backslash C \tag{6.19}$$

où $\mathbb{E}[V(X_1)|X_0 = x]$ représente l'espérance de $V(X_1)$ sous la mesure de probabilité P(x,.). Alors la chaîne est récurrente au sens de Harris.

L'équation 6.19 est appelée la **condition de drift** pour la récurrence.

Ergodicité géométrique

Propriété 8 *Une chaîne de Markov est géométriquement ergodique s'il existe une constante $r > 1$ telle que*

$$\sum_{t=1}^{\infty} r^t \|P^t(x,.) - \pi\| < \infty \quad \forall x \in E \tag{6.20}$$

L'ergodicité géométrique est une propriété plus forte que l'ergodicité. Elle impose que la convergence de P^t vers π se fasse avec une vitesse géométrique puisque l'équation 6.20 implique l'inéquation suivante :

$$\|P^t(x,.) - \pi\| \leq M r^{-t} \tag{6.21}$$

avec $M = \sum_{t=1}^{\infty} r^t \|P^t(x,.) - \pi\|$. L'équation 6.20 peut être démontrée par une condition de drift géométrique consistant à démontrer l'existence d'une fonction $V :\to [1,\infty[$, des constantes $b < \infty$ et $\lambda < 1$, et un ensemble petit $C \in \mathcal{B}$ tel que :

$$\mathbb{E}[V(X_1)|X_0 = x] \leq \lambda V(x) + b\mathbb{1}_C(x) \quad \forall x \in E \backslash C \tag{6.22}$$

Le principal intérêt de la propriété d'ergodicité géométrique réside dans l'établissement du Théorème de la Limite Centrale, particulièrement utile pour contrôler la convergence des algorithmes de type MCMC.

Théorème de la limite centrale

Soit $g(X)$, une statistique d'intérêt telle que g soit π intégrable :

$$\mu_g = \mathbb{E}_\pi[g(X)] = \int g(x) d\pi(x) \tag{6.23}$$

où π est la mesure stationnaire d'une chaîne de Markov récurrente au sens de Harris.
Soit également une estimée $\widehat{g_t}$ obtenue à partir d'une réalisation tronquée quelconque $X_1, ..., X_t$:

$$\widehat{g_t} = \frac{1}{t} \sum_{i=1}^{t} g(X_i) \tag{6.24}$$

Si la fonction g satisfait la condition de Lyapunov :

$$\int |g(x)|^{2+\varepsilon} d\pi(x) < \infty \tag{6.25}$$

avec $\varepsilon > 0$, alors le Théorème de la Limite Centrale implique :

$$\sqrt{t}(\widehat{g_t} - \mu_g) \rightarrow \mathcal{N}(0, \gamma^2) \tag{6.26}$$

où :

$$\gamma^2 = Var_\pi(g(X)) + 2\sum_{i=1}^{\infty} Cov_\pi(g(X_t)g(X_{t+i})) \tag{6.27}$$

Le Théorème de la Limite Centrale permet de contrôler la convergence des algorithmes MCMC dans la mesure où, lorsque $\gamma^2 > 0$, on peut vérifier si les moyennes $\widehat{g_t}$ convergent vers la quantité μ_g. La principale difficulté de ce résultat est l'estimation de la quantité γ^2. Par ailleurs, on ne dispose pas d'indication concernant le nombre d'échantillons nécessaires pour réaliser un calcul fiable.

D - Calcul des noyaux de proposition

La bibliothèque $\mathcal{M}$ est composée de 26 modèles. Parmi ces modèles, il y a 12 jeux de paramètres différents. Il s'agit donc de déterminer $12^2 = 144$ bijections et paramètres de complétion associés. Cela est long et fastidieux, c'est pourquoi nous détaillons dans cette annexe quelques cas variés de calculs.

Pour reprendre le problème, on considère deux modèles $\mathcal{M}_m$ et $\mathcal{M}_n$ et une perturbation d'un objet $x = (m, \theta)$ par $\widehat{x} = (n, \widehat{\theta})$. Le principe est alors de créer une bijection entre l'espace des paramètres des modèles $\mathcal{M}_m$ et $\mathcal{M}_n$. Pour cela, θ est complété en simulant $u_{mn} \sim \varphi_{mn}(.)$ à travers (θ, u_{mn}), et $\widehat{\theta}$ en simulant $v_{nm} \sim \varphi_{nm}(.)$ à travers $(\widehat{\theta}, v_{mn})$ de sorte que l'application Ψ_{mn} entre (θ, u_{mn}) et $(\widehat{\theta}, v_{mn})$ soit une bijection (c'est-à-dire $(\widehat{\theta}, v_{mn}) = \Psi_{mn}(\theta, u_{mn})$).

L'objectif est alors, connaissant les deux modèles, de déterminer la bijection permettant de passer paramétriquement de l'un à l'autre, afin notamment de calculer le jacobien associé présent dans le taux de Green. Cette bijection dépend de la manière dont on a défini les paramètres de complétion u_{mn} et v_{mn}. Les paramètres utilisés dans les modèles ont été définis de manière à simplifier au maximum le calcul des noyaux (leur domaine de définition est principalement l'intervalle unitaire $[0,1]$).

Considérons, dans un premier cas, un saut du modèle $(\mathcal{F}_{11}, \mathcal{V}_-)$ (modèle noté $\mathcal{M}_1$) vers le modèle $(\mathcal{F}_{21}, \mathcal{V}_-)$ (noté $\mathcal{M}_2$). Paramétriquement, on saute de $\theta = H_g$ vers $\widehat{\theta} = (H_g, H_t, \phi_1)$. Le paramètre H_g étant commun aux deux modèles, il s'agit uniquement de compléter le modèle $\mathcal{M}_1$ par $u_{12} = (H_t, \phi_1)$. On a ainsi $\widehat{\theta} = \Psi_{12}(\theta, u_{12})$ avec Ψ_{12} correspondant à la fonction identité : le déterminant du jacobien vaut 1. Il faut noter que, grâce à la condition de réversibilité de la chaîne, le taux de Green du saut inverse se calcule simplement en prenant l'inverse du taux de Green du saut normal (cela divise par deux le nombre de calculs des bijections).

Considérons maintenant un second cas dans lequel on effectue un saut du modèle terrasse $(\mathcal{F}_{12}, \mathcal{V}_-)$ (noté $\mathcal{M}_1$) vers le modèle bi-plan dissymétrique avec attaque en V $(\mathcal{F}_{22}, \mathcal{V}_V)$ (noté $\mathcal{M}_2$). On a $\theta = (H_g, H_t, \xi_1, \xi_2, \xi_3, \xi_4)$ et $\widehat{\theta} = (H_g, H_t, \zeta, \phi_2, \eta)$. Les paramètres communs sont H_g et H_t. De plus, il existe, quelque soit le noyau utilisé parmi les trois possibles, une relation directe identitaire entre les paramètres ξ_1, ξ_2 d'un côté et ζ, η de l'autre. Il n'est, par contre, pas possible de trouver une bijection entre ξ_3 et ϕ_2. Ainsi les paramètres de complétion sont $u_{12} = \phi_2$ et $v_{21} = (\xi_3, \xi_4)$. On a $(\widehat{\theta}, v_{21}) = \Psi_{12}(\theta, u_{12})$ où Ψ_{12} correspond, comme dans le premier cas, à la fonction identité et dont le déterminant du jacobien vaut 1. Dans le cas du noyau uniforme par exemple, on aura dans le taux de Green $\varphi_{21}^{(1)}(v_{21}) = \mathcal{U}_{[0,1]^2}(v_{21}) = 1$ en considérant que ξ_3 et ξ_4 sont indépendants, et $\varphi_{12}^{(1)}(u_{12}) = \frac{1}{4}$.

E - Processus 3D après simplification séquentielle des supports 2D

Cette annexe présente un modèle simplifié de la méthode de reconstruction présentée dans le chapitre 3. L'idée consiste à considérer chaque bâtiment comme une association de séquences d'objets (voir figure 3.10). Pour ce faire, nous utilisons un principe similaire à celui utilisé dans l'annexe B-1. Pour un bâtiment donné, nous schématisons les relations de connexion entre supports par un graphe dont chaque noeud correspond à un quadrilatère. La méthode consiste à rechercher le plus long chemin simple dans ce graphe, lequel correspond à la première séquence extraite. On met alors à jour le nouveau graphe en considérant uniquement les noeuds non parcourus, puis on y extrait le plus long chemin simple. Et l'on itére le procédé jusqu'à ce que le graphe ne contienne plus de noeuds.

Dès lors, l'objectif est d'optimiser, suivant l'énergie U mise en place dans la partie 3.2, chaque séquence d'objets de manière indépendante. Pour chaque séquence de N supports, le problème est modélisé par une chaîne de Markov cachée homogène $(X_t, Y_t)_{t=1}^{N}$ dont le graphe de dépendance est donné par la figure 3.11. Les probabilités de transition de la chaîne $P(X_{t+1} = x_{t+1} | X_t = x_t)$ correspondent aux intéractions de cardinal deux de l'énergie U, et les probabilités des observations conditionnellement aux états $P(Y_t = y_t | X_t = x_t)$ sont données par les vraisemblances locales associées à U :

$$P(y_t/x_t) = \frac{1}{Z_t^l} \exp -\Gamma^{\alpha}_{(t)}(\mathcal{S}_{x_t}, y_t) \tag{6.28}$$

$$P(x_{t+1}|x_t) = \frac{1}{Z_t^p} \exp -\mathbb{1}_{\{x_{t+1} \sim_a x_t\}} \beta g(x_{t+1}, x_t) \tag{6.29}$$

où Z_t^l et Z_t^p sont des constantes de normalisations.

L'algorithme de Viterbi [Viterbi, 1967], qui est une méthode d'optimisation déterministe adaptée aux problèmes causaux, permet de trouver la séquence optimale d'objets de la chaîne de Markov cachée. Il peut être utiliser avec des distributions non normalisées, ce qui évite le calcul des constantes de normalisation Z_t^l et Z_t^p pour chaque élément t de la chaîne. Notons $\delta_t(i)$, la probabilité du chemin optimal qui permet d'atteindre l'état i du t^{ieme} support de la séquence, connaissant les observations $\mathcal{Y}$ ($\delta_t(i)$ représente la probabilité du MAP de la sous-chaîne finissant à t avec X_t fixé).

$$\delta_t(i) = \max_{x_1,\ldots,x_{t-1}} P(y_1,\ldots,y_{t-1}, X_1 = x_1, \ldots, X_{t-1} = x_{t-1}, X_t = i) \tag{6.30}$$

Par récurrence, δ_{t+1} peut être déterminé à partir de δ_t :

$$\begin{aligned}
\delta_{t+1}(j) &= \max_{x_1,\ldots,x_t} P(y_1,\ldots,y_t, X_1 = x_1, \ldots, X_t = x_t, X_{t+1} = j) \\
&= \max_{x_t} \left[P(X_{t+1} = j | X_t = x_t) \max_{x_1,\ldots,x_{t-1}} P(y_1,\ldots,y_t, X_1 = x_1, \ldots, X_t = x_t) \right] \\
&= \max_{i} \left[a_{ij} P(y_t | X_t = i) \max_{x_1,\ldots,x_{t-1}} P(y_1,\ldots,y_{t-1}, X_1 = x_1, \ldots, X_{t-1} = x_{t-1}, X_t = i) \right]
\end{aligned}$$

Nous obtenons ainsi la relation suivante :

$$\delta_{t+1}(j) = \max_i \left[a_{ij} \mathcal{L}_i(y_t) \delta_t(i) \right] \tag{6.31}$$

où $\mathcal{L}_i(y_t) = P(y_t | X_t = i)$ représente la vraisemblance locale du t^{ieme} elément de la chaîne, et a_{ij}, la probabilité de transition de l'état i vers l'état j. On note $\Psi_{t+1}(j)$, l'argument qui correspond à la valeur maximale de $\delta_{t+1}(j)$:

$$\Psi_{t+1}(j) = \arg\max_i \left[a_{ij} \mathcal{L}_i(y_t) \delta_t(i) \right] \tag{6.32}$$

L'algorithme de Viterbi consiste à :

- A $t = 1$: initialiser $\delta_1(i) = P(X_1 = i)$, $\forall i$ (dans notre cas, on prend les $P(X_1)$ équiprobables).
- Pour $t = 2...N$: calculer et mémoriser dans une table le couple $(\delta_t(j), \Psi_t(j))$, $\forall j$.
- calculer le MAP donné par : $\widehat{x}_N = \arg\max_i \mathcal{L}_i(y_N) \delta_N(i)$
- Lire le chemin optimal dans la table par : $\widehat{x}_t = \Psi_{t+1}(\widehat{x}_{t+1})$, pour $t = N-1, ..., 1$.

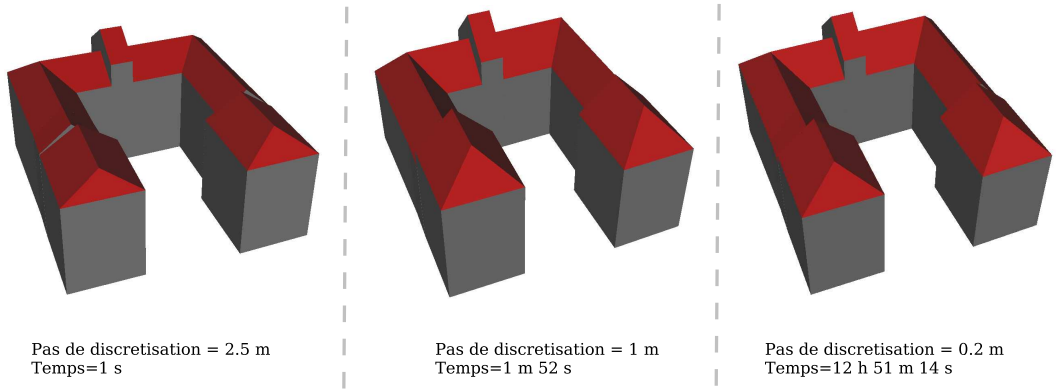

FIG. 6.12 – Résultats sur l'hôtel de ville d'Amiens avec différents pas de discrétisation de l'espace d'état

L'agorithme de Viterbi procéde par une recherche exhaustive de la solution en parcourant l'ensemble des solutions pour chaque élément de la chaîne. Les temps de calcul dépendent donc de la taille de l'espace des configurations. Il est important de discrétiser pertinement cet espace afin d'avoir des temps de calcul raisonnables et d'éviter les problèmes de saturation de mémoire. La figure 6.12 montre les résultats obtenus sur l'hôtel de ville d'Amiens avec différents pas de discrétisation. La valeur du pas de discrétisation est directement liée au degré de précision (en mètre) de la modélisation par rapport aux données. Sur la figure de droite, le pas de discrétisation est très fin ($0,2$ mètre), et par conséquent le résultat est très précis. Cependant, les temps de calcul sont extrèmement élevés (environ 13 heures). Comparativement aux temps de calcul mis à travers l'optimisation par MCMC, ce résultat n'est clairement pas intéressant, tout comme celui du centre de la figure. Par contre,

sur le résultat de gauche, on obtient une solution, certes très générale (2.5 mètres de précision), mais avec des temps de calcul très faibles (de l'ordre d'une seconde). Il faut noter que, quel que soit le pas choisi, les résultats sont très bien régularisés.

Pour résumer, cet algorithme est intéressant pour obtenir rapidement une modélisation approchée (i.e. avec un pas de discrétisation de l'ordre de deux mètres) sous des temps de calcul très faibles (de l'ordre d'une seconde pour un bâtiment). Ces performances pourraîent être clairement améliorées par une restriction intelligente de l'espace des configurations (par exemple, en estimant les hauteurs des objets par un procédé analogue à celui présenté dans l'annexe B-2. Le pas de discrétisation ne doit cependant pas être trop important (supérieur à trois mètres) sinon on risque d'obtenir des modélisations très érronées avec notamment une mauvaise reconnaissance des formes de toit.

On peut notamment imaginer utiliser les résultats de cet algorithme comme une bonne initialisation pour un modèle déterministe plus complexe. Les résultats permettent, en effet, de donner le type de modèles des objets ainsi qu'une estimée des valeurs des paramètres.

F - Résultats supplémentaires

Cette annexe présente des résultats subsidiaires relatifs à différentes parties du manuscrit. La figure 6.14 illustre les resultats de la méthode d'extraction des caricatures de bâtiments par agencement de rectangles (présentée dans la partie 2.1) sur les zones test A, B et C. La figure 6.13 montre un résultat du processus automatique complet sur une zone d'Amiens non localement plate à partir d'un MNE PLEIADES. Les autres figures présentent des résultats du processus de reconstruction 3D par extraction manuelle des supports 2D à partir de MNE PLEIADES à $0,7$ mètre (Figures 6.15, 6.16 et 6.17), de MNE aériens à $0,25$ mètre (Figures 6.18 et 6.19) et de MNE aériens à $0,1$ mètre permettant la reconstruction des superstructures (Figure 6.20).

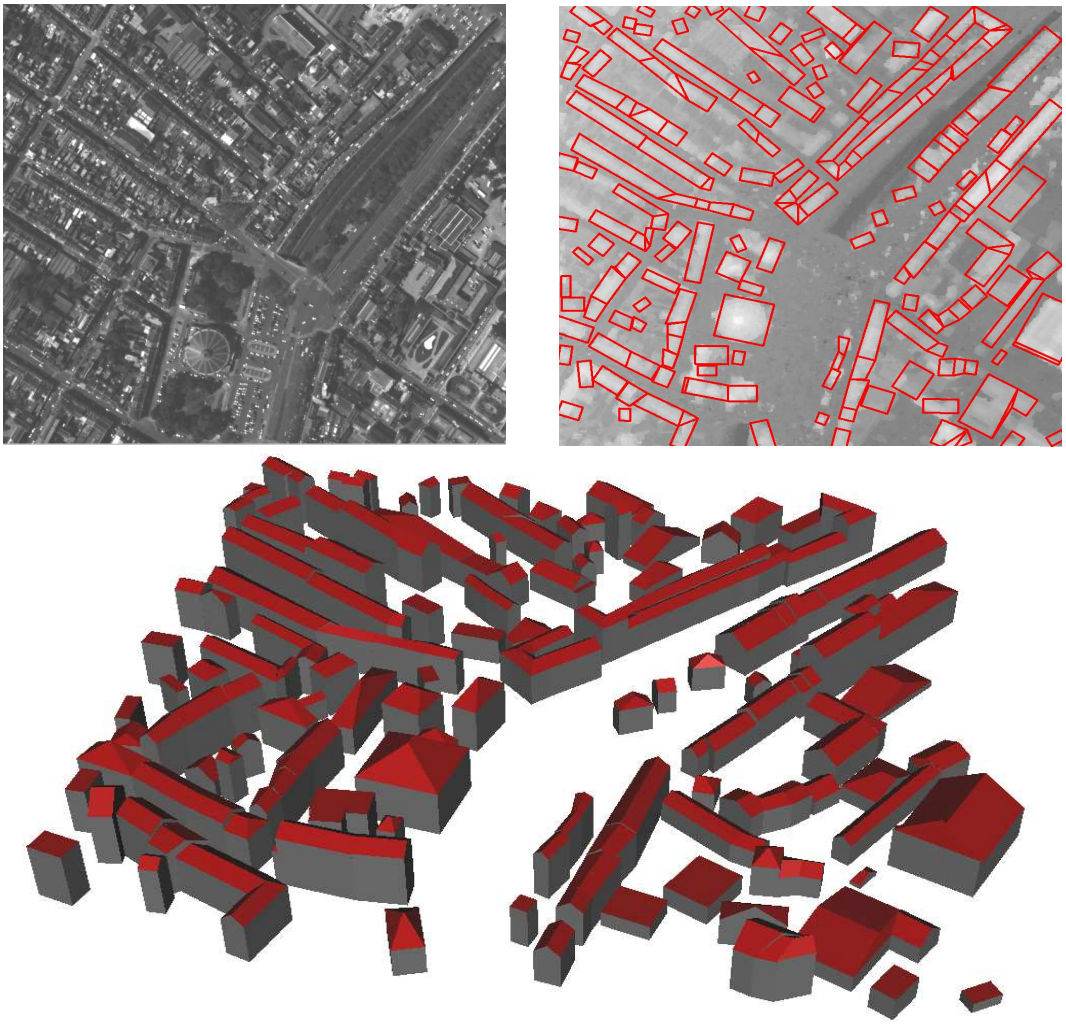

FIG. 6.13 – Résultat du processus automatique complet sur une zone d'Amiens non localement plate à partir d'un MNE PLEIADES

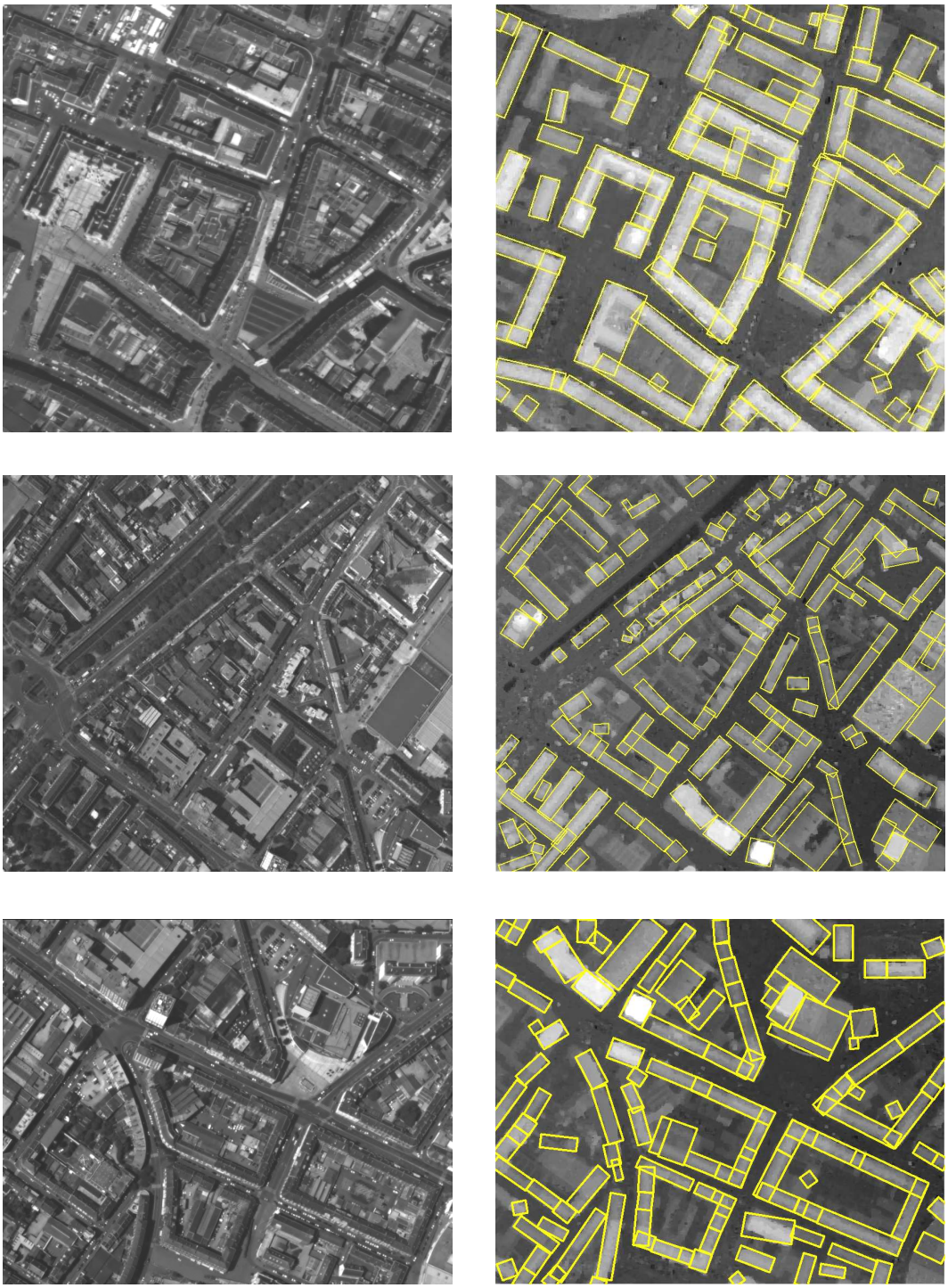

FIG. 6.14 – Résultats de l'extraction des caricatures des bâtiments par agencement de rectangles sur les zones A, B et C.

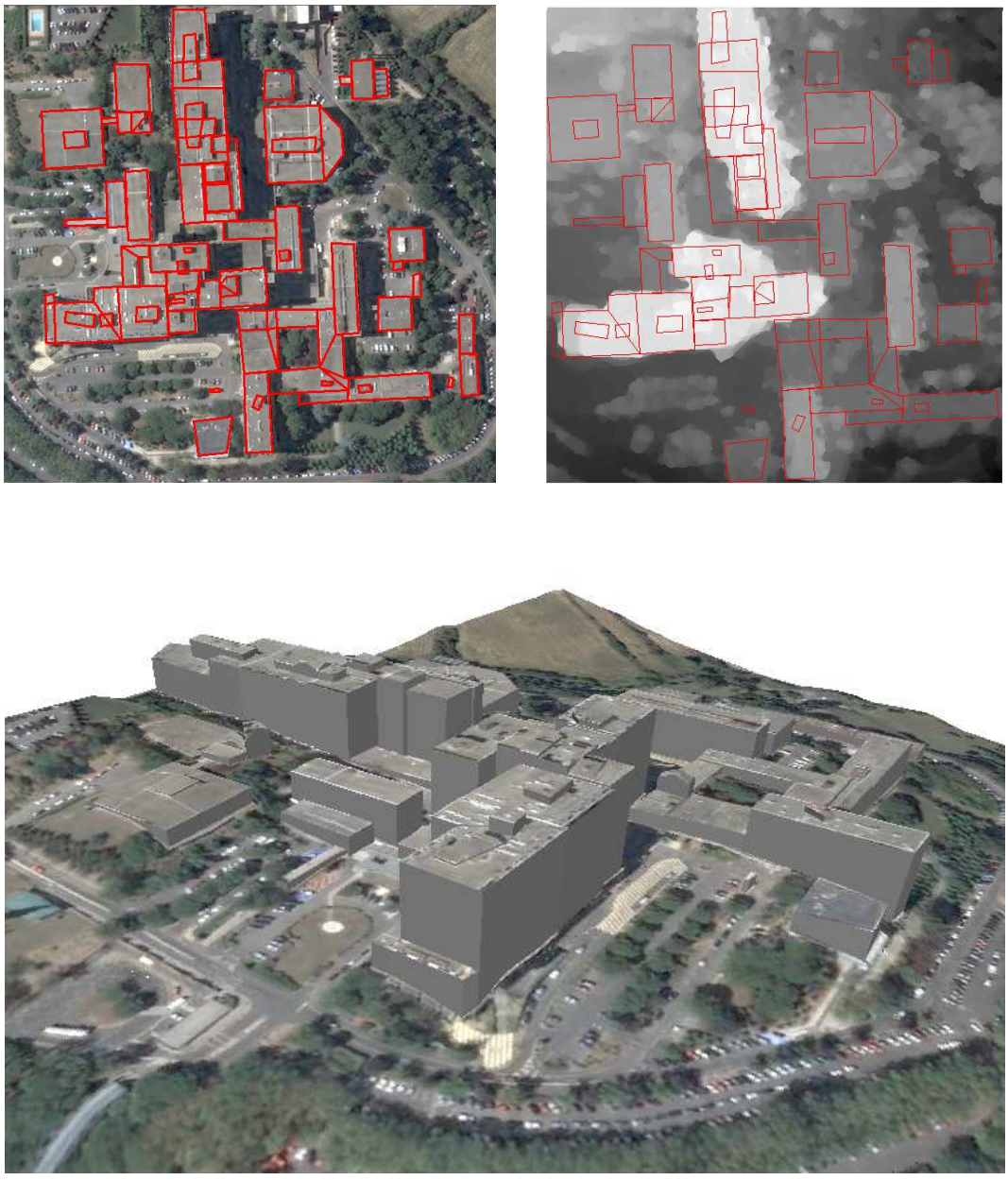

FIG. 6.15 – Reconstruction 3D du CHU de Rangueil (Toulouse) à partir de simulations PLEIADES et d'une extraction manuelle des supports 2D.

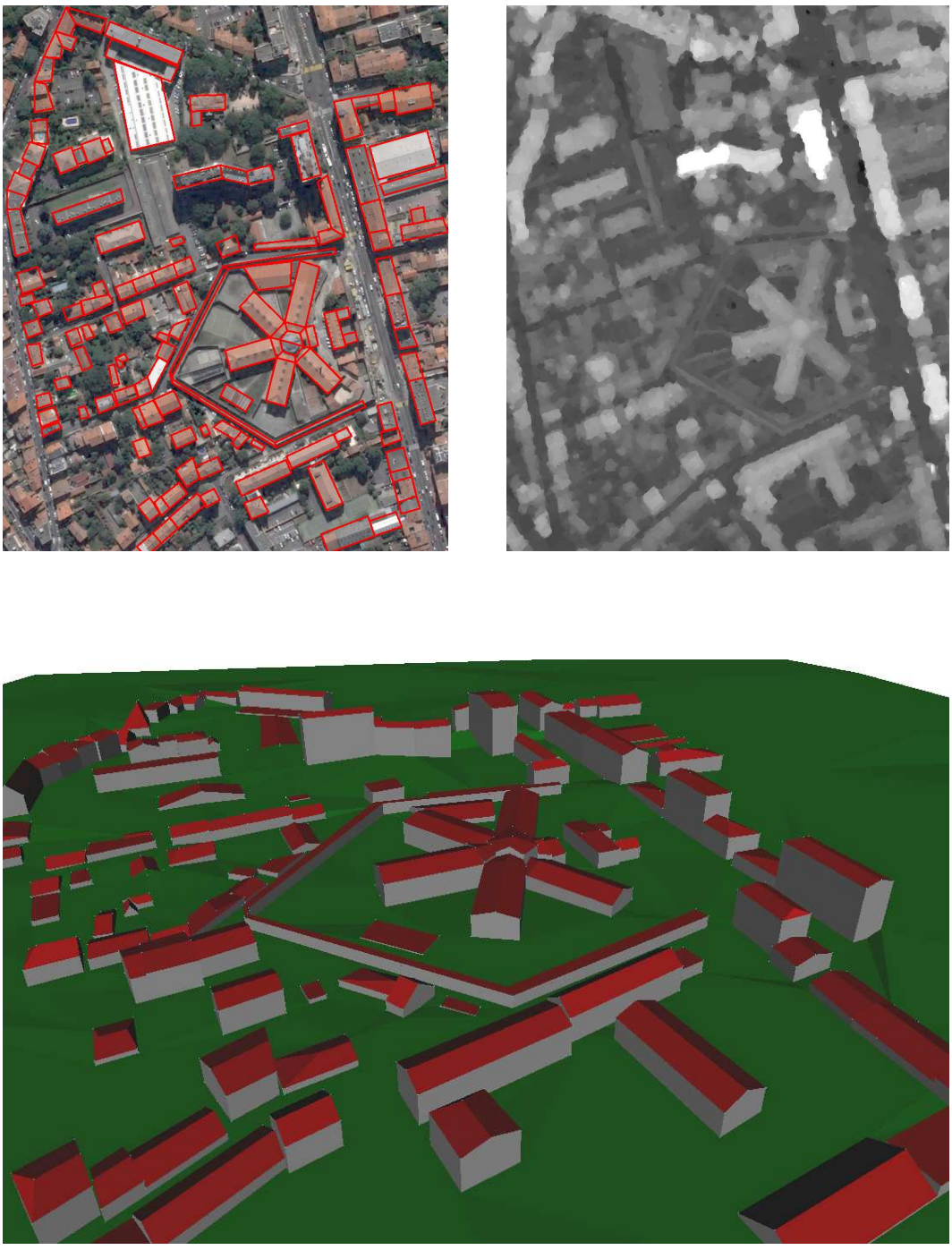

FIG. 6.16 – Reconstruction 3D du quartier Saint-Michel de Toulouse à partir de simulations PLEIADES et d'une extraction manuelle des supports 2D.

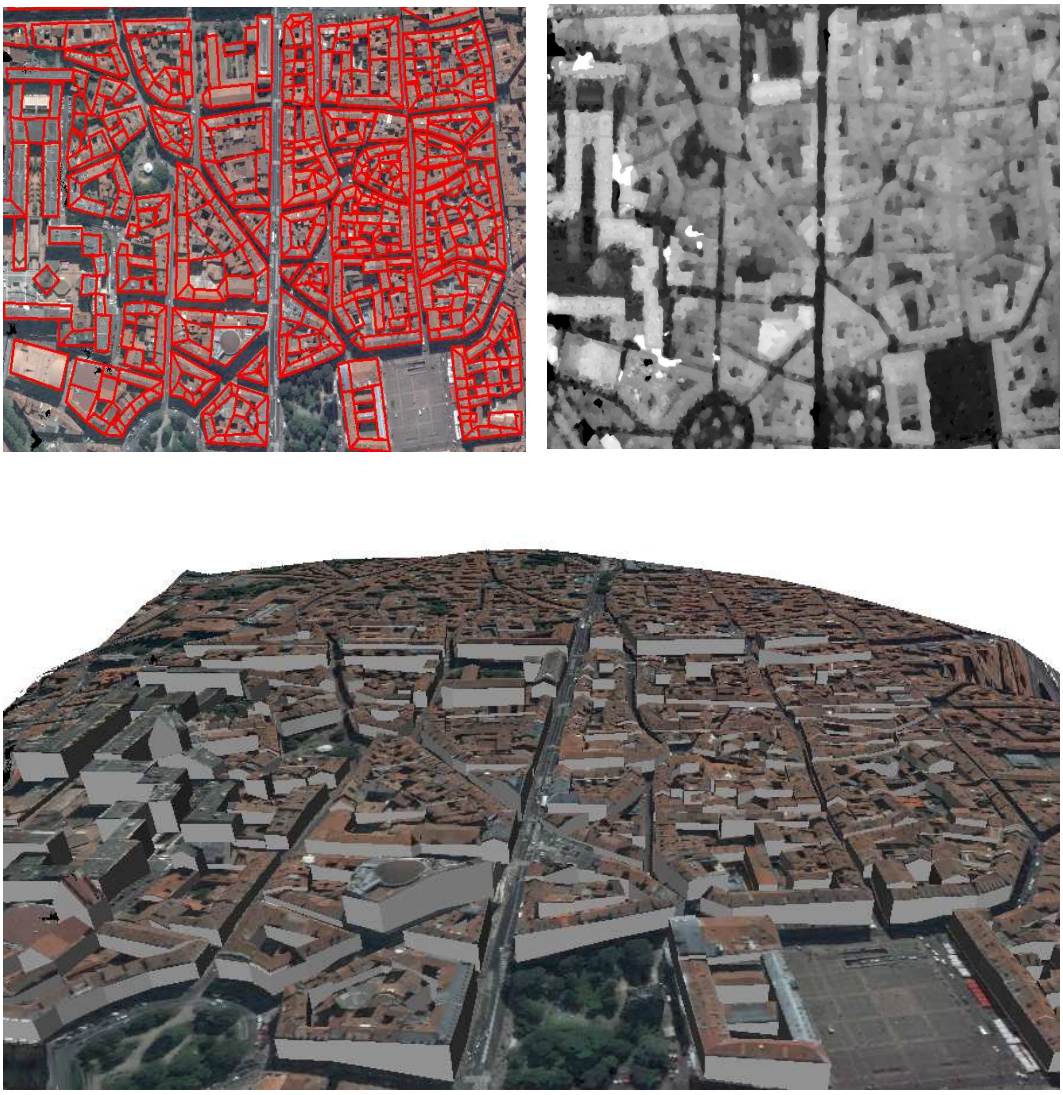

FIG. 6.17 – Reconstruction 3D du centre-ville de Toulouse (zone du Capitole) à partir de simulations PLEIADES et d'une extraction manuelle des supports 2D.

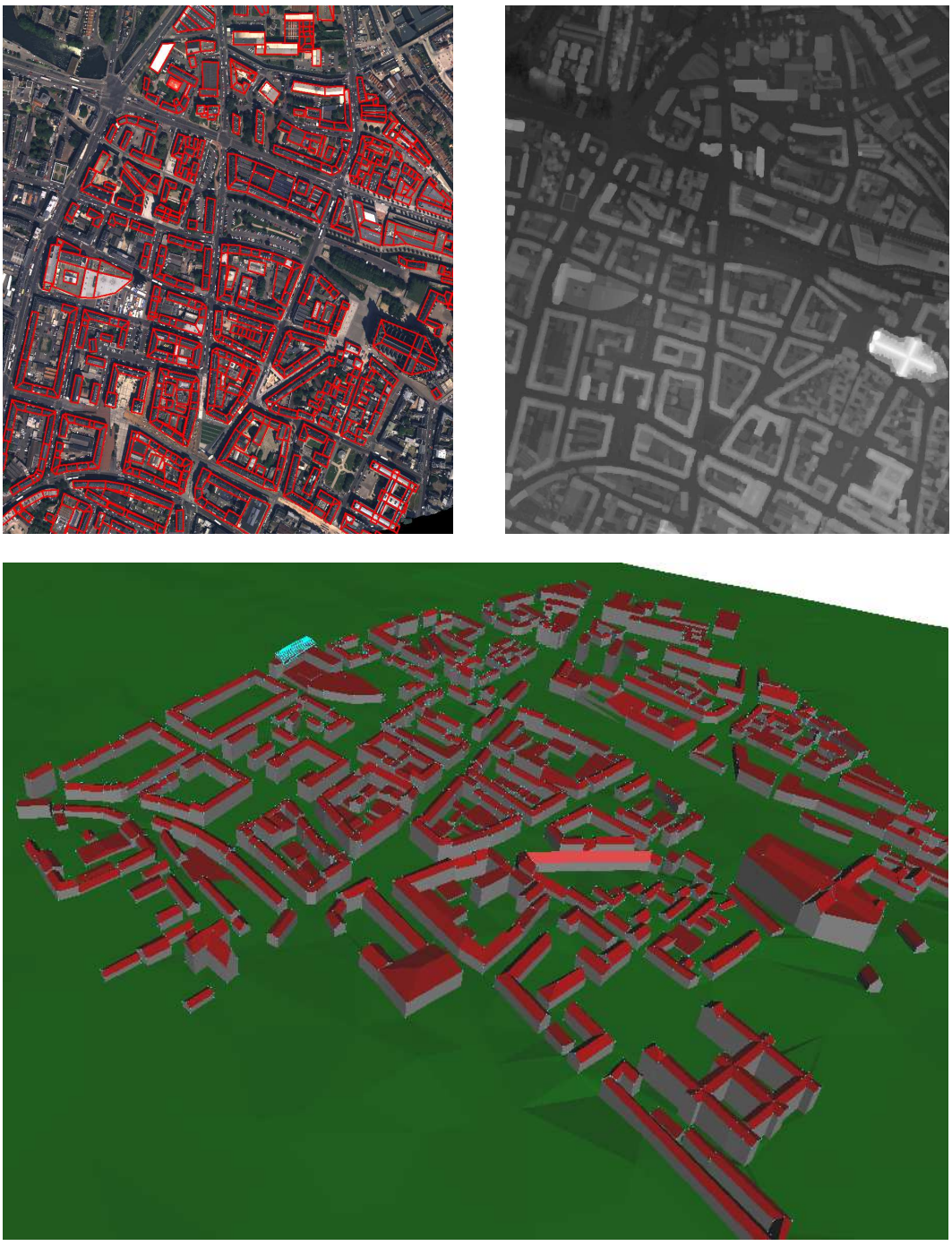

FIG. 6.18 – Reconstruction 3D du centre ville d'Amiens à partir d'images aériennes ($0,25$ mètre) et d'une extraction manuelle des supports 2D.

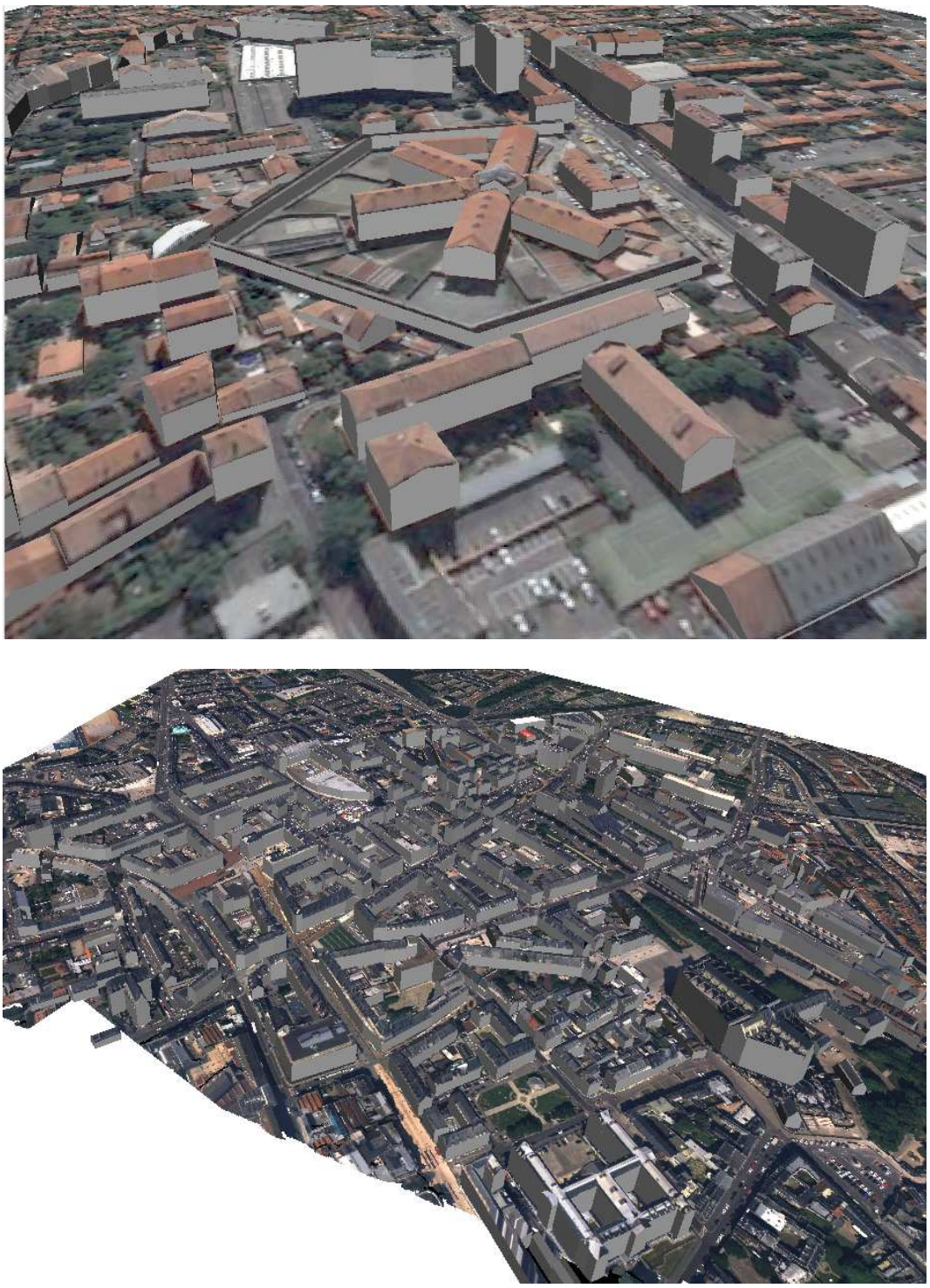

FIG. 6.19 – Reconstruction 3D avec texturation des résultats présentés sur les figures 6.16 et 6.18.

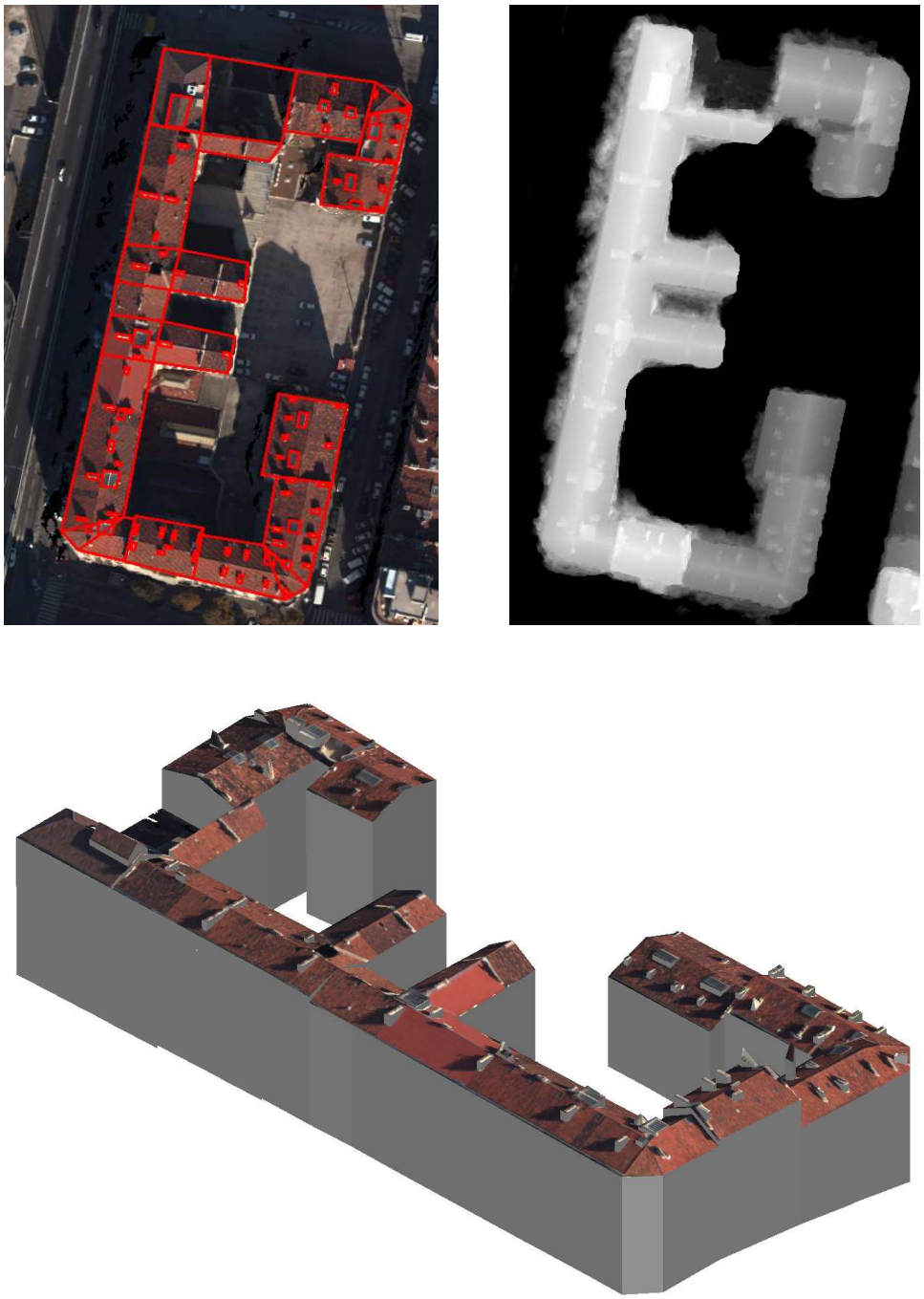

FIG. 6.20 – Reconstruction 3D d'un bâtiment et de ses superstructures à partir d'images aériennes (0,1 mètre) et d'une extraction manuelle des supports 2D.

RESUME

Cette thèse aborde le problème de la reconstruction tridimensionnelle de zones urbaines à partir d'images satellitaires très haute résolution. Le contenu informatif de ce type de données est insuffisant pour permettre une utilisation efficace des nombreux algorithmes développés pour des données aériennes. Dans ce contexte, l'introduction de connaissances a priori fortes sur les zones urbaines est nécessaire. Les outils stochastiques sont particulièrement bien adaptés pour traiter cette problématique.

Nous proposons une approche structurelle pour aborder ce sujet. Cela consiste à modéliser un bâtiment comme un assemblage de modules urbains élémentaires extraits d'une bibliothèque de modèles 3D paramétriques. Dans un premier temps, nous extrayons les supports 2D de ces modules à partir d'un Modèle Numérique d' Elévation (MNE). Le résultat est un agencement de quadrilatères dont les éléments voisins sont connectés entre eux. Ensuite, nous reconstruisons les bâtiments en recherchant la configuration optimale de modèles 3D se fixant sur les supports précédemment extraits. Cette configuration correspond à la réalisation qui maximise une densité mesurant la cohérence entre la réalisation et le MNE, mais également prenant en compte des connaissances a priori telles que des lois d'assemblage des modules. Nous discutons enfin de la pertinence de cette approche en analysant les résultats obtenus à partir de données satellitaires (simulations PLEIADES). Des expérimentations sont également réalisées à partir d'images aériennes mieux résolues.

Mots clé : *reconstruction 3D, zones urbaines, images satellitaires, approche structurelle, recuit simulé, MCMC, Modèle Numérique d'Elévation.*

ABSTRACT

This thesis tackles the problem of the 3D building reconstruction from very high resolution satellite images. The information provided by this kind of data are not accurate enough to allow an efficient use of the varied algorithms developped in an aerial framework. In this context, it is necessary to introduce strong prior knowledge related to the urban areas. The stochastic tools are especially well adapted to deal with this problem.

A structural approach is proposed to address this topic. It consists in modeling a building through an assembling of basic urban structures which are extracted fom a library of 3D parametric models. First, the 2D supports of these structures are extracted from a Digital Elevation Model (DEM). The result is a quadrilateral layout of which the neighboring elements are connected. Then, the buildings are reconstructed by finding the optimal configuration of 3D models which are fixed onto the extracted supports. This configuration corresponds to the realization which maximizes a density. The last one measures the coherence between the realization and the DEM, and takes into account prior knowledge such as the assembling law of the structures. Finally, we discuss on the relevance of this approach by analysing the obtained results from satellite data (PLEIADES simulations). Experiments are also carried on from higher resolution aerial images.

Keywords : *3D reconstruction, urban areas, satellite images, structural approach, simulated annealing, MCMC, Digital Elevation Model.*