THÈSE

présentée pour obtenir le titre de docteur de l'École Polytechnique

Spécialité : Mathématiques

Franck TAILLANDIER

Reconstruction du bâti en milieu urbain : une approche multi-vues

Soutenue le 12/10/2004 devant le jury composé de :

Pr. Georges STAMON	Président
Pr. Henri MAITRE	Rapporteurs
Pr. Jean-Pierre COCQUEREZ	
Dr. Hervé LE MEN	Examinateurs
Dr. Michel FLORENZANO	
Dr. Rachid DERICHE	Directeur de thèse

Rien n'est jamais perdu, tant qu'il reste quelque chose à trouver.
Pierre Dac

Remerciements

Je tiens en premier lieu à exprimer ma gratitude à *M. Deriche* qui a accepté d'encadrer cette thèse. Malgré la distance, il a su par de nombreuses remarques et références pertinentes m'orienter dans mes recherches. Je lui suis tout particulièrement reconnaissant pour sa confiance et ses conseils patients lors de la rédaction d'articles.
Je remercie également *M. Le Men* pour l'intérêt qu'il a manifesté pour mes travaux, son encadrement et sa relecture minutieuse du manuscrit.

Je remercie par ailleurs tous les membres du jury :
M. Stamon qui m'a fait l'honneur de présider mon jury de thèse.
MM. Maître et Cocquerez qui ont accepté d'être rapporteurs de cette thèse, pour le temps qu'il ont consacré à la relecture du manuscrit et pour les innombrables remarques qui ont permis de l'améliorer.
MM. Stamon, Le Men et Florenzano pour m'avoir fait l'honneur de bien vouloir participer au jury.

Mes remerciements vont ensuite à toute l'équipe du MATIS, permanents, doctorants ou stagiaires et plus particulièrement à *Marc* et *Didier*, chefs successifs du laboratoire MATIS pour leur soutien, *Frank* qui par son enthousiasme m'a donné très tôt l'envie de m'engager dans la reconstruction 3D de bâtiments, *Laurent* et *Hassan* pour les innombrables discussions et aides techniques, sans lesquelles bien des travaux exposés ici n'auraient abouti, *Dandan*, *Gilles*, *Didier* et *Roger* pour le côté "gourou Unix", dépannage en tous genres, *Grégoire* et *David* pour la mise à disposition du formidable outil qu'est la maquette et leur disponibilité, *Nicolas* pour sa gentillesse et son attention, *François* pour sa disponibilité, sa gentillesse et son impressionnante connaissance de l'IGN et de ses rouages...
Je suis tout particulièrement reconnaissant envers *Franck*, pour ses compétences techniques et sa gentillesse, et surtout pour son écoute attentive et sa confiance inébranlable dans les moments de doute et je n'oublie pas non plus tous mes collègues de bureau : *Grégoire*, *Olivier*, *Nesrine* et *Frédéric* pour leur soutien et la très très bonne ambiance...

J'exprime la plus profonde reconnaissance à tous les membres de ma famille et tout particulièrement à mon épouse et mon fils qui, par leur affection, m'ont soutenu au cours de ces années.
Enfin, à tous ceux, et ils sont nombreux, qui, un jour, m'ont demandé : "alors cette thèse, c'est pour quand ?". Ça y est, elle est là.

Summary

Despite advances in Photogrametry and Computer Vision, automatic scene reconstruction from photographs seems a formidable task, far from being reached. In this thesis, we focus on building reconstruction from multiple aerial images. A new approach is presented, that uses a very generic modelling of buildings as polyhedral shapes with no overhang, in which external knowledge is introduced through a priori constraints on base primitives.

The general strategy is subdivided in four main steps. In the first step, three types of 3D primitives are extracted from images : 3D segments, planar patches and oriented portions of facades. They will be of primary importance to infere buildings hypotheses and to select the best one. In this step, multiscopy context will help as far as exhaustivity and precision are concerned. In a second step, from the arrangement of planes deduced from plane primitives the system builds up a 3D graph of facets and then a so-called "compatibility graph" where the nodes are the initial facets of the 3D graph and edges between two nodes state that both facets belong to at least one common hypothese of building and thus are "compatible". In our scheme, buildings are modelled, in a very generic way, as polyhedral volumes with no overhang and it is shown that maximal cliques in the compatibility graph supply all the hypotheses of buildings that can be deduced from the arrangement of planes. Before effectively computing all the solutions by brute force search, some simplifications in the initial 3D and compatibility graphs are made, based on DEM and focalisation masque, in order to reduce the crippling combinatory that arise from the maximal cliques problem. In the set of all the hypotheses finally extracted from the compatibility graph, the choice of the final model is done, in a third step, through a bayesian formulation that enables different kinds of observations to be taken into account. Model complexity and a priori constraints on base primitives such as orthogonalities, parallelisms, symetries, horizontalities are also naturally introduced to balance caricature needs and data adequacy. The last step consists of effectively enforcing constraints in the building reconstruction. An implicit parametrization of the building enables to ensure that constraints are exactly verified.

Results are provided and show the validity of the approach that stays very generic on the contrary to model-based methods while bringing external architectural information through geometric constraints, which generally lacks in data-driven algorithms. The system can handle arbitrary complex buildings, even with internal altimetric discontinuities and overcome focussing errors. Despite its generality, the algorithm remains very tolerant to errors in primitive detections and shows promising results for the modelling of periurban scenes.

Keywords : *Buildings Reconstruction, Aerial Images, Multi-view Context, Bayesian Formulation, Digital Elevation Model, Hypothesize-and-verify Strategies.*

Résumé

En dépit des progrès réalisés en vision par ordinateur et en photogrammétrie, la reconstruction tridimensionnelle de scènes à partir de photographies reste un problème extraordinairement complexe et loin d'être résolu. Dans cette thèse, nous nous intéressons au domaine plus restreint de la reconstruction automatique de bâtiments à partir d'images aériennes en multi-recouvrement. En décalage par rapport aux méthodes à base de modèles prédéfinis, notre approche utilise une modélisation très générique des bâtiments comme polyèdres sans surplomb. Pour pallier le manque de robustesse inhérent à la plupart des approches génériques, des contraintes a priori sont introduites dans la sélection de la meilleure représentation favorisant ainsi, sans les imposer, des choix correspondant à des situations récurrentes dans un environnement urbain : parallélisme, horizontalité, orthogonalité, symétrie verticale de certaines facettes. Par ce raisonnement, le système concilie à la fois les problématiques de généricité et de généralisation.

La stratégie générale se compose de quatre étapes majeures. La première consiste à extraire des images trois types de primitives tridimensionnelles qui serviront à construire les modèles de bâtiments puis à sélectionner la meilleure représentation : segments 3D, facettes planes, discontinuités altimétriques. Dans un deuxième temps, l'algorithme extrait, à partir de l'arrangement de plans déduit des primitives planes, toutes les hypothèses de bâtiments modélisés comme surfaces polyédriques continues et sans surplomb. Il est prouvé, à cet effet, que cette procédure se ramène à la recherche exhaustive des cliques maximales d'un graphe adéquat. Dans l'ensemble de toutes ces hypothèses, la sélection du modèle final se fait, dans un troisième temps, par le biais d'une formulation bayésienne qui prend en compte différents types d'observation. La complexité du modèle de bâtiment et les contraintes a priori sur les primitives telles qu'orthogonalité, parallélisme, symétrie, horizontalité sont naturellement introduites dans cette formulation pour concilier adéquation aux données et simplicité de formes. Enfin, les contraintes détectées sur les primitives sont effectivement introduites dans la reconstruction. L'algorithme utilise alors une paramétrisation implicite du bâtiment qui assure l'application des contraintes sur le modèle final.

La méthodologie de recherche de l'ensemble des formes admissibles par cliques maximales qui est une extension des travaux de H. Jibrini au cas des surfaces présentant des discontinuités altimétriques, constitue, avec l'intégration de contraintes dans le processus de choix de la représentation et de recalage du modèle, les contributions majeures du système présenté. Des résultats et une évaluation sur images aériennes prouvent la validité de cette approche qui permet de traiter des bâtiments aux formes arbitrairement complexes et de pallier, le cas échéant, les défauts de focalisation. Malgré cette généricité, le système se révèle tolérant aux erreurs des détecteurs de primitives et présente des caractéristiques intéressantes pour la modélisation de scènes péri-urbaines.

Mots clés : *Reconstruction de bâti, images aériennes, multi-vues, formulation bayésienne, modèle numérique d'élévation, stratégies hypothèses-et-vérifie.*

Table des matières

Liste des Illustrations **v**

I Introduction **1**

1 Introduction **3**
1.1 Contexte . . . 3
1.1.1 Contexte général . . . 3
1.1.2 Les recherches à l'IGN en reconstruction de bâtiments . . . 4
1.2 Cadre spécifique de la thèse . . . 4
1.3 Objectifs . . . 5
1.3.1 Que reconstruire ? . . . 5
1.3.2 Comment le décrire ? . . . 5
1.3.3 Avec quelle précision ? . . . 6
1.4 Plan du manuscrit et contributions . . . 6

2 Analyse et Stratégie **9**
2.1 La reconstruction du bâti : une problématique complexe . . . 9
2.1.1 Généricité et généralisation . . . 9
2.1.2 Résolution des images . . . 10
2.2 Les données . . . 11
2.2.1 Principe de la multiscopie . . . 11
2.2.2 Images . . . 12
2.2.3 Produits dérivés : MNE . . . 12
2.3 État de l'art en reconstruction de bâti . . . 13
2.3.1 Contexte . . . 14
2.3.2 Données sources . . . 17
2.3.3 Complexité des bâtiments reconstruits . . . 18
2.3.4 Stratégie générale . . . 22
2.3.5 Bilan . . . 22
2.4 Stratégie . . . 22
2.4.1 Analyse . . . 22
2.4.2 Démarche . . . 23
2.4.3 Récapitulatif . . . 25
2.5 Conclusion . . . 26

II Détection de primitives **27**

3 Détection de segments 3D **29**
3.1 Introduction . . . 29
3.1.1 État de l'art . . . 29
3.1.2 Algorithme . . . 30
3.2 Mise en correspondance dans l'Espace Objet . . . 30
3.2.1 Techniques multi-images . . . 30
3.2.2 Mise en correspondance . . . 31
3.3 Reconstruction itérative de segments 3D . . . 34
3.3.1 Droites de l'espace . . . 34
3.3.2 Reconstruction de droites . . . 35
3.3.3 Qualification du résultat . . . 38
3.3.4 Reconstruction des extrémités . . . 38
3.4 Algorithme global . . . 38
3.4.1 Extraction de segments 2D . . . 38
3.4.2 Élagage de Ω, ensemble des hypothèses de droites 3D . . . 39
3.5 Résultats . . . 40
3.5.1 Adéquation à la loi du χ^2 . . . 40
3.5.2 Simulations . . . 40
3.5.3 Images réelles . . . 42
3.6 Extensions . . . 43
3.6.1 Restriction de l'Espace de Recherche . . . 43
3.6.2 Extension à 2 images . . . 44
3.7 Évaluation . . . 44
3.7.1 Précision théorique . . . 44
3.7.2 Précision en fonction des configurations . . . 46
3.7.3 Taux de recouvrement en fonction du nombre d'images . . . 47
3.7.4 Surdétections / Sousdétections . . . 47
3.7.5 Comparaison avec la reconstruction à partir des coordonnées de Plücker . . . 47
3.8 Conclusion . . . 49
3.8.1 Apports . . . 49
3.8.2 Discussion et perspectives . . . 50

4 Détection de facettes planes **51**
4.1 Introduction . . . 51
4.1.1 Contexte . . . 51
4.1.2 Objectifs . . . 51
4.2 État de l'art . . . 51
4.2.1 Approches image . . . 51
4.2.2 Approches segmentation de MNE . . . 52
4.3 Schéma Général . . . 52
4.3.1 Formulation . . . 52
4.3.2 Segmentation multi-échelles et hiérarchie . . . 53
4.3.3 Hiérarchie d'escalade . . . 53
4.3.4 D'un arbre binaire à un arbre n-aire . . . 53
4.4 Modèle de facettes planes . . . 55
4.4.1 Estimation robuste de plans . . . 55

4.4.2 Codage du modèle . . . 55
4.4.3 Codage des données : codage par rapport au plan . . . 55
4.4.4 Modèle global . . . 56
4.5 Modèle de facettes planes et homogènes . . . 56
4.5.1 Orthoimage de gradients . . . 56
4.5.2 Modèle de facettes homogènes . . . 56
4.5.3 Modèle global . . . 57
4.5.4 Adjonction de contraintes . . . 57
4.6 Obtention d'une caricature . . . 57
4.6.1 Post-traitements . . . 58
4.6.2 Mise en application sur un masque déterminé . . . 58
4.7 Détection de façades . . . 59
4.8 Évaluation et résultats . . . 59
4.9 Conclusion . . . 60
4.9.1 Apports . . . 60
4.9.2 Discussions . . . 63
4.9.3 Perspectives et utilisation . . . 64

III Reconnaissance de bâtiments 65

5 Reconnaissance automatique de bâtiments 67
5.1 Méthodologie générale . . . 67
5.1.1 Problématique . . . 67
5.1.2 Hypothèses . . . 67
5.1.3 Schéma général . . . 67
5.2 Recherche des formes acceptables . . . 68
5.2.1 Création d'un graphe 3D . . . 68
5.2.2 Définitions . . . 69
5.2.3 Propriétés . . . 73
5.2.4 Simplifications du graphe . . . 75
5.2.5 Recherche des solutions . . . 76
5.3 Choix de la meilleure représentation . . . 79
5.3.1 Modélisation bayésienne . . . 79
5.3.2 Modèle . . . 82
5.3.3 Observations . . . 84
5.4 Filtrages et extraction des bâtiments . . . 85
5.4.1 Filtrages . . . 85
5.4.2 Extraction des toits de bâtiments . . . 86
5.4.3 Corrections topologiques . . . 86
5.5 Optimisation de la recherche . . . 86
5.5.1 Rendre le problème additif . . . 86
5.5.2 Heuristiques d'accélération de la recherche . . . 87
5.5.3 Remarques sur les techniques d'Hassan Jibrini . . . 88
5.6 Résultats . . . 89
5.7 Conclusion . . . 89

6 Recalage de bâtiments avec contraintes **91**
6.1 Introduction . . . 91
6.1.1 Motivations . . . 91
6.1.2 État de l'art . . . 91
6.1.3 Schéma général . . . 92
6.2 Inférence des contraintes . . . 93
6.2.1 Topologie et géométrie : notations . . . 93
6.2.2 Clustering des directions . . . 93
6.2.3 Graphe des contraintes . . . 94
6.2.4 Graphe de dépendances . . . 95
6.2.5 Calculs via le graphe de dépendances . . . 97
6.3 Application des contraintes au modèle . . . 98
6.3.1 Paramétrisation géométrique . . . 98
6.3.2 Schéma de minimisation itératif . . . 100
6.4 Gestion des surcontraintes et des dégénérescences . . . 101
6.4.1 Détection . . . 101
6.4.2 Gestion . . . 102
6.5 Résultats . . . 102
6.5.1 Protocole d'évaluation . . . 102
6.5.2 Contraintes a priori . . . 104
6.5.3 Contraintes a posteriori . . . 104
6.6 Conclusion . . . 104
6.6.1 Apports . . . 104
6.6.2 Extension . . . 104
6.6.3 Perspectives . . . 107

IV Évaluation **109**

7 Résultats **111**
7.1 Introduction . . . 111
7.2 Quelques résultats qualitatifs . . . 112
7.2.1 Quelques bâtiments à la loupe . . . 112
7.2.2 Bâtiments avec discontinuités altimétriques internes . . . 112
7.2.3 Bâtiments complexes . . . 113
7.2.4 Niveau de généralisation . . . 113
7.2.5 Contraintes . . . 115
7.3 Évaluation structurelle . . . 115
7.3.1 Protocole d'évaluation . . . 115
7.3.2 Influence du paramètre beta . . . 117
7.3.3 Sources d'erreurs . . . 117
7.4 Évaluation de la précision . . . 118
7.4.1 Protocole d'évaluation . . . 118
7.4.2 Résultats . . . 118
7.5 Conclusion . . . 119

8 Conclusion **121**
8.1 Analyse . . . 121
8.1.1 Inconvénients . . . 121
8.1.2 Apports . . . 122
8.2 Perspectives . . . 123
8.2.1 Optimisation . . . 123
8.2.2 Introduction d'une librairie de modèles . . . 123
8.2.3 Introduction des emprises planimétriques . . . 123
8.3 Bilan . . . 124

Annexes **125**

A Reconstruction avec la représentation de Plücker **127**
A.1 Coordonnées homogènes et notations . . . 127
A.2 Reconstruction algébrique . . . 128
A.3 Reconstruction optimale . . . 129

B Recalage de jonctions 3D dans les images **131**
B.1 Recalage de coins . . . 131
B.2 Recalage de jonctions 3D . . . 132
B.3 Résultats . . . 132

C Incertitude géométrique de facettes planes **135**
C.1 Points indépendants . . . 135
C.2 Points issus de corrélation . . . 136
C.3 Résultats . . . 136

D L'algorithme en quelques planches **137**

E Zones de travail **145**
E.1 Zone 1 . . . 146
E.1.1 Référence . . . 146
E.1.2 Résultats aux différents β . . . 147
E.2 Zone 2 . . . 152
E.2.1 Référence . . . 152
E.2.2 Résultats aux différents β . . . 153

Publications **158**

Glossaire **161**

Bibliographie **165**

Liste des Illustrations

2.1 Exemples de généralisation 9
2.2 Un bâtiment et une caricature de ce bâtiment 10
2.3 Principe de la multiscopie 12
2.4 Exemple de 6 images aériennes d'une même scène 13
2.5 Exemple de MNE dense obtenu en multiscopie 14
2.6 Différentes descriptions des bâtiments 19
2.7 La géométrie induit la topologie. 24
2.8 Exemples de MNE, de zone et de masque de focalisation 25
2.9 Méthodologie générale 26

3.1 Principe de l'appariement dans l'Espace Objet. 31
3.2 Technique du balayage de plans. 32
3.3 Calcul de la taille des cellules de la grille terrain. 32
3.4 Calcul du pas de balayage. 32
3.5 Projection d'un pixel sur la grille terrain. 34
3.6 Représentation d'une droite dans l'image. 36
3.7 Adéquation des résidus avec la loi du χ^2, représentation $\mathbf{abpq}$ 41
3.8 Modèle filaire du bâtiment utilisé comme référence dans toutes les simulations. 41
3.9 Simulations. 42
3.10 Résultat de la reconstruction. 43
3.11 Principe de la nappe englobante 43
3.13 Configuration spatiale et schématisation en vue verticale. 44
3.12 Résultats sur images réelles. 45
3.14 Évolution de la précision en fonction du nombre d'images. 46
3.15 Évolution de la précision en fonction des configurations. 46
3.16 Comparaison du taux de recouvrement en fonction du nombre d'images. 48
3.17 Évolution du taux de détection. 48
3.18 Évolution du taux de surdétection. 48
3.19 Evolution du taux de détection. 48
3.20 adéquation des résidus avec la loi du χ^2, représentation de Plücker. 49
3.21 Comparaison avec l'approche à partir des coordonnées de Plücker. 49
3.22 Comparaison avec l'approche à partir des coordonnées de Plücker (2). 49

4.1 Principe de la construction de la hiérarchie d'escalade. 54
4.2 Segmentations implicites dans la hiérarchie. 54
4.3 Région non pertinente. 54
4.4 Exemple de contraintes rajoutées sur l'orthoimage de gradients et watershed correspondant. . 58
4.5 Exemple de segmentations avec focalisation sur un masque. 59

4.6 Exemples de discontinuités altimétriques. 60
4.7 Modèle de facettes planes. Hiérarchie et exemples de segmentations. 61
4.8 Modèle de facettes planes et homogènes. Hiérarchie et exemples de segmentations. 62
4.9 Vue 3D d'une caricature de scène 63

5.1 Algorithme de reconnaissance de bâtiments 69
5.2 Les différentes étapes de l'algorithme appliqué à un exemple. 70
5.3 Intersection de plans et Graphe 3D issu de l'arrangement - vue schématique. 71
5.4 Intersection de plans et Graphe 3D issu de l'arrangement - vue 3D. 71
5.5 Orientation induite pour un polytope. 71
5.6 Cohérence de l'orientation de 2 polytopes. 71
5.7 Exemples de facettes non verticales et verticales cohérentes. 72
5.8 Exemples de facettes non verticales et verticales incohérentes. 72
5.9 Une facette réalise l'altitude maximale sur tous ses points intérieurs. 73
5.10 Deux configurations possibles. 74
5.11 Surfaces enveloppes. 74
5.12 Graphe 3D issu de l'arrangement de plans. 74
5.13 Toute surface admissible est sans surplomb. 75
5.14 Différentes étapes de simplification - vue schématique. 76
5.15 Différentes étapes de simplification - vue 3D. 76
5.16 Calcul des compatibles. 78
5.17 Exemple de graphe élagué et de graphe d'incompatibilités. 79
5.18 Le graphe après simplifications suivi de 11 solutions sur les 3646 de l'exemple présenté. . . 80
5.19 Principe de la démonstration de l'équivalence entre clique maximale et surface admissible. . 81
5.20 Évolution du nombre de cliques maximales en fonction du nombre de noeuds. 81
5.21 Comparaison des algorithmes de [Jib02] et [BK73] 81
5.22 Schéma d'inférence des contraintes et représentation du graphe de contraintes. 83
5.23 Principe du calcul de $P(I/M)$. 85
5.24 Contre-exemple à l'algorithme "worst-out". 88
5.25 Contre-exemple à l'algorithme "worst-out". 89

6.1 Schéma général. 93
6.2 Bâtiment 3D approximatif et sa représentation schématique. 94
6.3 Résultat du clustering de directions sur l'exemple précédent. 94
6.4 Graphe de contraintes et graphe de dépendances. 97
6.5 Étude d'un cas de surcontrainte. 102
6.6 Différentes vues de quelques bâtiments avec $\sigma = 0.5m$. 105
6.7 Gain apporté par le recalage avec contraintes. 106
6.8 Gain apporté par le recalage avec contraintes dans les images. 108

7.1 Bâtiments complexes. 113
7.2 Rôle du paramètre β dans le niveau de généralisation. 114
7.3 Rôle des contraintes dans le choix du meilleur modèle. 116
7.4 Importance des contraintes dans le recalage. 117
7.5 Influence du paramètre β. 118
7.6 Évaluation de la précision géométrique. 119

A.1 Représentation de Plücker d'une droite 3D. 128

B.1 Recalage de jonctions en 2D. . . . 133
B.2 Recalage de jonctions 3D sur images réelles. . . . 133
B.3 Autres exemples de recalage. . . . 134

C.1 Points projetés avec incertitude. . . . 136

D.1 Les différentes étapes de l'algorithme appliqué à un exemple. . . . 138
D.2 Les différentes étapes de l'algorithme appliqué à un exemple. . . . 139
D.3 Les différentes étapes de l'algorithme appliqué à un exemple. . . . 140
D.4 Les différentes étapes de l'algorithme appliqué à un exemple. . . . 141
D.5 Les différentes étapes de l'algorithme appliqué à un exemple. . . . 142
D.6 Les différentes étapes de l'algorithme appliqué à un exemple. . . . 143
D.7 Les différentes étapes de l'algorithme appliqué à un exemple. . . . 144

E.1 Zone 1 avec bâtiments numérotés. . . . 146
E.2 Référence pour la zone 1. . . . 146
E.3 Résultats zone 1, $\beta = 0.01$. . . . 147
E.4 Résultats zone 1, $\beta = 10$. . . . 147
E.5 Résultats zone 1, $\beta = 25$. . . . 148
E.6 Résultats zone 1, $\beta = 50$. . . . 148
E.7 Résultats zone 1, $\beta = 75$. . . . 149
E.8 Résultats zone 1, $\beta = 100$. . . . 149
E.9 Résultats zone 1, $\beta = 250$. . . . 150
E.10 Résultats zone 1, $\beta = 500$. . . . 150
E.11 Résultats zone 1, $\beta = 100000$. . . . 151
E.12 Zone 2 avec bâtiments numérotés. . . . 152
E.13 Référence pour la zone 2. . . . 152
E.14 Résultats zone 2, $\beta = 0.01$. . . . 153
E.15 Résultats zone 2, $\beta = 10$. . . . 153
E.16 Résultats zone 2, $\beta = 25$. . . . 154
E.17 Résultats zone 2, $\beta = 50$. . . . 154
E.18 Résultats zone 2, $\beta = 75$. . . . 155
E.19 Résultats zone 2, $\beta = 100$. . . . 155
E.20 Résultats zone 2, $\beta = 250$. . . . 156
E.21 Résultats zone 2, $\beta = 500$. . . . 156
E.22 Résultats zone 2, $\beta = 100000$. . . . 157

Première partie

Introduction

Chapitre 1

Introduction

1.1 Contexte

1.1.1 Contexte général

Posséder une description de l'environnement urbain est indispensable pour de nombreuses applications. C'est pourquoi la constitution de bases de données urbaines *tridimensionnelles* représente un enjeu important dans des domaines aussi variés que les études environnementales - étude de propagation de bruit, d'écoulement d'eau, d'implantation d'espaces verts -, l'aménagement urbain - étude d'impact d'implantation d'ouvrages d'art dans un environnement donné, simulation d'éclairement - , l'architecture - modélisation de bâtiments de prestige, visite virtuelle de zones urbaines -, les propagations d'ondes électromagnétiques - placement optimisé des antennes pour les opérateurs de téléphonie -, les jeux vidéo - réalisme accru des simulateurs - ou les applications militaires - simulation de combats urbains, guidage de missiles -. La multiplication des projets de recherche au niveau mondial pour l'automatisation des processus de production de ces bases de données témoigne de l'intérêt que les communautés scientifique et industrielle portent à ce produit.

Au niveau national, l'Institut Géographique National (IGN), traditionnel fournisseur de cartes et de bases de données géographiques, s'est rapidement intéressé aux méthodes de production de ces bases de données tridimensionnelles. Une première chaîne de production, la chaîne TRAPU®, entièrement manuelle, basée sur des techniques de saisie photogrammétrique traditionnelle sur appareils analogiques a ainsi été mise au point dans les années 80. Cependant, la réalisation manuelle de telles bases de données, si elle s'avère possible, constitue un processus délicat et surtout extrêmement coûteux ce qui a conduit à l'abandon de cette chaîne de production. Depuis lors, l'IGN mène d'actives recherches sur l'automatisation des processus de production de ces bases tridimensionnelles[1]. Ces recherches, menées au sein du laboratoire MATIS (Méthodes d'Analyse et de Traitement d'Images pour la Stéréorestitution) de l'IGN se focalisent plus spécifiquement sur le thème cartographique "bâti" désignant l'ensemble des bâtiments[2] et s'articulent au sein d'un axe de recherche très actif : BRIGE (Bâtiments Reconstruits sur Images à Grande Échelle).

[1] L'intérêt pour ces recherches a été réaffirmé dans le dernier Contrat d'Objectifs et de Moyens signé entre l'IGN et l'État en septembre 2003. Ce contrat, qui définit les engagements de l'IGN sur une période de 4 ans et la part de financement de l'État, souligne en effet dans l'article 6.2.2 l'importance des recherches sur l'automatisation de la constitution de bases de données 3D.

[2] Contrairement au terme "bâti" qui se réfère au thème cartographique, le terme "bâtiments" désigne l'ensemble des objets réels. Nous confondrons cependant couramment *et abusivement* les deux termes dans la suite de ce manuscrit.

1.1.2 Les recherches à l'IGN en reconstruction de bâtiments

La reconstruction de bâtiment constitue un axe de recherche majeur à l'IGN, en tant que tel comme nous venons de le voir pour la constitution de bases de données tridimensionnelles mais également car ces recherches sur la production de données urbaines se placent en amont d'un certain nombre d'autres travaux comme la production d'orthophotographies avec bâtiments redressés [Bol02], ou à plus long terme la modélisation fine de façades ou de mobilier urbain à l'aide de clichés pris à partir d'un vecteur terrestre, initiant ainsi des recherches sur le couplage terrestre/aérien [BPPDH03, Pen06].
En ce qui concerne les méthodes de production de données 3D, plusieurs axes sont à l'étude. Responsable, d'après ses missions statutaires de la couverture périodique du territoire national par des clichés aériens, l'Institut a tout naturellement orienté son action sur l'utilisation de ces clichés haute résolution pour l'extraction de l'information 3D. Les efforts se sont, dans un premier temps, portés sur le contexte stéréoscopique, contexte habituel de la saisie manuelle [Fuc01], mais les scènes urbaines sont difficiles à analyser compte tenu de la complexité et de l'hétérogénéité des objets, des fortes variations de relief, des phénomènes d'occultation et d'ombres portées entre bâtiments, de la diversité des formes. Face à ces difficultés et aux problèmes classiques inhérents à la stéréoscopie simple que nous développerons en section 2.3.2, des contextes différents sont à l'étude.
Cette thèse se place ainsi dans un contexte multiscopique c'est-à-dire en utilisant au moins trois images en multi-recouvrement afin, entre autres, d'étudier l'impact du multi-vues sur l'exhaustivité et la précision de la reconstruction. Nous détaillerons le contexte spécifique de la thèse plus avant dans cette introduction et présenterons les objectifs fixés dans le cadre de cette étude.
[BPDR03] s'intéresse à l'utilisation d'informations 3D directement issues de méthodes actives comme le laser aéroporté. Cette technique permet d'obtenir une information 3D directe sans passer par une étape de corrélation et de mise en correspondance qui génère inévitablement des erreurs. Ces techniques d'acquisition active suscitent un vif intérêt dans la communauté scientifique (section 2.3.2) même si l'échantillonnage irrégulier reste une problématique majeure de ce genre d'acquisition et si le géoréférencement par rapport à des clichés aériens constitue un problème délicat.
[CJPDS03] se place dans un contexte d'images satellitaires à moyenne résolution. À terme, le satellite devrait en effet fournir des images de qualité équivalente aux clichés aériens avec une répétitivité largement supérieure. En revanche, les différences entre les deux systèmes de prise de vue nécessitent une étude approfondie pour déterminer les adaptations à mettre en oeuvre sur les algorithmes actuels.
Enfin, soulignons tout particulièrement les recherches menées avec utilisation des limites cadastrales [Jib02, FMJ03]. Cette information prend en effet une importance croissante avec la mise en place du Référentiel à Grande Échelle (RGE) [COM03] qui intègre entre autres les limites planimétriques des bâtiments. Compte tenu du gain qu'apporte la délimitation même partielle des bâtiments [Jib02], *il paraît essentiel que tout système puisse intégrer cette information qui sera disponible à terme sur l'ensemble du territoire*. Cette remarque est d'autant plus pertinente que d'autres pays ont à leur disposition ce type d'informations [BHF01, VS01].

1.2 Cadre spécifique de la thèse

Comme nous l'avons déjà évoqué, cette thèse qui s'est déroulée au sein du laboratoire MATIS, a pour objectif général de proposer un système de reconstruction de bâti à partir d'images aériennes en multi-recouvrement. L'accent sera mis ici sur *l'étude du multi-vues* dans les méthodes de détection de primitives et sur la mise en place d'un *système ouvert*, capable d'intégrer des données externes telles que les limites cadastrales ou des points laser.
Le dernier objectif se justifie aisément dès que l'on prend en compte les orientations actuelles de l'ensemble des équipes travaillant sur ce thème (voir section 2.3). La possibilité d'introduire aisément les contraintes

cadastrales ou des données laser est une contrainte imposée dès le départ à la chaîne mise en place, de façon à assurer une certaine pérennité au système et à pouvoir intégrer des données disponibles d'ici quelques années.
Il convient là encore d'insister sur le contexte opérationnel de la recherche à l'IGN : l'objectif est de fournir un outil d'aide à la production à moyen terme. Nous avons toujours travaillé avec cette idée à l'esprit ce qui a influencé quelques-uns de nos choix. Nous définirons plus précisément dans la section suivante ce que nous entendons par reconstruction de bâti et les objectifs applicatifs que nous nous sommes fixés.

1.3 Objectifs

1.3.1 Que reconstruire ?

Il convient tout d'abord de bien préciser les objectifs du système que nous voulons mettre en place et dans un premier temps l'information que l'on cherche à extraire des images, c'est-à-dire la *reconstruction* que l'on souhaite obtenir.
Il est clair, mais encore cela mérite-t-il d'être rappelé, que nous ne rechercherons que la reconstruction des pans de toits des bâtiments. Le contexte aérien, bien qu'en multiscopie, ne nous permet pas d'espérer une description précise des façades de bâtiments (balcons, fenêtres) qui fait l'objet de recherches connexes [Pen06]. Nous nous limiterons donc à la description de la structure des toits (pourtour et structure) en excluant *a priori* les superstructures de toits (chiens assis, cheminées) qui sont difficiles à reconstruire compte tenu de la précision des images aériennes utilisées (section 2.1.2). Nous verrons que malgré ces précautions, la problématique reste complexe en raison des questions de généralisation et de résolution des images (voir section 2.1.1).
Notre étude se limitera également à des reconstructions de bâtiments isolés. Nous supposerons l'étape de focalisation déjà effectuée. Différents algorithmes pour le découpage automatique d'une scène globale en *zones* ou *masques* de focalisation seront détaillés en section 2.4.3. La reconstruction automatique s'effectuera donc non pas sur une ville entière mais sur un îlot contenant un, voire quelques, bâtiments accolés ou séparés. Le système devra prendre en compte des erreurs dans le processus de focalisation et donc pouvoir gérer le cas de plusieurs bâtiments sur une seule zone. Sur cette zone, nous rechercherons, non pas une description purement topographique, c'est à dire sans recherche d'interprétation mais une description vectorielle (nous verrons ce que l'on entend par là en section 1.3.2) de chaque bâtiment individuel.

1.3.2 Comment le décrire ?

Nous avons utilisé plusieurs fois le terme de reconstruction de bâti sans préciser le sens exact que nous lui attribuons. Nous reprenons pour cela la définition donnée par [Fuc01] : une reconstruction est une représentation *vectorielle* d'une scène. Nous entendons par là un ensemble d'entités *géométriques* : sommets 3D, arêtes 3D et faces 3D reliées entre elles par des informations *topologiques* permettant d'énoncer des relations d'adjacence entre ces primitives : un point appartient à une arête, une arête borde une face... La reconstruction que nous recherchons regroupe donc l'ensemble de ces informations qui permet de décrire précisément la surface polyédrique formée par les pans de toits. C'est un modèle classique qui s'intègre naturellement dans des bases de données géographiques [SVP+96]. On insiste bien ici sur l'importance des deux types d'information : *géométrie* et *topologie*.
Concrètement, dans le cas de la reconstruction de bâti à partir d'images aériennes, les faces qui sont des portions de plan correspondent aux pans de toits, les arêtes aux gouttières, faîtes et arêtes de rive de toits et les points 3D aux intersections de ces différentes arêtes.
On notera que le passage des images à la représentation vectorielle que nous venons d'énoncer est un pro-

cessus classique de vectorisation qui permet de compresser l'information et de l'intégrer naturellement dans une base de données vecteur. Ce processus de passage d'une information raster à une information vecteur est le coeur d'un très grand nombre de travaux en 2D ou en 3D. C'est le cas, en particulier, à l'IGN pour la vectorisation automatique du cadastre : partant de planches de cadastre scannées, des outils automatiques permettent d'obtenir une représentation vectorielle de la zone. Ce passage raster->vecteur est *toujours* une étape délicate à réaliser.

1.3.3 Avec quelle précision ?

La précision recherchée est $\cdots$ la meilleure possible ! ! Compte tenu de la résolution et de la qualité des images (section 2.2.2), une précision de l'ordre de 25cm peut être envisagée. Nous verrons toutefois que les principes de généralisation peuvent nous amener à définir des sommets qui ne correspondent pas à des objets physiques réels. Sur ces sommets, la précision recherchée est la *même* en ce sens que le sommet doit être positionné le plus précisément possible par rapport au sommet *virtuel* qui se trouverait là si on considérait un bâtiment généralisé.

1.4 Plan du manuscrit et contributions

Le chapitre 2 de ce mémoire récapitule les difficultés intrinsèques à la reconstruction de bâtiments en zone urbaine, les méthodes les plus couramment employées et justifie en la présentant de façon générale la méthodologie adoptée dans cette thèse. Il constitue une contribution intéressante de cette thèse dans la mesure où nous nous sommes efforcés de comparer les approches les plus connues en fonction de différents critères et de synthétiser cette comparaison dans un tableau récapitulatif.
La suite s'articule en 3 parties bien distinctes, la première partie étant consacrée à la détection de primitives tridimensionnelles. Dans un premier temps, on propose un algorithme de détection de primitives linéaires (chapitre 3) qui a fait l'objet de deux publications [TD02a, TD02b] et qui a été utilisé dans différentes caricatures de scènes urbaines [MPT02, FJM$^+$02, PMT$^+$01]. La détection de facettes planes fait l'objet du chapitre suivant (chapitre 4) et a donné lieu à une publication [TGD03]. Dans ces deux algorithmes, un effort est fait, autant que faire se peut, sur la quantification du gain du multi-vues par rapport à la stéréoscopie simple.
La deuxième partie concerne les algorithmes "haut-niveau" de génération d'hypothèses de bâtiments à partir des primitives planes et de choix du meilleur polyèdre sans surplomb en utilisant en particulier des contraintes a priori (chapitre 5). Cette étape qui constitue la contribution majeure de la thèse permet la reconstruction d'un bâtiment générique tout en autorisant l'introduction de données externes architecturales par le biais de contraintes privilégiées en milieu urbain. Nous verrons que ces deux aspects permettent de combiner à la fois les impératifs de généricité et de généralisation du problème que nous exposerons dans la suite. Ce chapitre a fait l'objet de deux publications [TD04b, TD04a]. Dans le chapitre suivant est décrit l'algorithme de recalage de modèles avec application de contraintes (chapitre 6) qui constitue le prolongement naturel de la phase de détection/reconstruction. L'approche proposée dans ce chapitre peut toutefois s'étendre à des modélisations plus complexes, tout automatiques ou semi-automatiques et revêt donc une importance supplémentaire pour des applications ultérieures en reconstruction de bâti.
Enfin la dernière partie contient une évaluation du système et la conclusion (chapitres 7 et 8). Le lecteur trouvera également en fin de document, outre les annexes et la bibliographie habituelles, une liste des publications relatives aux travaux présentés ainsi qu'un glossaire des principaux termes employés. Ce glossaire reprend la définition des termes techniques utilisés habituellement en photogrammétrie et en vision par ordinateur et s'adresse donc essentiellement au lecteur non initié. Nous insistons sur le fait que les termes plus spécifiques à nos travaux ne sont pas définis dans ce glossaire mais à l'intérieur du document.

La contribution majeure de cette thèse repose sur le développement d'un système de reconstruction automatique de bâti basé sur l'introduction de contraintes dans les étapes de choix de la meilleure représentation et de recalage du modèle a posteriori. À mi-chemin entre les approches classiques à base de données qui sont générales mais peu robustes et les approches à base de modèles, robustes mais peu générales, ce système, nous en sommes convaincus, constitue une alternative sérieuse pour la reconstruction automatique en milieu urbain. Nous espérons que ce manuscrit ralliera le lecteur à notre opinion...

Chapitre 2

Analyse et Stratégie

2.1 La reconstruction du bâti : une problématique complexe

2.1.1 Généricité et généralisation

La reconstruction de bâtiments à partir d'images aériennes est un problème extrêmement complexe. L'ensemble des approches se heurte à deux problématiques contradictoires : la nécessité de pouvoir reconstruire n'importe quel bâtiment et donc la *généricité* de l'approche et la possibilité de faire intervenir dans le processus des simplifications ou caricatures de forme (*généralisation*) bien connues des cartographes pour éliminer les petits détails (chiens-assis, lucarnes, cheminées, décrochements faibles...) au profit de structures plus importantes.

Le principe de généralisation permet d'accepter pour un même bâtiment plusieurs représentations. Ainsi, sur la figure 2.1, les représentations du même bâtiment, qui ne différent que par le niveau de détail, sont *toutes* acceptables en fonction du niveau de généralisation toléré. Ce principe de généralisation rend le problème de reconstruction de bâti extrêmement mal posé puisque plusieurs modélisations peuvent être solutions. En revanche, nous verrons que ce même principe est essentiel pour pallier les déficiences des détecteurs de primitives tout en proposant un modèle valide de la zone.

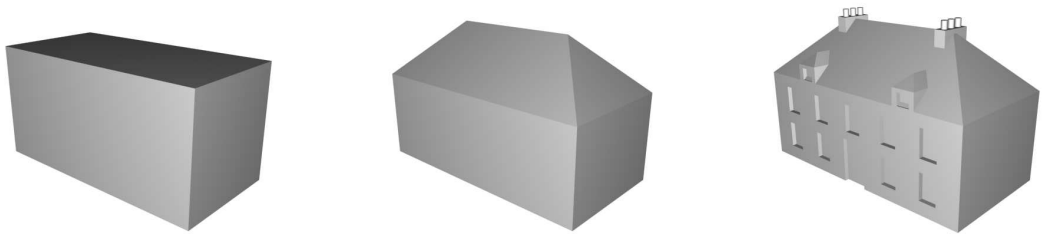


Figure 2.1 - Exemples de généralisation. Trois représentations à des niveaux de détail différents du même bâtiment.

Nous définissons comme *niveau de généralisation* la taille maximale du détail qui peut être omis ou caricaturé dans une modélisation représentative de la réalité. Ainsi, avec un niveau de généralisation toléré de 1m, les *deux* représentations de la figure 2.2 sont acceptables. On préférera toujours la représentation la plus précise, bien entendu (le processus de généralisation pouvant avoir lieu après la reconstruction) mais la représentation la plus caricaturée dans ce niveau de généralisation sera aussi considérée comme une

représentation acceptable de la forme. En revanche, si le niveau de généralisation accepté est inférieur au mètre, dans le cas de la figure, seule la première représentation est acceptable. Si le niveau de généralisation est nécessaire pour définir quelles représentations sont acceptables, il n'est pas suffisant. Il faut en effet que les représentations soient localement cohérentes. Par exemple, sur le bâtiment de la figure 2.2, si le décrochement est représenté en biseau et non avec un angle droit, la représentation n'est pas acceptable. Ici, le rôle des spécifications est donc primordial pour définir le niveau de caricature voulu pour l'application.

Par conséquent, l'objectif d'un système de reconstruction est d'obtenir une *caricature* (répondant à un ensemble de spécifications) d'un bâtiment *générique*.

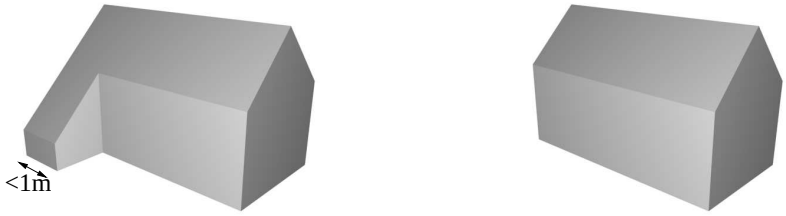


Figure 2.2 - Un bâtiment et une caricature de ce bâtiment. En fonction du niveau de généralisation toléré, la forme caricaturée est acceptée ou ne l'est pas (voir texte pour détail).

On notera que cette notion de généralisation a permis de justifier les approches de reconstruction de bâti à partir de modèles de bâtiments prédéfinis [Fuc01]. Implicitement, on suppose dans ce cas que la plupart des bâtiments peuvent, à un certain niveau de généralisation, se conformer à un des modèles. Cependant, la diversité des formes, y compris en incluant ce degré de caricature reste très importante. De plus, avec l'augmentation de la résolution des images et le besoin croissant d'une reconstruction la plus fidèle possible à la réalité, le niveau de généralisation toléré tend à diminuer ce qui rend ce type d'approches très délicat ou amène à accroître la bibliothèque de modèles nécessaire de façon exponentielle. On verra toutefois que le principe de généralisation est essentiel pour tenir compte de certaines erreurs des détecteurs et proposer ainsi une caricature du modèle qui soit une représentation acceptable mais non exacte de la réalité (sections 2.1.2 et 2.4.2).

Dans notre contexte, sans entrer dans une description précise et exhaustive de spécifications qui serait hors de propos, nous adopterons le même principe un peu flou que [Fuc01] en tolérant pour notre part un niveau de généralisation de 1,5m. A l'intérieur d'un toit, toute structure (décrochement, chien assis...) dont la taille caractéristique est supérieure à 1,5 mètre devra donc être représentée. Des détails de taille caractéristique inférieure seraient quoi qu'il en soit difficiles à reconstruire compte tenu de la résolution des images utilisées (section 2.1.2). C'est ainsi le cas de la presque totalité des chiens-assis ou des cheminées que nous ne chercherons pas, de fait, à modéliser.

2.1.2 Résolution des images

Nous avons abordé précédemment les niveaux de représentation qui mettent en exergue la notion de généralisation. Définir une échelle de travail est indissociable de la résolution des données avec lesquelles nous travaillons : un niveau de généralisation inférieur à la résolution des images n'a pas de sens. Il est ainsi impensable, comme nous l'avons déjà évoqué, de pouvoir obtenir une représentation aussi précise

que la troisième représentation dans la figure 2.1 compte tenu des données dont nous disposons qui ne nous permettent pas d'observer des détails aussi fins que les cheminées. La résolution des données limite le niveau de détail accessible et le niveau de généralisation toléré doit évidemment prendre en compte ce facteur. Le niveau de 1,5m que nous avons choisi tient compte à la fois de la résolution des images dont nous disposons (section 2.2.2) et du degré de caricature que nous acceptons.

Si la notion de généralisation rend le problème mal posé, elle permet en revanche une certaine latitude par rapport aux échecs inévitables des détecteurs de primitives 3D à partir des images (sur ou sous détection) dont le système doit tenir compte. La tolérance à un certain degré de caricature permet en effet de pallier une éventuelle absence de primitives dans les images par l'utilisation de *règles de construction*[1] qui devraient fournir *in fine* une caricature de bâtiment dans les limites de la tolérance.

2.2 Les données

L'étude se situe explicitement dans un contexte multi-vues ; c'est-à-dire que chaque point du terrain est vu en moyenne par strictement plus de 2 images (en pratique entre 4 et 15). L'apport du multi-vues sur la précision et l'exhaustivité des détecteurs est d'ailleurs une étude au coeur de cette thèse. Nous nous autorisons à utiliser toute donnée automatiquement issue des images.

2.2.1 Principe de la multiscopie

Cette section présente les principes fondamentaux des mécanismes de multiscopie. On désigne par ce terme une situation dans laquelle une scène est observée suivant n points de vue différents, avec $n > 2$, en opposition à la stéréoscopie où $n = 2$. Ce cas s'obtient dans le cas d'images aériennes lorsque, en projetant l'emprise des images au niveau de la scène, cette dernière appartient à la zone commune à toutes les emprises, encore appelée zone de recouvrement (figure 2.3).
Un point $\mathbf{M}$ de la zone de recouvrement se projette dans *chaque* image en un point $\mathbf{m}_i$. Pour effectuer cette projection, il est toutefois nécessaire de connaître les paramètres de projection pour chaque cliché i. Ces éléments sont constitués des paramètres intrinsèques et extrinsèques de la caméra liée à l'image i. Les paramètres intrinsèques regroupent l'ensemble des données issues de la calibration interne de la caméra (focale, position du point principal dans l'image, taille des pixels, distorsion éventuelle). Les paramètres extrinsèques sont liés à la position de la caméra dans l'espace au moment de la prise de vue et regroupent donc les paramètres de translation et de rotation définissant la position du sommet de prise de vue ($\mathbf{C}_i$) et l'orientation de la caméra (roulis, tangage, lacet) dans le référentiel global. Les paramètres extrinsèques constituent la calibration externe de la caméra et sont déterminés par des techniques d'aérotriangulation que nous ne détaillerons pas dans ce manuscrit (voir [Kra93, BFJW01]). Nous supposerons dans la suite que *tous* ces éléments de calibration sont connus et nous dirons donc que les images sont parfaitement géoréférencées. En supposant que nous sommes dans un contexte de caméra sténopé (hypothèse que nous adopterons dans toute la suite), on sait donc parfaitement calculer la position des points images $\mathbf{m}_i(u_i, v_i)$ à partir du point 3D $\mathbf{M}(X, Y, Z)$ par une relation faisant intervenir la matrice 3x4 de transformation perspective $\mathcal{T}_i$:

$$s \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \mathcal{T}_i \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{2.1}$$

[1] elles peuvent être globales dans des modèles prédéfinis ou locales par des contraintes sur les primitives et portant sur des configurations privilégiées.

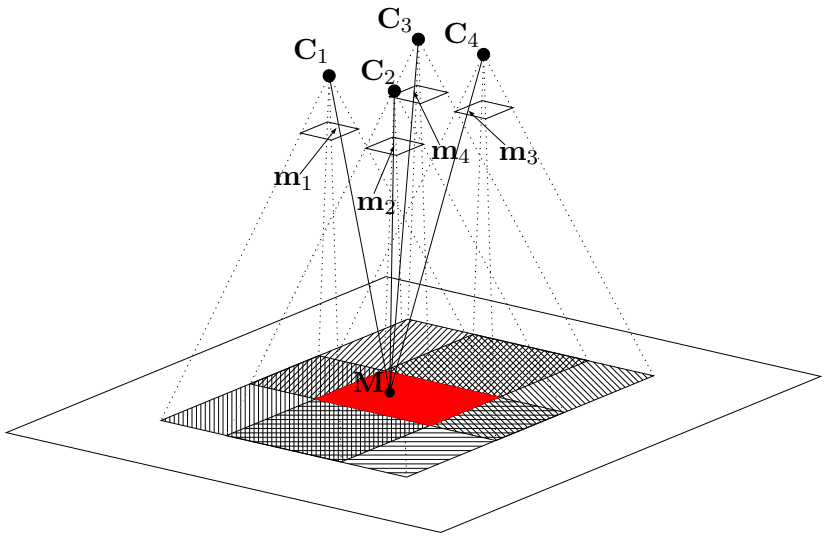


Figure 2.3 - Principe de la multiscopie. La zone en rouge correspond à la zone de recouvrement dans laquelle un point se projette dans toutes les images.

où $s \in \mathbb{R}^*$ est un facteur d'échelle arbitraire. De cette relation, on peut facilement déduire les valeurs de u_i et v_i en éliminant le paramètre s. Les matrices $\mathcal{T}_i$ contiennent les paramètres intrinsèques et extrinsèques des caméras et sont donc parfaitement connues.
Inversement si les points $\mathbf{m}_i$ sont connus, alors la position du point 3D correspondant $\mathbf{M}$ s'obtient en minimisant la distance aux rayons perspectifs $\mathbf{C}_i\mathbf{m}_i$. Toute la difficulté consiste à mettre en correspondance automatiquement ces points $\mathbf{m}_i$, c'est-à-dire à déterminer qu'ils correspondent tous à la projection du *même* point 3D. Ce problème de mise en correspondance peut se transposer aux segments (chapitre 3), aux régions ou à tout autre objet.
Dans notre contexte, comme nous l'avons déjà dit, nous nous placerons systématiquement dans des conditions de prise de vue qui assurent que tout point de la scène est vu par un minimum de 3 images.

2.2.2 Images

Les images utilisées sont des images noir et blanc issues de la caméra numérique de l'IGN dont la résolution est d'environ 25cm (figure 2.4)[2], ce qui signifie qu'un pixel représente, au sol, une dalle carrée d'en moyenne 25cmx25cm. On notera que par rapport aux travaux antérieurs [Fuc01], la résolution est moins bonne (8cmx8cm). Toutefois, la qualité des images numériques permet d'obtenir un bien meilleur rapport signal sur bruit que pour les clichés argentiques scannés utilisés dans son contexte.

2.2.3 Produits dérivés : MNE

Tous les produits automatiquement dérivés de ces images pourront être utilisés dans le cadre de cette thèse. C'est en particulier le cas des Modèles Numériques d'Élévation (MNE) obtenus par corrélation. On appelle MNE une représentation maillée de *toute* la surface topographique, c'est à dire de l'ensemble des objets vus depuis les clichés (sol, bâtiments, végétation, ponts, etc compris) [3]. Cette représentation est la

[2]Bien que les caméras numériques de l'IGN fournissent maintenant des images couleur avec la même résolution, nous n'étudierons pas l'impact de cette information dans nos travaux, étant donné que la couleur n'était pas disponible au début de nos recherches.

[3]On fera ici l'amalgame entre MNE et Modèle Numérique de Surface (MNS). Dans la communauté "photogrammétrique", à strictement parler, un MNE ne s'applique *qu'aux* objets du sursol tels que bâtiments ou végétation à l'exclusion du sol. Le MNS

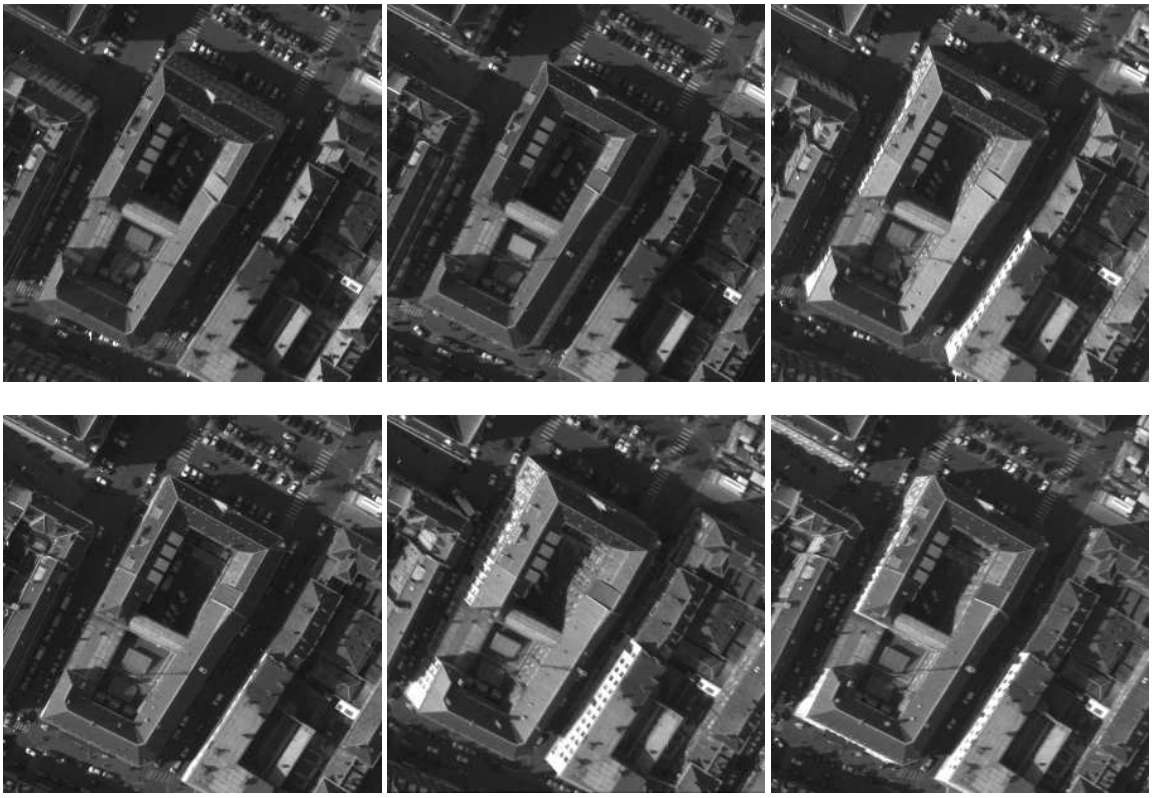


Figure 2.4 - Exemple de 6 images aériennes d'une même scène. Un pixel représente au sol une dalle d'environ 25cmx25cm.

plupart du temps une représentation 2D1/2 sous forme d'une grille d'altitude fournissant pour chaque (u, v) une altitude z. Le géoréférencement de la grille intrinsèque au MNE permet alors de définir le point 3D (X, Y, Z) correspondant.
Beaucoup de recherches ont été menées pour automatiser la génération de MNE, entre autres à partir de techniques de mise en correspondance de points entre images par corrélation de vignettes. Nous ne nous étendrons pas sur ce domaine très vaste, mais signalons juste que nous utiliserons les MNE de corrélation obtenus par l'algorithme de C. Baillard [Bai97] à partir d'un couple stéréoscopique. Cet algorithme utilise une programmation dynamique le long des lignes épipolaires du couple stéréoscopique et tient compte des parties cachées inévitables en stéréoscopie. La fusion des différents couples dans le cas du multi-vues permet d'assurer la densité du MNE résultant que nous utiliserons à plusieurs reprises (voir figure 2.5).

2.3 État de l'art en reconstruction de bâti

De nombreux systèmes automatiques ou semi-automatiques existent et se développent pour détecter et reconstruire des formes de plus en plus complexes de bâtiments *à partir d'images aériennes*. Dans la suite, nous ne chercherons pas à faire un inventaire exhaustif des systèmes existants dont une très bonne synthèse peut être trouvée dans [Ves00] ou [May99], mais nous nous efforcerons d'approfondir certaines stratégies afin d'en cerner les forces et les faiblesses. Cette étude permettra de dégager et de justifier l'approche choisie

concerne au contraire toute la surface topographique, sol compris. En revanche, dans la communauté "Vision par Ordinateur", on emploie le terme MNE (Digital Elevation Model en anglais) pour désigner une représentation maillée de *toute* la surface. Afin de ne pas créer de confusion, nous avons donc choisi d'utiliser cette définition dans la suite.

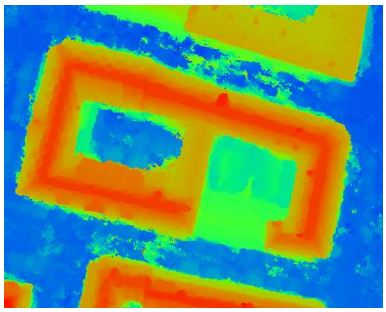


Figure 2.5 - Exemple de MNE dense obtenu en multiscopie sur la zone d'exemple illustrée par les images. Le MNE est référencé par rapport à la scène topographique et non pas aux images, ce qui explique le changement d'orientation.

dans le cadre de cette thèse pour la modélisation des bâtiments.
Étudier les systèmes de reconstruction automatique de bâtiments peut se faire en fonction de différents paramètres :

- le contexte : automatique, automatique avec données cadastrales, semi-automatique (intervention de l'opérateur),
- les données source : MNE laser, monoscopie, stéréoscopie, multiscopie,
- la complexité des bâtiments reconstruits et leur description sous forme : générique, structurelle, paramétrique, prismatique,
- la stratégie générale : bottom-up, top-down, hypothesize-and-verify (bottom-up puis top-down).

Un récapitulatif des différents systèmes et de leurs caractéristiques apparaît dans le tableau 2.1. La dernière ligne de ce tableau récapitule les caractéristiques de l'approche que nous avons mise en place et permet de la comparer aux travaux antérieurs. Une description et une comparaison plus fines sont proposées en section 2.4.2.

2.3.1 Contexte

Le contexte se différencie suivant les degrés d'intervention de l'opérateur et les données sources utilisées. Nous différencierons ainsi les méthodes automatiques sans données cadastrales en entrée, les méthodes automatiques utilisant des données cadastrales en entrée et les méthodes semi-automatiques dans lesquelles l'opérateur fournit tout ou partie de l'information. Les limites cadastrales sont, encore une fois, mises en avant en raison de l'information structurante très forte qu'elles apportent et des différences d'approche qu'elles ont suscitées dans les travaux : il n'y a guère d'intérêt à les comparer directement à des méthodes tout automatiques. Bien que les deux derniers contextes soient explicitement exclus du cadre que nous nous sommes fixés, il est intéressant d'étudier quelques-uns des systèmes utilisés dans ces contextes pour la modélisation des bâtiments qu'ils proposent.

Approches semi-automatiques

CC modeler (ETHZ) [GW98] est certainement le système semi-automatique le plus intéressant. L'opérateur saisit sur un appareil de restitution analytique classique les points du bord du bâtiment dans un ordre précis (trigonométrique ou anti-trigonométrique) puis dans un ordre dispersé les points internes au bâtiment. Le système est alors capable de reconstituer la structure topologique du bâtiment. L'immense avantage de l'algorithme est l'absence de données externes utilisées dans le processus : pas de MNE, pas de limites cadastrales. En outre, *n'importe quelle forme polyédrique* peut ainsi être représentée. La formulation topologique d'une forme polyédrique est très prometteuse et de portée très générale. Les auteurs ont pu montrer

Recherche	Contexte	Données sources	Modèles de bâtiments	Stratégie
[Gru98]	Semi-automatique	stéréoscopie	Paramétrique	-
[GW98]	Semi-automatique	stéréoscopie	Polyédrique	-
[Koe02]	Semi-automatique	stéréoscopie	Polyédrique	-
[VS01]	Automatique-cadastre	multiscopie	Paramétriques (CSG)	Top-Down
[FMJ03]	Automatique-cadastre	MNE	Polyédrique (modèles depuis cadastre)	Top-Down
[Jib02]	Automatique-cadastre	stéréoscopie	Polyédrique	Bottom-Up
[VS01]	Automatique-cadastre	Laser	Polyédrique	Bottom-Up
[BHF01]	Automatique-cadastre	Laser	Polyédrique	Bottom-Up
[LN98]	Automatique	monoscopie	Prismatique (parallélépipèdes)	Hypothesize-And-Verify
[VC02]	Automatique	MNE + orthoimage	Prismatique (parallélépipèdes)	Top-Down
[ODZ03]	Automatique	MNE	Paramétrique	Top-Down
[WF95]	Automatique	MNE	Paramétrique et prismatique	Bottom-Up ou Top-Down
[RB02]	Automatique	Laser	Polyédrique	Bottom-Up
[NN01]	Automatique	multiscopie	Prismatique (parallélépipèdes)	Hypothesize-And-Verify
[CJC+98]	Automatique	multiscopie	Prismatique	Bottom-Up
[JHRS97]	Automatique	multiscopie	Paramétriques	Hypothesize-And-Verify
[Fuc01]	Automatique	stéréoscopie	Structurels	Top-Down
[FKL+98]	Automatique	multiscopie	Structurels	Hypothesize-And-Verify
[BZ99]	Automatique	multiscopie	Polyédrique	Bottom-Up
[HLF00]	Automatique	multiscopie	Polyédrique	Bottom-Up
[AF00]	Automatique	multiscopie	Polyédrique	Bottom-Up
[Hen98]	Automatique	multiscopie	Polyédrique	Bottom-Up
[WA96]	Automatique	multiscopie	Polyédrique	Hypothesize-And-Verify
[SMVG02]	Automatique	multiscopie	Polyédrique	Hypothesize-And-Verify
[TD04a]	Automatique	multiscopie	Polyédrique / Structurels	Hypothesize-And-Verify

Table 2.5 - Récapitulatif des différents systèmes et de leurs caractéristiques.

l'intérêt de cette modélisation pour toute forme y compris très complexe. Récemment, ils ont également mis en avant une modélisation géométrique des bâtiments en introduisant dans le système la possibilité d'utiliser des contraintes géométriques telles que coplanarité, colinéarité... En revanche, l'algorithme ne permet pas de gérer des points faux qui sont inévitables dans le cas d'une recherche de processus automatique : les points saisis par l'opérateur sont *supposés fournir une description exhaustive et juste* du bâtiment. On notera que ce système est une version plus générique d'un système antérieur utilisant une batterie de modèles : TOBAGO [Gru98].
La chaîne de production TRAPU® qui existait à l'IGN dans les années 90 reposait sur la saisie par l'opérateur des informations topologiques et géométriques. Seules les hauteurs des façades étaient déduites automatiquement à partir des informations géométriques saisies et d'un Modèle Numérique de Terrain (MNT)[4]. C'est la même stratégie qui est employée à l'INSA [Koe02].

Approches automatiques à partir de limites cadastrales

Les limites cadastrales définissent les emprises au sol des bâtiments et sont de plus en plus utilisées pour faciliter la reconstruction de bâtiments.
A l'institut photogrammétrique de Stuttgart, Claus Brenner [BHF01] a développé un système original basé sur une classification des différentes jonctions triples. Partant d'une emprise cadastrale, l'auteur infère un ensemble de points de jonction triple par intersection des plans. Les plans sont, dans un premier temps tous censés avoir la même pente ce qui limite l'intérêt de l'approche, mais fournit des résultats encourageants. Dans un deuxième temps, s'appuyant sur un Modèle Numérique de Surface, l'auteur développe une autre approche à base de règles qui permet d'étiqueter de façon cohérente les différents plans détectés sur le MNE. La reconstruction fournie est alors très encourageante, mais il n'y a aucune remise en question des limites cadastrales et aucune gestion de discontinuités altimétriques éventuelles au sein de l'emprise de travail.
Dans le même ordre d'idées, Hassan Jibrini à l'IGN [Jib02] a développé un système automatique qui utilise l'emprise cadastrale pour la recherche de la forme des bâtiments. L'ensemble des plans s'appuyant sur les bords du cadastre est détecté par une technique de transformée de Hough 3D. L'intersection de tous ces plans fournit un ensemble de facettes parmi lesquelles il faut choisir le sous-ensemble qui soit le plus proche du MNE initial et qui ait une complexité de structure minimale. La contrainte imposée est que le modèle soit une surface polyédrique continue, qui puisse s'écrire z=f(x,y) et qui s'appuie sur les hypothèses de plans. La technique la plus simple dans le cas où le nombre de facettes n'est pas trop important est de parcourir l'ensemble des modèles possibles par recherche des cliques maximales sur un graphe de compatibilité. Dans le cas contraire, Jibrini restreint les surfaces possibles par des heuristiques simples avant de choisir la forme définitive. Cette approche a l'avantage de former un ensemble de modèles génériques. Ici, l'emprise cadastrale permet de déterminer *a priori* les orientations des plans et donc de guider le processus. Là encore, aucune remise en cause des limites de cadastre n'est possible et les discontinuités altimétriques éventuelles ne sont pas gérées. En revanche, la formulation générique de la surface polyédrique permet de contrer les surdétections de plans et d'intégrer des notions de simplicité de forme.
David Flamanc et Grégoire Maillet [FMJ03] utilisent une stratégie de découpage a priori de l'emprise cadastrale. Un ensemble d'hypothèses de bâtiments est généré à partir d'un pseudo squelette de cette emprise et de différentes hypothèses de terminaison (attaque en V, attaque droite). Le meilleur modèle est ensuite choisi par adéquation avec le MNE. Cette stratégie de découpage se retrouve chez [HB98]. Dans ce cas, on perd bien entendu un peu de généralité : tous les modèles de bâtiments ne peuvent être reconstruits. En outre, seuls les modèles avec pans de toits symétriques sont gérés dans [FMJ03].

De façon générale, dans toutes ces applications, le cadastre n'est pas remis en cause alors que les déli-

[4]contrairement au MNE, un MNT ne représente que la surface du sol, à l'exclusion des objets constituant le sursol que sont la végétation, les bâtiments en particulier.

mitations cadastrales sont parfois incertaines. En outre, aucun système ne permet de gérer de discontinuités internes qui apparaissent lorsque l'emprise ne délimite qu'un îlot et non un bâtiment individuel. En revanche, le cadastre apporte une information structurante très importante, en particulier quant à l'orientation des pans de toit.

Approches tout automatiques

Toutes les autres approches sont tout automatiques et seront étudiées plus avant dans la suite de cet état de l'art.

2.3.2 Données sources

Les données sont de différents types et peuvent être mélangées. On notera ainsi l'utilisation de données laser, de la multiscopie, stéréoscopie ou monoscopie. [May99] étudie aussi les différents systèmes en fonction de la résolution des données. Nous nous cantonnerons aux images dont la résolution oscille entre $8cm$ et $50cm$ et excluons *de facto* dans notre étude toutes les recherches à partir de données satellitaires dont la résolution actuelle ne permet pas d'obtenir une finesse de reconstruction satisfaisante.

Monoscopie

Les systèmes à base de monoscopie tendent à disparaître. La monoscopie est un contexte beaucoup trop pauvre pour espérer une reconstruction tridimensionnelle réaliste d'un ensemble de bâtiments. Tous les systèmes [LN98, MS94] s'appuient de fait sur une modélisation extrêmement simple des bâtiments en prisme à base rectangulaire. Ils proposent tous une approche par groupement de primitives parmi lesquelles une recherche de parallélogramme est effectuée. La validation du modèle se fait en utilisant les informations d'ombres projetées.

MNE et/ou orthophotographie

On citera les travaux de [VC02, ODZ03, WF95] qui n'utilisent comme simple hypothèse qu'un MNE pour l'extraction et la reconstruction de bâtiments. [VC02] effectue a posteriori un recalage dans une orthoimage pour affiner les paramètres de ses bâtiments. Dans les deux premières approches, les auteurs utilisent des modélisations paramétriques des bâtiments qui sont les modélisations les plus simples et donc les plus robustes dans des conditions très délicates, la troisième approche utilise des modèles paramétriques pour les bâtiments les plus simples (emprise rectangulaire) et un modèle prismatique dans les autres cas.

Stéréoscopie

La stéréoscopie est le contexte habituel de saisie d'informations 3D par un opérateur. Les recherches en automatisation du processus ont donc conduit les premières thèses à s'orienter vers ce contexte. La stéréoscopie, si elle permet d'obtenir une information 3D fiable pose néanmoins de sérieux problèmes d'occultations qui s'avèrent rédhibitoires dans un contexte urbain dense. De fait, la plupart des travaux utilisent des modèles de bâtiments simples : segmentation en facettes planes sans recherche de topologie [CJC01], bâtiments parallélépipédiques chez [Dan94]. Toutefois, quelques auteurs se sont penchés sur des modèles plus complexes : structurels chez [Fuc01] qui tire parti d'une formalisation sous forme d'isomorphisme de graphes avec distance d'édition pour pallier les sous-détections de primitives ou encore prismatiques pour [PCJC98]. Là encore, les premiers résultats ont prouvé la difficulté d'automatiser un processus de production qui, lorsqu'il est réalisé manuellement, repose beaucoup sur les capacités d'interprétation de l'opérateur et sa connaissance du milieu urbain.

Multiscopie

Face aux difficultés de la stéréoscopie, beaucoup de travaux se sont tournés vers la multiscopie qui apporte une précision accrue et permet d'espérer des modèles plus complexes de bâtiments. Ainsi, si [JHRS97, NN01] travaillent sur des modèles paramétriques de bâtiments, [CJC+98] sur des modèles prismatiques et [FKL+98] sur des bâtiments définis structurellement, la plupart des autres travaux [WA96, HLF00, AF00, SMVG02, BZ99, Hen98] proposent une modélisation générique de bâtiments.

Données laser

Les données laser fournissent une information altimétrique non dense mais précise et fiable. La plupart des équipes (en particulier allemande, hollandaise et française) se tournent vers ces données pour la segmentation en facettes planes [VS01, HB98, RB02]. En revanche se pose alors le problème de recalage des données image et laser [BPDR03].

2.3.3 Complexité des bâtiments reconstruits

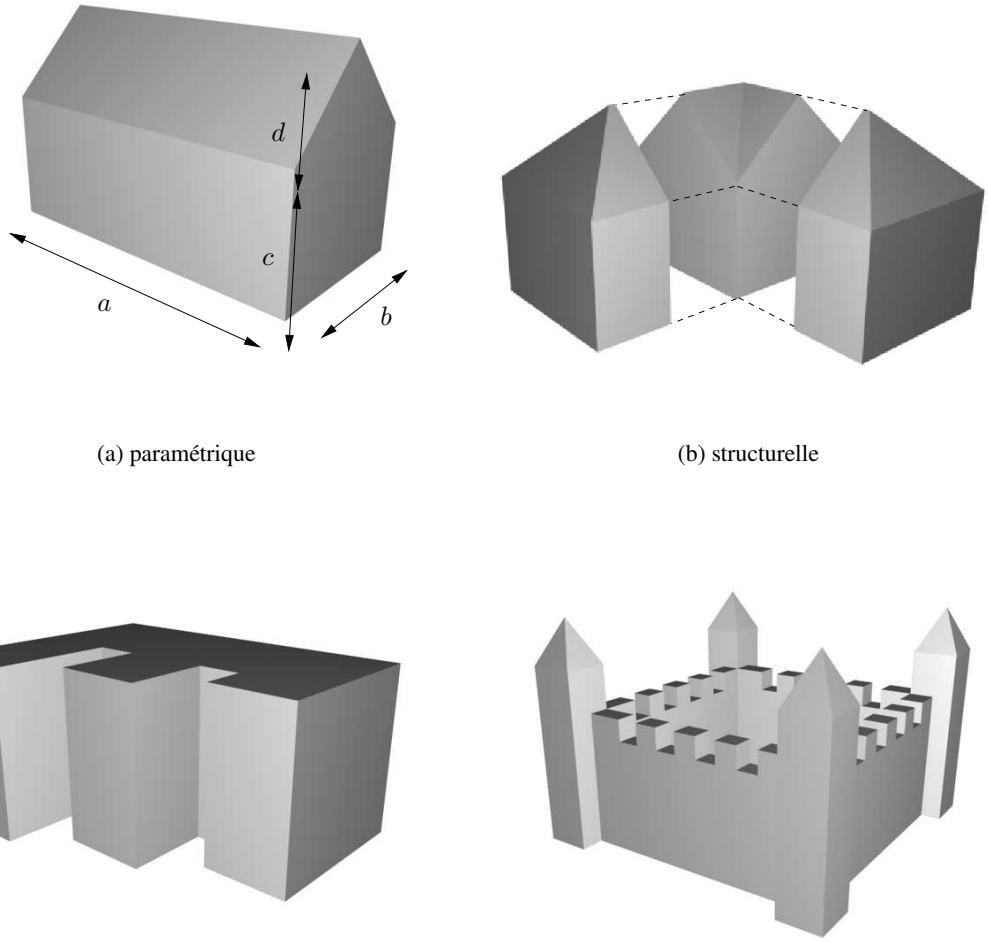
Tous les auteurs à quelques rares exceptions près comme [JHRS97] supposent généralement que les bâtiments sont polyédriques c'est-à-dire qu'ils peuvent être décrits par la juxtaposition de facettes *planes*. Cette définition reste suffisamment générale pour décrire la quasi-totalité des bâtiments. En revanche, dans cette classe, beaucoup de différences apparaissent. On distingue ainsi du moins contraint au plus contraint les modèles génériques susceptibles de représenter une forme polyédrique quelconque, les modèles prismatiques qui supposent une forme planimétrique polygonale quelconque mais un toit plat, les représentations structurelles qui décrivent un bâtiment comme assemblage de modules élémentaires simples, et les représentations paramétriques beaucoup plus contraintes où la structure topologique du bâtiment est prédéterminée et où seuls les paramètres géométriques de la forme, dont le nombre est fixé, sont à déterminer (figure 2.6).

Caricature de surface

Parmi les approches automatiques, on peut citer les approches à base de segmentation de MNE, telles que [Ves00, CJC01]. Tous ces travaux séparent les objets de la scène en zones de sol/sursol puis détourent les objets du sursol par polygonisation. Ensuite ils segmentent un MNE en facettes planes par des techniques MDL ou bayésienne. Dans [Ves00], une attention particulière est portée au détourage des bâtiments pour des raisons applicatives. Dans ces systèmes, on notera qu'aucune topologie n'est reconstruite. Ces systèmes s'apparentent plus à des caricatures de surface qu'à une reconstruction vectorielle de la scène au sens où nous l'entendons.

Modèles paramétriques

D'autres travaux utilisent des formes extrêmement simples dans lesquelles les modèles non extensibles ne permettent pas la description de scènes urbaines voire péri-urbaines complexes. Ce sont en général des travaux utilisant des images faible résolution ou travaillant en monoscopie. Dans ce cas, l'apport d'une contrainte externe forte permet de contrer le manque d'informations. On notera dans ce cadre les travaux de [LN98, NN01, VC02, ODZ03] qui se limitent à des détections de modèles à base rectangulaire. Pour les modèles de toit, [LN98, VC02] utilise un modèle de toit plat, d'autres ont une modélisation paramétrique de bipan [NN01] voire un peu plus complexe dans le cas de [ODZ03]. L'approche de [ODZ03] paraît un peu plus intéressante par les possibilités de caricature de scènes urbaines très complexes qu'elle propose et par la possibilité d'intégration de modèles plus élaborés dans le schéma utilisé. En revanche, la combinatoire reste une difficulté importante dans ces approches "Markov objet " qui, intrinsèquement, ne peuvent



(a) paramétrique

(b) structurelle

(c) prismatique

(d) générique

Figure 2.6 - Différentes descriptions des bâtiments.

gérer que des formes paramétriques. Dans le cadre de formes relativement simples, on notera également le projet ASCENDER [CJC+98]. Une première version permettait par groupement perceptuel d'extraire des formes prismatiques à base rectangulaire. L'amélioration apportée dans la deuxième version [JHRS97] a permis d'étendre ce modèle à des formes paramétriques plus complexes. En particulier, c'est un des rares cas avec l'exemple du Grand Palais dans [ML92] où une modélisation (paramétrique) de formes arrondies est envisagée.

Modèles structurels

Un autre axe de recherche utilise des modèles de bâtiments définis non pas de façon paramétrique mais structurellement pour restreindre la bibliothèque de formes à reconstruire [FKL+98, Fuc01]. L'apport d'une information externe forte permet là encore de contrer les problèmes de sous-détection de primitives (complétion de formes basées sur les modèles) ou de sur-détection (élimination de primitives générant des

formes qui ne sont pas dans la bibliothèque). En revanche, contrairement aux modèles paramétriques, la souplesse apportée par la modélisation structurelle du bâtiment apporte une réponse partielle au manque de généricité inhérent à l'utilisation d'une base préétablie de modèles.
[FKL+98] n'utilise qu'un seul type de primitives (coins) pour générer les hypothèses de bâtiments et l'ensemble des images pour choisir le bon modèle, Les auteurs cherchent à contrer le manque de généricité par des "règles" heuristiques consistant à agglomérer des parties simples de bâtiments (les "plugs" : attaque en V, terminaison de bipans) de façon à générer des formes beaucoup plus complexes. [Fuc01] est le premier à avoir formalisé l'introduction des modèles dans la reconstruction de bâti par une approche à base d'isomorphisme de sous-graphes avec tolérance d'erreur. Par recherche d'isomorphisme de sous-graphes, l'auteur cherche à déterminer le modèle qui correspond le mieux à un ensemble de primitives 3D détectées. Pour faire face aux sous-détections inévitables dans la recherche de primitives en stéréoscopie, l'auteur s'appuie sur un formalisme de contrôle d'édition permettant d'obtenir une certaine tolérance aux erreurs. En outre, grâce à l'introduction de grammaire de graphes, la base de modèles peut naturellement s'étendre et englober des modèles non définis au départ. Cette formalisation structurelle permet d'accroître naturellement et automatiquement la base des modèles envisageables. Cette technique est très intéressante par la formalisation du problème qu'elle propose ainsi que par l'approche à partir d'un ensemble de primitives disparates que l'on va confronter pour avoir le maximum d'informations. En revanche, les résultats fournis sont assez pauvres en terme de géométrie et, de plus, la méthode, malgré l'introduction de grammaires de graphes, ne peut générer que des formes relativement simples de bâtiments. Enfin, malgré un effort algorithmique évident, la complexité reste importante et le temps augmente avec l'extension de la base de modèles : plus les formes sont génériques, plus le temps de calcul devient rédhibitoire.

Les deux dernières remarques sont les reproches généraux que l'on peut faire à ce type d'approches : d'une part le manque de généricité et d'autre part un problème combinatoire évident lorsque la bibliothèque de modèles est étendue pour contrer ce manque de généricité.

Modèles prismatiques

Les systèmes utilisant une modélisation primastique des bâtiments sont assez rares. Nous citerons entre autres [CJC+98] qui est l'aboutissement du projet ASCENDER-I. La suite donnée à ce projet s'est orientée vers la recherche de structures 3D réelles permettant une description des formes internes des toits et témoigne du besoin d'obtenir une véritable reconstruction 3D plutôt qu'une délimitation planimétrique attribuée d'une seule altitude.

Modèles génériques

Parmi les systèmes les plus intéressants, plusieurs [Hen98, BZ99, HLF00, AF00] ont choisi de ne pas limiter a priori les formes à reconstruire. La plupart de ces approches utilisent un seul type de primitives (segments, coins ou plans) et gèrent difficilement les cas d'erreurs ou de sous-détection.

Henricsson [Hen98] utilise des attributs colorimétriques pour générer des modèles très complexes de toits à partir de segments 3D obtenus par mise en correspondance stéréoscopique. Les résultats proposés sont très encourageants, mais la méthode ne gère pas de cas de sous-détection de segments et les images utilisées ont une très grande résolution (8cm).
Baillard et Zisserman [BZ99] propose une méthode basée exclusivement sur les segments 3D. Autour de ces segments, l'algorithme détermine des hypothèses de facettes par corrélation dans les images et finalement reconstitue le toit par intersection des plans ainsi définis. Cet algorithme ne constitue pas un système à proprement parler dans la mesure où il n'est pas capable de gérer les cas de sous-détection et qu'il ne mise

que sur la détection des segments pour inférer des hypothèses de plan. Aucune modélisation explicite d'un bâtiment n'intervient.
L'approche de Heuel et al [HLF00] intègre à la fois des informations géométriques et des informations topologiques. Les coins sont au centre de l'algorithme : à partir des résultats dans l'image de Feature Extraction (FEX) [Fuc95], les auteurs infèrent les relations topologiques de diverses primitives des images entre elles. Ils reconstituent ensuite un ensemble de coins 3D d'ordre 3 (3 arêtes attenantes) qui constituent les éléments de base de leur structure topologique. Les coins sont en effet reliés entre eux si ils le sont dans les images ou si les directions de 2 arêtes sont compatibles. Cette technique est la première qui intègre pleinement une information topologique et géométrique tout en restant générique puisque n'importe quel polyèdre peut être reconstruit. Un certain nombre d'heuristiques basées sur des techniques de groupement perceptuel permettent de pallier quelques sous détections. L'inconvénient majeur de ce type d'approche réside dans le peu de fiabilité que l'on peut accorder aux relations topologiques dans les images ainsi que dans l'utilisation exclusive des coins comme primitives.
Ameri et Fritsch [AF00] utilisent une formalisation très générique de bâtiments comme l'agglomération de patchs plans permettant de reconstruire une forme polyédrique extrêmement générale. Sur une zone de focalisation, une segmentation est réalisée sur une image. Elle fournit des surfaces 2D qui sont ensuite plongées en 3D par régression sur des points d'un MNE. Les surfaces 3D ainsi obtenues sont ensuite intersectées si nécessaire en utilisant leurs relations de voisinage obtenues par un diagramme de Voronoï dans l'image. À l'issue de ce processus, les auteurs obtiennent un modèle grossier de bâtiment présentant de nombreux artefacts en particulier au niveau des frontières. Un recalage sur les gradients image permet ensuite de tenir compte de contraintes internes (coplanarité, alignement, orthogonalité) pour obtenir un modèle raffiné qui se conforme aux informations image. Les auteurs ne précisent pas clairement si les contraintes sont détectées automatiquement ou si elles ont fournies par un opérateur. L'algorithme dans son ensemble présente l'intérêt de pouvoir gérer n'importe quel type de forme et d'intégrer un ensemble de contraintes qui semblent contrecarrer une partie des défauts des détecteurs de primitives. En revanche, l'information "image" n'est pas remise en cause : aucun cas de sous-détection n'est géré. De plus, comme pour [HLF00], on peut s'interroger sur la fiabilité des relations topologiques détectées dans les images et sur lesquelles repose en grande partie la construction de la surface. Enfin, les contraintes ne sont pas intégrées dans le processus de choix du modèle de bâtiment et on ne sait pas si elles proviennent d'une information externe ou pas. On retrouve un système présentant de nombreuses similitudes chez [RB02] qui extrait toutefois ses patchs plans à partir de MNE Laser.

On notera également l'apparition de modélisations géométriques *et* sémantiques des bâtiments au contraire des précédentes approches qui ne reposent que sur une modélisation purement géométrique (les modèles de bâtiments ne sont définis qu'à partir des relations géométriques que les primitives possèdent entre elles). [WA96, SMVG02] proposent ainsi une modélisation dans laquelle les primitives acquièrent très tôt une valeur sémantique (arête de rives, faîte de toit, arête interne) et offrent une modélisation complète des bâtiments pour contrer les sous-détections des primitives. Ces systèmes peuvent gérer n'importe quel bâtiment respectant la définition, mais ils restent très complexes et d'après les premiers résultats relativement sensibles.

Comme on l'a déjà dit, les principaux reproches que l'on peut faire à ce type d'approche résident dans leur manque de robustesse aux erreurs des détecteurs de primitives, que ce soit en termes de sous-détection ou en termes de sur-détection. En outre, la plupart ne tirent avantage que d'un seul type de primitives sans utiliser la complémentarité des primitives que l'on retrouve dans [Fuc01].

2.3.4 Stratégie générale

Deux processus bien connus en algorithmie lorsqu'une information doit être extraite de données initiales sont utilisés dans le cas de la reconstruction de bâti : les approches bottom-up et les approches top-down. Les approches bottom-up sont plus génériques et partent des primitives pour inférer la reconstruction finale sans connaissance a priori du modèle que l'on cherche à reconstruire. Les problèmes de sous-détection de primitives sont, en général, traités localement par groupement perceptuel. Les primitives sont ensuite groupées par contraintes locales et raisonnements ad hoc. Dans le cas du raisonnement top-down, le modèle de bâtiment est connu à l'avance et des évidences de présence de ce modèle sont recherchées dans les informations à disposition. Les approches top-down sont intrinsèquement limitées par la librairie de modèles qu'elles sont capables de gérer. Elles sont toutefois en général plus robustes. L'intégration des deux processus constitue une stratégie hypothesize-and-verify où les hypothèses sont d'abord formulées à partir d'un ensemble de primitives par un processus bottom-up puis vérifiées dans l'ensemble des informations par une stratégie top-down.

Les stratégies top-down sont *toujours* mises en oeuvre dans les travaux faisant intervenir des bibliothèques de modèles. Les stratégies bottom-up sont *toujours* mises en oeuvre dans les travaux ne faisant pas intervenir de modèle a priori, dans leur première phase (inférence des formes possibles). Les stratégies mixtes ont un intérêt immense : elles ne limitent pas la forme a priori mais permettent de choisir le modèle suivant des critères locaux.

2.3.5 Bilan

On constatera que les stratégies sont, comme on pouvait s'y attendre, très dépendantes du contexte et des données utilisées. Dans les systèmes les plus pauvres en termes de résolution des images et de données utilisées (monoscopie ou stéréoscopie), les auteurs utilisent des batteries de modèles afin d'apporter l'information externe indispensable pour contrer les manques d'information. On obtient alors des résultats très caricaturés. La forme choisie pour ces modèles va de la plus simple (modèles paramétriques) à des modélisations structurelles assurant une certaine extensibilité des modèles de bâtiments.
Avec l'apparition de la multiscopie et l'augmentation de la résolution des données, les systèmes cherchent à modéliser de plus en plus finement les bâtiments et à pouvoir reconstruire n'importe quelle structure. Le problème devient alors de pouvoir contrer une information manquante ou une information superflue sans modèle restrictif.
On remarquera également que les systèmes les plus opérationnels en termes de production automatique de bases de données urbaines sont tous basés sur les limites cadastrales. Compte tenu des objectifs opérationnels de l'IGN qui vise, *à moyen terme*, à la mise à disposition d'un outil d'aide à la saisie de ces bases, il est essentiel, au moment de la conception d'un système, de prévoir la possibilité d'intégrer de telles informations. Bien que notre contexte n'inclue pas les limites planimétriques des bâtiments, nous avons gardé à l'esprit cette contrainte opérationnelle qui a influencé notre choix de stratégie.

2.4 Stratégie

2.4.1 Analyse

D'après ce que l'on a vu, deux tendances se distinguent nettement. La première est la recherche d'une modélisation générique des bâtiments. Avec l'apparition de missions aériennes (seul vecteur aérien susceptible de fournir des images à des résolutions suffisantes pour le moment) couvrant des zones à des résolutions de plus en plus élevées et en multiscopie, les modèles prédéfinis de bâtiments ne sont plus satisfaisants pour décrire de façon suffisamment fidèle la grande diversité des structures existantes.

La deuxième tendance est une recherche d'intégration de données externes. On notera en particulier l'apparition des données laser qui fournissent une information altimétrique non dense mais fiable. Et dans un contexte opérationnel, il nous faut encore une fois insister sur les données cadastrales. Les systèmes utilisant ces données en entrée sont les seuls systèmes automatiques dont les performances permettent d'espérer une mise en production proche. Dans notre contexte IGN où l'objectif à moyen terme est une mise en production d'un outil d'aide à la fabrication de bases de données tridimensionnelles, nous gardons à l'esprit cette problématique et, bien que notre contexte ne l'intègre pas, notre système sera conçu pour, à terme, pouvoir intégrer cette information.

2.4.2 Démarche

Un système générique, ouvert et contraint ! !

Compte tenu des remarques faites précédemment, en imposant un ensemble de modèles dès le début de la stratégie, il paraît difficile, voire impossible, de générer l'ensemble des formes visibles, même caricaturées, en zone urbaine ou péri-urbaine ; ensemble qui peut présenter des formes très variées.En outre, en général, la combinatoire augmente avec la taille de la bibliothèque de modèles ce qui réduit l'intérêt de ce type d'approches. D'un autre côté, les approches exclusivement à base des données peinent à extraire la bonne représentation et à contrer les problèmes des détecteurs de primitives.
Nous avons donc opté pour une méthodologie intermédiaire suivant une architecture hypothesize-and-verify, c'est à dire, dans un premier temps, partant des observations pour construire les modèles puis dans un deuxième temps, choisissant le meilleur modèle parmi l'ensemble des données. À partir de primitives 3D, l'objectif est de générer un ensemble de modèles de bâtiments répondant à une modélisation extrêmement générique de polyèdre sans surplomb puis de choisir, parmi ces modèles, le bâtiment qui vérifie le plus les données tout en étant le plus simple. Dans le processus de choix, des règles simples de caricature sont introduites par le biais de contraintes sur les primitives. En ce sens, la stratégie permet d'apporter une information "modèle" sans toutefois empêcher la généricité. La contrepartie évidente *semble être* de ne pouvoir gérer les cas de sous-détection des primitives. Nous comptons ici sur le multi-vues pour éviter ces problèmes. Cependant, de très simples règles de groupement perceptuel pourraient, le cas échéant, générer de nouvelles primitives. Cela ne sera pas traité dans le cadre de cette thèse. En outre, nous verrons dans la section suivante qu'une stratégie reposant sur l'information *géométrique* permet de contrer des cas de sous-détections de primitives par l'utilisation du principe de généralisation. Reprenant l'idée de F. Fuchs [Fuc01], nous comptons utiliser également la complémentarité des primitives pour créer ou choisir les modèles.

La géométrie induit la topologie

Reprenons l'exemple de la figure 2.2. Supposons que les (petites) façades de l'extension n'aient pas été détectées par les extracteurs de primitives et que toutes les primitives n'aient pas la bonne élongation, cas de figure qu'il nous faut prendre en compte, car ces sous-détections seront inévitables dans la phase de détection de primitives (figure 2.7). En utilisant un raisonnement topologique seul, il paraît difficile, voire impossible de reconstruire un modèle de bâtiment. En revanche, en utilisant un pur raisonnement géométrique, par intersection, on peut reconstruire un modèle de bâtiment qui soit *acceptable* avec le niveau de généralisation toléré. Ainsi, cette stratégie d'intersection, si elle ne permet pas de reconstruire le modèle parfait fournit des modèles acceptables. Si les primitives *les plus importantes* sont présentes en géométrie (c'est-à-dire si l'équation définissant la primitive est *à peu près* valide), alors un modèle acceptable peut être généré à partir de cette seule géométrie. Nous dirons ainsi que la *géométrie induit la topologie*. Cette stratégie d'intersection est à la base de notre système de génération d'hypothèses de bâtiments (première partie du hypothesize-and-verify). Elle est basée plus précisément sur l'intersection des primitives planes. De fait, l'analyse nous a conduits à adopter et étendre le modèle de [Jib02] reposant sur une modélisation analogue, mais n'incluant

pas les pans verticaux. Cela permettra également in fine d'intégrer naturellement les informations cadastrales dans le système.

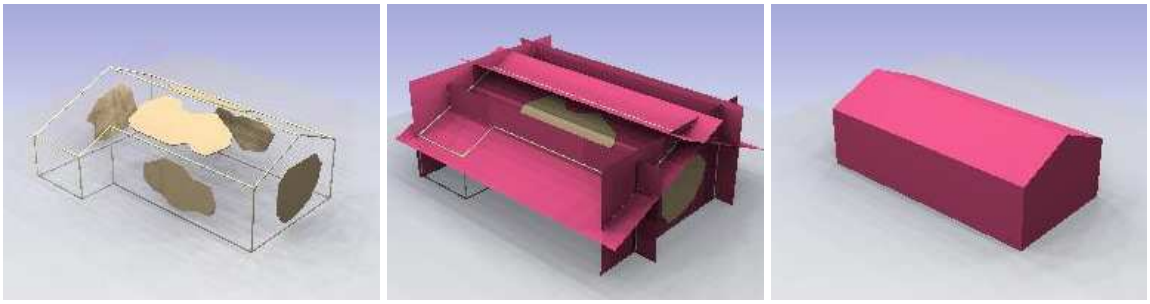


Figure 2.7 - Exemple du bâtiment de la figure 2.2. Les primitives planes détectées sont les patchs présents sur les plans dans la figure de gauche, le contour parfait est en filaire. Les intersections possibles des différents plans sont représentées sur la figure du milieu et permettent de reconstruire la topologie du bâtiment. Le résultat final est donc le bâtiment de droite correspondant à une représentation valide du modèle avec le niveau de généralisation toléré.

Choix des contraintes

Comme nous l'avons dit, nous introduisons, dans le processus de choix du meilleur modèle, des contraintes a priori pour choisir un modèle de bâtiment qui respecte quelques contraintes architecturales simples. Cette introduction a pour objectif essentiel d'aider au choix du bâtiment le plus proche d'un modèle intuitif lorsque les seuls modèles extraits du générateur d'hypothèses sont des modèles généralisés c'est-à-dire caricaturés et ne représentant donc pas exactement la réalité.
Nous avons choisi des contraintes "de bon sens". Elles ne reposent pas sur une réflexion poussée sur les caractéristiques architecturales mais sur des observations simples : les façades sont généralement parallèles ou perpendiculaires, les arêtes (de faite, de rives) sont très souvent horizontales, les pans de toit vérifient assez souvent une contrainte de symétrie verticale, certains toits sont plats... Nous inclurons donc dans notre système l'ensemble de ces contraintes qui guideront le choix du modèle et seront vérifiées au moment de la reconstruction proprement dite.

Complémentarité des primitives

Un des grands avantages de la méthode exposée par [Fuc01] est la complémentarité des primitives. [Fuc01] gère en effet trois types de primitives détectées *indépendamment* les unes des autres. Il tire ainsi avantage des différences de ces primitives. Ainsi, les primitives planes possèdent l'indéniable avantage d'être denses, c'est-à-dire de fournir une information en tout point de la scène. En outre, les équations définissant ces primitives sont en général bonnes car elles sont définies sur des zones importantes et peuvent donc tirer parti de cette relative étendue pour rejeter les grosses erreurs. Enfin, il paraît aisé d'obtenir une sursegmentation à partir de ces primitives. En revanche, les frontières sont en général mal localisées. Au contraire, les primitives linéaires (segments 3D) forment une information peu dense mais structurante et précise (section 3) et facilitent donc le découpage du modèle au niveau des frontières de ses facettes.
Dans notre approche, nous tirerons parti de la densité et de la sursegmentation des primitives planes pour générer les modèles possibles de bâtiment. Ces deux caractéristiques doivent nous permettre d'obtenir *toutes* les structures. Les primitives linéaires auront leur utilité dans le choix du meilleur modèle.

2.4.3 Récapitulatif

L'algorithme général est récapitulé sur la figure 2.9. Le découpage en modules traduit la volonté du travail dont ce rapport est l'écho de découper les problèmes et d'être capable de fournir des entités plus ou moins indépendantes. À défaut d'avoir un système global ayant un taux de succès de 100%, ce qui paraît utopique, chaque bloc peut être réutilisé dans un programme indépendant (ce qui, d'ailleurs, a été le cas pour l'algorithme de détection et de reconstruction de segments 3D [MPT02]). Un soin tout particulier sera donc apporté, tout au long du manuscrit, à l'évaluation de ces modules intermédiaires de façon à ce que ces parties puissent être réutilisées en toute connaissance de cause.
Comme indiqué, le système s'appuie sur un Modèle Numérique d'Élévation qui sert de base à la focalisation et à bon nombre d'algorithmes. Nous avons choisi, par commodité et disponibilité immédiate, le MNE de C. Baillard [BD00] qui a démontré sa robustesse et un bon comportement au niveau des discontinuités altimétriques. La fusion des MNE issus de couples stéréoscopiques assure la densité du MNE final. À noter que d'autres MNE sont envisageables (par exemple, une généralisation de l'algorithme de Roy [RC98] au cas multi-vues). La zone de focalisation ou polygone de focalisation est un polygone convexe définie manuellement et située à une altitude $z = z_g$. Il convient de bien préciser qu'en aucun cas, les contours précis du bâtiment ne sont donnés : l'utilisateur précise une vague zone dans laquelle se trouve un ou plusieurs bâtiments. Un seuillage d'altitude permet alors d'isoler les zones de sursol constituant le masque de focalisation. Un exemple d'une telle zone est fourni sur la figure 2.8. À noter que, sur ces zones peuvent se trouver des agrégats de bâtiments, des bâtiments fermés aux frontières non connexes (présence de cours intérieures) ou tout simplement des bâtiments indépendants. L'algorithme peut en effet gérer des défauts de focalisation. Idéalement et théoriquement, en faisant l'hypothèse de terrain plat, il peut traiter une ville entière, sans focalisation particulière. Pratiquement, la combinatoire serait rédhibitoire et le système nécessite, de fait, ces masques de focalisation.

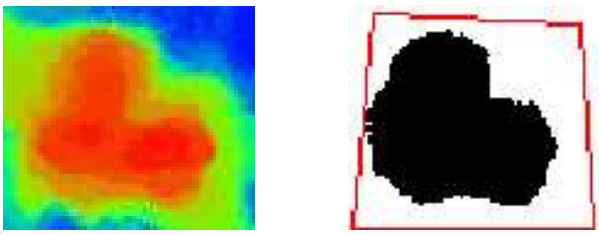


Figure 2.8 - Exemples de MNE, de zone et de masque de focalisation.

Le système mis en place est une extension de la stratégie de Jibrini [Jib02]. À partir de l'intersection de zones planes (façades comprises) détectées, un graphe 3D est généré, regroupant l'ensemble des facettes de l'arrangement de plans. De celui-ci est déduit un graphe de compatibilité à partir duquel l'ensemble des polyèdres sans surplomb est extrait par recherche de cliques maximales, reprenant en cela la modélisation de Jibrini (section 5.2). Dans cet ensemble de bâtiments possibles, la sélection est faite en fonction de l'adéquation du modèle aux données externes d'une part, et de la simplicité du modèle d'autre part, par le biais d'une modélisation probabiliste qui facilite l'introduction de nouvelles données externes. Pour l'instant, seuls les segments 3D, les images, le MNE et le masque de focalisation sont pris en compte, mais on peut très bien imaginer les points lasers, des points GPS, des saisies manuelles... Dans cette partie est introduite l'information a priori par les contraintes sur les primitives de base qui influe sur le choix du modèle via un terme de simplicité de modèle (section 5.3). Le système s'achève par l'introduction effective des contraintes pour la reconstruction du modèle (chapitre 6). En annexe D, le lecteur trouvera 7 exemples de fonctionnement de l'algorithme sur zones réelles faisant apparaître les étapes les plus importantes de la procédure.

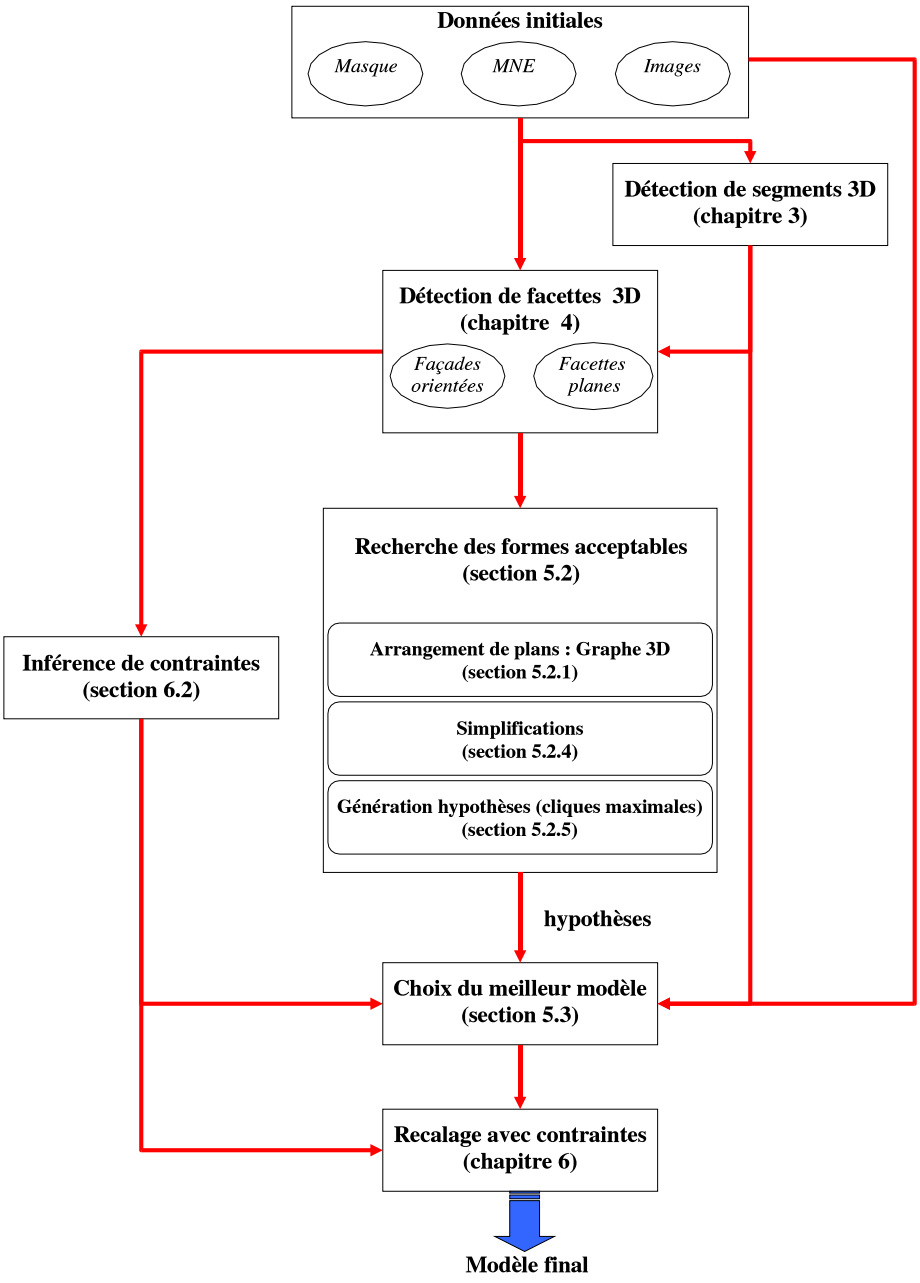


Figure 2.9 - Méthodologie générale. Chaque étape renvoie à la section dans laquelle elle est décrite.

2.5 Conclusion

Ce chapitre a présenté le contexte général de la reconstruction de bâti et plus spécifiquement les notions essentielles de généralisation et de généricité, le contexte plus restreint de travail dans le cadre de notre thèse et en particulier les données utilisées. Ensuite, après une analyse de l'existant, nous avons justifié notre approche et l'avons située par rapport aux travaux antérieurs, avant de présenter le canevas général de l'algorithme proposé.

Deuxième partie

Détection de primitives

Chapitre 3

Détection de segments 3D

3.1 Introduction

L'apport de segments 3D dans le cadre d'une modélisation de scènes urbaines paraît essentiel [BSZF99, WA96, NN01]. L'intérêt des segments 3D réside dans la fiabilité des primitives géométriques extraites ainsi que dans leur capacité de structuration d'un environnement. Ces primitives sont en effet *compactes*, car elles permettent une modélisation précise des objets tout en réduisant la complexité des algorithmes de mise en correspondance, *intrinsèques* car elles correspondent à la projection d'objets physiques, *robustes* au bruit, *discriminantes* de par leurs caractéristiques géométriques et *précisément situées* [Aya89]. La mise en correspondance de segments est un problème délicat essentiellement en raison des défauts de segmentation dans les images : bien que l'orientation de la droite puisse être déterminée avec une bonne précision, les extrémités ne sont pas fiables et les relations topologiques sont souvent perdues. De plus, en raison de phénomènes d'occultation ou de mauvaise segmentation, un segment 3D peut avoir des aspects différents d'une image à l'autre (morcellement, taille réduite...) voire totalement disparaître. En revanche, les contraintes géométriques liées à l'appariement de droites sont beaucoup plus fortes et permettent de discriminer efficacement les appariements potentiels. En outre, une fois les segments 2D mis en correspondance entre les différentes images, la qualité de la reconstruction est bien meilleure. Enfin, la faible densité de l'information à traiter permet de limiter la combinatoire et accélère considérablement les algorithmes. L'objectif ici est d'utiliser le maximum d'informations image afin de fiabiliser les appariements et d'affiner la reconstruction. Le contexte multi-vues permettra de pallier les défauts de segmentation dans différentes images. Un des éléments clés est d'éviter une combinatoire excessive et de ne privilégier aucune image.

3.1.1 État de l'art

La détection et la reconstruction de segments 3D est un problème qui a reçu une grande attention depuis de nombreuses années dans la communauté scientifique.

Stéréoscopie

La difficulté dans le cas de 2 images est que la contrainte géométrique est réduite à une contrainte de recouvrement. Pour pallier ce problème, les auteurs ont développé deux types de stratégies différents. La première repose sur la mise en correspondance de groupes de segments, ce qui permet d'introduire des contraintes très fortes. Ces méthodes de mise en correspondance de graphes [AF87, HS89] sont a priori plus fiables, mais sont souvent sensibles aux erreurs de segmentation qui modifient les apparences des graphes d'une image à l'autre. De plus, elles ont une grande complexité, ce qui les rend difficilement généralisables au

cas multi-vues. La deuxième stratégie repose sur l'introduction d'attributs géométriques divers tels qu'orientation, longueur, recouvrement [MN85,ZF92,GBB98] ou radiométriques [SZ97]. L'ensemble de ces attributs est souvent sensible aux erreurs de segmentation dans les images, mais également aux conditions de prises de vue. Ainsi, dans un contexte urbain, le voisinage radiométrique risque d'être très variable en fonction des prises de vue et des parties cachées. Cette remarque est particulièrement vraie pour les gouttières qui bordent les façades, compte tenu des textures ou couleurs très différentes entre le toit et la façade et entre la façade et le sol.

Multiscopie

Dans un contexte multi-vues, les approches se sont développées avec l'apparition du tenseur trifocal qui assure une contrainte géométrique forte [Har95b, Sha94, TZ97, PF98]. Des résultats prometteurs sont fournis par [SZ97] qui prouvent, en outre, que l'introduction de cette nouvelle contrainte permet de lever un grand nombre d'ambiguïtés. Ce tenseur pourrait également se généraliser aux cas utilisant strictement plus de 3 images avec les généralisations proposées par [FLP01, Moo96]. Les auteurs qui se penchent sur les séquences d'images ont développé des méthodes originales basées sur les filtres de Kalman [ZF92] qui font apparaître les notions de variance des segments reconstruits. D'autres auteurs font de la fusion de données stéréoscopiques en utilisant des chemins fortement connectés à l'intérieur de graphes [GBB98] ou de la fusion de segments résultant d'une détection dans plusieurs triplets d'images [BSZF99]. La plupart de ces techniques sont confrontées à des problèmes :

- *de combinatoire* : il faut étudier tous les triplets de segments,
- *de fusion d'information* : pour fusionner les différents triplets obtenus,
- *de dissymétrie* : car on part *toujours* d'une image maîtresse.

3.1.2 Algorithme

L'algorithme abordé ici présente une approche de reconstruction de segments 3D basée sur un appariement point à point dans l'espace-objet, totalement multi-image et symétrique (sans aucune *image maîtresse*) et qui génère un ensemble d'hypothèses d'appariement entre segments 2D (section 3.2). Cette approche permet de limiter la complexité dans le cas du multi-vues et de travailler en $\mathcal{O}(N)$ où N est le nombre d'images. Dans un deuxième temps, on présente une technique de reconstruction de segments 3D qui prend en compte l'incertitude sur chacun des segments 2D. La reconstruction fournit ainsi un score statistique utilisable *quel que soit le nombre d'images* qui permet d'invalider ou de valider rapidement un appariement donné en tenant compte de sa cohérence géométrique (section 3.3). Les appariements potentiels détectés dans l'Espace Objet sont ainsi élagués sur ce critère de cohérence géométrique mais également sur un critère d'unicité (section 3.4). Les résultats obtenus sur simulation ou données réelles prouvent la résistance de l'algorithme au bruit et à des défauts de segmentation dans plusieurs images (section 3.5).

3.2 Mise en correspondance dans l'Espace Objet

3.2.1 Techniques multi-images

La première étape de reconstruction 3D de primitives, réputée la plus difficile, est la mise en correspondance des primitives 2D entre elles. Ce problème, traité en profondeur dans le cas stéréoscopique reste difficile lorsque l'on est en multi-recouvrement car la recherche exhaustive de tous les appariements possibles possède alors une combinatoire rédhibitoire. Notre objectif ici est d'avoir une méthode qui respecte les 3 principes du "vrai" multi-images édictés par Collins [Col96] :

- la méthode doit être généralisable à un nombre *quelconque* (plus grand que 2) d'images,

– la complexité algorithmique doit être en $\mathcal{O}(N)$ où N est le nombre d'images,
– toutes les images doivent être traitées de la même façon (approche symétrique).

On pourra se référer à cet article pour plus de bibliographie concernant les techniques multi-images développées dans la littérature respectant ou ne respectant pas les conditions évoquées.
L'algorithme retenu pour la mise en correspondance se rapproche de la méthode décrite par l'auteur, qui consiste à déterminer des points de contours 3D par des mises en correspondance de points de contours 2D dans l'Espace Objet en considérant tous les voxels (volume élémentaire de l'espace 3D) successivement dans l'espace discrétisé (figure 3.1). L'intérêt de cette approche, outre qu'elle répond aux trois critères du vrai multi-images, réside dans l'importance accordée à la géométrie de prise de vue ainsi que dans la simplicité de mise en oeuvre. Le problème de cette méthode est qu'aucun lien explicite n'est fait entre les primitives mises en correspondance. Collins utilise un critère statistique pour définir s'il y a bien un point de contour 3D à la position (x,y,z) considérée en fonction du nombre de rayons perspectifs qui intersectent le voxel, sans relier explicitement les points de contours 2D entre eux. L'algorithme proposé étend cette méthode pour stocker les appariements de segments potentiels (c'est-à-dire qui ont au moins un point d'intersection). Ces appariements potentiels seront ensuite discriminés sur des critères de cohérence géométrique et d'unicité.

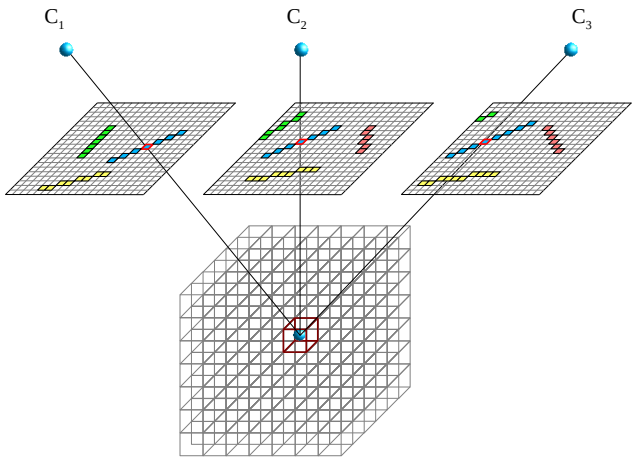


Figure 3.1 - Principe de l'appariement dans l'Espace Objet.

3.2.2 Mise en correspondance

Images de segments

Une extraction de contours suivie d'une polygonisation est tout d'abord effectuée sur chacune des images. On obtient ainsi, pour chaque cliché, une image de labels que l'on nommera *image de segments* (un point de contour appartient à un segment numéroté). Plus formellement, on définit $\mathcal{C}_i$ comme l'ensemble des contours labellisés de l'image i et $\mathcal{C}_i^{\infty} = \mathcal{C}_i \cup \{\Phi\}$ où Φ désigne l'absence de contours. Chaque pixel $\mathbf{p}$ d'une image de segments I_i vérifie donc $I_i(\mathbf{p}) \in \mathcal{C}_i^{\infty}$.

Approche par balayage de plans

La mise en correspondance se fait par balayage d'un plan dans l'espace 3D discrétisé. Un plan est déplacé le long d'une ligne qui lui est orthogonale. Dans le cas d'images aériennes, on choisira un plan horizontal partitionné en cellules alignées avec les axes X-Y de la scène, que l'on nommera *grille terrain*. Chaque cellule de la grille définissant un voxel (x, y, z_k), le balayage de z_{min} à z_{max} permet ainsi de parcourir

l'ensemble des voxels du volume. La taille "optimale" des cellules de la grille terrain ainsi que le pas de

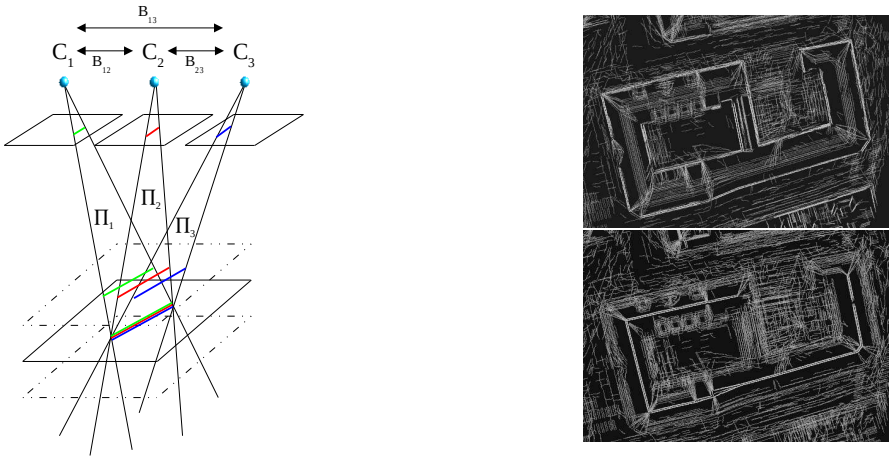


Figure 3.2 - 1ère colonne : technique du balayage de plans. 2ème colonne : 2 grilles terrain à des altitudes différentes.

balayage en z peuvent être fixés par l'opérateur pour tenir compte de la vitesse d'exécution ou des erreurs dans les paramètres d'aérotriangulation (paramètres de caméra). On peut néanmoins définir ces paramètres de façon automatique à partir d'informations connues sur le capteur ou la position du sommet de prise de vue au moment du cliché. La taille des cellules carrées de la grille t_g peut se définir par une relation simple entre la taille t *réelle* en mètres du pixel sur le capteur, la focale f et l'altitude de vol de l'avion h_i (figure 3.3) par rapport à la scène photographiée i dont l'altitude moyenne est $\bar{z}$:

$$t_g = \min_i \left[t \frac{h_i - \bar{z}}{f} \right]$$

Le pas de balayage Δ_z entre les altitudes z_{min} et z_{max} peut également être déterminé assez simplement (figure 3.4) grâce à une relation photogrammétrique classique reliant t_g et les bases B_{ij} entre les clichés i et j (distance entre les sommets de prise de vue) :

$$\Delta_z = \min_i \left[\frac{h_i - \bar{z}}{\max_j B_{ij}} t_g \right]$$

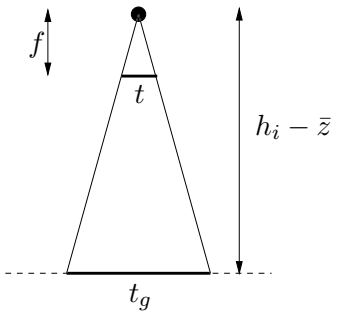


Figure 3.3 - Calcul de la taille des cellules de la grille terrain.

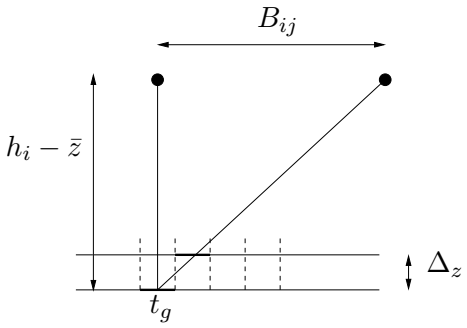


Figure 3.4 - Calcul du pas de balayage.

Grilles terrain

Pour une altitude donnée $z = z_k$, les images de segments sont projetées sur la grille terrain. Chaque point de contour labellisé vote dans un ensemble de cellules entourant l'intersection du rayon perspectif

et de la grille terrain. Pour déterminer l'ensemble de ces cellules, on considère le cône de sommet le centre de prise de vue et de section carrée équivalente au niveau de l'image à la taille du pixel. L'intersection de ce cône avec la grille terrain fournit l'ensemble des cellules dans lesquelles le point de contour vote (figure 3.5). Ce mécanisme permet de tenir compte de l'incertitude de localisation des points de contour, du fait que les points de contour sont localisés plus finement par les images proches que par les images lointaines et permet de rendre l'algorithme indépendant de la taille de la cellule de la grille terrain.

Nous donnons ici la méthode de calcul du nombre de cellules de la grille terrain intersectées par le cône. Considérons la transformation projective d'un point $\mathbf{M} = (x, y, z)$ de l'Espace-Objet en $\mathbf{m} = (u_i, v_i)$ dans l'image i. D'après l'équation 2.1, on a l'expression non linéaire :

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \frac{\mathcal{T}_{11}x+\mathcal{T}_{12}y+\mathcal{T}_{13}z+\mathcal{T}_{14}}{\mathcal{T}_{31}x+\mathcal{T}_{32}y+\mathcal{T}_{33}z+\mathcal{T}_{34}} \\ \frac{\mathcal{T}_{21}x+\mathcal{T}_{22}y+\mathcal{T}_{23}z+\mathcal{T}_{24}}{\mathcal{T}_{31}x+\mathcal{T}_{32}y+\mathcal{T}_{33}z+\mathcal{T}_{34}} \end{bmatrix} \tag{3.1}$$

où les indices i ont été omis pour plus de lisibilité. L'expression précédente se simplifie dans le cas d'un point 3D situé à une altitude donnée $z = z_k$ et se réduit à une fonction vectorielle à deux paramètres x et y. On sait alors que, au point (x, y), le jacobien de cette fonction $J_i(x, y)$ équivaut au rapport des surfaces élémentaires de l'image et de la grille terrain $z = z_k$. En conséquence, le cône de sommet le centre de prise de vue et de section carrée équivalente au niveau de l'image à la taille du pixel intersecte dans la grille terrain une surface carrée de surface $J_i(x, y)$.
On fera en outre l'approximation que pour une altitude et une image donnée, le jacobien est *constant*. Ceci revient à supposer que l'application projective est affine c'est à dire que les dénominateurs de l'équation 3.1 sont constants et ne dépendent pas de x, y ou z. Dans un contexte aérien, cette approximation paraît raisonnable. Les variations d'altitude au niveau des objets sont en effet faibles par rapport à la hauteur de l'avion au moment de la prise de vue ce qui permet de considérer que nous sommes proches d'une situation de faible perspective, cas particulier d'une caméra affine (voir [XZ96, section 2.2.1]).
Cette approximation simplifie considérablement les calculs. Soit les 4 coins c_k de l'image i et leurs projections sur la grille terrain a_k. On notera $\mid \{c_1, c_2, c_3, c_4\} \mid$ le nombre de pixels compris dans le quadrilatère formé par c_1, c_2, c_3 et c_4 et $\mid \{a_1, a_2, a_3, a_4\} \mid$ le nombre de cellules de la grille terrain incluses dans a_1, a_2, a_3 et a_4. On approxime alors le jacobien pour l'image i par la formule ([Col96]) :

$$J_i = \frac{\mid \{c_1, c_2, c_3, c_4\} \mid}{\mid \{a_1, a_2, a_3, a_4\} \mid} \tag{3.2}$$

Associations et Ω

En déplaçant la grille dans l'Espace Objet par balayage en z, on réalise ainsi une mise en correspondance point à point des segments 2D. On définit une hypothèse d'appariement par un N-uplet :

$$h = (c_1, c_2, \ldots, c_N) \text{ où } c_i \in \mathcal{C}_i^{\infty}$$

L'ensemble des hypothèses d'appariement déterminé par l'algorithme dans un voxel $\mathbf{V} = (x, y, z)$ donné est noté $\mathcal{H}_{\mathbf{V}}$. L'ensemble des appariements obtenus par balayage de plans constitue donc un ensemble d'hypothèses qu'il faut valider ou invalider. On obtient par cette méthode *tous* les appariements compatibles au moins en un point. Cette méthode d'appariement point à point permet également de garder une trace du nombre d'occurrences $v(h)$ d'une hypothèse d'appariement donnée, qui, dans le cas d'un appariement correct est proportionnel au recouvrement des segments 2D dans l'espace :

$$v(h) = \mid \{\mathbf{V}/h \in \mathcal{H}_{\mathbf{V}}\} \mid$$

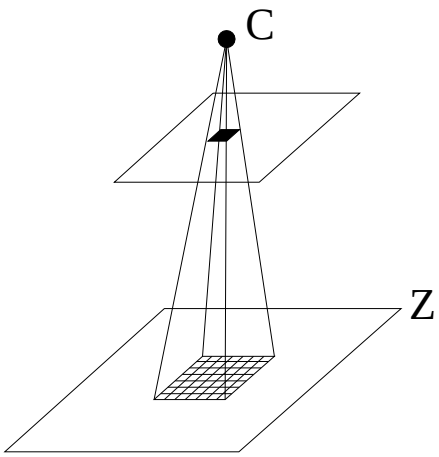


Figure 3.5 - Projection d'un pixel sur la grille terrain.

Cette valuation sera utilisée dans la suite pour choisir les différentes hypothèses valides. On définit alors une *association* ω par un couple $(h_\omega, v_\omega) = (h, v(h))$. L'ensemble des associations valides est donc $\Omega = \bigcup \omega$. À chaque association est assigné un ordre qui correspond en fait au nombre de contours véritablement mis en correspondance dans les images. Plus formellement :

$$ord(\omega) = ord(h_\omega) = Card\{c_i, i = 1 \dots N / c_i \neq \Phi\}$$

L'ensemble des associations calculées par l'algorithme définit un ensemble d'hypothèses qu'il faut garder ou rejeter. On présente dans la section 3.3 une méthode de reconstruction qui fournit un critère statistique dont on se sert pour élaguer Ω. Un ordre total est également défini permettant de ne garder que les associations maximales vérifiant un critère d'unicité (section 3.4).

3.3 Reconstruction itérative de segments 3D

3.3.1 Droites de l'espace

Représentation de Roberts

Parmi les représentations minimales de la droite, on peut citer la représentation de Roberts qui représente une droite non-orientée par un vecteur à 4 dimensions :

$$\left[\theta, \phi, x', y'\right]^T$$

où θ et ϕ sont les composantes angulaires du vecteur directeur de la droite et x', y' sont les coordonnées de l'intersection de la droite avec le plan passant par le point O (arbitraire) et orthogonal à la droite. Ces dernières sont exprimées dans un repère purement local, ce qui est le désavantage principal de ce type de représentation. Une méthode de calcul de droites 3D à partir de segments 2D en utilisant cette représentation est donnée dans [Cho00]. Le principal défaut de ce type de représentation est que la reconstruction ne peut se faire en une seule passe étant donné le caractère local des deux dernières coordonnées.

Représentation $\mathbf{abpq}$

Une autre représentation minimale est la représentation $\mathbf{abpq}$. La droite est alors représentée comme l'intersection de 2 plans :

$$\begin{cases} x = az + p \\ y = bz + q \end{cases} \tag{3.3}$$

Cependant, l'équation 3.3 ne peut pas représenter les droites parallèles au plan XY rendant la représentation incomplète. On utilise généralement donc la représentation $\mathbf{abpq}$ complète donnée par les 3 cartes :

$$\begin{cases} \text{Carte 1 :} & x = az + p, \quad y = bz + q \\ \text{Carte 2 :} & y = ax + p, \quad z = bx + q \\ \text{Carte 3 :} & x = ay + p, \quad z = by + q \end{cases} \tag{3.4}$$

Représentation de Plücker

La représentation de Plücker est une représentation non minimale à 6 paramètres couramment utilisée en géométrie projective. La difficulté pour l'utilisation de cette représentation vient de l'introduction de contraintes dans le processus de reconstruction en raison de la non minimalité de la représentation. Cette représentation ainsi qu'une technique de reconstruction sont détaillées en annexe A. Nous ferons, dans la partie résultats, une comparaison de notre méthode avec cette approche.

3.3.2 Reconstruction de droites

L'objectif est de calculer la droite $\mathcal{D}$ intersection de n_s plans Π_i où Π_i est défini comme le plan passant par un segment 2D S_i de l'image i et le centre de prise de vue correspondant C_i (figure 3.2).
Dans ce qui suit, après un rappel des relations liant l'équation des droites dans les images aux équations de plan Π_i, on décrit une méthode qui tient compte de l'incertitude sur les paramètres des droites portant les $S_i : (\theta_i, \rho_i)$ et permet d'obtenir un critère valide de qualification de la reconstruction. Nous décrivons cette méthode en utilisant la représentation $\mathbf{abpq}$ complète décrite précédemment.

Des droites 2D aux plans Π_i

On peut représenter une droite $\mathcal{D}_i$ dans le plan par les paramètres θ_i et ρ_i et l'équation :

$$u \sin\theta_i - v \cos\theta_i + \rho_i = 0 \tag{3.5}$$

où u et v sont les coordonnées d'un pixel dans l'image (figure 3.6). Dans tout ce qui suit, on supposera connues les matrices de variance covariance $\Lambda_{\mathcal{D}_i}$ liées à l'incertitude de ces paramètres qui peuvent être déterminées après les étapes de polygonisation (section 3.4.1). Les paramètres de prises de vue étant donnés par hypothèse, on connaît, pour chaque caméra, la matrice de transformation perspective 3x4 $\mathcal{T}$:

$$\mathcal{T} = \begin{bmatrix} \mathcal{T}_{11} & \mathcal{T}_{12} & \mathcal{T}_{13} & \mathcal{T}_{14} \\ \mathcal{T}_{21} & \mathcal{T}_{22} & \mathcal{T}_{23} & \mathcal{T}_{24} \\ \mathcal{T}_{31} & \mathcal{T}_{32} & \mathcal{T}_{33} & \mathcal{T}_{34} \end{bmatrix} \tag{3.6}$$

En considérant que la caméra suit un modèle simple sténopé, on peut alors relier un point 3D $\mathbf{M}(x, y, z)$ à sa projection dans une image $\mathbf{m}(u, v)$ par l'équation :

$$\tilde{\mathbf{m}} = s\mathcal{T}\tilde{\mathbf{M}} \quad \text{soit} \quad \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = s\mathcal{T} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad \text{avec} \quad s \neq 0 \tag{3.7}$$

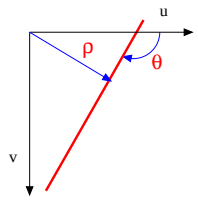


Figure 3.6 - Représentation d'une droite dans l'image.

En substituant u et v de l'équation (3.7) dans l'équation (3.5), on obtient alors :

$$\begin{aligned}&(\mathcal{T}_{11}.x + \mathcal{T}_{12}.y + \mathcal{T}_{13}.z + \mathcal{T}_{14})\sin\theta \\ -&(\mathcal{T}_{21}.x + \mathcal{T}_{22}.y + \mathcal{T}_{23}.z + \mathcal{T}_{24})\cos\theta \\ +&(\mathcal{T}_{31}.x + \mathcal{T}_{32}.y + \mathcal{T}_{33}.z + \mathcal{T}_{34})\rho = 0\end{aligned} \tag{3.8}$$

ce qui peut encore s'écrire

$$\tilde{\mathbf{M}}^T\mathbf{v} = 0 \tag{3.9}$$

avec

$$\mathbf{v} = \begin{bmatrix}\mathcal{T}_{11}\sin\theta - \mathcal{T}_{21}\cos\theta + \mathcal{T}_{31}\rho \\ \mathcal{T}_{12}\sin\theta - \mathcal{T}_{22}\cos\theta + \mathcal{T}_{32}\rho \\ \mathcal{T}_{13}\sin\theta - \mathcal{T}_{23}\cos\theta + \mathcal{T}_{33}\rho \\ \mathcal{T}_{14}\sin\theta - \mathcal{T}_{24}\cos\theta + \mathcal{T}_{34}\rho\end{bmatrix} \tag{3.10}$$

Pour chaque segment 2D S_i, on peut donc représenter le plan Π_i correspondant par un vecteur à 4 paramètres $\mathbf{v_i} = [\alpha_i, \beta_i, \gamma_i, \delta_i]^T$ définis dans l'équation (3.10) ; tout point 3D appartenant au plan vérifie alors l'équation (3.9).

Reconstruction itérative

Pour déterminer la droite $\mathcal{D}$ intersection des plans Π_i, on utilisera ici la représentation **abpq** complète décrite en section 3.3.1. La reconstruction doit être menée parallèlement dans les 3 cartes pour pouvoir reconstruire toute droite solution. Nous ne décrivons toutefois que les calculs pour la carte 1, le raisonnement étant identique dans les autres cartes. Tout point sur la droite $\mathcal{D}$ représentée dans cette carte vérifie alors :

$$[az + p, bz + q, z, 1]\mathbf{v_i} = 0 \qquad \forall i \in [1..n_s] \tag{3.11}$$

Pour les n_s segments, on en déduit $2n_s$ équations :

$$\begin{aligned}f_{2i} &= a\alpha_i + b\beta_i + \gamma_i = 0 \\ f_{2i+1} &= p\alpha_i + q\beta_i + \delta_i = 0 \qquad \forall i \in [1..n_s]\end{aligned} \tag{3.12}$$

Pour déterminer les 4 inconnues a, b, p, q, on est donc ramené à chercher $\mathbf{X} = [a, b, p, q]^T$ solution du système linéaire $\mathbf{AX} = \mathbf{b}$ avec $\mathbf{A}$ matrice de taille $2n_s$x4 et $\mathbf{b}$ vecteur de taille $2n_s$ définis par :

$$\mathbf{A} = \begin{bmatrix}\vdots & \vdots & \vdots & \vdots \\ \alpha_i & \beta_i & 0 & 0 \\ 0 & 0 & \alpha_i & \beta_i \\ \vdots & \vdots & \vdots & \vdots\end{bmatrix} \quad \text{et} \quad \mathbf{b} = \begin{bmatrix}\vdots \\ -\gamma_i \\ -\delta_i \\ \vdots\end{bmatrix} \tag{3.13}$$

La résolution aux moindres carrés de ce système, qui a déjà été abordée dans [ZF92], traduit la minimisation de la somme des carrés des résidus :

$$\mathcal{S} = \sum_i \ldots + f_{2i}^2 + f_{2i+1}^2 + \ldots \tag{3.14}$$

La solution est classiquement $\widehat{\mathbf{X}} = \left(\mathbf{A^T A}\right)^{-1} \mathbf{A^T b}$. Le problème de la méthode précédente est que l'on suppose implicitement la même variance pour chaque résidu. La solution des moindres carrés n'est en fait optimale qu'à cette condition. On donne donc la même importance à chacun des plans alors que la variance Λ_{v_i} dépend beaucoup de la variance sur les paramètres de droite dans les images. De façon à tenir compte de cette incertitude, on emploie une technique de moindres carrés pondérés et on cherche à minimiser :

$$\tilde{\mathcal{S}} = \sum_i \ldots + \frac{f_{2i}^2}{\sigma_{f_{2i}}^2} + \frac{f_{2i+1}^2}{\sigma_{f_{2i+1}}^2} + \ldots \tag{3.15}$$

Cette résolution revient donc à la résolution du problème des moindres carrés pondérés $\mathbf{KAX} = \mathbf{Kb}$ avec

$$\mathbf{K} = diag(\frac{1}{\sigma_{f_0}}, ... \frac{1}{\sigma_{f_{2*n+1}}})$$

Lorsque K est constant, on a simplement $\widehat{\mathbf{X}} = (\mathbf{A}^T\mathbf{K}^T\mathbf{K}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{K}^T\mathbf{K}\mathbf{b}$. Dans notre cas, toutefois, les paramètres de variance sur les résidus σ_{f_j}, qui se calculent facilement à l'aide des équations (3.16) [ZF92, XZ96] en supposant les variances $\Lambda_{\mathcal{D}_i}$ connues, *dépendent de* $\mathbf{X}$.

$$\begin{aligned}
\sigma_{f_{2i}}^2(X) &= \frac{\partial f_{2i}}{\partial \mathbf{v_i}}^T \Lambda_{\mathbf{v_i}} \frac{\partial f_{2i}}{\partial \mathbf{v_i}} \\
&= \frac{\partial f_{2i}}{\partial \mathbf{v_i}}^T \frac{\partial \mathbf{v_i}}{\partial \mathcal{D}_i}^T \Lambda_{\mathcal{D}_i} \frac{\partial \mathbf{v_i}}{\partial \mathcal{D}_i} \frac{\partial f_{2i}}{\partial \mathbf{v_i}} \\
\sigma_{f_{2i+1}}^2(X) &= \frac{\partial f_{2i+1}}{\partial \mathbf{v_i}}^T \Lambda_{\mathbf{v_i}} \frac{\partial f_{2i+1}}{\partial \mathbf{v_i}} \\
&= \frac{\partial f_{2i+1}}{\partial \mathbf{v_i}}^T \frac{\partial \mathbf{v_i}}{\partial \mathcal{D}_i}^T \Lambda_{\mathcal{D}_i} \frac{\partial \mathbf{v_i}}{\partial \mathcal{D}_i} \frac{\partial f_{2i+1}}{\partial \mathbf{v_i}}
\end{aligned} \tag{3.16}$$

On utilise donc finalement le schéma itératif suivant (Algorithme 1) :

Algorithme 1: Reconstruction de segments 3D par minimisation itérative.

Entrée : Ensemble de n_s plans et leurs variances
Sortie : Segment 3D en représentation $\mathbf{abpq}$ complète
$r = 0$
résolution de $\mathbf{AX}^{(0)} = \mathbf{b}$
répéter
  calcul de $\sigma_{f_l}^{(r)}(\mathbf{X}^{(r)})$ avec (3.16)
  résolution de $\mathbf{KAX}^{(r+1)} = \mathbf{Kb}$ avec $\mathbf{K} = diag(\frac{1}{\sigma_{f_0}^{(r)}}, \frac{1}{\sigma_{f_1}^{(r)}}, \ldots, \frac{1}{\sigma_{f_{2n+1}}^{(r)}})$
  si $(X^{(r+1)} - X^{(r)})^2 < seuil)$ **alors** convergence=vraie
  sinon convergence=faux
  $r = r + 1$
jusqu'à *(convergence)* $\vee$ *(*$r > N_{\max}$*)*

Choix de la carte

Comme évoqué précédemment, un des problèmes de la reconstruction $\mathbf{abpq}$ est le choix de la carte de représentation définitive du segment. Dans notre cas, les calculs sont menés en parallèle dans les 3 cartes jusqu'à la dernière itération. Nous choisissons alors la carte qui fournit la reconstruction la plus stable. Intuitivement, il s'agit de la carte telle que le segment 3D est défini par deux plans formant un angle le plus proche d'un angle droit. Il est aisé de montrer que dans n'importe quelle carte, le cosinus de l'angle entre les deux normales ($\mathbf{p}_1$ et $\mathbf{p}_2$) aux plans définissant la droite par intersection est donné par :

$$c_m = \frac{\mathbf{p}_1 . \mathbf{p}_2}{\|\mathbf{p}_1\| . \|\mathbf{p}_2\|} = \frac{ab}{\sqrt{(1+a^2)(1+b^2)}} \tag{3.17}$$

La carte choisie sera donc celle qui minimise $\mid c_m \mid$.

3.3.3 Qualification du résultat

Un des points essentiels de cette approche est que les résidus sont *normalisés*. En supposant que les erreurs suivent une loi gaussienne, la somme au carré $\tilde{\mathcal{S}}$ de ces résidus suit donc une loi du χ^2 à $d = 2 * n_s - 4$ degrés de liberté (on a en effet $2 * n_s$ équations indépendantes et 4 paramètres indépendants), que nous noterons χ^2_d. Ce résultat permet de qualifier la reconstruction et de valider ou de rejeter un appariement donné. On utilisera cette donnée dans l'algorithme final pour sélectionner les appariements valides.

On notera également que, puisque la résolution est aux moindres carrés, on peut déterminer la matrice de variance covariance de la droite porteuse des segments [XZ96], ce qui permet d'avoir une idée de la précision de localisation des paramètres et peut servir dans l'utilisation de tests d'égalité, de parallélisme ou autres.

3.3.4 Reconstruction des extrémités

Les extrémités du segment 3D sont finalement reconstruites en projetant chacune des extrémités des segments 2D sur la droite 3D reconstruite. On choisit comme [XZ96] une stratégie d'union qui consiste à prendre comme segment reconstruit l'union de tous les segments 2D projetés sur la droite 3D.

3.4 Algorithme global

3.4.1 Extraction de segments 2D

Dans un premier temps, une extraction de contours est effectuée sur chaque image en utilisant un opérateur classique [Der87] suivie d'une détection des maxima locaux dans la direction du gradient et d'un seuillage hystérésis à deux paramètres. Les contours sont chaînés et polygonisés. On utilise pour ce faire un processus de fusion itératif basé sur le résidu maximum de la régression orthogonale : on fusionne d'abord les polylignes dont la fusion fournit un résidu maximum minimal. Une tolérance sur la polygonisation permet d'arrêter l'algorithme lorsque la fusion envisagée fournit un résidu maximum supérieur à un seuil donné par l'utilisateur s_{pol}. Une fois la polygonisation effectuée, les paramètres des droites portant les segments sont estimés en utilisant les résultats de [DVF92]. La régression étant effectuée sur n points (x_i, y_i) dont le

barycentre est $(\bar{x}, \bar{y})$, on note :

$$a = \sum_{i=1}^{n} (x_i - \bar{x})^2$$
$$b = 2 \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$
$$c = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

On a alors

$$\theta = \frac{1}{2} \arctan\left(\frac{b}{a-c}\right) \tag{3.18}$$
$$\rho = \bar{y}\cos(\theta) - \bar{x}\sin(\theta) \tag{3.19}$$

On estime aussi la matrice de covariance des paramètres de la droite en supposant que les points détectés par l'opérateur de Canny-Deriche [Der87] ont une variance donnée par

$$\Lambda = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}$$

On notera que le paramètre σ peut être estimé à partir du rapport signal sur bruit de l'image et du paramètre du filtre de contours α. Avec cette hypothèse, on obtient :

$$\Lambda_{\mathcal{D}_i} = \Lambda_{\rho,\theta} = \frac{\sigma^2(a+c)}{(a-c)^2 + b^2} \begin{bmatrix} 1 & -d \\ -d & d^2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{\sigma^2} \end{bmatrix} \tag{3.20}$$

Cette expression est directement utilisée dans les équations 3.16 pour le calcul itératif de la meilleure droite 3D.

3.4.2 Élagage de Ω, ensemble des hypothèses de droites 3D

Élagage géométrique

En utilisant la technique d'appariement décrite en section 3.2, on génère un ensemble d'associations qui respectent toutes, au moins en un point, les contraintes épipolaires. Une association n'est considérée comme valide que si le score résultant de la reconstruction vérifie le test du χ^2 avec une probabilité $\mathcal{P}$ définie par l'utilisateur.
Ainsi, pour une hypothèse d'appariement ω donnée, connaissant le nombre de degrés de liberté $d_\omega = 2v_\omega - 4$ associé à cette association, on sait que théoriquement, $\tilde{\mathcal{S}}$ suit la loi du $\chi^2_{d_\omega}$. On détermine alors un niveau seuil $a_{d_\omega,\mathcal{P}}$ qui vérifie $P(\chi^2_{d_\omega} > a_{d_\omega,\mathcal{P}}) = \mathcal{P}$. Si $\tilde{\mathcal{S}} > a_{d_\omega,\mathcal{P}}$, l'hypothèse est rejetée. Dans le cas contraire, l'association est acceptée. Ce test du χ^2 assure que l'on *accepte quasi certainement (probabilité de $\mathcal{P}$) l'appariement dans le cas où cette hypothèse est vraie.* Ce test est conduit sur Ω qui est donc élagué via cette contrainte géométrique.

Élagage sur critère d'unicité

On définit ensuite dans Ω une relation d'ordre total $\prec$:

$$\omega \prec \omega\prime \Leftrightarrow \begin{cases} ord(\omega) < ord(\omega\prime) \\ \text{ou sinon } v_\omega < v_{\omega\prime} \\ \text{ou sinon } \tilde{S}_{h_\omega} < \tilde{S}_{h_{\omega\prime}} \end{cases}$$

$\tilde{S}$ représente la somme des résidus au carré normalisés de la reconstruction. Cette relation d'ordre total privilégie les associations qui ont un grand nombre de segments mis en correspondance et ensuite les associations qui ont un nombre important de points de contours mis en correspondance. Cette relation d'ordre permet de faire du "winner takes all" *tout en gardant la symétrie du problème et sans privilégier aucune image*. À partir de l'ensemble Ω et de la relation définis précédemment on peut classer les associations entre elles. On choisit donc itérativement les associations maximales (représentant les "segments les plus sûrs") puis on "nettoie" l'ensemble Ω en vérifiant la contrainte d'unicité : un segment dans une image ne peut appartenir qu'à une seule association. On obtient ainsi l'ensemble des associations finales acceptées comme segments 3D qui sont alors reconstruits.

3.5 Résultats

3.5.1 Adéquation à la loi du χ^2

L'élagage géométrique reposant sur un test d'adéquation de la somme des carrés des résidus normalisés avec une loi du χ^2, il est important de valider le comportement de la reconstruction avec cette loi. Pour ce faire, nous avons effectué le test aléatoire suivant :

1- projeter un segment 3D échantillonné dans 6 images,

2- ajouter un bruit gaussien de paramètre σ sur chaque point 2D dans les images,

3- reconstruire le segment 3D en calculant la somme des carrés des résidus normalisés $\tilde{\mathcal{S}}_t$ pour le test t.

En répétant un très grand nombre de fois les étapes 2 et 3, on estime la proportion $g(x)$ des tests pour lesquels la somme normalisée des carrés des résidus $\tilde{\mathcal{S}}_t$ est supérieure à un nombre x donné :

$$g(x) = \frac{Card\{t/\tilde{\mathcal{S}}_t > x\}}{Card\{t\}}$$

Si $\tilde{\mathcal{S}}_t$ suit une loi du χ^2 à $2 * 6 - 4 = 8$ degrés de liberté, g doit suivre la loi de probabilité $\mathcal{P}(\chi^2_{[8]} > x)$. On peut montrer que la densité de probabilité de la loi du χ^2_d est $c_d \exp(-y/2) y^{(d/2)-1}$ où c_d est une constante de normalisation [Nev96]. g doit donc suivre la loi donnée par l'équation 3.21 :

$$\mathcal{P}(\chi^2_{[8]} > x) = \frac{\int_x^\infty u^{(8/2)-1} * \exp(-u/2) du}{\int_0^\infty u^{(8/2)-1} * \exp(-u/2) du} \tag{3.21}$$

Les résultats récapitulés sur les figures 3.7 prouvent que l'adéquation de $g(x)$ avec cette loi est relativement bonne. On voit que la courbe théorique et celle obtenue par le test statistique sont proches. Les écarts observés sont vraisemblablement dus aux approximations du premier ordre faites dans les calculs des variances aux différents stades de la reconstruction [DVF92, XZ96]. En outre, nous n'utilisons pas de procédé itératif avec réévaluation des observations pour compenser ces approximations [För01]. La proximité des courbes justifie néanmoins l'utilisation du test statistique du χ^2 pour rejeter les appariements non valides.

3.5.2 Simulations

Des simulations ont été effectuées avec un modèle de bâtiment issu de la BD TRAPU®. Ce modèle est constitué d'un ensemble de faces polygonales dont on extrait les arêtes. On les projette ensuite dans les différentes images (6 dans notre test). Afin de tester les capacités d'appariement de l'algorithme indépendamment des erreurs de segmentation dans les images, *nous n'avons pas effectué de détection de contours* dans ces simulations. En d'autres termes, les segments 2D sont directement les projections des segments 3D existants. Dans toutes les simulations, on utilise une valeur du paramètre $\mathcal{P} = 0.9$ et un volume "terrain" de 160x170x40 m^3 pour la phase d'appariement.

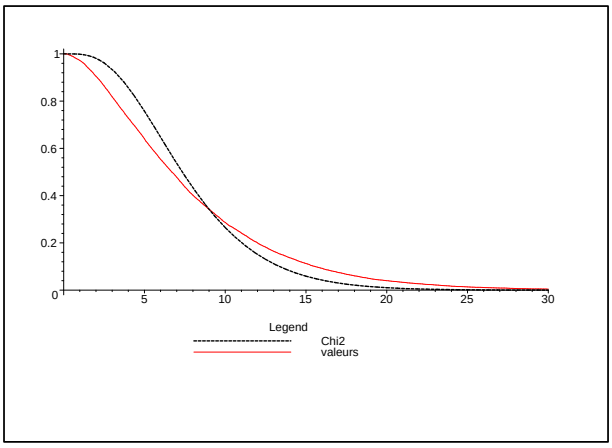


Figure 3.7 - Adéquation des résidus avec la loi du χ^2, représentation $\mathbf{abpq}$. $\sigma = 2$, nombre de tests : 10000. La courbe noire représente la valeur théorique du χ^2 à 8 degrés de liberté. La courbe rouge matérialise les valeurs obtenues par simulation (voir texte).

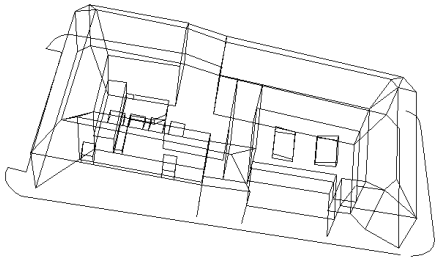


Figure 3.8 - Modèle filaire du bâtiment utilisé comme référence dans toutes les simulations.

Sans bruit

Les résultats (figure 3.9(a)) montrent que la reconstruction en partant de segments non bruités correspond exactement à la réalité. Tous les segments ont été reconstruits avec une très bonne localisation.

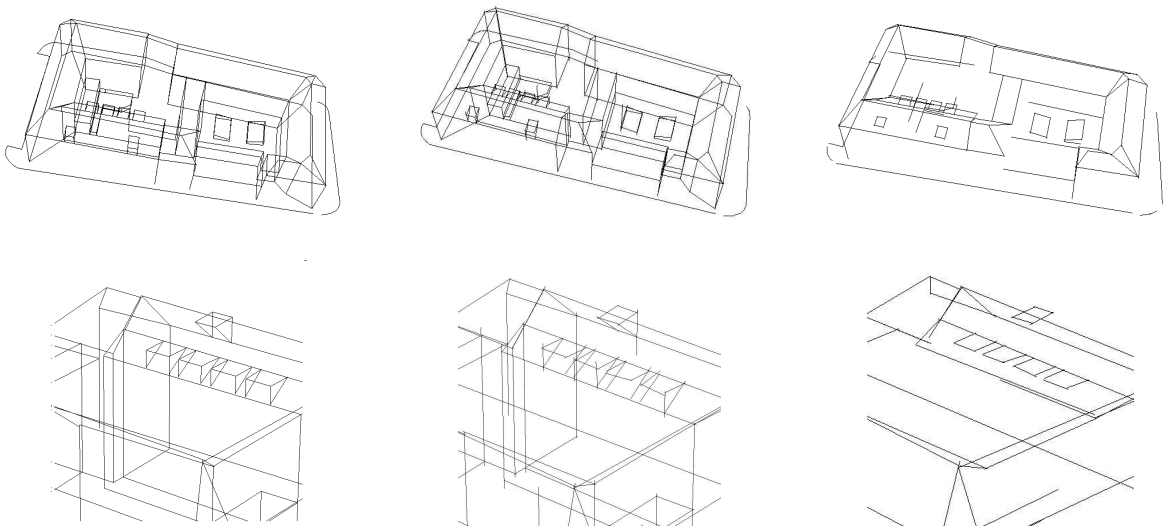
Avec bruit

La même expérience a été reconduite (figure 3.9(b)) mais en bruitant les points projetés dans les images. On cherche ainsi à valider la précision de la reconstruction y compris en présence de bruit. Les points de contours constitutifs des segments ont ainsi été bruités avec un bruit gaussien isotrope d'écart type 0.5. C'est ce même écart-type qui a, bien sûr, servi à estimer les variances des paramètres des droites. Les résultats sont, là encore, très bons. Tous les appariements sont corrects et la quasi-totalité des segments sont appariés. Les segments non appariés ont été rejetés sur des considérations géométriques (le choix de $\mathcal{P} \neq 1$ induit inévitablement quelques rejets). À noter la dégradation visuelle évidente de la précision de reconstruction.

Avec bruit et faces cachées

Dans cette simulation (figure 3.9(c)), on projette les segments dans les images en tenant compte des parties cachées, et en bruitant les points. Il s'agit ici de prouver les capacités de l'algorithme à traiter les

cas où les segments sont visibles dans un nombre quelconque (>=3) d'images. La reconstruction est, encore une fois, juste dans la mesure où les appariements sont tous corrects. L'algorithme a bien géré les cas où le segment est vu dans un nombre d'images inférieur à $N = 6$. Dans le cas d'un détecteur de segments parfait, on est donc assuré que la génération d'hypothèses d'appariement est correcte et que l'algorithme extrait bien la bonne association indépendamment du nombre d'images.



(a) sans bruit (b) avec bruit (c) avec bruit et faces cachées

Figure 3.9 - Simulations.

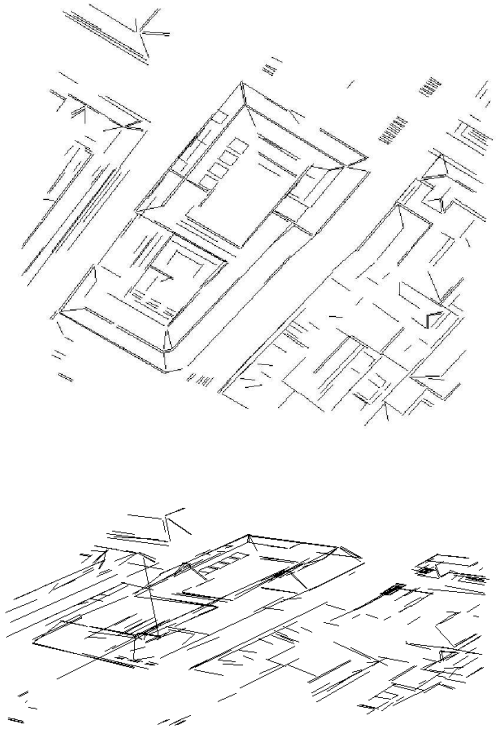
3.5.3 Images réelles

Dans le cas d'images réelles, l'expérience a été conduite sur 6 images représentant un bâtiment. On utilise des images venant de la caméra numérique de l'IGN qui possède un bon rapport signal sur bruit (figure 3.10(a)). Un traitement spécifique est néanmoins nécessaire pour tenir compte des défauts dus à la distorsion. Les paramètres utilisés sont :

- taille de la zone : 160x170x40 m^3
- $pas_{XY} = 20cm$ et $pas_Z = 50cm$
- suréchantillonnage de l'image : 2
- alpha du filtre Canny-Deriche : 1.5
- seuils hystérésis : s_b=0 s_h=5
- tolérance de polygonisation : 1
- taille minimum d'un segment : 20 pixels
- nombre d'images requises pour une association : 4
- $\mathcal{P} = 0.9$

Les résultats (figure 3.10(b)) montrent que 413 segments 3D ont été restitués. Compte tenu du nombre de segments (figure 3.10(a)) détectés dans chacune des images, ce résultat est très satisfaisant. Les résultats montrent une bonne restitution de l'ensemble des détails, y compris de lignes très proches et parallèles. Il n'y a d'autre part aucun faux appariement. On remarquera un segment vertical au centre du bâtiment qui semble être une erreur. Il s'agit en fait d'un mât dont la hauteur est estimée par l'algorithme à 67,5 m et qui culmine en fait à 69 m !

(a) 2 des 6 images utilisées. En haut : segments détectés dans les images (1012 et 845). En bas : segments reconstruits projetés dans les images.



(b) vues en 3D. En haut : vue de dessus. En Bas : vue latérale

Figure 3.10 - Résultat de la reconstruction.

3.6 Extensions

3.6.1 Restriction de l'Espace de Recherche

Pour réduire l'Espace de Recherche et également éviter de faux appariements, un volume de recherche peut être dérivé de la dilatation d'un MNE. On peut ainsi réduire l'ensemble des voxels valides et forcer les segments reconstruits à avoir leurs extrémités dans le volume de recherche (figure 3.11).

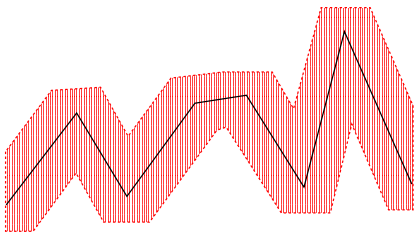


Figure 3.11 - Principe de la nappe englobante. En rouge la nappe Min/Max, au centre le MNE

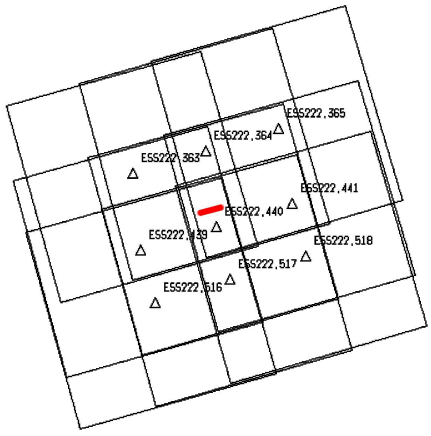
3.6.2 Extension à 2 images

Pour éviter la plupart des problèmes de sous-segmentation, l'algorithme peut facilement être étendu pour gérer des hypothèses d'appariement à 2 segments. Dans ce cas, naturellement, la contrainte géométrique est caduque, on utilise à la place une contrainte de recouvrement (50%). La figure 3.12 prouve que beaucoup de segments sont reconstruits en utilisant ces 2 extensions. Néanmoins, de faux appariements apparaissent, certainement dus à la faiblesse de la contrainte géométrique dans le dernier cas. Ces 2 extensions permettent de traiter des régions où le nombre d'images est relativement faible ou des régions étendues tout en gardant un temps de calcul raisonnable comme dans la figure 3.12, où 3500 segments 3D ont été reconstruits à partir de 5 images en utilisant un MNE pour réduire l'Espace de Recherche.

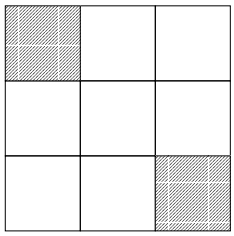
3.7 Évaluation

3.7.1 Précision théorique

Dans un premier temps, nous avons évalué l'apport du multi vues pour la reconstruction de droites. Un segment 3D de référence d'environ 50 mètres est projeté dans 9 images puis les 2 extrémités de chaque segment 2D sont bruitées en utilisant un bruit gaussien d'écart type $\sigma = 2$ (0,5m). La configuration spatiale est celle représentée sur la figure 3.13 où sont représentés le segment 3D en projection et l'emprise des images projetées au niveau du sol. On retrouve schématiquement un damier 3x3.



(a) Les emprises au sol des 9 images utilisées sont représentées ainsi que les sommets de prises de vue (numérotés). Le segment apparaît en rouge au centre des emprises.



(b) Schématisation signifiant que 2 images sont utilisées : en haut à gauche et en bas à droite correspondant aux images "ESS222-363" et "ESS222-518".

Figure 3.13 - Configuration spatiale et schématisation en vue verticale.

Le segment 3D est ensuite reconstruit en utilisant successivement $3, 4 \cdots N$ images (dans notre cas $N = 9$). On mesure alors la norme de la différence entre les 2 vecteurs 3D normalisés (reconstruit-référence) La figure 3.14 récapitule le gain obtenu sur une batterie de 10000 essais. L'augmentation du nombre d'images accroît sensiblement la précision de la reconstruction puisqu'un gain de 33% est obtenu entre la reconstruction utilisant 3 images et la reconstruction en utilisant 9. Les "sauts" de la courbe s'expliquent par des passages à des configurations qui stabilisent considérablement la reconstruction (voir section 3.7.2).

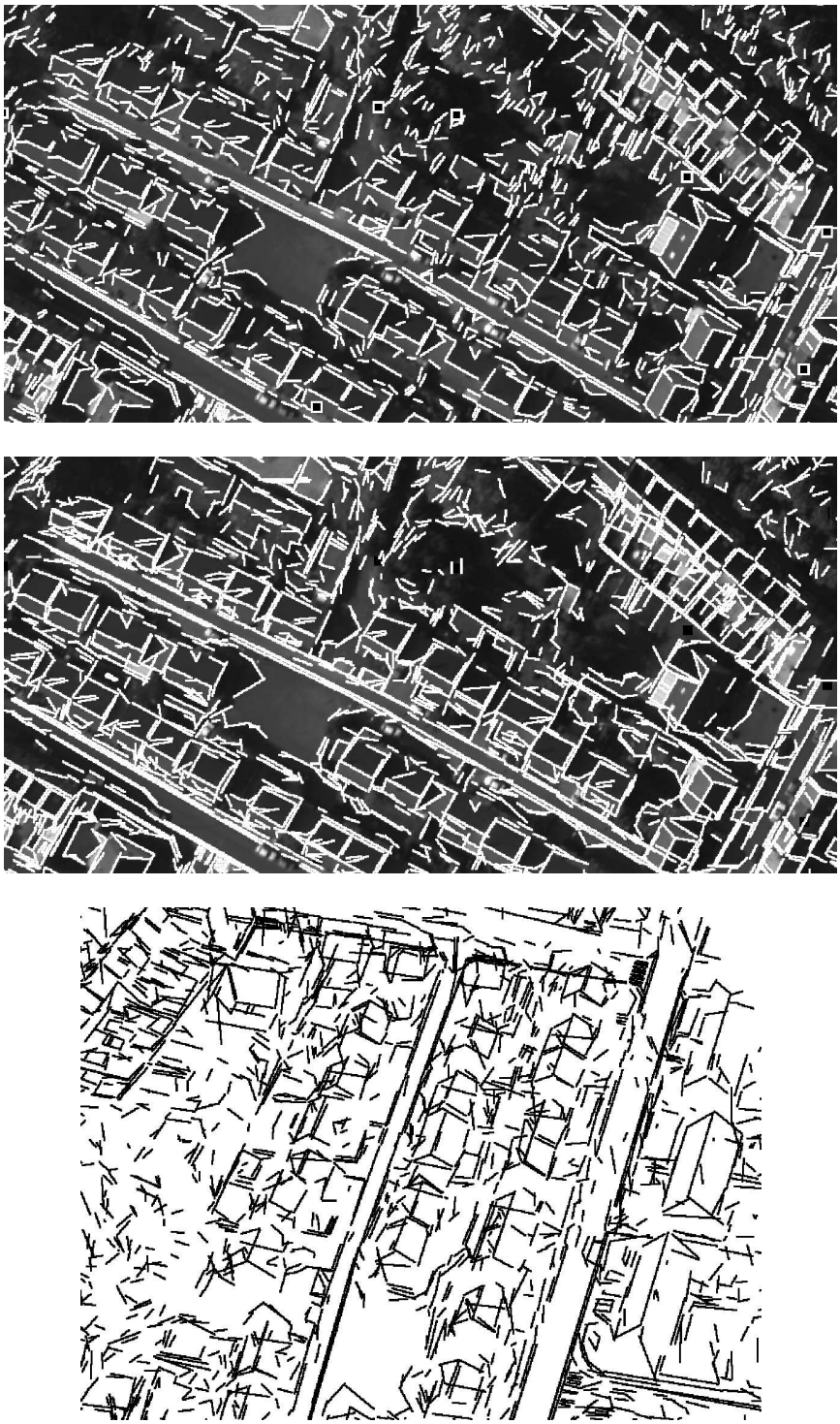


Figure 3.12 - Résultats sur images réelles.2 extraits des 5 images utilisées dans cet essai. au-dessus : segments détectés. milieu : segments reconstruits projetés. en-dessous : vue latérale de la reconstruction

3.7.2 Précision en fonction des configurations

Ce test met en évidence l'influence des configurations sur la précision de la reconstruction. Le segment 3D a été reconstruit en utilisant 3 images sur 9 disponibles et en testant les 84 configurations possibles. Le graphique 3.15 dans lequel les configurations sont classées par ordre croissant de précision illustre bien l'importance du choix des images et en particulier le passage à des configurations dégénérées (haut de la courbe) qui induit une perte considérable de précision (les configurations dégénérées sont ici celles pour lesquelles les sommets de prise de vue sont dans l'alignement du segment 3D). On démontre bien ainsi que des points de vue écartés sont essentiels pour une bonne reconstruction. On notera que pour l'évaluation de la précision (section précédente), nous sommes partis de la configuration à 3 images optimale pour ne pas biaiser le raisonnement. En regard du graphique 3.14, on remarquera que les "sauts" de précision n'apparaissent que lors de l'introduction d'une image de la bande du haut ou du bas, ce qui justifie encore le peu d'apports de la bande médiane sur la précision de la reconstruction.

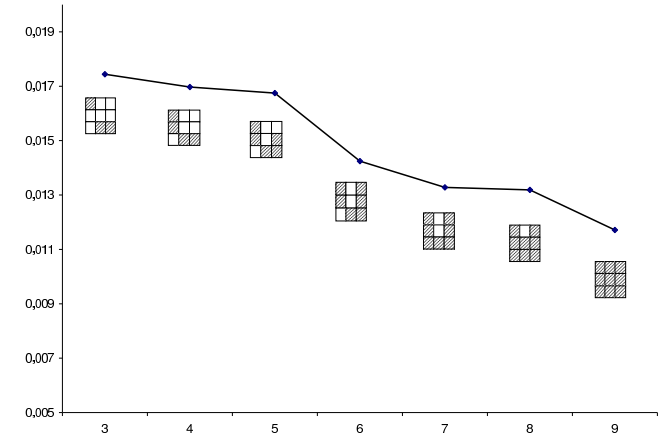


Figure 3.14 - Évolution de la précision en fonction du nombre d'images.

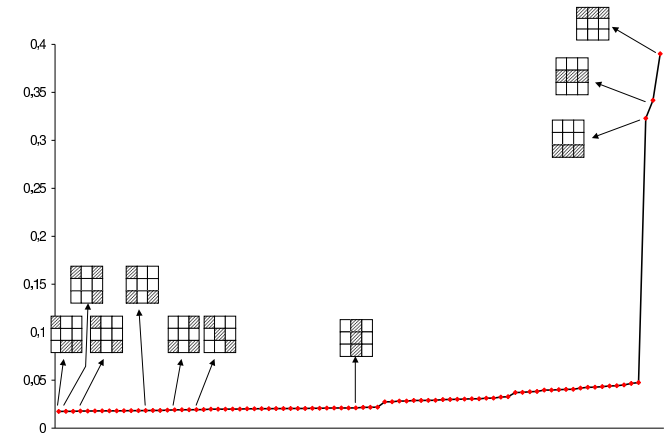


Figure 3.15 - Évolution de la précision en fonction des configurations.

3.7.3 Taux de recouvrement en fonction du nombre d'images

On se place ici dans une configuration réelle où 9 images ont été utilisées pour reconstruire les segments 3D de la scène. Disposant d'une référence 3D saisie à la main par un opérateur (figure 3.8), on compare la reconstruction et la référence. Pour ce faire, les segments reconstruits sont appariés à la référence. L'ensemble des segments reconstruits sera noté $\mathcal{S}$, les segments références $\mathcal{A}$. Un segment $s \in \mathcal{S}$ est apparié à un segment $a \in \mathcal{A}$ si et seulement si l'écart angulaire est inférieur à un paramètre $\theta_s = 15^\circ$ et si l'écart en distance est inférieur à un autre paramètre $d_s = 1,5m$. Pour chaque segment de la référence a, on définit alors la longueur de recouvrement par l'ensemble des segments de $\mathcal{S}$ appariés à a : $r_a(\mathcal{S})$ qui traduit, en quelque sorte, l'"explication" du segment a par $\mathcal{S}$. Pour l'ensemble des segments $\mathcal{A}$, on obtient donc le taux de recouvrement par

$$r_{\mathcal{A}}(\mathcal{S}) = \sum_{a \in \mathcal{A}} \frac{r_a(\mathcal{S})}{\|a\|} \tag{3.22}$$

Cet indicateur permet de mesurer le taux de recouvrement de la scène par les segments reconstruits. Il sera d'ailleurs réutilisé dans l'évaluation d'un modèle de bâtiment par les segments 3D détectés (section 5.3.3). Ce taux a été mesuré en fonction du nombre d'images et du nombre minimal d'images utilisées pour la reconstruction (figure 3.16). On constate un accroissement très net jusqu'à 5 images (80% de la scène expliqués) puis une relative stabilité, ce qui amène à penser qu'en terme d'explication de la scène, 4 ou 5 images sont un bon rapport exhaustivité/nombre d'images. On constate également que le nombre minimal d'images a peu d'influence sur le taux de recouvrement dès que le nombre d'images est supérieur à 4. À noter que nous n'avons pas utilisé la référence pour déterminer la précision de la reconstruction pour des raisons d'insuffisance notoire de cette référence en termes de localisation.

3.7.4 Surdétections / Sousdétections

L'article [MS94] définit une technique d'évaluation de détecteurs de primitives dont nous allons nous inspirer pour étudier les sur et sous détections. On dit qu'un segment a référence est correctement détecté ("True Positive" : TP) si $r_a(\mathcal{S}) > 0.33$. Dans le cas contraire, a est "True Negative" : TN. Si un segment reconstruit s est tel que $r_s(\mathcal{A}) < 0.33$, on considère que s est une surdétection soit "False Positive" : FP. On définit alors une série de mesure de qualité :

- **le taux de détection** $\frac{TP}{TP+TN}$
- **le taux de surdétection** $\frac{FP}{TP+FP}$

On constate d'après les résultats (figures 3.16, 3.17 et 3.18) que, au delà de 6 images, l'extension permettant de gérer les intersections à 2 plans devient relativement inutile compte tenu du gain pour la détection et les surdétections générées. On constate aussi que tous les segments longs sont correctement détectés (figure 3.19).

3.7.5 Comparaison avec la reconstruction à partir des coordonnées de Plücker

Une méthode de reconstruction à partir des coordonnées de Plücker a été proposée pendant la durée de cette thèse par l'équipe de Bonn [För01]. Nous rappelons en annexe A les principes de cette méthode qui présente énormément de similitudes avec notre approche et nous comparons ici les résultats des deux techniques. La méthode de mise en correspondance reste identique dans le cadre de ce test.

Adéquation à la loi du χ^2

En utilisant le protocole expérimental détaillé plus haut, nous avons comparé l'adéquation des deux méthodes de reconstruction à la loi du χ^2. Comme l'atteste la figure 3.20, les courbes, dans le cas de la

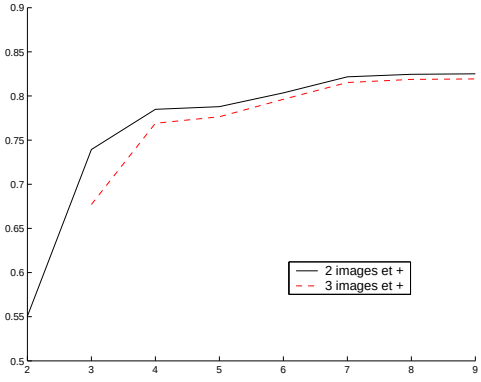


Figure 3.16 - Comparaison du taux de recouvrement en fonction du nombre d'images (abscisses) et du nombre d'images minimal pour la reconstruction (2 ou 3).

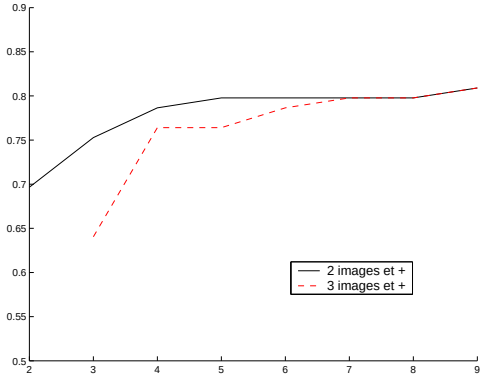


Figure 3.17 - Évolution du taux de détection en fonction du nombre d'images et du nombre d'images minimal (2 ou 3).

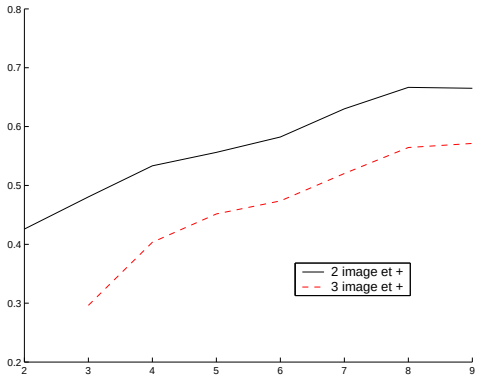


Figure 3.18 - Evolution du taux de surdétection en fonction du nombre d'images et du nombre d'images minimal (2 ou 3).

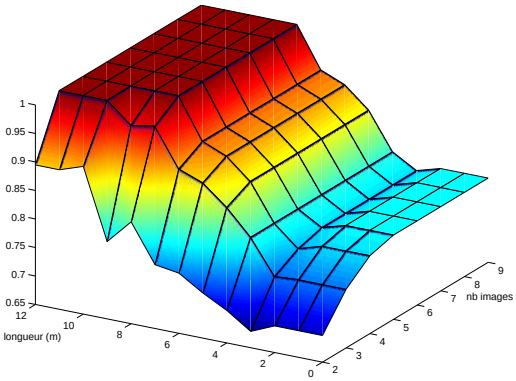


Figure 3.19 - Evolution du taux de détection en fonction du nombre d'images et de la longueur des segments.

reconstruction avec les coordonnées de Plücker, sont quasiment confondues. Comme nous l'avons déjà souligné, l'écart avec la représentation **abpq** vient de ce que l'on ne réévalue pas les *observations* en fonction de l'estimée comme dans le modèle de Gauss-Helmert [För01].

Image réelles

Sur images réelles, les résultats (figures 3.21 et 3.22) différent peu, excepté lorsque le nombre d'images est relativement faible (3 ou 4). Dans ce contexte, notre méthode paraît plus exhaustive et plus stable probablement du fait du non-ajustement des plans en fonction de la variance (annexe A) qui diminue le nombre de paramètres à ajuster dans notre cas. De fait, même si pour un nombre d'images élevé, les résultats sont similaires en terme de recouvrement et de détection, le nombre de segments reconstruits à partir d'un grand nombre d'images est plus important dans le cas de notre méthode, ce qui est un gage de précision.

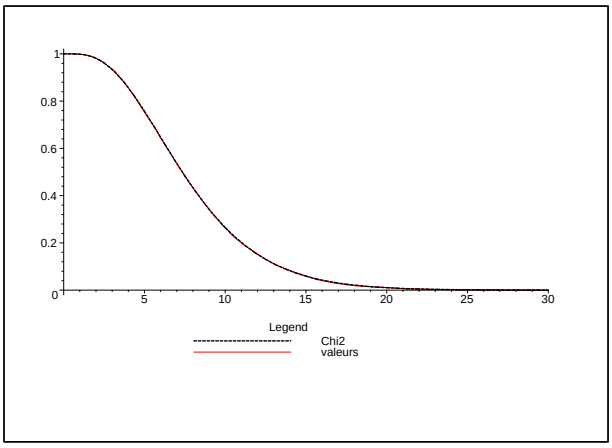


Figure 3.20 - adéquation des résidus avec la loi du χ^2, représentation de Plücker. $\sigma = 2$, nombre de test : 10000. La courbe noire représente la valeur théorique du χ^2 à 8 degrés de liberté. La courbe rouge matérialise les valeurs obtenues par simulation (voir texte).

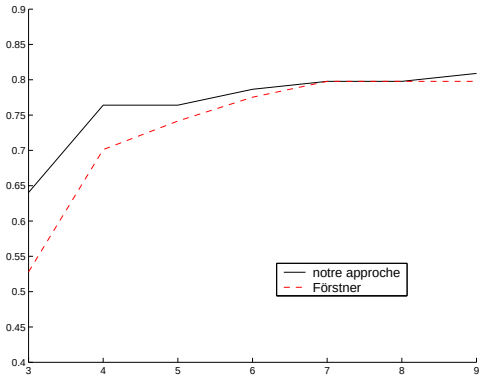


Figure 3.21 - Comparaison du taux de détection entre notre approche et l'approche à partir des coordonnées de Plücker.

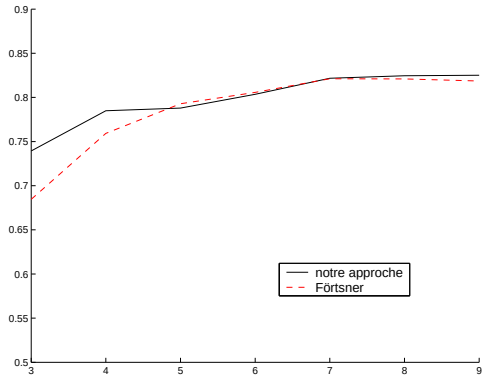


Figure 3.22 - Comparaison du taux de surdétection entre notre approche et l'approche à partir des coordonnées de Plücker.

3.8 Conclusion

3.8.1 Apports

On a présenté un algorithme qui permet un appariement et une reconstruction multi-images de segments 2D. L'algorithme fait l'appariement de contours dans l'Espace Objet et n'accorde ainsi aucun rôle particulier à une image précise. Il est, en ce sens, parfaitement symétrique. L'algorithme cherche à garder les appariements qui assurent des recouvrements maximum. En outre, la mise en correspondance est en $\mathcal{O}(N)$ et évite ainsi les problèmes de combinatoire habituels tout en tenant compte de l'ensemble des segments présents dans les images. On utilise également une méthode de reconstruction de segments 3D qui tient compte de l'incertitude sur la détermination des paramètres des droites 2D portant les segments. Cette méthode fournit des résidus normalisés qui permettent de qualifier la reconstruction.
Il est à noter que les résultats de cet algorithme ont déjà été utilisés avec des points issus de corrélation dans le cadre d'une caricature de scène urbaine qui fournit des résultats très prometteurs [PMT+01, MPT02].

3.8.2 Discussion et perspectives

L'algorithme actuel ne gère pas les incertitudes sur les paramètres de caméra qui devraient être modélisés pour avoir une propagation d'incertitude globale. Un autre défaut est de ne pas utiliser le voisinage radiométrique. Cette information est utile dans le cas d'images dont le point de vue change peu (rapport base sur hauteur faible en imagerie aérienne) et dans le cas de couples stéréoscopiques pour lesquels la contrainte géométrique ne peut pas être utilisée.
En revanche, l'algorithme fournit un ensemble de segments fiables, géométriquement bien localisés avec un taux de surdétection raisonnable. Ces caractéristiques permettent une utilisation dans le cadre de caricature de scènes comme nous l'avons dit et, de façon générale, en complément d'autres primitives qu'ils peuvent conforter ou au contraire invalider. Nous verrons que les segments 3D seront utilisés dans la suite pour le choix d'hypothèses de bâtiments, la présence d'un segment près d'une arête déduite d'une intersection de plans renforçant la probabilité qu'un bâtiment possédant cette arête existe.

Chapitre 4

Détection de facettes planes

4.1 Introduction

4.1.1 Contexte

Dans le cadre de la reconstruction de scènes urbaines, comme nous l'avons déjà dit, l'utilisation de primitives différentes est cruciale pour faciliter la génération d'hypothèses de bâtiments et réduire la combinatoire sous-jacente au problème. Dans ce contexte, les primitives linéaires fournissent une information structurante essentielle [TD02a, SZ97, BSZF99]. En revanche, leur manque de densité et leur sous-segmentation rendent la génération de modèles de bâtiments délicate à partir de ces seules primitives. Plusieurs auteurs ont mis en avant l'utilisation de primitives différentes pour la modélisation de bâtiments [FLM99, BSZF99].

4.1.2 Objectifs

L'objectif de ce chapitre est d'obtenir un ensemble de primitives planes à partir d'un MNE dense. L'approche proposée est basée sur un processus de grossissement de régions et conduit à l'obtention d'une hiérarchie de segmentations du MNE à différentes échelles. On utilise pour ce faire les outils de segmentation multi-échelles de L. Guigues [GLMC03]. Nous tenons à remercier ce dernier pour les innombrables et riches discussions qui ont permis l'aboutissement de ce travail. On montrera que le processus est général et peut segmenter un MNE quelconque, qu'il fournit un ensemble de régions concurrentes, et qu'il permet d'obtenir naturellement des segmentations du MNE à différentes échelles. Enfin, en s'appuyant sur une orthoimage de gradients résultant des travaux de D. Boldo [Bol02], on verra que cette hiérarchie de segmentations s'appuie sur des contours images et évite ainsi les délocalisations fréquentes du MNE tout en gardant la possibilité d'introduire des contraintes linéaires (segments, limites cadastrales de bâtiments).

4.2 État de l'art

4.2.1 Approches image

L'extraction de facettes planes à partir de deux images est un problème qui a reçu une grande attention dans la communauté. On peut citer les travaux de Tarel, Vezien, Randriamasy et plus globalement de l'ex-équipe SYNTIM de l'INRIA [Vez95, Tar96]. Cette approche se caractérise par l'utilisation d'attributs peu sensibles aux petites erreurs de segmentation pour apparier des régions dans un couple d'images préalablement segmentées. En revanche, la méthode est sensible aux occultations et aux différences importantes de segmentation entre les images. En outre, le problème de l'appariement reste entier dans le cas du multi-images où la mise en correspondance pose un problème combinatoire important.

Récemment, Ishikawa [IJ01] a fourni un algorithme capable d'extraire des régions communes à partir de plusieurs images. Cet algorithme est basé sur un appariement des contours des régions dans un espace $\mathbb{R}^{2n}$ en utilisant divers critères (gradient, maxima de flots optiques...) et en recherchant les contours qui s'apparient le mieux par une technique de minimisation de cycle dans un graphe. L'auteur avance des résultats encourageants, mais sur quelques régions seulement. En outre, si le processus fournit *de facto* un indicateur de qualité (l'ordre dans lequel les régions apparaissent), il ne gère pas explicitement de processus multi-échelles dont F. Fuchs a montré l'importance dans le cadre d'une modélisation et d'une caricature de scènes urbaines [Fuc01].

4.2.2 Approches segmentation de MNE

La segmentation d'images de profondeur est la technique qui se rapproche le plus du travail présenté ici. De nombreux auteurs se sont penchés sur ce problème. Hoover et al dans [HJBJ$^+$96] comparent quelques-uns des algorithmes de segmentation (voir cet article pour plus de bibliographie dans le domaine).

La plupart des algorithmes existants utilisent des techniques de "region-growing" et d'estimation robuste pour déterminer les régions pertinentes. Dans le cas d'images aériennes [Ves00] propose une approche en 3 étapes : sélection des modèles plans par un algorithme du RANSAC, fusion des hypothèses basée sur le F-Test et sélection des hypothèses sur un critère MDL. D'autres techniques utilisent le critère MDL [FLM99], des critères bayésiens [CJC01] ou la transformée de Hough pondérée [JPPDM00].

Toutes les approches travaillant directement sur des MNE issus de corrélation utilisent des critères robustes pour pallier les erreurs inévitables de la corrélation. S'appuyant toutes exclusivement sur les MNE, elles sont toutefois sujettes aux délocalisations et les techniques proposées ne permettent pas d'intégrer simplement des données images exceptée la méthode évoquée dans [Bol02], très proche de notre travail.

Enfin, si la plupart proposent des partitionnements du MNE, peu donnent la possibilité de travailler avec des régions concurrentes. Le travail de [Fuc01] se rapproche le plus de notre travail avec la gestion du multi-échelles et la génération d'hypothèses concurrentes. Nous verrons d'ailleurs qu'un cas particulier de notre segmentation permet d'obtenir les segmentations obtenues par sa méthode, mais notre travail se place sur une échelle continue là où lui n'obtenait qu'un ensemble discret de segmentations. Dans notre cas, l'approche utilisée permet une représentation explicite de la hiérarchie des partitions (ce qui facilite le calibrage des paramètres) et aussi l'obtention des partitions du MNE constituées de régions présentes à des échelles différentes.

4.3 Schéma Général

4.3.1 Formulation

Les fonctions de coût utilisées dans notre travail sont, comme la plupart des fonctions de coût utilisées en segmentation d'images, région-séparables, c'est-à-dire que l'énergie globale d'une segmentation s'écrit comme la somme des énergies des régions prises individuellement. Nous considérons en outre la classe des énergies séparables multi-échelles qui peuvent s'écrire pour une partition donnée $\mathbf{P}$ d'une image de profondeur I :

$$\mathbf{E}_\lambda(\mathbf{P}) = \sum_{R \in \mathbf{P}} \mathbf{E}_\lambda(R) = \sum_{R \in \mathbf{P}} \lambda \mathbf{C}(R) + \mathbf{D}(R) \tag{4.1}$$

où la somme est sur toutes les régions R de la partition et $\lambda \in \mathbb{R}^+$ est un paramètre d'échelle qui définit le niveau de caricature. L'objectif de la segmentation multi-échelles est de déterminer la partition qui minimise l'énergie pour chaque paramètre d'échelle λ. Le problème est insoluble en général compte tenu de la complexité de la formulation. Cependant, nous expliquons en section 4.3.2 comment construire *en une étape* une hiérarchie fournissant implicitement pour chaque échelle une solution sous-optimale à ce problème étant

données *deux fonctions d'énergie* $\mathbf{C}$ *et* $\mathbf{D}$ *sur les régions*. À partir de la section 4.4, nous explorons différents choix de fonctions basés sur une formulation classique qui met en balance un terme de complexité de description de modèle et de description des données connaissant le modèle :

$$\mathbf{C}(R) = \mathcal{L}(M) \qquad \mathbf{D}(R) = \mathcal{L}(I \mid M) \tag{4.2}$$

Dans ce schéma, λ règle le niveau de caricature de la segmentation.

4.3.2 Segmentation multi-échelles et hiérarchie

Nous rappelons ici brièvement quelques-uns des principes de la segmentation multi-échelles développée en détail dans [GLMC03]. L'idée, partant d'une sursegmentation initiale de l'image de profondeur, est de construire une hiérarchie de régions dont les coupes minimales pour $\mathbf{E}_\lambda$ sont des approximations des minima de $\mathbf{E}_\lambda$ sur le treillis complet des partitions $\mathbb{P}$. [GLMC03] montre que si $\mathbf{C}$ est une fonction décroissante sur $\mathbb{P}$, c'est à dire vérifie

$$\forall R_1, R_2 \qquad \mathbf{C}(R_1 \cup R_2) < \mathbf{C}(R_1) + \mathbf{C}(R_2) \tag{4.3}$$

alors les coupes minimales de la hiérarchie pour $\mathbf{E}_\lambda$ forment une séquence de *partitions ordonnées* suivant λ : de la plus fine à la plus grossière.

4.3.3 Hiérarchie d'escalade

L'idée est alors de construire la hiérarchie du bas vers le haut par une stratégie de fusion de régions qui simule une croissance progressive de λ. Toute paire de régions adjacentes est une hypothèse de fusion et pour chaque hypothèse de fusion de R_1 et R_2, on note *l'échelle d'apparition*, la valeur minimale λ_{app} de λ à partir de laquelle $\mathbf{E}_\lambda(R_1 \cup R_2) < \mathbf{E}_\lambda(R_1) + \mathbf{E}_\lambda(R_2)$, désignant ainsi l'échelle *à partir de laquelle la fusion des deux régions fait décroître l'énergie globale et est donc pertinente*. Une méthode naturelle pour la construction de la hiérarchie est alors de partir d'une partition fine de I qui constitue les *noeuds de base* de la hiérarchie et de réaliser itérativement la fusion de l'hypothèse qui a l'échelle d'apparition minimale (figure 4.1). Un nouveau noeud est créé dans la hiérarchie pour représenter cette nouvelle région qui est un ensemble connecté de noeuds de base de I. Cette méthode appelée "escalade d'échelle" est sans paramètre et fournit une hiérarchie valuée dans laquelle chaque noeud est valué par son énergie d'apparition. Cette hiérarchie H peut être représentée par un arbre dont l'axe vertical est l'axe des échelles (figures 4.2 et 4.7). On note λ_{dis} l'échelle de disparition d'un noeud n qui n'est autre que l'échelle d'apparition du père de ce noeud dans la hiérarchie. $\Lambda_n = [\lambda_{app}, \lambda_{dis}[$ fournit alors l'intervalle d'échelles pendant lequel la région représentée par le noeud n appartient à une coupe optimale de H pour $\mathbf{E}_\lambda$. Quelques-unes de propriétés importantes de H sont :

- la base est la partition initiale, le haut l'image complète.
- de bas en haut, H fournit des ensembles de plus en plus grands de régions décrivant des modèles de plus en plus grossiers,
- toute coupe dans H donne une partition de I [HP76, GLMC03]. Les coupes horizontales fournissent des segmentations à différentes échelles (figures 4.2 et 4.7),
- la génération de H est robuste à toute transformation linéaire sur $\mathbf{C}$ et $\mathbf{D}$.

4.3.4 D'un arbre binaire à un arbre n-aire

Au cours du processus de construction de la hiérarchie d'escalade, il peut arriver que l'échelle d'apparition d'un noeud soit supérieure à celle de son père dans la hiérarchie. L'intervalle d'échelles correspondant au noeud Λ_n est vide et de fait, la région correspondante ne peut apparaître à aucune échelle, on dit qu'elle n'est pas pertinente. L'étape de fusion binaire correspondante est alors inutile et doit être fusionnée avec l'étape

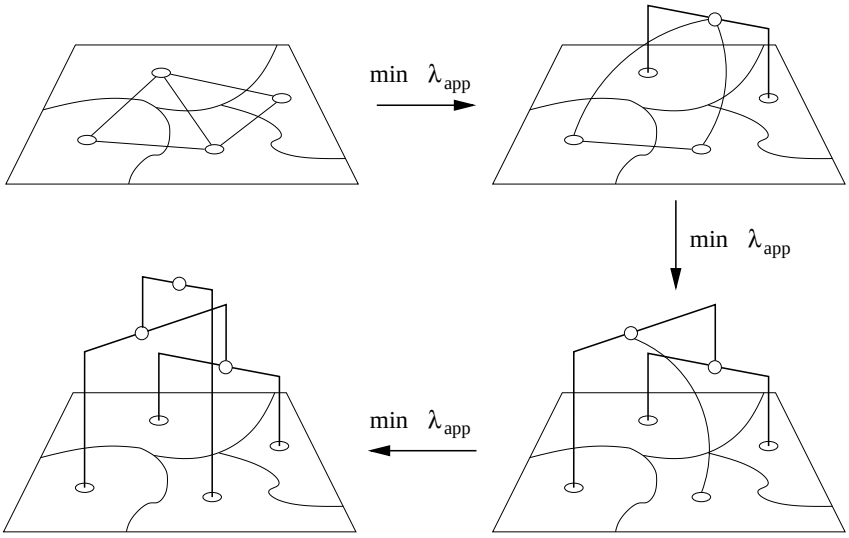


Figure 4.1 - Principe de la construction de la hiérarchie d'escalade.

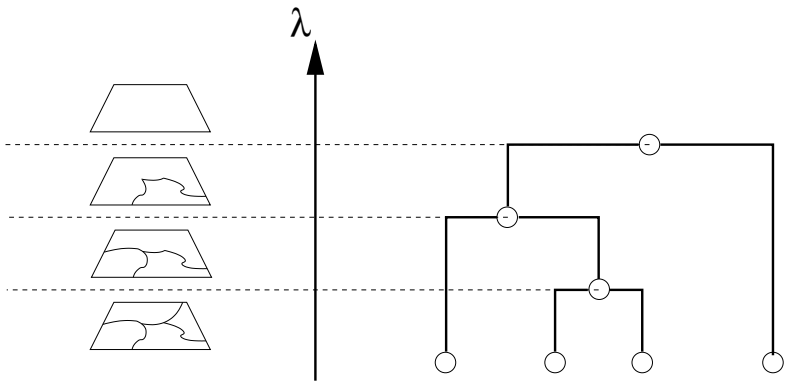


Figure 4.2 - Segmentations implicites dans la hiérarchie.

du père ce qui conduit à la création à l'échelle d'apparition du père d'un noeud n-aire comme l'illustre la figure 4.3.

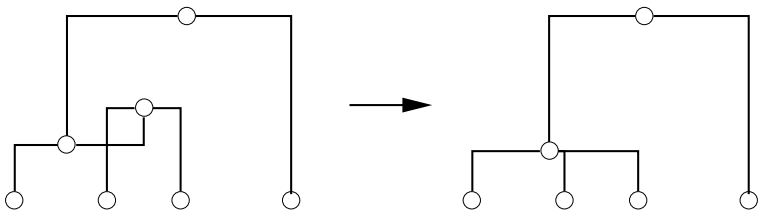


Figure 4.3 - Région non pertinente.

4.4 Modèle de facettes planes

4.4.1 Estimation robuste de plans

Pour déterminer l'écart d'une région R au modèle plan qui sert de référence (voir section 4.4.3), il est nécessaire pour chaque région de déterminer le meilleur plan $\mathcal{P}$ estimé sur l'ensemble des points de la région, défini par 3 paramètres et l'équation $z = \mathcal{P}(x, y) = ax + by + d$. L'estimation des trois paramètres a, b, d se fait par régression sur l'ensemble des points intérieurs à la région par le biais d'un estimateur robuste L_p. Nous avons choisi le M-estimateur $L_{1.2}$ qui présente un bon compromis entre complexité de calcul et stabilité [ZF92]. Par rapport à d'autres estimateurs (RANSAC, Least Median of Squares), cet estimateur n'est probablement pas le plus robuste et il est connu pour diverger lorsque le pourcentage d'erreurs est trop important. En revanche, si *l'initialisation* est bonne, cet estimateur prend en compte les erreurs résiduelles autour de la solution tout en éliminant les erreurs importantes. Or dans notre cas, comme nous avons un processus de fusion de régions, on peut avoir une bonne initialisation du plan solution à partir de la plus grande des deux régions filles dans la hiérarchie.

4.4.2 Codage du modèle

Dans le cas d'environnement urbain, nous recherchons des facettes planes aux formes simples. La simplicité de forme de la région est évaluée par la longueur de code de la séquence de Freeman de la frontière pixellaire donné par $\mid \partial R \mid \log 8$.
Comme rappelé auparavant, le modèle de plan est défini par 3 paramètres réels supposés indépendants et obtenus à partir de $\mid R \mid$ points. Rissanen [Ris89] dériva pour encoder cette information l'expression $(3/2) \log \mid R \mid$ en optimisant la précision à laquelle les paramètres sont encodés (voir aussi [KDN+95]). L'expression finale pour décrire le modèle de facettes planes devient alors :

$$\mathbf{C}_{\mathcal{P}} = 3|\partial R| + \frac{3}{2} \log |R| \tag{4.4}$$

4.4.3 Codage des données : codage par rapport au plan

Considérant qu'un plan $z = \mathcal{P}(x, y) = ax + by + d$ a été évalué sur une région donnée R, la fonction utilisée pour encoder les erreurs introduites par cette approximation pour chaque pixel $\mathbf{x} \in R$ est déduite du théorème de Shannon [KDN+95] qui donne la longueur de code minimale en supposant la distribution connue : $\mathcal{L}(\mathcal{E}(\mathbf{x}) \mid \mathcal{P})) = -\log p_1(\mathcal{E}(\mathbf{x}) \mid \mathcal{P})$ où $\mathcal{E}(\mathbf{x}) = I(\mathbf{x}) - \mathcal{P}(\mathbf{x})$ définit l'erreur introduite par l'approximation planaire pour un pixel $\mathbf{x}$, c'est à dire la distance du point à son projeté vertical sur le plan estimé. En utilisant un estimateur L_p, nous avons implicitement supposé une distribution probabiliste des erreurs donnée par [TMHF00] :

$$p_1(\mathcal{E}(\mathbf{x} \in R) \mid \mathcal{P}) = F \exp^{-\frac{|\mathcal{E}|^p}{p\sigma^p}} \tag{4.5}$$

où F est une facteur de normalisation et σ^2 est la variance de l'erreur. En supposant les points indépendants, la longueur de description pour une facette devient alors

$$\mathbf{D}_{\mathcal{P}}(R) = -|R| \log(F) + \sum_{\mathbf{x} \in R} \frac{|\mathcal{E}(\mathbf{x})|^p}{p\sigma^p} \tag{4.6}$$

Le premier terme est constant lorsque l'on somme sur l'ensemble des pixels de I et n'a donc aucune influence sur l'énergie globale de la partition à une échelle donnée ; on le négligera dans la suite.

4.4.4 Modèle global

Le modèle final pour chaque région intègre les deux arguments et peut ainsi s'écrire

$$\mathbf{E}_\lambda(R) = \lambda(3|\partial R| + \frac{3}{2}\log|R|) + \sum_{\mathbf{x}\in R} \frac{|\mathcal{E}(\mathbf{x})|^p}{p\sigma^p} \tag{4.7}$$

Partant d'une segmentation initiale fruste en régions carrées de taille arbitraire (au moins 2x2 pixels), la hiérarchie est créée en suivant le schéma décrit dans 4.3 et l'énergie rappelée ci-dessus. La section 4.8 présente quelques résultats. À noter que le terme $C_\mathcal{P}$ remplit les conditions de croissance spécifiées dans la section 4.3.2.

4.5 Modèle de facettes planes et homogènes

La plupart des MNE de corrélation présente des erreurs, des délocalisations importantes au niveau des discontinuités altimétriques dues, principalement, à la corrélation par vignette qui est perturbée par les changements de texture d'une image à l'autre au niveau de ces sauts d'altitude. Le phénomène de lissage observé est, quant à lui, lié à la taille de la vignette utilisée : plus la vignette est importante, plus le lissage l'est. Au contraire, avec des vignettes plus petites, le lissage apparaît moins mais les erreurs grossières (de type "poivre et sel") sont plus importantes. Peu de travaux tentent de contrer ces artefacts directement dans l'espace objet en utilisant l'information multi-vues. On peut citer entre autres [ADWS02, CPJ98] qui tiennent compte des discontinuités image pour le calcul du MNE. Dans la suite, nous montrons comment l'orthoimage de gradients [Bol02] peut être utilisée à l'intérieur de notre schéma pour intégrer les contraintes images.

4.5.1 Orthoimage de gradients

On désigne comme orthoimage de gradients l'image générée par l'accumulation des modules de gradient dans l'Espace Objet comme indiqué dans [Bol02]. Chaque pixel (x, y) de l'orthoimage de gradients correspondant à un pixel de MNE est calculé par la fonction :

$$G(x,y) = \max_{z\in[z_{min},z_{max}]} \sqrt[n]{\prod_{i=1}^{i=n} \nabla I_i(\mathcal{T}_i(x,y,z))} \tag{4.8}$$

où $\nabla I_i(\mathcal{T}_i(x,y,z))$ est la valeur du module de gradient de l'image $i \in [1..n]$ à la projection du point (x,y,z), z_{min} et z_{max} sont l'altitude minimale et maximale déterminées à partir du MNE par une bande de tolérance autour de ce dernier (figure 3.11). Cet opérateur permet de gérer les discontinuités altimétriques et donne plus de poids aux gradients cohérents en 3D à travers l'équation 4.8. Une étude plus approfondie est faite dans [Bol02] pour justifier le choix de cette fonction.
Cette orthoimage de gradients sera la base des contraintes image introduites dans le processus. Dans un premier temps, un algorithme de ligne de partage des eaux est réalisé sur cette orthoimage de gradients, il permet d'obtenir l'ensemble des régions élémentaires formant la sursegmentation de base de la hiérarchie. Chaque frontière de la segmentation est une limite potentielle et la valeur moyenne de l'orthoimage de gradients le long de cette frontière modélise sa "force".

4.5.2 Modèle de facettes homogènes

On peut considérer l'orthoimage de gradients sur laquelle s'appuie la segmentation en ligne de partage des eaux comme une modélisation des changements d'informations radiométriques des images placées dans un référentiel commun. Nous voulons imposer, pour chaque région, que la surface soit la plus homogène

possible. Ce modèle d'homogénéité, noté $\mathcal{H}$ n'a pas besoin d'être codé dans la mesure où il est *a priori pour toutes les régions* : $\mathbf{C}_{\mathcal{H}} = 0$. Pour forcer les régions à s'appuyer sur des gradients élevés tout en étant aussi homogènes que possible et pour éviter de franchir les gradients importants dans le processus de fusion, on code les gradients le long des frontières internes comprises dans la région. Pour coder l'inhomogénéité, chaque frontière f *interne* à la facette R est codée par ses pixels pondérés par la distribution *a priori* sur le gradient. En supposant les points indépendants, cela se réduit à

$$\sum_{f \subset R} \sum_{\mathbf{x} \in f} -\log(p_2(G(\mathbf{x}) \mid \mathcal{H})) \tag{4.9}$$

La distribution p_2 doit suivre une loi décroissante pour modéliser l'écart à la valeur 0 (modèle homogène). Après plusieurs essais en particulier avec la distribution $p_2(x) = \frac{1}{x}$, nous avons finalement opté pour une fonction à décroissance rapide $p_2(x) = \exp(-x)$ qui pénalise fortement les gradients incohérents c'est-à-dire les faibles valeurs de G. Le modèle devient alors :

$$\mathbf{D}_{\mathcal{H}}(R) = \sum_{f \subset R} \sum_{\mathbf{x} \in f} G(\mathbf{x}) \tag{4.10}$$

4.5.3 Modèle global

Le modèle global est le modèle de *facettes simples, planes et homogènes* et reprend les 2 précédents termes :

$$\begin{aligned} \mathbf{E}_{\lambda}(R) &= \lambda \mathbf{C}_{\mathcal{P}}(R) + (\mathbf{D}_{\mathcal{P}}(R) + \beta \mathbf{D}_{\mathcal{H}}(R)) \\ &= \lambda(3|\partial R| + \frac{3}{2}\log|R|) \\ &+ \sum_{\mathbf{x} \in R} \frac{|\mathcal{E}(\mathbf{x})|^p}{p\sigma^p} + \beta \sum_{f \subset R} \sum_{\mathbf{x} \in f} G(\mathbf{x}) \end{aligned} \tag{4.11}$$

La segmentation initiale découle cette fois de l'application d'un watershed sur l'orthoimage de gradients de façon à s'arrêter sur les contours images. β règle le ratio entre le modèle de facettes planes et de facettes homogènes. Ce modèle de facettes planes et homogènes permet la segmentation *simultanée* de l'image de profondeur et de l'orthoimage de gradient.

4.5.4 Adjonction de contraintes

Dans le cadre d'une segmentation en facettes planes, on veut souvent introduire des contraintes déduites d'autres détecteurs de primitives tels que des détecteurs de segments 3D ou des détecteurs de façades. Compte tenu du modèle choisi (segmentation sur une représentation 2D1/2 de la scène), ces contraintes seront ici introduites en 2D1/2 en projetant les segments 3D dans l'orthoimage de gradients et en affectant une valeur de gradient forte à ces segments sur lesquels nous voulons que la segmentation s'arrête. L'adjonction de ces segments permet de contraindre la sursegmentation initiale et de forcer la fusion au cours du processus grâce à la fonction décrite en section 4.5.2. On peut en particulier constater que le watershed initial respecte mieux les discontinuités avec l'adjonction de contraintes (figure 4.4).

4.6 Obtention d'une caricature

Nous présentons ici différentes extensions pour l'obtention d'une caricature de scène à partir d'une partition à une échelle donnée. Ce travail se place un peu en marge de l'objectif initial de reconstruction de bâti et s'apparente plus à une recherche de caricatures de MNE telles que l'on peut en trouver chez [Ves00, CJC01].

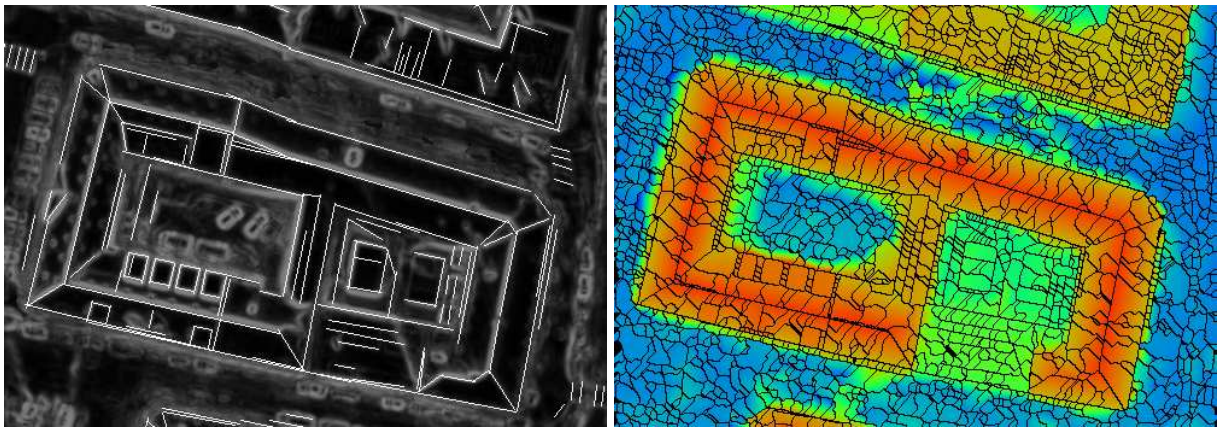


Figure 4.4 - Exemple de contraintes rajoutées sur l'orthoimage de gradients et watershed correspondant.

4.6.1 Post-traitements

À l'issue de la coupe, pour une mise en 3D réaliste à partir d'un MNE de corrélation, donc en utilisant la segmentation sur orthoimage de gradients, plusieurs étapes de post-traitement sont nécessaires.

Suppression des facettes trop étroites

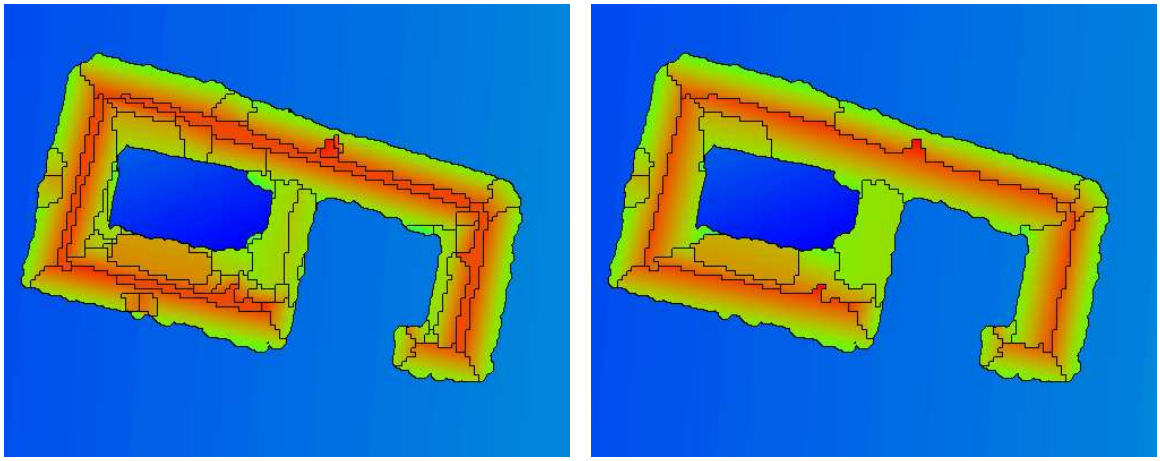
La création de l'orthoimage de gradients par projection de gradients sur un MNE dilaté génère un grand nombre de petites régions quasi pixellaires aux gradients assez forts (essentiellement au niveau des façades) qui bloquent la fusion comme le montre [Bol02]. En outre, dans le calcul de l'incertitude géométrique, on peut montrer que comme pour les régressions de droites à partir de points [DVF92], la variance des paramètres du plan dépend des moments d'ordre 2 de la région. Pour éviter d'avoir des régions trop incertaines et pour éliminer les régions trop petites, les régions dont un des moments d'inertie est inférieur à un seuil s_{MI} sont supprimées. De façon à garder la propriété de segmentation de l'image (utile dans le cas d'une caricature), ces régions sont fusionnées avec la région voisine qui a le gradient moyen minimum le long de la frontière.

Suppression des facettes trop verticales

Une autre étape de post-traitement consiste à éliminer les régions trop "verticales" peu intéressantes dans le cas d'une segmentation sur le MNE qui est une représentation 2D1/2 de la surface. Il est à noter que ces régions répondent souvent au premier critère de filtrage car elles sont en bordure de façade. Ces régions sont fusionnées avec leur voisine la plus "basse".

4.6.2 Mise en application sur un masque déterminé

La mise en application de l'algorithme sur un masque déterminé permet de se focaliser sur des zones obtenues par segmentation sol/sursol [Bai97] ou sur des zones de focalisation issues du cadastre [Jib02]. Une propriété intéressante est la gestion immédiate des différents niveaux de représentation de la zone concernée qui permet de faire apparaître des bâtiments à des niveaux de généralisation différents (figure 4.5).



(a) coupe à $\lambda = 1.16$ (b) coupe à $\lambda = 2.25$

Figure 4.5 - Exemple de segmentations avec focalisation sur un masque.

4.7 Détection de façades

À une échelle donnée, une segmentation fournit une partition de l'image de profondeur en facettes planes. Les limites entre facettes sont autant d'hypothèses de discontinuités altimétriques. Chaque pixel d'une limite possède en effet, virtuellement, deux altitudes calculées par projection de sa position planimétrique sur les plans respectifs déterminés sur les deux facettes. Nous considérerons qu'une limite entre deux facettes est une discontinuité altimétrique si la moyenne de la valeur absolue des altitudes de ses pixels est supérieure à un seuil Δ_f. Dans le cas d'une limite ne bordant qu'une facette (facette en bord de masque de focalisation), la limite est automatiquement classée comme discontinuité.
A l'issue de cette classification, une polygonisation des discontinuités altimétriques est effectuée en utilisant l'algorithme décrit en section 3.4.1 et un seuil c_f. Rappelons que la géométrie est primordiale pour l'algorithme. De fait, ce seuil sera choisi faible pour limiter les caricatures et donc la perte des qualités géométriques des limites.
À l'issue du processus, toutes les discontinuités altimétriques sont déterminées, les équations des plans verticaux qu'elles définissent sont calculées. Nous verrons toutefois que l'algorithme de génération d'hypothèses de bâtiments a besoin d'une orientation des normales des facettes planes et des discontinuités altimétriques. Pour déterminer cette orientation, une altitude moyenne est déterminée pour chaque côté de la façade et la normale est choisie de façon à pointer vers l'extérieur, c'est-à-dire du côté le plus haut vers le côté le plus bas. La figure 4.6 présente quelques exemples de détection de façades sur une orthophotographie avec orientation des normales.

4.8 Évaluation et résultats

Les résultats sont présentés sur un MNE dense généré par la fusion de plusieurs MNE de corrélation obtenus à partir de paires stéréoscopiques [BD00]. La figure 4.7 présente une hiérarchie générée à partir d'un MNE et d'une segmentation initiale en dalles carrées 2x2 ainsi que des résultats obtenus à différentes échelles de représentation. $\sigma = 0,5m$, $p = 1,2$ et $\beta = 100$ furent utilisés dans nos tests. Il est à noter que la valeur de σ n'a toutefois aucune influence sur la *structure* de la hiérarchie $\mathcal{H}$ en raison de la robustesse

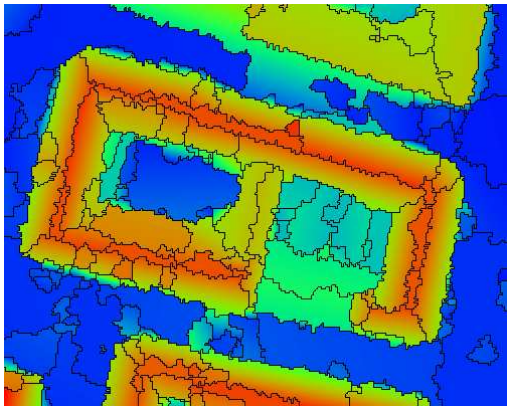
Figure 4.6 - Quelques exemples de discontinuités altimétriques orientées projetées sur une orthophotographie.

de la génération de $\mathcal{H}$ à des transformations linéaires de $\mathbf{D}$. La figure 4.8 montre quelques exemples de segmentations obtenues à partir du même MNE et de contraintes images incluant des contraintes ajoutées par projection de segments 3D. La comparaison entre les deux figures montre le gain obtenu en ce qui concerne la délimitation des facettes planes. Dans le cas de la segmentation simultanée de l'orthoimage de gradients et du MNE, les contours épousent beaucoup plus les formes réelles et la fusion s'arrête sur ces détails. D'autres contraintes (limites cadastrales, linéons projetés [JTP02]) pourraient être appliquées en lieu et place des segments 3D qui sont utilisés ici.
La figure 4.9 montre un exemple de caricature obtenue sur la zone étudiée à l'aide de l'ensemble des contraintes.

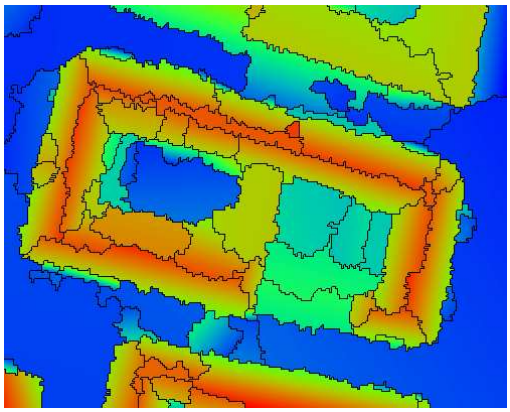
4.9 Conclusion

4.9.1 Apports

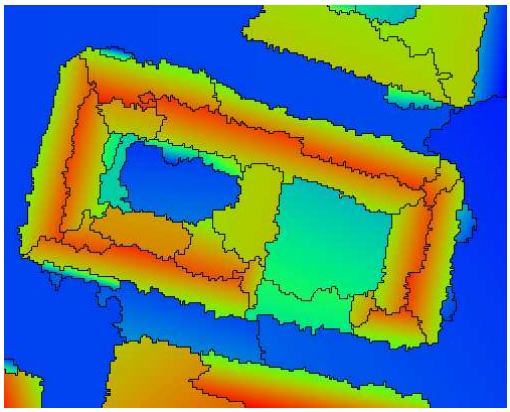
En conclusion, nous avons présenté un algorithme capable de générer un *ensemble* de segmentations de MNE. Cet algorithme peut fonctionner sur un MNE seul ou utiliser une information externe provenant des images en intégrant l'information dans la fonction de coût. Nous avons montré également qu'il pouvait naturellement intégrer des contraintes planimétriques. Le choix d'une coupe permet d'obtenir immédiatement une caricature de la zone sous forme d'un ensemble de facettes planes.
Par rapport à l'algorithme développé par Boldo [Bol02], la méthode présentée ici est plus générale car elle s'étend à n'importe quel MNE, n'utilise qu'une seule fonction de coût gérant la fusion *et* la segmentation optimale, permet d'obtenir *en une seule passe* toutes les coupes à des échelles λ différentes ou de faire des coupes utilisant des régions apparaissant à des échelles variables. Enfin, la fusion hiérarchique est guidée *à la fois* par la planéité de la surface et par la présence ou l'absence de gradient interne. En outre, dans le cas d'un générateur brut d'hypothèses de plans, on peut extraire, sur une même zone, des régions qui apparaissent à des échelles différentes.
En annexe C, nous présentons une étude sur l'évaluation de l'incertitude sur les paramètres décrivant chacun des plans ce qui peut être utile pour effectuer des tests entre primitives provenant de détecteurs différents [FBH00].



$\lambda = 0.5$

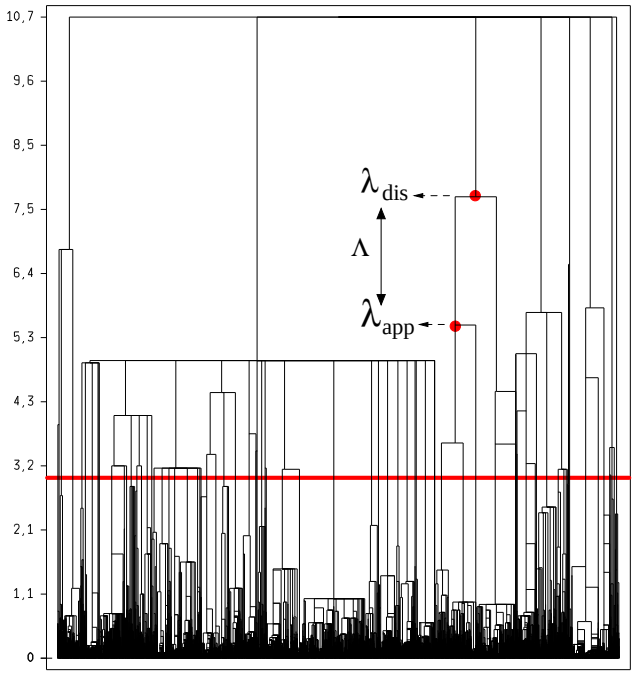


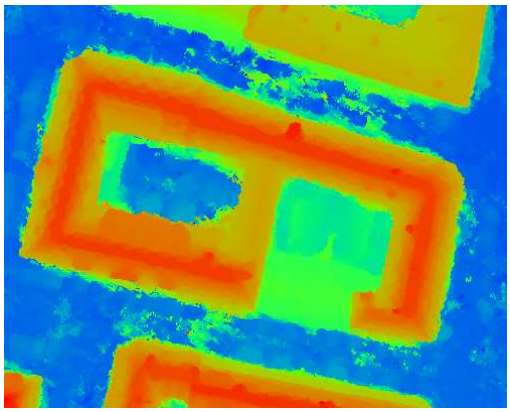
$\lambda = 1$



$\lambda = 3$

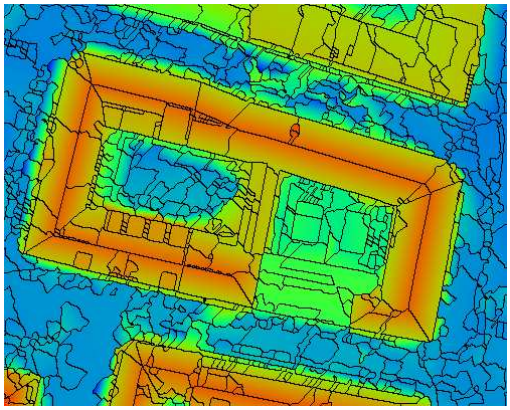
axe des échelles



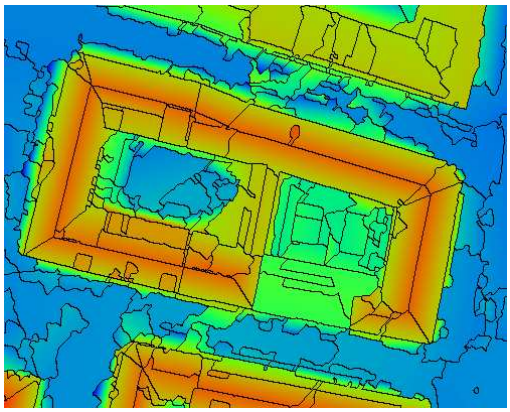


MNE initial

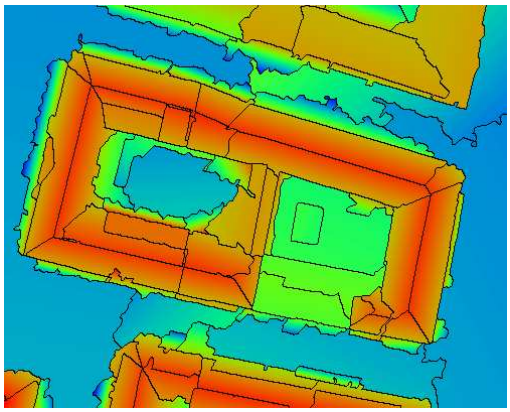
Figure 4.7 - Modèle de facettes planes. Hiérarchie et exemples de segmentations.



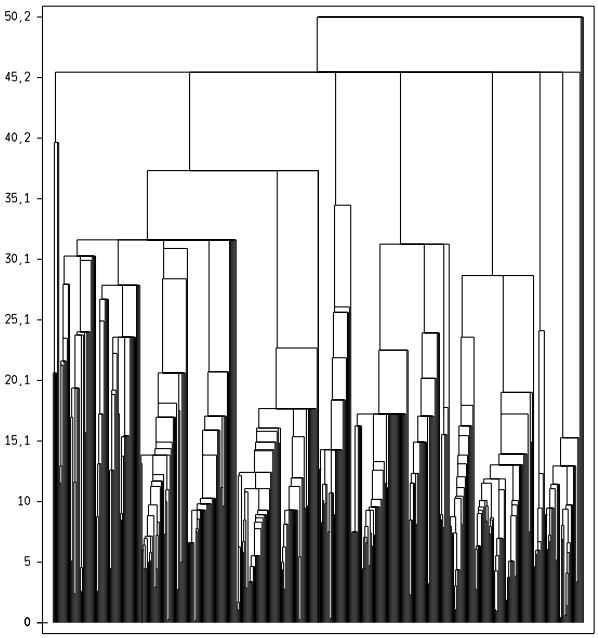
$\lambda = 1$

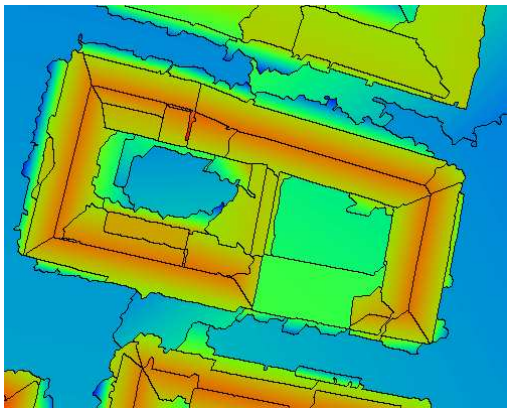


$\lambda = 2$



$\lambda = 4$





$\lambda = 5$

Figure 4.8 - Modèle de facettes planes et homogènes. Hiérarchie et exemples de segmentations.

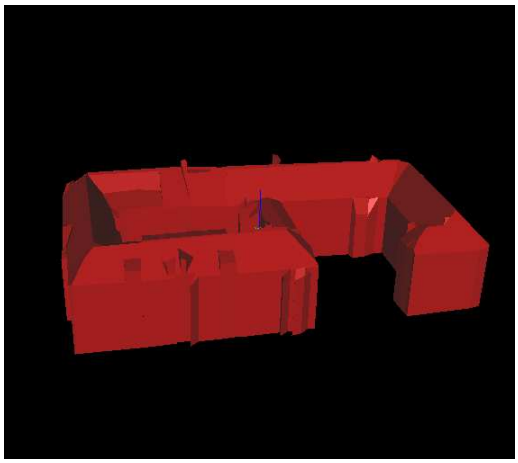


Figure 4.9 - Vue 3D d'une caricature de scène issue d'une segmentation sur une orthoimage de gradients avec contraintes.

4.9.2 Discussions

L'algorithme n'intègre pour l'instant pas d'informations radiométriques qui pourraient provenir d'une orthophotographie par exemple et conforteraient le modèle : facettes planes et homogènes. Une étude plus approfondie est de plus nécessaire pour le choix de la pondération β. La caricature obtenue est directement utilisable dans certaines applications nécessitant peu de qualité visuelle. En revanche, les discontinuités inévitables au niveau des frontières de facettes posent des problèmes de modélisation (pas de topologie reconstruite). Tous les essais effectués laissent, de plus, perplexe quant à la possibilité d'obtenir "la bonne segmentation". En fait, il s'avère que, si on est capable d'obtenir toutes les hypothèses de plans à un endroit donné dans la hiérarchie *ou par une sursegmentation*, définir une segmentation parfaite reste délicat : le choix du λ idéal ou d'une fonction de coût définissant une coupe unique permettant d'obtenir le niveau de représentation voulu est utopique. Nous pensons que, quelle que soit la stratégie employée, aucune ne permettra d'obtenir cette segmentation parfaite et qu'il faut passer soit par un processus de sélection de régions concurrentes à l'intérieur de la hiérarchie soit par une "faible" sursegmentation pour garantir la présence des plans principaux (dont les limites seront toutefois imprécises voire erronées mais dont l'équation, donc la *géométrie* sera correcte) et éliminer les plans erronés en faisant intervenir d'autres primitives ou d'autres critères. La première idée qui est celle mise en application par F. Fuchs a l'avantage de prendre en compte *toutes* les régions de la hiérarchie. En revanche, elle oblige à définir des nappes $[\lambda_{min}, \lambda_{max}]$ d'action et, pour éviter des redondances, des critères complexes de choix de régions (on pourrait imaginer ici faire intervenir les notions de persistance de régions [Gui03]). La deuxième idée pour laquelle nous avons opté, ne fait intervenir qu'un paramètre (un λ de sursegmentation). Le choix d'une sursegmentation permet en outre de définir les façades à partir des limites des facettes obtenues. C'est cette approche qui a été choisie pour la détection de discontinuités altimétriques. Nous sommes pleinement conscients des limites et de la relative faiblesse de ce dernier détecteur de discontinuités. Nous n'avons pas porté d'attention particulière à ce détecteur pour des raisons de temps et de choix. Avec la disponibilité croissante, même si cela est hors du cadre de cette thèse, de limites cadastrales, qui à défaut de fournir toutes les discontinuités, définissent au moins des orientations privilégiées, nous avons jugé plus utile de nous consacrer aux autres types de primitives.

4.9.3 Perspectives et utilisation

Comme annoncé plus haut, l'algorithme sera utilisé en faible sursegmentation de façon à garantir la présence de *tous* les plans principaux (plans non verticaux et verticaux). La coopération entre les différentes primitives (segments et hypothèses de plans) permettra de ne garder que les plans pertinents dans la modélisation polyédrique.

Troisième partie

Reconnaissance de bâtiments

Chapitre 5

Reconnaissance automatique de bâtiments

5.1 Méthodologie générale

5.1.1 Problématique

Nous avons choisi de modéliser un bâtiment par une forme polyédrale sans surplomb dont le contour externe est constitué d'un ensemble de plans verticaux. Cette définition très générique permet de représenter la quasi-totalité des bâtiments présents en agglomération urbaine (à l'exception des rares bâtiments aux formes arrondies). L'objectif de cette partie est de pallier les erreurs dues aux détecteurs de primitives présentés dans les chapitres précédents, d'être capable de générer *à partir des primitives* un ensemble de modèles de bâtiments plausibles (section 5.2), d'avoir un système ouvert capable, par une modélisation simple, de gérer aisément diverses observations externes et de pouvoir intégrer une certaine forme de généralisation par le biais de contraintes géométriques sur les primitives (orthogonalité, parallélisme...) pour choisir le bon modèle (section 5.3).

5.1.2 Hypothèses

Nous supposerons ici que toutes les primitives planes importantes sont présentes en géométrie, c'est-à-dire que même si les pourtours des régions planes ne sont pas corrects, nous supposerons que les équations sont approximativement justes, que les directions, c'est à dire les normales (pour les zones planes non verticales comme pour les façades) sont donc relativement - ce relativement entrera en compte dans le choix de certains paramètres - correctes et surtout sont *toutes* présentes. Cette hypothèse, forte dans certains cas (stéréoscopie par exemple), est assez facilement vérifiée dans notre contexte de multiscopie (chapitre 4).
En ce qui concerne les segments 3D, nous considérerons qu'ils forment un ensemble de primitives correctement localisées. En revanche, aucune hypothèse n'est nécessaire cette fois-ci sur le taux de sous et sur détection.

5.1.3 Schéma général

La figure 5.1 qui est un rappel de la figure 2.9 récapitule le fonctionnement schématique général de l'algorithme de reconnaissance de bâtiments. La figure 5.2 permet de suivre les grandes étapes de cet algorithme sur un exemple réel. Le lecteur trouvera, en annexe D, d'autres exemples auxquels il est fortement invité à se référer pour suivre la progression du système. Cette section donne un aperçu du fonctionnement de la technique mise en oeuvre et renvoie le lecteur aux sections suivantes pour une description beaucoup plus détaillée des différentes étapes.
Comme nous l'avons déjà dit, l'algorithme fonctionne sur des régions d'intérêt ou zones de focalisation

sélectionnés manuellement ou détectés par un algorithme idoine [BM99] (figure 2.8). Dans une zone donnée, plans et façades sont extraits par les techniques présentées précédemment. Ils constituent les primitives de base utilisées dans un arrangement qui génère un graphe 3D de facettes $\mathcal{G}$ (section 5.2.1). De fait, ces primitives seront considérées pour leur géométrie : ce sont les équations de plans de ces primitives qui interviendront dans cette partie. Nous rappelons également que toutes les primitives planes sont orientées : vers le haut pour les facettes planes extraites par segmentation du MNE, vers l"'extérieur" pour les plans verticaux ou façades.
À l'intérieur de ce graphe 3D, après une simplification initiale, on montrera que la recherche des hypothèses de bâtiments se ramène à la recherche d'un ensemble de surfaces polyédriques du graphe 3D, appelées *surfaces admissibles*. Les sections 5.2.2 et 5.2.3 fournissent une étude approfondie de ces surfaces et de leurs propriétés qui permet d'aboutir à la démonstration de l'équivalence entre recherche de surfaces admissibles et recherche des hypothèses de bâtiments (section 5.2.4).
Plusieurs étapes de simplifications sur $\mathcal{G}$ sont nécessaires de façon à réduire le nombre de facettes et par là même le nombre d'hypothèses de bâtiments (section 5.2.4). La première étape utilise le bandeau de tolérance autour du MNE. La deuxième étape, de façon analogue, filtre les facettes en utilisant le masque de focalisation érodé.
Après ces simplifications, on décrit en section 5.2.5 une procédure de recherche des surfaces admissibles et donc de recherche de *l'ensemble des formes acceptables de bâtiments* $\mathcal{M}$. Cet algorithme fonctionne par énumération des cliques maximales à l'intérieur d'un graphe de compatibilité et reprend en cela les idées de [Jib02].
À l'issue du processus, le meilleur modèle de $\mathcal{M}$ est déterminé en utilisant une modélisation bayésienne faisant intervenir des informations externes (images, masque, segments 3D...) et une modélisation de la complexité d'un modèle dans laquelle des contraintes a priori sur les primitives (orthogonalité, parallélisme, horizontalité de certaines arêtes) peuvent être intégrées (section 5.3). Ces contraintes ne sont pas limitatives, mais assurent l'introduction d'une certaine forme de généralisation dans le processus de choix. Ces contraintes seront ensuite appliquées dans l'étape de reconstruction proprement dite qui sera détaillée dans le chapitre 6.
Le chapitre se clôt par une section énumérant quelques éléments pour optimiser la recherche de l'ensemble des formes acceptables qui est, comme on le verra, un problème combinatoire.

5.2 Recherche des formes acceptables

Comme nous l'avons déjà signalé, cette partie est une extension des travaux de [Jib02] au cas des discontinuités altimétriques. Néanmoins, ces façades introduisent de profonds changements dans la méthode et en particulier dans les démonstrations qui seront décrites ici.

5.2.1 Création d'un graphe 3D

Dans la suite, on appellera *polytope* tout polygone bidimensionnel plan (les points sont tous dans un même plan : le plan porteur) convexe. Soit $\mathcal{H}$ un ensemble de H plans de $\mathbb{R}^3$. On rappelle que l'intersection de ces plans définit un ensemble de polytopes d'intérieurs disjoints, d'arêtes et de points dans $\mathbb{R}^3$, qui constituent l'arrangement de plans de $\mathcal{H}$ [Ede87, BY95]. Dans la suite, par abus de langage, nous appellerons facette ou facette élémentaire tout polytope généré par l'arrangement de plans. Dans notre cas, l'ensemble des primitives planes détectées (plans et façades) constituant $\mathcal{H}$ est plongé dans un arrangement de plans qui permet de générer un graphe 3D fournissant toutes les relations topologiques d'adjacence entre les points, les arêtes et les facettes. Ce graphe est confiné dans une forme prismatique convexe s'appuyant sur le polygone de focalisation et délimitée par une altitude minimum $z = z_g$ (le "sol") et maximum (le "ciel") (figure 5.4). Le polygone de focalisation convexe délimitant cette forme prismatique au niveau du sol est intégré à

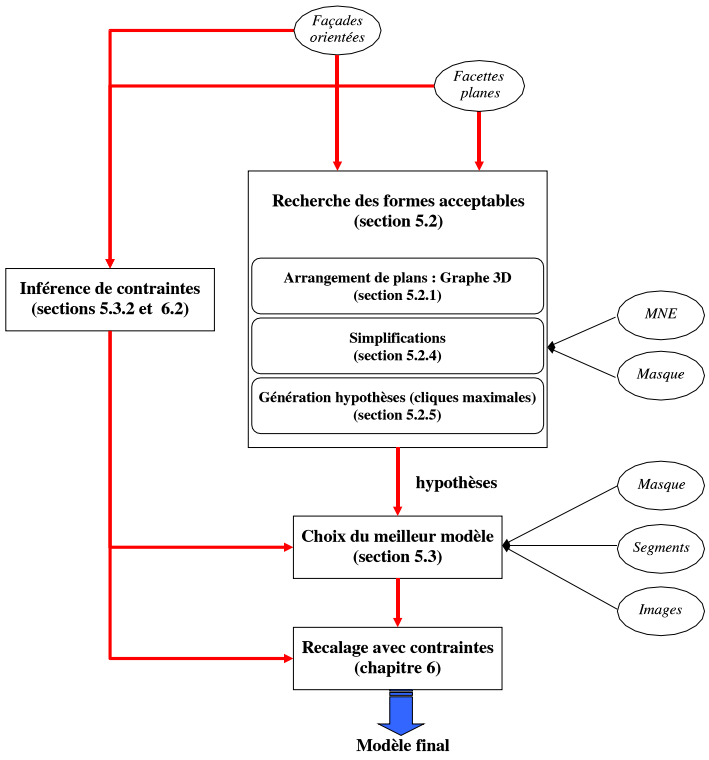


Figure 5.1 - Algorithme de reconnaissance de bâtiments. Principe général.

l'arrangement de plans et donc découpé en facettes élémentaires appartenant au graphe 3D. Dans la suite, on dénommera "face-sol" $\mathcal{F}_s$ ce polygone situé au niveau du sol, c'est à dire, dans le graphe 3D initial, l'ensemble des facettes incluses dans le plan horizontal au niveau du sol[1]. On notera que toutes les facettes du graphe *ont une orientation* : vers le haut par défaut pour les facettes qui s'appuient sur un plan non vertical, et dans le même sens que leur plan porteur dans le cas des portions de façades.

5.2.2 Définitions

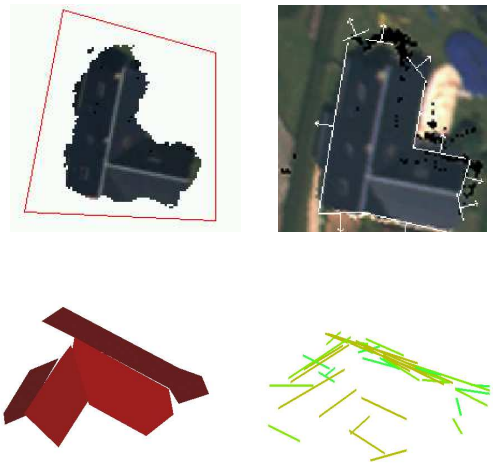
Orientations des facettes

On rappelle dans cette section quelques définitions issues pour la plupart de [BY95].
Orienter un polytope revient à choisir un ordre circulaire sur ses sommets. La donnée de l'orientation du plan porteur d'un polytope et d'un sens de rotation arbitraire induit une orientation du polytope. Dans la suite, on choisira, sans perte de généralité, une orientation arbitraire *directe* pour définir l'orientation des polytopes à partir des plans porteurs.
Les orientations de deux polytopes adjacents sont *cohérentes* si elles induisent sur l'arête commune des orientations opposées (figure 5.6). On peut alors montrer le lemme 1 qui permet de tester aisément la cohérence de deux polytopes adjacents.

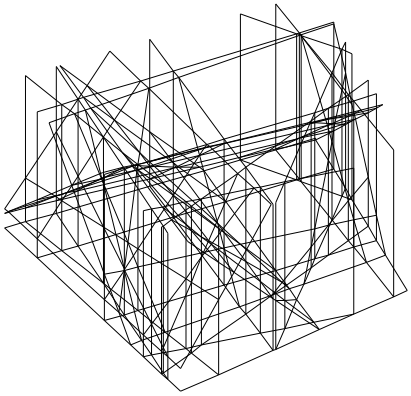
Lemme 1
2 polytopes adjacents ont des orientations cohérentes autour d'une arête $\mathbf{s}$ *si et seulement si, avec les notations de la figure 5.6,*

$$[\mathbf{s}, \mathbf{v_1}, \mathbf{n_1}].[\mathbf{s}, \mathbf{v_2}, \mathbf{n_2}] < 0$$

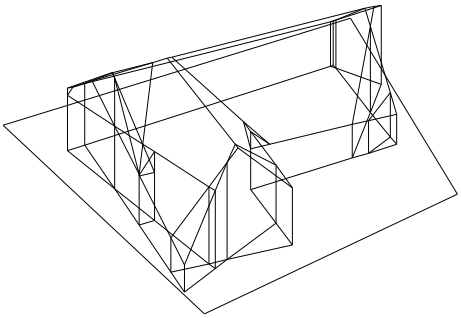
[1] La notation sous forme calligraphique traduit bien le fait que $\mathcal{F}_s$ est un ensemble de facettes élémentaires



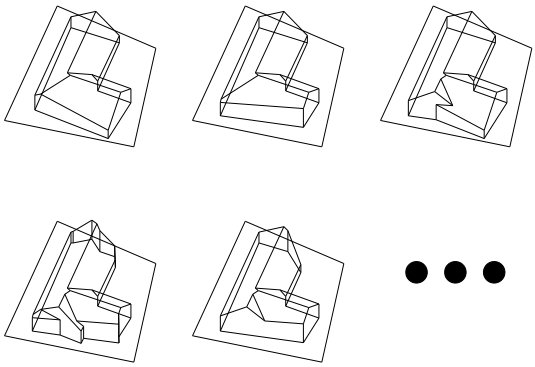
(a) Masque de focalisation, façades, facettes et segments.



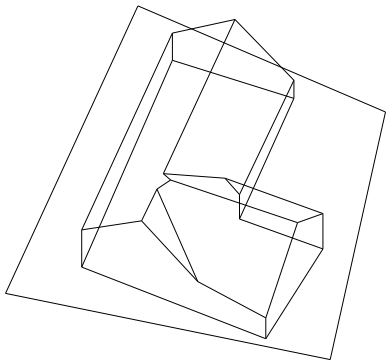
(b) Arrangement de plans.



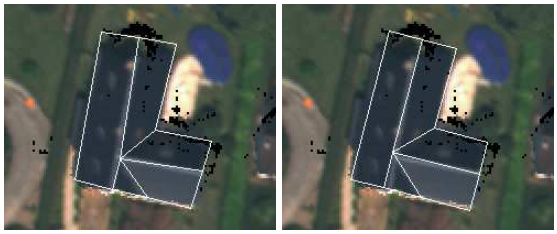
(c) Après filtrages.



(d) Recherche de l'ensemble des solutions.



(e) Choix de la meilleure solution.



(f) La solution avant et après recalage avec contraintes.

Figure 5.2 - Les différentes étapes de l'algorithme appliqué à un exemple.

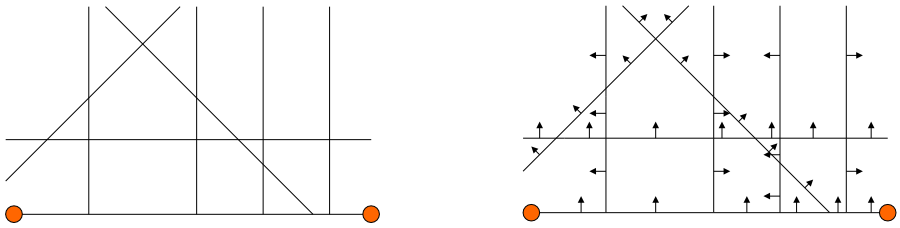


Figure 5.3 - Intersection de plans et Graphe 3D issu de l'arrangement (découpage en facettes) - vue schématique. Les points rouges en bord de figure délimitent les arêtes de bord ou arêtes de la face-sol.

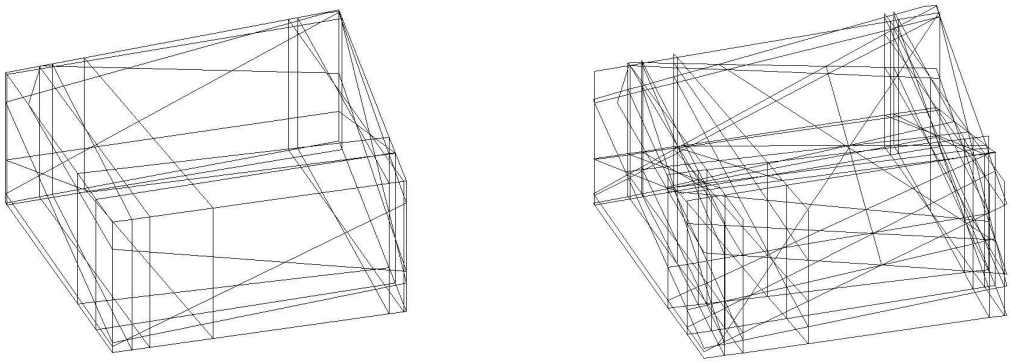


Figure 5.4 - Intersection de plans et Graphe 3D issu de l'arrangement (découpage en facettes) - vue 3D.

Preuve du lemme 1. Considérons un polytope orienté p et la normale de son plan porteur $\mathbf{n}$. Soit $\mathbf{s}$ une de ses arêtes bordantes, ainsi que le milieu de cette arête $\mathbf{c}$. Soit $\mathbf{g}$ le centre de gravité du polytope, on définit le vecteur $\mathbf{v} = \mathbf{cg}$ (non colinéaire à $\mathbf{s}$ si p est non dégénéré). L'arête $\mathbf{s}$ est orientée dans le sens induit par l'orientation du plan porteur p si et seulement si $[\mathbf{s}, \mathbf{v}, \mathbf{n}] > 0$ (figure 5.5).
Considérons, avec les notations de la figure 5.6 l'arête $\mathbf{s}$ orientée dans le sens induit par le polytope p_1. On a donc $[\mathbf{s}, \mathbf{v_1}, \mathbf{n_1}] > 0$. Les deux polytopes ont des orientations cohérentes autour de $\mathbf{s}$ si et seulement si, par définition, $-\mathbf{s}$ est orientée dans le sens induit par p_2 soit $[-\mathbf{s}, \mathbf{v_2}, \mathbf{n_2}] > 0$ ce qui amène le résultat. En supposant $\mathbf{s}$ orientée dans le sens opposé pour p_1, les rôles sont intervertis entre p_1 et p_2 mais le résultat reste le même. ■

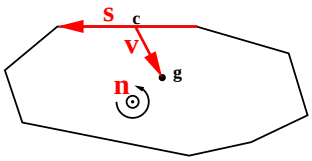


Figure 5.5 - Orientation induite pour un polytope.

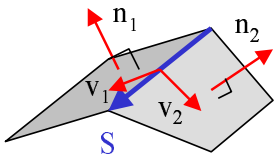


Figure 5.6 - Cohérence de l'orientation de 2 polytopes.

Polyèdres et surfaces polyédriques

Une *surface polyédrique* est un ensemble de polytopes arrangés dans $\mathbb{R}^3$ tel que

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

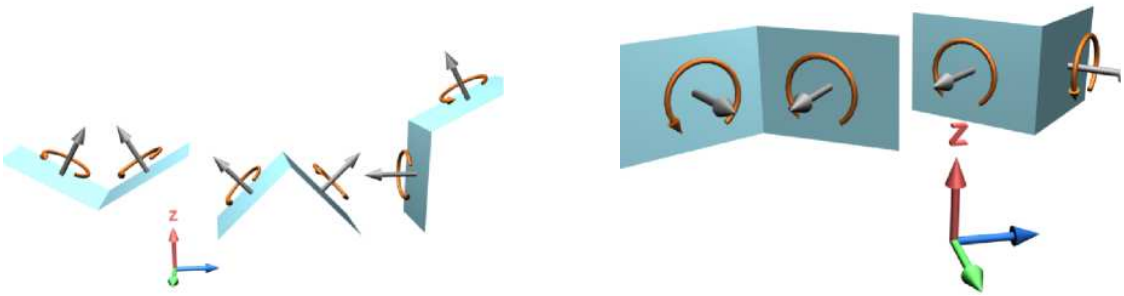


Figure 5.7 - Exemples de facettes non verticales et verticales cohérentes.

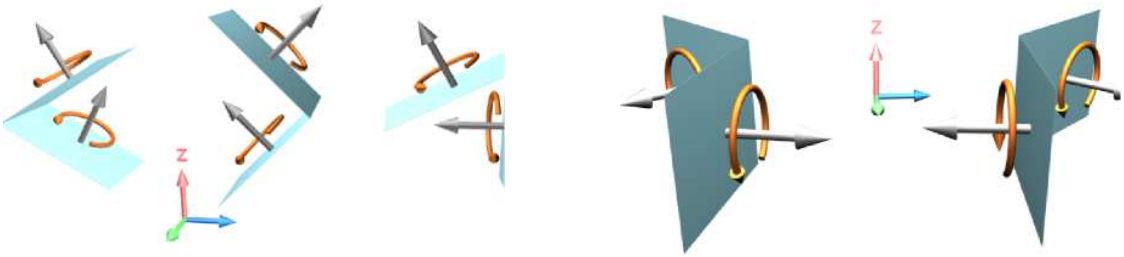


Figure 5.8 - Exemples de facettes non verticales et verticales incohérentes.

– tous les polytopes adjacents ont des orientations cohérentes.

Cette définition reprend la définition de triangulation orientable de [BY95] en l'étendant au cas des polytopes.
Le *bord* d'une surface polyédrique Σ est défini comme l'union des côtés des polytopes qui n'appartiennent qu'à un seul polytope. Il est noté $\partial\Sigma$.
Une surface polyédrique *fortement connexe* est une surface polyédrique Σ telle que, quels que soient deux points de Σ, il existe un arc sur Σ les reliant, en ne traversant que des arêtes de Σ (et pas de sommets, à la différence de simplement connexe). On exclut ainsi les configurations avec sommet singulier [BY95].
Un polyèdre est une surface polyédrique fortement connexe sans bord ($\partial\Sigma = \emptyset$).

Admissibilité et compatibilité

Une *surface admissible* est une surface polyédrique fortement connexe dont le bord recouvre les arêtes de la face-sol et dont chaque polytope est une facette élémentaire de l'arrangement de plans.
Une arête est localement admissible si elle appartient au bord de la face-sol ou si il existe au moins deux facettes aux orientations cohérentes ayant cette arête en commun.
Une facette est localement admissible si elle est localement admissible autour de chacune de ses arêtes bordantes.
Une facette est admissible si au moins une surface admissible la contient (une facette admissible est nécessairement localement admissible).
Un graphe 3D est dit localement admissible si toutes ses facettes le sont.
Un graphe 3D est dit admissible si toutes ses facettes le sont.
Deux facettes sont *compatibles* si elles appartiennent à au moins une surface admissible commune. On notera $\mathcal{C}(f)$ l'ensemble des facettes compatibles avec la facette f et $\bar{\mathcal{C}}(f) = \mathcal{G}\backslash\mathcal{C}(f)$ les facettes incompatibles avec f.

5.2.3 Propriétés

Surfaces enveloppes

Dans cette section, on mentionne différentes propriétés des surfaces et graphes admissibles. La figure 5.12 présente en particulier les notions de surfaces enveloppes.
Dans la suite, on dira qu'une facette non verticale f réalise l'altitude maximale (resp. minimale) en un point $\mathbf{X}$ lorsque aucune facette n'intersecte la verticale passant par $\mathbf{X}$ au dessus (resp. en dessous) de $\mathbf{X}$. On a alors le théorème :

Théorème 1
Si $\mathcal{G}$ est un graphe localement admissible, toute facette élémentaire non verticale f réalisant l'altitude maximale (resp. minimale) en un point la réalise en tous ses points intérieurs. On dira alors par abus de langage que f réalise l'altitude maximale.

Preuve du théorème 2. Supposons le contraire et soit Z une zone intérieure de f ne réalisant pas l'altitude maximale et $\mathbf{a}$ une arête de Z. Cette arête existe car, par hypothèse, Z ne couvre pas toute la facette élémentaire f. Par définition, $\mathbf{a}$ est la projection d'une arête $\mathbf{A}$ *au dessus* (ou en dessous) de f. $\mathbf{A}$ étant localement admissible, soit la surface verticale attenante à $\mathbf{A}$ est localement inadmissible, soit elle intersecte f (facette élémentaire) ce qui est impossible dans les deux cas. ■

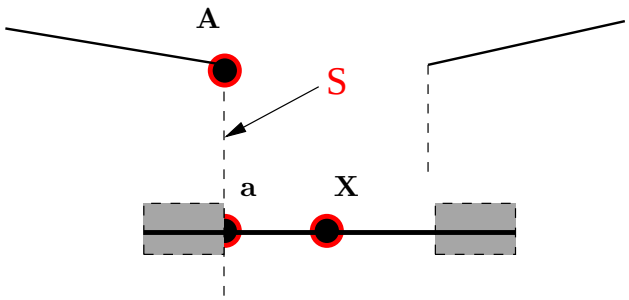


Figure 5.9 - Une facette réalise l'altitude maximale sur tous ses points intérieurs. La zone grisée représente Z dans la démonstration.

Théorème 2
Si $\mathcal{G}$ est un graphe 3D localement admissible, son emprise planimétrique est celle de $\mathcal{F}_s$.

Preuve du théorème 2. Supposons le contraire et soit $\mathbf{a}$ une arête 2D dans le plan de $\mathcal{F}_s$ limitant cette emprise. Soit $\mathbf{c}$ le milieu de $\mathbf{a}$. On définit par $\mathbf{e}$ le vecteur unitaire, orthogonal à $\mathbf{a}$, dans le plan de $\mathcal{F}_s$, pointant à partir de $\mathbf{c}$ vers l'extérieur de l'emprise c'est-à-dire vers la partie que l'emprise de $\mathcal{G}$ ne recouvre pas. Soit $\mathcal{V}_\mathbf{a}$ l'ensemble des facettes verticales qui se projettent partiellement ou entièrement sur $\mathbf{a}$.
Dans le cas où $\mathcal{V}_\mathbf{a} = \emptyset$, $\mathbf{a}$ est le projeté d'un côté d'une facette non verticale (on dira porteuse) f. De fait, il est évident que f n'est pas localement admissible ce qui est contradictoire. On supposera donc $\mathcal{V}_\mathbf{a} \neq \emptyset$. Soit $P_\mathbf{a}$ le plan porteur de ces facettes, $\mathbf{n}$ son orientation. 2 cas récapitulés dans la figure 5.10 sont envisageables :
$\mathbf{e} \cdot \mathbf{n} < 0$. Considérons la facette f_h de $\mathcal{V}_\mathbf{a}$ possédant l'arête bordante la plus haute. f_h est localement inadmissible.
$\mathbf{e} \cdot \mathbf{n} > 0$. À l'inverse, soit f_b la facette de $\mathcal{V}_\mathbf{a}$ possédant l'arête bordante la plus basse. f_b est localement inadmissible.
Dans les 2 cas, on contredit l'hypothèse initiale. L'emprise planimétrique est donc celle de la face-sol. ■

Théorème 3
Soit $\mathcal{G}$ un graphe 3D localement admissible. Soit $\mathcal{P}^+$ (respectivement $\mathcal{P}^-$) l'ensemble des facettes non *verticales réalisant les altitudes maximales (resp. minimales) sur l'emprise planimétrique. Il existe une unique surface admissible $\mathcal{Z}^+$ (resp. $\mathcal{Z}^-$) passant par $\mathcal{P}^+$ (resp. $\mathcal{P}^-$). $\mathcal{Z}^+$ et $\mathcal{Z}^-$ sont dénotées surfaces enveloppes.*

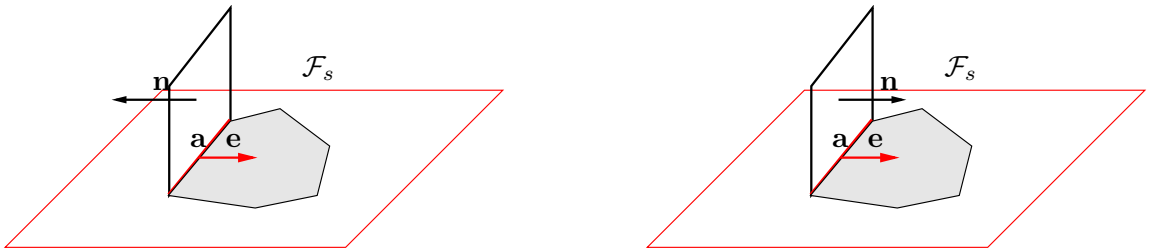


Figure 5.10 - Deux configurations possibles.

Preuve du théorème 3. Étant donné que les facettes non verticales sont toutes orientées vers le haut, il est immédiat de voir que toute surface continue et connexe de $\mathcal{P}^+$ est une surface polyédrique (de fait $\mathcal{P}^+$ est également une surface polyédrique). Par admissibilité locale du graphe, l'ensemble de ces surfaces sont reliées entre elles par un ensemble de facettes verticales $\mathcal{V}^+$ qui forment une surface polyédrique et sont localement admissibles avec les surfaces continues et connexes de $\mathcal{P}^+$. $\mathcal{P}^+$ et $\mathcal{V}^+$ forment donc une unique surface polyédrique continue dont l'emprise planimétrique est celle de la face-sol. Cette surface $\mathcal{Z}^+$ est donc admissible. ■

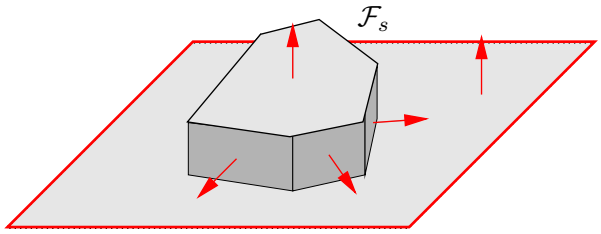


Figure 5.11 - En grisé clair apparaît $\mathcal{P}^+$. En grisé plus foncé, $\mathcal{V}^+$. L'ensemble constitue $\mathcal{Z}^+$.

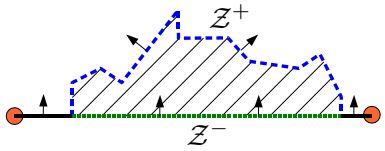


Figure 5.12 - Graphe 3D issu de l'arrangement de plans et localement admissible. Les surfaces enveloppes sont respectivement en bleu/tiret et en vert/point. Les parties attenantes aux bords de la face-sol, en noir appartiennent aux deux surfaces. Le graphe 3D est situé entre les 2 enveloppes.

Théorème 4

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

Preuve du théorème 4. Soit $\partial\Sigma$ le bord de la surface Σ qui, par définition, s'appuie sur le bord de $\mathcal{F}_s$. On considère le polyèdre construit en rajoutant des "parois latérales" et un "plancher" correctement orientés comme indiqué sur la figure 5.13. La surface admissible étant une surface polyédrique fortement connexe, il est aisé de voir que la surface ainsi générée est un polyèdre. D'après [BY95], un polyèdre est homéomorphe à une sphère avec anses. Dans notre cas, l'absence de polytope orienté "vers le bas" interdit la constitution d'une structure homéomorphe à une anse. Le polyèdre est donc homéomorphe à une sphère topologique et la surface Σ n'a ni anse ni surplomb (pas de facettes orientées vers le bas). ■

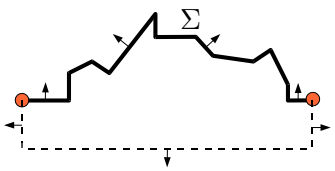


Figure 5.13 - La surface Σ étudiée dans le théorème 4 est représentée en noir trait fort, les parties en pointillés représentent les "parois latérales" et le "plancher" virtuels utilisés pour la démonstration du théorème.

5.2.4 Simplifications du graphe

Algorithme récursif

Considérons la procédure de suppression récursive de facettes (Algorithme 2) qui utilise la formule donnée dans le lemme 1 pour déterminer si une facette est localement inadmissible. Elle renvoie l'ensemble des facettes qui récursivement deviennent localement inadmissibles du fait de la suppression initiale de f. À l'issue de cet algorithme, il est clair que, si initialement le graphe est localement admissible, le graphe résultant est localement admissible.

Algorithme 2: Algorithme de suppression récursive.

Entrée : facette f, graphe 3D $\mathcal{G}$
Sortie : $\mathrm{RecSup}(f, \mathcal{G})$: ensemble des facettes devenues localement inadmissibles par suppression de f
début
  $\mathcal{R}(f) = f$
  Supprimer f
  Soit $\mathcal{A}_f$ les facettes adjacentes à f
  pour chaque *facette* $f' \in \mathcal{A}_f$, f' *localement inadmissible* **faire** $\mathcal{R}(f) = \mathcal{R}(f) \cup \mathrm{RecSup}(f', \mathcal{G})$
  retourner $\mathcal{R}(f)$
fin

Suppression de facettes initiales

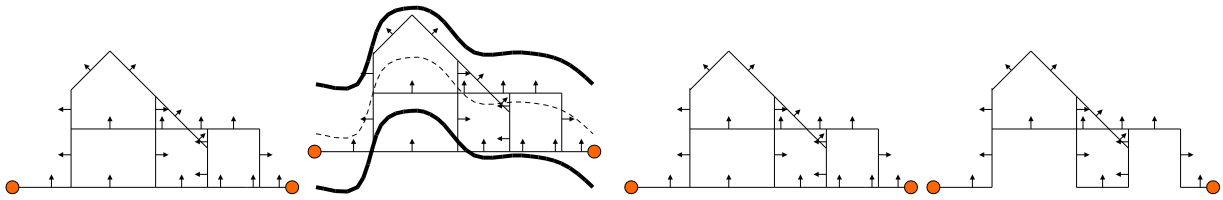
Dans un premier temps, en utilisant l'Algorithme 2, on supprime récursivement toutes les facettes qui, dans le graphe 3D, sont localement inadmissibles. Il s'agit des facettes touchant les bords de la forme prismatique dans laquelle est confiné le graphe et qui ne touchent pas (aucune de leurs arêtes n'est incluse dans) la face-sol $\mathcal{F}_s$. Il en résulte un graphe localement admissible. On supprime ensuite récursivement les facettes non verticales attenantes au sol (figure 5.15(a)). ***On remarquera alors que, d'après le théorème 4, les surfaces admissibles du graphe 3D ainsi élagué sont telles que, sur toute surface admissible, tout ensemble connexe de facettes ne possédant aucune facette incluse dans $\mathcal{F}_s$ est un bâtiment***. La recherche des surfaces admissibles permet, de fait, de récupérer toutes les hypothèses de bâtiment sur la zone de focalisation et, le cas échéant, de *pouvoir modéliser plusieurs bâtiments sur une seule zone*. On peut ainsi pallier les erreurs de focalisation.

Bandeau de tolérance autour du MNE

Avant de rechercher de façon explicite les surfaces admissibles, il est indispensable de réduire le nombre des facettes du graphe. La recherche des solutions est en effet un problème NP-complet en fonction du nombre de facettes du graphe. Pour réduire ce nombre, on supprime toute facette non incluse dans $\mathcal{F}_s$ qui n'est pas entièrement contenue dans une "nappe" autour du MNE [TD02b] (figures 5.14(b), 5.15(b) et 5.15(c)).

Zone de tolérance autour du masque

De la même façon, on supprime de façon récursive toute facette attenante à la face-sol qui touche le masque érodé et toute facette verticale qui se trouve hors du masque dilaté (figures 5.15(d) et 5.14(c)).



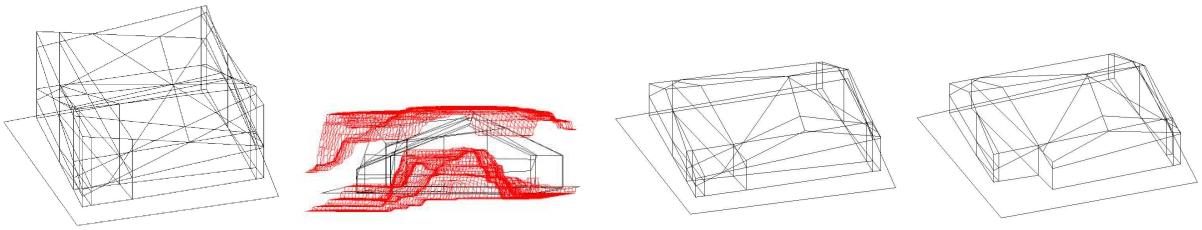
(a) Après simplification initiale

(b) Avec la "nappe" issue du MNE

(c) Avec le masque

(d) Après simplifications successives

Figure 5.14 - Différentes étapes de simplification - vue schématique.



(a) Après simplification initiale

(b) Avec la "nappe" issue du MNE

(c) Après simplification du MNE

(d) Après simplification du masque

Figure 5.15 - Différentes étapes de simplification - vue 3D.

5.2.5 Recherche des solutions

Comme montré auparavant, pour déterminer les hypothèses de bâtiments, il faut rechercher les surfaces admissibles du graphe élagué. Nous allons maintenant détailler la technique de recherche de ces surfaces. On utilise ici le principe de recherche des cliques maximales dans un graphe de compatibilité [Jib02].

Calcul du graphe de compatibilité

Le calcul des facettes compatibles reprend l'idée de [Jib02] en l'étendant au cas des facettes verticales. Ce calcul se fait en deux temps à partir d'un graphe 3D $\mathcal{G}$ (figure 5.16 et algorithme 3) :
Pour toute facette non verticale f, on supprime récursivement et virtuellement les facettes non verticales dont le projeté sur le plan porteur de f est d'intersection non vide avec f. D'après le théorème 4, ces facettes ne sont pas compatibles avec f. Si le graphe virtuel $\tilde{\mathcal{G}}$ (localement admissible) résultant est non vide (ou ce qui est équivalent si f n'est pas "incompatible" avec elle-même), toute facette *admissible* du graphe virtuel est compatible avec f qui est nécessairement admissible puisque les surfaces enveloppes de $\tilde{\mathcal{G}}$ sont des surfaces admissibles et passent par f par construction (théorème 5). Dans le cas contraire, f n'est pas admissible et est supprimée récursivement du graphe $\mathcal{G}$.
Ensuite, on utilise la symétrie de la relation d'incompatibilité et pour toute facette verticale v, on supprime virtuellement toute facette non verticale incompatible avec v déterminée d'après le calcul précédent. On peut là aussi montrer que si le graphe résultant est non vide, v est admissible et toute facette *admissible* du graphe est compatible avec v.

Algorithme 3: Calcul des incompatibles d'une facette.

Entrée : Un Graphe 3D : $\mathcal{G}$
Sortie : Pour chaque facette, l'ensemble de ses incompatibles $\bar{\mathcal{C}}(f)$
début
  Soit $\mathcal{F}_{nv}$ (resp $\mathcal{F}_v$) les facettes non verticales (resp verticales) de $\mathcal{G}$
  Pour $f \in \mathcal{F}_{nv}$, on note $\mathcal{V}(f)$ les facettes de $\mathcal{F}_{nv}$ dont le projeté sur le plan porteur de f est d'intersection non vide f
  pour chaque $f \in \mathcal{F}_{nv}$ **faire**
    $\tilde{\mathcal{G}} = \mathcal{G}$
    $\forall f' \in \mathcal{V}(f), \bar{\mathcal{C}}(f) = \bar{\mathcal{C}}(f) \cup \mathrm{RecSup}(f', \tilde{\mathcal{G}})$
  fin
  pour chaque $f \in \mathcal{F}_v$ **faire**
    $\tilde{\mathcal{G}} = \mathcal{G}$
    $\bar{\mathcal{C}}(f) = \{f' \in \mathcal{F}_{nv} / f \in \bar{\mathcal{C}}(f')\}$
    $\forall f' \in \bar{\mathcal{C}}(f), \bar{\mathcal{C}}(f) = \bar{\mathcal{C}}(f) \cup \mathrm{RecSup}(f', \tilde{\mathcal{G}})$
  fin
fin

Théorème 5
À l'issue de l'algorithme 3, pour toute facette f, si le graphe virtuel résultant $\tilde{\mathcal{G}}$, localement admissible, est non vide (ou ce qui est équivalent si f n'est pas "incompatible" avec elle-même), toute facette admissible *du graphe est compatible avec f qui est admissible.*

Preuve du théorème 5. Montrons que f est admissible. Pour une facette non verticale, par construction, f appartient aux deux surfaces enveloppes qui sont admissibles donc f est admissible et on a $\mathcal{Z}^+_{\tilde{\mathcal{G}}} \leq f \leq \mathcal{Z}^-_{\tilde{\mathcal{G}}}$[2]. Pour une facette verticale, on peut montrer, en utilisant un raisonnement analogue à la démonstration du théorème 2 que, là encore, f appartient aux deux surfaces enveloppes ce qui amène $\mathcal{Z}^+_{\tilde{\mathcal{G}}} \leq f \leq \mathcal{Z}^+_{\tilde{\mathcal{G}}}$.
Soit f' facette admissible du graphe $\tilde{\mathcal{G}}$. Par construction $\mathcal{Z}^-_{\tilde{\mathcal{G}}} \leq f' \leq \mathcal{Z}^+_{\tilde{\mathcal{G}}}$. Soit $\mathcal{S}$ surface admissible passant par f'. Considérons le graphe $\mathcal{G}_1$ constitué des facettes de $\mathcal{Z}^+_{\tilde{\mathcal{G}}}$ et de $\mathcal{S}$. $\mathcal{G}_1$ est admissible donc localement admissible donc sa surface enveloppe inférieure $\mathcal{Z}^-_{\mathcal{G}_1}$ est admissible. Par construction $\mathcal{Z}^-_{\mathcal{G}_1} \leq \mathcal{Z}^+_{\tilde{\mathcal{G}}}$ et $f' \in \mathcal{Z}^-_{\mathcal{G}_1}$. Considérons le graphe $\mathcal{G}_2$ constitué des facettes de $\mathcal{Z}^-_{\tilde{\mathcal{G}}}$ et de $\mathcal{Z}^-_{\mathcal{G}_1}$. $\mathcal{G}_2$ est admissible donc localement admissible donc sa surface enveloppe supérieure $\mathcal{Z}^+_{\mathcal{G}_2}$ est admissible. On obtient donc d'après ce qu'on vient de montrer $\mathcal{Z}^-_{\tilde{\mathcal{G}}} \leq \mathcal{Z}^+_{\mathcal{G}_2} \leq \mathcal{Z}^+_{\tilde{\mathcal{G}}}$ donc on obtient par encadrement que $f \in \mathcal{Z}^+_{\mathcal{G}_2}$. Comme par hypothèse $f' \in \mathcal{Z}^+_{\mathcal{G}_2}$, $\mathcal{Z}^+_{\mathcal{G}_2}$ est une surface admissible passant par f et f' qui sont donc compatibles. ■

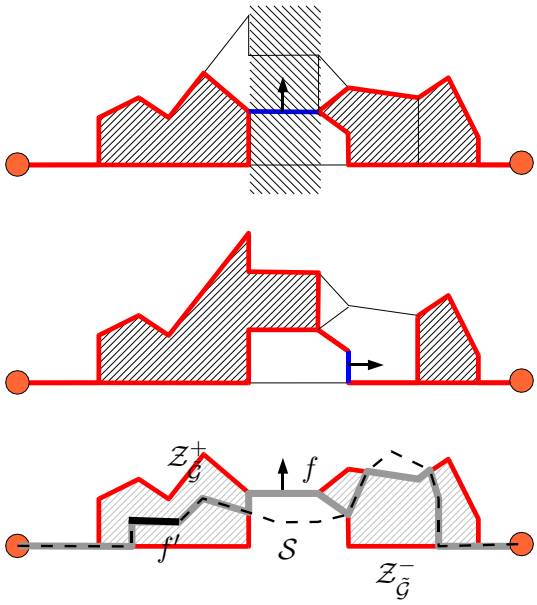


Figure 5.16 - Calcul des compatibles. Les deux premières figures illustrent les calculs des facettes compatibles pour une facette non verticale et verticale. Toutes les facettes admissibles contenues dans les zones hachurées sont compatibles avec la facette en bleu. Les traits gras définissent les enveloppes des graphes virtuels (voir texte pour détails). La dernière figure illustre la démonstration du théorème 5.

Graphe de compatibilité

À l'issue du processus, on supprime dans $\mathcal{G}$ toute facette non admissible et l'ensemble des facettes restantes dans $\mathcal{G}$ est donc admissible. De plus, on a déterminé les couples de facettes compatibles. On peut donc construire le graphe de compatibilité $\mathcal{G}_c$ dans lequel chaque noeud est une facette admissible et chaque arête entre 2 noeuds représente la relation de compatibilité. En suivant le raisonnement de [Jib02], on peut alors démontrer (théorème 6) que *les surfaces admissibles du graphe sont les cliques maximales du graphe de compatibilité* ou, ce qui est équivalent, que les surfaces admissibles sont les ensembles stables maximaux du graphe d'incompatibilité $\bar{\mathcal{G}}_c$ (figures 5.17 et 5.18). Cette recherche de cliques maximales est un problème NP-complet (la figure 5.20) montre l'accroissement exponentiel du nombre de solutions en fonction du nombre de facettes dans le graphe 3D). On peut montrer que par rapport à l'algorithme utilisé par [Jib02], l'algorithme de Bron-Kerbosch optimisé [BK73] permet une étude beaucoup plus approfondie dans le cas de nos graphes où la densité est relativement élevée (figure 5.21).

[2]La notation $f \leq \mathcal{Z}$ est ici évidente et traduit le fait qu'une facette se trouve en dessous de la surface $\mathcal{Z}$.

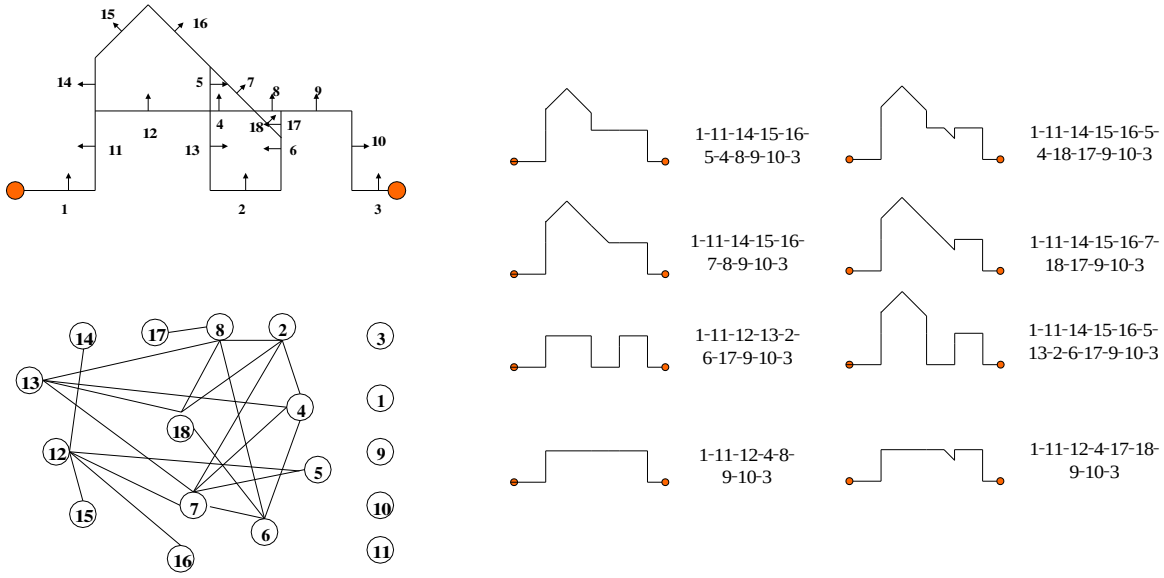


Figure 5.17 - Exemple de Graphe élagué et de graphe d'incompatibilités $\bar{\mathcal{G}}_c$ ainsi que les solutions sous forme de surface admissible et sous forme d'ensemble stable maximal.

Théorème 6

Les cliques maximales du graphe de compatibilité sont les surfaces admissibles du graphe 3D.

Preuve du théorème 6. Il est évident de voir que toute surface admissible est une clique maximale. C'est une clique car toutes les facettes sont compatibles deux à deux par définition de la compatibilité et de la surface admissible. Elle est maximale car la surface admissible couvre toute l'emprise. Le rajout d'une facette créera donc nécessairement une incompatibilité.
Réciproquement, toute clique est une surface polyédrique passant par des facettes de l'arrangement, en particulier une clique maximale. Soit $\mathcal{K} = (f_1, f_2, \cdots, f_n)$ une clique maximale. Par récurrence, et en utilisant le même raisonnement que pour le théorème 5, on montre que toute clique $(f_1, \cdots, f_n)$ appartient à une surface admissible (figure 5.19). Pour $\mathcal{K}$, on est donc assuré qu'il existe une surface admissible $\mathcal{S}$ passant par toutes les facettes de $\mathcal{K}$. Mais d'après ce qu'on a vu, une surface admissible est une clique maximale donc $\mathcal{K} = \mathcal{S}$ [3]sinon on contredit la maximalité de $\mathcal{K}$. ■

5.3 Choix de la meilleure représentation

5.3.1 Modélisation bayésienne

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

$$\hat{M} = \underset{M \in \mathcal{M}}{\operatorname{argmax}} P(M \mid \mathcal{D}) \tag{5.1}$$

soit en utilisant le théorème de Bayes

$$P(M \mid \mathcal{D}) = \frac{P(\mathcal{D} \mid M) \cdot P(M)}{P(\mathcal{D})} \tag{5.2}$$

[3]cette notation est abusive car la clique fait référence à des noeuds dans le graphe de compatibilité et la surface admissible à une surface dans le graphe 3D. Toutefois, la correspondance bijective entre les deux espaces rend l'égalité relativement intuitive.

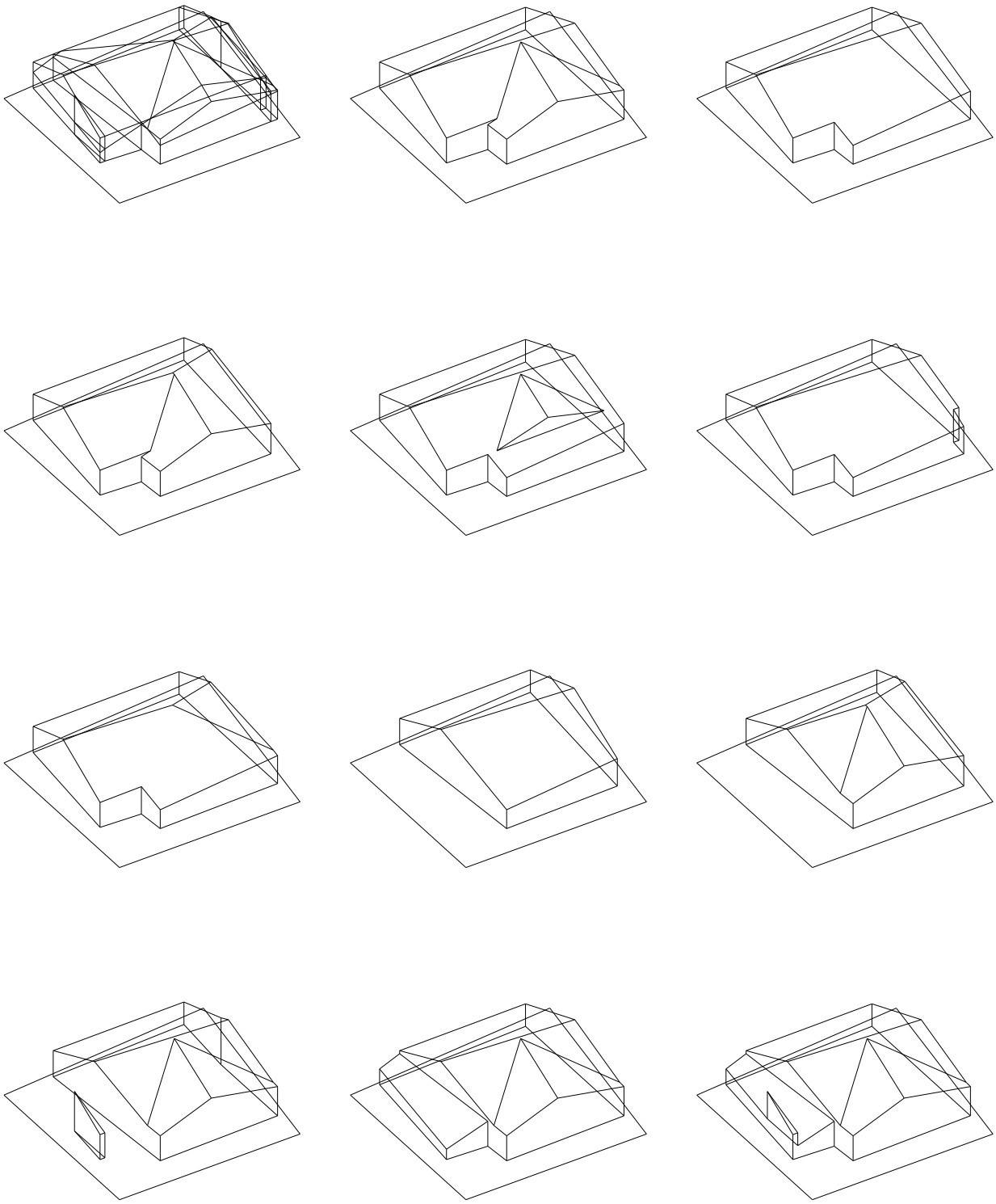


Figure 5.18 - Le graphe après simplifications suivi de 11 solutions sur les 3646 de l'exemple présenté.

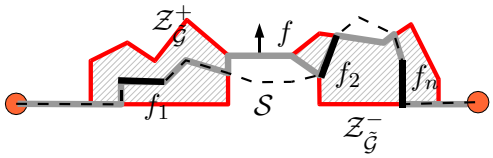


Figure 5.19 - Principe de la démonstration de l'équivalence entre clique maximale et surface admissible.

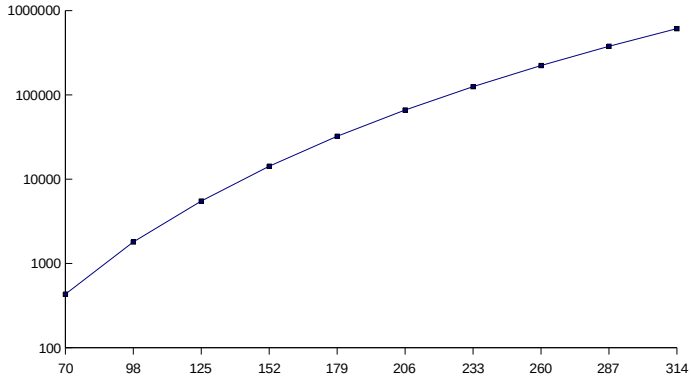


Figure 5.20 - Évolution du nombre de cliques maximales en fonction du nombre de noeuds. L'échelle des ordonnées est logarithmique.

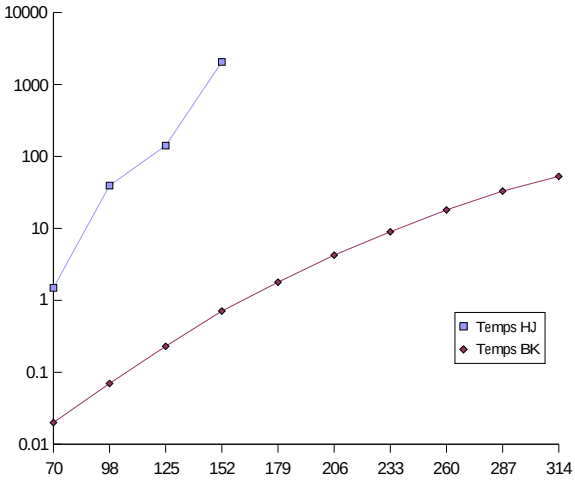


Figure 5.21 - Comparaison du temps de recherche des cliques maximales par les algorithmes utilisés par [Jib02] et celui de [BK73]. L'échelle des ordonnées est logarithmique en seconde. L'axe des abscisses représente le nombre de noeuds du graphe de compatibilité.

$P(\mathcal{D})$ ne dépendant pas du modèle, on est ramené à la recherche de $\hat{M} = \text{argmax}_{M \in \mathcal{M}} P(\mathcal{D} \mid M) \cdot P(M)$. Tout type d'observation peut être géré par cette modélisation qui permet donc d'intégrer des données hétérogènes. La difficulté est, bien entendu, de déterminer pour chaque type d'observation la fonction de probabilité et de régler les fonctions les unes par rapport aux autres de façon à ce qu'elles aient le même poids. Dans la suite, nous utiliserons comme observations : les images I, les segments 3D S et le masque de

focalisation M_a. En supposant les observations indépendantes, on obtient finalement :

$$\hat{M} = \underset{M \in \mathcal{M}}{\operatorname{argmax}} P(S \mid M) \cdot P(I \mid M) \cdot P(M_a \mid M) \cdot P(M) \tag{5.3}$$

La probabilité du modèle dépend de la complexité du modèle de bâtiment. A priori, on cherche en effet à obtenir la "modélisation" la plus simple possible compte tenu des observations. S'inspirant de la relation de Shannon [KDN+95], on relie $P(M)$ à la longueur minimale de description par la formule 5.4 :

$$P(M) = C \exp^{\frac{-\mathcal{L}(M)}{\beta}} \tag{5.4}$$

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

5.3.2 Modèle

Inférence des contraintes

Comme dit auparavant, les contraintes sur les primitives influent sur le choix du modèle, permettant ainsi une intégration d'une forme de régularisation du problème. Les contraintes sont déterminées sur les orientations des primitives de base : plans et façades. On suit ici la logique du système à base de contraintes développé par [Gro02] car c'est celui qui est adapté à notre problématique pour le recalage sous contraintes. Cette description n'est qu'un bref résumé de la technique d'inférence de contraintes développée en section 6.2.3. Nous en rappelons ici brièvement les principes pour assurer la continuité de la lecture. L'algorithme d'inférence de contraintes utilise deux paramètres angulaires $\sigma_{/\!/}$ et $\sigma_{\sphericalangle}$. Les normales des primitives planes sont tout d'abord regroupées par un algorithme itératif de clustering qui regroupe les normales dont l'écart angulaire est faible. À l'issue du processus, pour chaque groupe de normales, on définit une *direction*, moyenne des normales de chaque groupe. L'algorithme de clustering assure que la distance angulaire minimale entre ces directions est supérieure à $\sigma_{/\!/}$ et, par le regroupement des normales, permet de gérer le parallélisme. On définit ensuite sur ces directions un graphe de contraintes. Les noeuds du graphe sont les directions (on distingue 2 types de noeuds : façades et plans), les arêtes matérialisent différentes relations (figure 5.22) : orthogonalité, horizontalité du produit vectoriel (donc horizontalité de l'arête intersection des 2 plans, relation très importante en milieu urbain), symétrie verticale. Chaque arête est créée si la relation est vérifiée dans la tolérance angulaire $\sigma_{\sphericalangle}$. Chaque arête q est affectée d'un poids $C(q)$ donnant le nombre de degré de libertés que la contrainte supprime (exemple : 1 pour la contrainte d'orthogonalité, 2 pour la symétrie verticale). On notera que l'orthogonalité entre la direction verticale et les façades n'est pas testée : elle est implicite.

Longueur de description

On notera F_M, E_M et V_M le nombre de faces, d'arêtes et de points d'un modèle. On notera P_M et Fa_M le nombre de plans et de façades utilisés par le modèle, $D_M = P_M + Fa_M$ le nombre de directions générales et C_M le nombre de contraintes d'un modèle. La logique utilisée ici s'inspire des travaux de [Kol99]. On utilise le principe de transmission d'informations liées au modèle : topologique, géométrique et contraintes. À noter que la longueur de description initiale qui consiste à fournir (sur des entiers) le nombre maximum de faces, d'arêtes, de points, de façades, de plans et de contraintes du modèle est donnée par :

$$\mathcal{L}_g(f) = \log(\mid V \mid_{\max}) + \log(\mid F \mid_{\max}) + \ldots$$

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

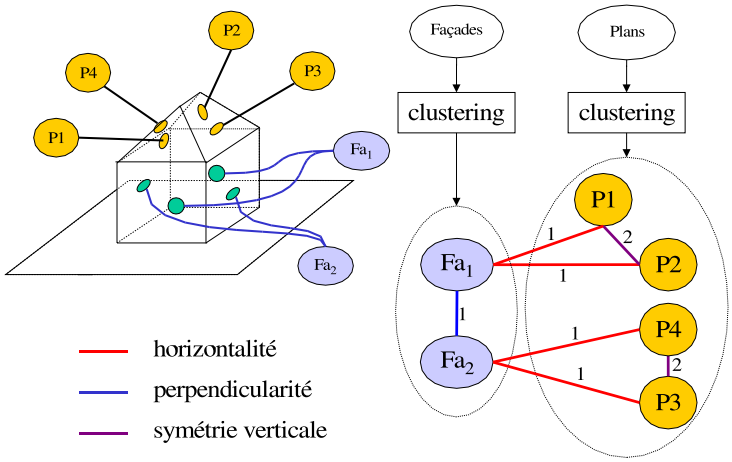


Figure 5.22 - Schéma d'inférence des contraintes et représentation du graphe de contraintes. Les façades parallèles sont regroupées par le clustering en deux directions Fa_1 et Fa_2. Les valuations sur les arêtes dépendent du type de contrainte.

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

$$\mathcal{L}(M) = \mathcal{L}_t(M) + \mathcal{L}_g(M) + \mathcal{L}_c(M) \tag{5.5}$$

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

$$\mathcal{L}_t(f) = \mid f \mid \log(V_M) + \log(D_M) \tag{5.6}$$

d'où en sommant sur l'ensemble des facettes et en tenant compte du fait que le nombre de points est égal au nombre d'arêtes pour une face fermée :

$$\mathcal{L}_t(M) = 2 \cdot E_M \log(V_M) + F_M \cdot \log(D_M) \tag{5.7}$$

géométrie La description géométrique utilise un modèle simple d'énumération des coordonnées des points et des directions. Chaque coordonnée est codée sur 12 bits ce qui fournit une précision d'un centimètre pour un point, en supposant un bâtiment d'environ 40 mètres, et une précision angulaire de $360/4092$ degré, ce qui est amplement suffisant. Chaque façade équivaut à 1 inconnue, chaque plan non vertical à 2 et chaque point à 3 inconnues. Chaque facette apporte implicitement $\mid f \mid -1$ équations de planarité du type $\mathbf{D_M}^T \cdot (\mathbf{V_i} - \mathbf{V_j}) = 0$. Chaque équation est supposée réduire d'un le nombre de degrés de liberté (pas de test de non indépendance des équations pour des raisons évidentes de temps de calcul) ce qui amène :

$$\begin{aligned} \mathcal{L}_g(M) &= (2 * P_M + Fa_M + 3V_M - \sum_f (\mid f \mid -1)) * 12 \\ &= (2 * P_M + Fa_M + 3V_M - 2E_M + F_M) * 12 \end{aligned} \tag{5.8}$$

La longueur géométrique est ainsi intimement liée au nombre de degrés de liberté du modèle ; propriété importante utilisée pour favoriser les modèles possédant un grand nombre de contraintes internes.

contraintes Chaque modèle ou surface admissible utilise un certain nombre de directions et définit donc un sous-graphe dans le graphe complet de contraintes. Toute arête présente dans ce sous-graphe représente une contrainte du modèle. Chaque contrainte est codée par : son type (orthogonalité, horizontalité ou symétrie verticale), les directions qu'elle met en relation et chaque contrainte réduit d'un nombre $C(q)$ lié au type c de la contrainte q le nombre de degrés de liberté du problème. On notera $\mid c \mid$ le nombre de types de contraintes (3 dans notre cas) et on suppose en outre dans l'évaluation du coût du modèle que les contraintes sont indépendantes entre elles, hypothèse forte, mais limitée par le fait que les normales sont déjà regroupées en directions. On obtient finalement :

$$\mathcal{L}_c(M) = \sum_q \log(\mid c \mid) + 2\log(D_M) - C(q) * 12 \tag{5.9}$$

L'équation 5.9 traduit bien l'importance accordée aux contraintes. Plus un modèle possède de contraintes internes, plus sa fonction de coût liée à la longueur de description du modèle diminue et donc plus la probabilité que ce modèle soit le "bon modèle" augmente.

5.3.3 Observations

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

Images

Pour calculer le score du modèle en fonction des images, l'idéal serait de calculer le score de corrélation des orthoimages correspondant au modèle de surface de toit, qui donne les meilleurs résultats. Toutefois, [Jib02] montre que le gain par rapport à la moyenne des scores de corrélation de l'ensemble des orthoimagettes horizontales régulièrement espacées et d'altitude donnée par le modèle de toit du bâtiment est marginal. En outre, cette dernière méthode présente d'indéniables avantages en termes de temps de calcul : le volume de corrélation peut-être précalculé avant la recherche des modèles (figure 5.23). L'évaluation de chaque modèle devient alors beaucoup plus rapide et l'on évite ainsi les étapes de projection des images sur le modèle. Nous avons donc choisi cette solution. Dans le cas du multi-images (n images), le score de corrélation utilisé est donné par l'équation 5.10 dans laquelle chaque fenêtre de corrélation de taille (f_u, f_v) centrée en (u, v) dans une image I_i est représenté par le vecteur $\mathbf{v_i}(u, v)$ et où (u_i, v_i) est le projeté du point (x, y, z) via la matrice de projection *connue* correspondante de I_i :

$$C(x, y, z) = \frac{\text{Var}\left(\sum_i \mathbf{v_i}(u_i, v_i)\right)}{n \cdot \sum_i \left(\text{Var}\left(\mathbf{v_i}(u_i, v_i)\right)\right)} \quad \in [0..1] \tag{5.10}$$

Il s'agit du score multi-images symétrique utilisé par [PTJ00] normalisé par le nombre d'images. La moyenne est calculée sur la surface planimétrique $\mid SP(M) \mid$ de l'ensemble $\mathcal{H}(M)$ des facettes non verticales du modèle qui n'appartiennent pas à la face-sol :

$$P(I \mid M) = \sum_{f \in \mathcal{H}(M)} \frac{C(x, y, M(x, y))}{\mid SP(M) \mid} \tag{5.11}$$

Segments 3D

Les segments 3D apportent une information de structure très forte, il semble donc logique de chercher à les apparier avec les arêtes du modèle. Pour chaque segment $\mathbf{s}$, on détermine les arêtes appariées, une arête $\mathbf{a}$

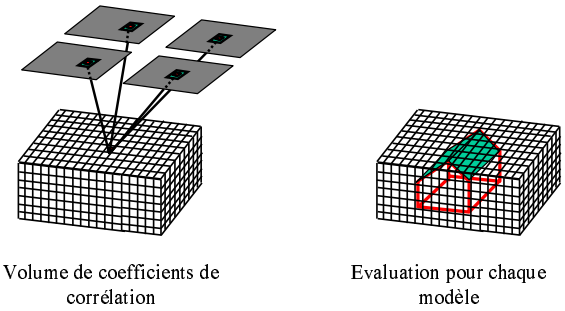


Figure 5.23 - Principe du calcul de $P(I/M)$.

du modèle est appariée avec $\mathbf{s}$ si et seulement si l'écart angulaire est inférieur à un paramètre θ_s et si l'écart en distance est inférieur à un autre paramètre d_s : on notera alors $\mathbf{a}\mathcal{R}\mathbf{s}$. Pour chaque arête appariée, on définit alors la longueur de recouvrement de $\mathbf{s}$ par $\mathbf{a}$: $r(\mathbf{a},\mathbf{s})$. L"'explication" du segment $\mathbf{s}$ *apparié* par le modèle M se traduit donc par :

$$P(\mathbf{s} \mid M) = \min\left(\sum_{\mathbf{a}\mathcal{R}\mathbf{s}} \frac{r(\mathbf{a},\mathbf{s})}{\|\mathbf{s}\|}, 1\right) \tag{5.12}$$

Pour prendre en compte les erreurs du détecteur de segments [Sch99], on affecte une valeur ϵ_s par défaut à un segment non apparié (que l'on suppose être, de fait, une surdétection). En pratique, nous avons choisi $\epsilon_s = 0,3$. L'équation précédente devient alors

$$P(\mathbf{s} \mid M) = \max\left(\min\left(\sum_{\mathbf{a}\mathcal{R}\mathbf{s}} \frac{r(\mathbf{a},\mathbf{s})}{\|\mathbf{s}\|}, 1\right), \epsilon_s\right) \tag{5.13}$$

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

$$P(S \mid M) = \prod_{\mathbf{s}\in S} P(\mathbf{s} \mid M) \tag{5.14}$$

Masque

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

$$P(M_a \mid M) = \frac{\mid SP(M) \cap M_a \mid}{\mid SP(M) \cup M_a \mid} \tag{5.15}$$

5.4 Filtrages et extraction des bâtiments

5.4.1 Filtrages

Dans la section précédente, la formulation bayésienne permettant d'évaluer chaque modèle et de déterminer la meilleure représentation de la zone en intégrant l'adéquation aux données et la simplicité du modèle a été explicitée. Au préalable, un filtrage des solutions est effectué sur 2 critères : un critère de taille et un critère d'angle. Toutes les solutions comportant une face dont la taille est inférieure à 1 m^2 ou comportant une

face possédant un angle inférieur à 10° sont supprimées. Ces solutions sont en effet notoirement aberrantes compte tenu de la résolution des images et de la forme relativement régulière des bâtiments. Nous pensons que ces étapes réduisent le nombre de solutions sans affecter le résultat final.

5.4.2 Extraction des toits de bâtiments

D'après ce que l'on a vu, il est aisé, une fois choisie la meilleure représentation pour la zone étudiée, d'extraire de cette solution les ensembles connexes de facettes ne touchant pas le plan $z = z_g$. On détermine ainsi les structures de toit des bâtiments présents sur la zone. Soulignons encore une fois que plusieurs bâtiments peuvent être présents dans la zone de focalisation comme l'atteste la figure D.7.

5.4.3 Corrections topologiques

L'ensemble des formes potentielles est déterminé par intersections de plans. Or, au niveau de la détection de ces plans, aucune contrainte n'est imposée, il est donc illusoire d'espérer obtenir un point intersection de 4 plans dans le graphe 3D. A priori donc, en raison des primitives, l'algorithme ne fournira pas de bâtiments présentant de points intersections de 4 plans. Ces bâtiments sont nombreux et pour pallier ce problème, un simple algorithme de "snapping" est mis en oeuvre sur le meilleur modèle déterminé. cet algorithme qui utilise un paramètre $\gamma_s = 1,5m$ directement relié au niveau de généralisation souhaité permet de corriger ces incohérences topologiques en snappant les points proches de moins de γ_s d'une arête et en fusionnant les arêtes dont la longueur est inférieure à γ_s.

5.5 Optimisation de la recherche

5.5.1 Rendre le problème additif

Une des problématiques soulevées par l'énumération exhaustive de toutes les solutions est la combinatoire. Cette combinatoire découle du nombre exponentiel de cliques maximales dans un graphe général. En d'autres termes, le nombre de solutions proprement dit est exponentiel. Par conséquent, l'évaluation de chacune de ces solutions représente un temps de traitement dont il faut tenir compte. Avoir un algorithme efficace de recherche de cliques maximales tout en ayant un temps d'évaluation important pour le score associé à chaque solution pénaliserait grandement l'algorithme. Nous montrons ici que tous les termes employés peuvent se mettre sous forme additive par rapport aux facettes constitutives des surfaces admissibles. L'évaluation d'une solution se ramène donc à des sommes sur l'ensemble des facettes de la surface (ou sur le sous graphe de contraintes associé à la surface). Cet aspect est également important en vue d'heuristiques pour le choix du meilleur modèle sans passer par une exploration combinatoire.

Paramètres de complexité du modèle

Soit $\mathcal{F}_\mathcal{G}(M)$ (resp $\mathcal{A}_\mathcal{G}(M)$) l'ensemble des facettes (resp arêtes) du graphe 3D constituant la surface admissible. Le nombre de faces du modèle n'est pas nécessairement égal au nombre de facettes élémentaires du modèle. En effet, une face peut être constituée de plusieurs facettes élémentaires qui doivent être fusionnées car elles sont adjacentes et appartiennent au même plan porteur. Pour évaluer les paramètres liés à la topologie du modèle, il est indispensable de tenir compte de cette distinction.
En parcourant $\mathcal{A}_\mathcal{G}(M)$, on détermine les arêtes fictives du modèle, c'est-à-dire les arêtes partagées entre deux facettes de *même* plan support. On construit dans le même temps le graphe des facettes élémentaires séparées par des arêtes fictives. Il est aisé de voir que le nombre de composantes connexes de ce graphe (algorithme en $O(\mathcal{F}_\mathcal{G}(M))$) détermine alors F_M. Les sommets réels du modèle sont les sommets du graphe 3D appartenant au modèle et définis par l'intersection d'au moins 3 arêtes non fictives. D'après les résultats

précédents, on obtient immédiatement V_M. Le nombre d'arêtes réelles du modèle est alors défini comme la moitié du nombre d'arêtes non fictives issues des sommets réels (chaque arête est comptée deux fois dans cette procédure). On a alors E_M. Les paramètres P_M et Fa_M se calculent de façon évidente par parcours des facettes du modèle. De même, il est aisé de déterminer l'ensemble des contraintes du modèle qui représentent un sous-graphe du graphe global des contraintes sur l'ensemble des primitives planes. Nous avons ainsi montré que l'ensemble des paramètres liés à l'évaluation de la complexité du modèle peut s'obtenir en un $\mathcal{O}(\mathcal{F}_{\mathcal{G}}(M))$.

Évaluation des données

L'évaluation du modèle en fonction des images $P(I \mid M)$ est par définition additive sur l'ensemble des facettes en supposant que nous ayons précalculé le score de corrélation pour chaque facette élémentaire non verticale et n'appartenant pas à la face-sol.
Pour l'évaluation du modèle en fonction du masque, on peut calculer pour chaque facette élémentaire les zones projetées au sol dans le masque $M_a(f)$ et hors du masque $\bar{M_a}(f)$. On a alors :

$$SP(M) \cap M_a = \coprod_{f \in \mathcal{H}(M)} M_a(f) \tag{5.16}$$

$$SP(M) \cup M_a = \left(\coprod_{f \in \mathcal{H}(M)} \left(M_a(f) \coprod \bar{M_a}(f) \right) \right) \coprod_{f \notin \mathcal{H}(M)} M_a(f) \tag{5.17}$$

ce qui, en terme de surface se traduit par

$$\mid SP(M) \cap M_a \mid = \sum_{f \in \mathcal{H}(M)} \mid M_a(f) \mid \tag{5.18}$$

$$\mid SP(M) \cup M_a \mid = \sum_{f \in \mathcal{H}(M)} \left(\mid M_a(f) \mid + \mid \bar{M_a}(f) \mid \right) + \sum_{f \notin \mathcal{H}(M)} \mid M_a(f) \mid \tag{5.19}$$

où $\coprod$ désigne l'union d'ensembles disjoints. Les facettes "au sol" donc non dans $\mathcal{H}(M)$ permettent de répertorier ce qui est dans le masque sans être dans le modèle de bâtiment proprement dit. L'additivité des opérations vient ici de la distributivité des opérations $\cup$ et $\cap$ et du fait que les facettes ne se recouvrent pas en planimétrie.
Le score lié aux segments se fait en précalculant pour les arêtes du graphe 3D les appariements et les recouvrements avec les segments 3D. La sommation sur l'ensemble des arêtes non fictives du modèle fournit alors les recouvrements pour chaque segment 3D. Ces recouvrements sont alors ramenés à la longueur du segment 3D pour fournir un score entre 0 et 1 (éventuellement majoré). Dans ce cas, l'algorithme fait un seul parcours des arêtes fictives et un parcours de chaque segment 3D.

5.5.2 Heuristiques d'accélération de la recherche

La combinatoire imposée par la stratégie d'énumération nous oblige à rechercher des solutions approchées. Nous ne ferons ici qu'ébaucher les principes pour le choix de bonnes heuristiques : nous n'avons pas eu le temps d'explorer les différentes stratégies. L'intuition est qu'il faut choisir itérativement les "meilleures" facettes au sens d'un poids relié à la fonction d'évaluation globale du modèle, mais il faut éviter des solutions aux formes trop identiques. De cette façon, le choix entre les modèles sera moins ambigu. De plus, dans le cas d'une interaction avec un opérateur, il est nécessaire de proposer des formes résolument distinctes et éviter d'avoir des modèles qui ne différent que par un léger changement de topologie. De fait, les stratégies de "Reactive Local Search" [BP01] paraissent bien adaptées à notre problématique. Elles évitent de retourner

dans la même branche de l'arbre des possibles par un mécanisme d'interdiction de choix de noeuds pendant un certain nombre d'itérations. Remarquons que le choix du poids élémentaire par facette devrait être facilité par les formes additives des fonctions de coût élaborées précédemment.

5.5.3 Remarques sur les techniques d'Hassan Jibrini

On peut s'interroger sur l'opportunité de réutiliser les techniques de "sequential greedy heuristics" proposées par [Jib02]. En fait Jibrini distingue dans son approche deux stratégies : "Best in" et "Worst out". La première consiste, comme son nom l'indique à ajouter à la clique courante le meilleur noeud possible[4] parmi les compatibles à tous les noeuds de la clique courante. Par définition, utilisée seule et jusqu'à son terme, cette stratégie fournit une clique maximale. La deuxième méthode consiste à supprimer du graphe le pire noeud. Jibrini fait alors remarquer que cette suppression doit se faire avec précaution car on risque dans ce cas de trouver des noeuds dans le graphe qui ne font plus partie d'aucune surface admissible. Jibrini prétend que ces noeuds sont supprimés par l'algorithme de suppression récursive appliqué au pire noeud. Le contre-exemple analysé en figure 5.24 montre que cet algorithme, y compris dans les conditions de travail de Jibrini (sans facettes verticales), ne garantit pas que cette suppression récursive préserve l'équivalence entre cliques maximales et surfaces admissibles. Jibrini propose une version "équivalente" de son algorithme "worst-out" sans retourner au graphe 3D, mais travaillant directement dans le graphe de compatibilité : avec le pire noeud w, il suffit de supprimer les noeuds compatibles ne possédant aucune arête avec un noeud incompatible à w. Là encore, l'analyse faite en figure 5.25 prouve que l'on n'est pas assuré que les cliques maximales sont encore des surfaces admissibles. En conclusion, la stratégie heuristique employée par Hassan Jibrini n'est pas assurée de préserver l'équivalence entre cliques maximales et surfaces admissibles. Force est de reconnaître toutefois que son algorithme n'a, semble-t-il, jamais rencontré de cas comme ceux décrits dans les figures de contre-exemple. La stratégie "worst-out" ne peut être employée qu'avant le calcul du graphe de compatibilité et ne doit pas être employée alors que ce dernier a déjà été calculé.

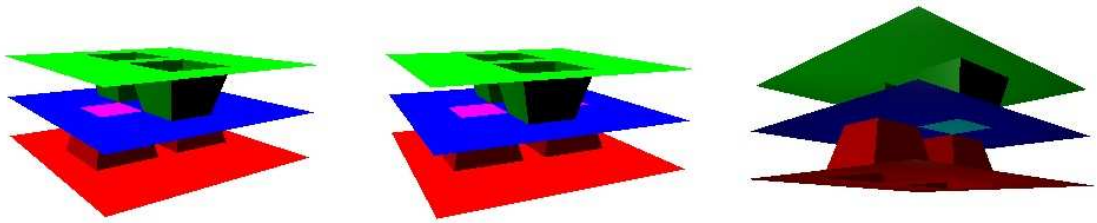


Figure 5.24 - Contre-exemple au "worst-out". Trois vues d'un même graphe. Chaque surface de couleur désigne une surface admissible au sens de Jibrini. Chaque facette n'appartient qu'à une seule surface admissible sauf les facettes au contact qui appartiennent à deux surfaces admissibles. Si on supprime récursivement une facette du bord de la surface bleue, la facette centrale du graphe (appartenant à la surface admissible bleue) ne peut être supprimée par l'algorithme. Or elle n'appartient qu'à une seule surface admissible. La clique maximale contenant cette facette ne contiendra qu'elle et de fait on n'a plus équivalence entre clique maximale et surface admissible.

[4] au sens d'un indicateur qui n'a que peu d'importance ici.

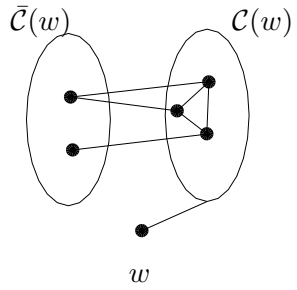


Figure 5.25 - Contre-exemple à l'algorithme "worst-out". On note $\mathcal{C}(w)$ les compatibles avec w et $\bar{\mathcal{C}}(w)$ ses incompatibles. Si on supprime w, les trois noeuds compatibles ne sont pas supprimés d'après l'algorithme "worst-out" de [Jib02]. Pourtant, ils forment une clique maximale qui ne correspond plus à une surface admissible.

5.6 Résultats

Comme nous l'avons signalé, plusieurs exemples sur des zones réelles apparaissent dans l'annexe D. On remarquera en particulier que sur les scènes 5 et 6 présentées dans cette annexe, l'algorithme a été capable de reconstruire les 2 bâtiments accolés et que sur la dernière scène, il a reconstruit les 2 bâtiments présents sur la zone de focalisation.
L'annexe E présente des résultats plus complets de l'algorithme appliqué sur plusieurs scènes et obtenus avec différents paramètres β ; paramètre au sujet duquel une discussion sera menée dans le chapitre 7. Tous ces résultats témoignent de la validité et de l'intérêt de l'approche pour la reconstruction de bâtiments.

5.7 Conclusion

L'algorithme décrit dans ce chapitre permet de générer à partir de primitives planes (façades et plans) un ensemble d'hypothèses de bâtiments génériques parmi lesquelles le choix est fait en utilisant le plus d'observations externes possibles ainsi que des contraintes a priori qui permettent l'introduction *souple*, dans le processus, d'une connaissance externe sur la structure des formes rencontrées.
Cet algorithme qui représente la contribution majeure de cette thèse peut en particulier reconstruire des bâtiments accolés ou des bâtiments présentant des discontinuités altimétriques internes. En outre, malgré sa généricité, l'algorithme est capable de gérer les cas d'erreurs des détecteurs. En revanche, il repose entièrement sur les caractéristiques géométriques des primitives et dépend donc beaucoup de la qualité géométrique des images et de leur calibrage ainsi que de la précision des algorithmes de détection de primitives.

Chapitre 6

Recalage de bâtiments avec contraintes

6.1 Introduction

Nous présentons dans ce chapitre une technique pour inférer et imposer des contraintes sur un modèle 3D approximatif. Cette méthode s'intègre naturellement dans la stratégie que nous avons adoptée. Les contraintes imposées sont celles détectées sur les primitives de base, le modèle 3D approximatif est celui issu de la reconnaissance de bâtiment. Toutefois, nous présentons ce chapitre comme une entité indépendante pour bien mettre en exergue les possibilités multiples d'utilisation : pour la reconstruction de bâtiments dans un système semi-automatique, ou automatique, dans une stratégie top-down ou bottom-up mais de façon encore plus générale pour la reconstruction d'une quelconque scène structurée présentant des contraintes géométriques fortes (environnement artificiel) .
Comparée à des travaux antérieurs, notre approche ne requiert aucune intervention manuelle pour la définition des contraintes, vérifie, de façon intrinsèque, les contraintes détectées et gère les cas de surcontraintes.

6.1.1 Motivations

La reconstruction d'objets d'environnements urbains ou tout au moins artificiels à partir d'un ensemble d'images est de première importance dans un grand nombre de domaines tels que la CAO, l'architecture assistée par ordinateur ou, dans notre contexte, la reconstruction automatique de bâtiments. Prendre en compte les propriétés géométriques dans ce type d'environnement peut grandement améliorer la qualité de la reconstruction [SHS98], est essentiel pour réduire le nombre de degrés de liberté du problème (section 5.3) et fournir une bonne modélisation de la scène.
Nous étudierons dans cette partie la reconstruction avec contraintes proprement dite, la topologie du modèle est supposée connue par les techniques décrites précédemment (chapitre 5). Bien que les résultats soient présentés dans le cadre de la reconstruction de bâtiments en continuité naturelle du contexte de cette thèse, l'algorithme, comme nous l'avons dit, peut tout à fait être utilisé dans la reconstruction d'une scène texturée quelconque et en particulier dans le domaine de la reconstruction architecturale.

6.1.2 État de l'art

Dans le cadre de la reconstruction de bâtiments et plus généralement pour la modélisation architecturale à partir d'un ensemble d'images géoréférencées, comme on l'a vu, les méthodes qui cherchent à intégrer des contraintes peuvent être divisées en deux catégories : "model-based" et "constraint-based". Dans les premières, les objets sont décomposés en primitives de base connues (parallélépipède, rectangles...) groupées pour construire des volumes plus complexes [FLM99, DTM96, SV00, vdHV97, FKL+98, JRH03, NN01, KHN01]. Une mise en correspondance est effectuée en général entre le modèle et les primitives images ou

3D et les contraintes connues *a priori* dans le modèle doivent donc ensuite être imposées sur les primitives mises en correspondance pour la reconstruction. Imposer ces contraintes est une tâche délicate, car les redondances et les surcontraintes ne sont en général pas gérées dans les définitions géométriques des modèles. les méthodes à base de contraintes sont plus générales et ne souffrent pas du manque de généricité des approches précédentes. Elles reposent sur des propriétés géométriques très classiques en environnement urbain, sans modèle a priori, pour reconstruire toute forme polyédrique. Plusieurs auteurs tirent parti de ces propriétés [SHS98, GSV02, Bon00, vdH99, AF00] mais en raison de la complexité du problème, ils dépendent tous d'une intervention manuelle pour la description topologique *et* géométrique. De façon générale, tous les systèmes de reconstruction 3D automatiques et semi-automatiques (tel que CC modeler [GW98]) bénéficieraient énormément de l'introduction de contraintes géométriques tels que coplanarité, horizontalité des arêtes, parallélisme, orthogonalité, symétrie. Cependant, la description des relations géométriques entre les primitives par un utilisateur est souvent rédhibitoire, en termes de temps, pour le travail sur des zones étendues.
En fonction des techniques utilisées pour imposer les contraintes sur le modèle 3D approximatif, les conditions géométriques sont exprimées de façon très différente. Certaines incluent les contraintes dans le formalisme de moindres carrés [vdH99, AF00, SHS98] ce qui conduit à une solution où il est difficile de déterminer quelles sont les contraintes exactement vérifiées. D'autres font une modélisation implicite avant le processus de minimisation [Bon00, GSV02]. Cela a l'avantage de vérifier *intrinsèquement* les contraintes fournies au préalable par l'utilisateur. Cependant, pour ces systèmes, bien que des techniques existent pour détecter les cas de surcontraintes, aucune solution automatique n'est proposée pour les gérer.
Tous les travaux utilisent la contrainte de coplanarité qui permet de vérifier les contraintes topologiques résultant de la description de l'objet (points appartenant à la même facette donc au même plan). [Bon00, GSV02] utilisent également des contraintes purement géométriques qui apparaissent entre les composantes géométriques de l'objet : orthogonalité, parallélisme...

6.1.3 Schéma général

L'algorithme abordé ici étend le modèle développé dans [GSV02] pour contourner les inconvénients rappelés dans la section 6.1.2. Dans un premier temps, les contraintes sont directement déduites de la forme initiale. Aucune intervention manuelle n'est nécessaire. Ensuite les contraintes sont imposées sur le modèle et les dégénérescences sont gérées en supprimant itérativement certaines d'entre elles. Nous nous sommes limités aux types de contraintes suivants, essentiels dans un environnement urbain : parallélisme, horizontalité, orthogonalité, horizontalité des arêtes, symétrie verticale.
Le schéma général est rappelé dans la figure 6.1 et suppose un modèle 3D initial approché dont les relations topologiques sont connues. Dans un premier temps, les normales des facettes sont regroupées (section 6.2.2), ce qui permet de traiter la contrainte parallélisme et le cas des plans horizontaux. Ensuite un graphe de contraintes est déduit de ces clusters, intégrant toutes les contraintes géométriques qui seront effectivement appliquées sur le modèle (sections 6.2.3 et 6.2.4). Dans un second temps, un ensemble d'équations est construit à partir d'équations de planarité entre les points et les normales des plans. Une paramétrisation implicite de cet ensemble permet de *forcer* les points à appartenir à leur plan respectif, assurant ainsi que les contraintes sont exactement vérifiées. Une minimisation itérative est ensuite appliquée par le biais d'un schéma de Levendberg-Marquardt pour trouver les points *vérifiant les contraintes géométriques* qui minimisent la distance avec le modèle initial. Surcontraintes et dégénérescences sont automatiquement détectées et gérées en supprimant itérativement certaines contraintes jusqu'à convergence. Une évaluation qualitative et quantitative montre la validité de l'approche contrôlée par deux paramètres $\sigma_{/\!/}$ et $\sigma_{\triangleleft}$ qui sont facilement réglés en fonction du bruit présent dans le modèle 3D initial.

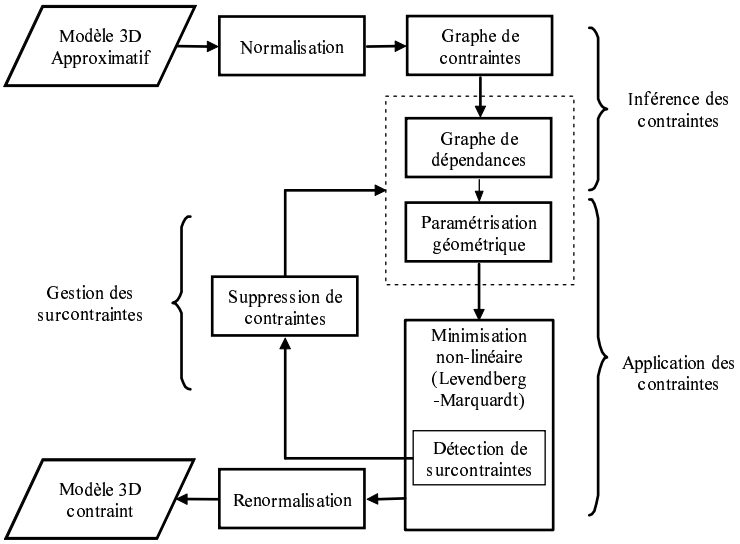


Figure 6.1 - Schéma général.

6.2 Inférence des contraintes

6.2.1 Topologie et géométrie : notations

On se borne ici à la description de la géométrie des toits. Les façades constituant les bords de bâtiments sont déduites de cette description. Elles sont essentielles pour imposer des contraintes sur les arêtes externes du bâtiment. En plus des facettes constituant les toits des bâtiments, une facette est donc ajoutée pour chaque arête externe ou arête de bord du bâtiment (en revanche, une discontinuité interne du bâtiment est gérée comme une facette normale). Par conséquent, ce que l'on appellera facette dans ce chapitre n'est pas nécessairement un polygone fermé, mais peut être constitué de *seulement deux points*.
Dans la suite, nous noterons $\mathcal{F}$ l'ensemble des facettes dont le cardinal est F et $\mathcal{X}$ l'ensemble des points dont on connaît une *position initiale approximative*. Un point $\mathbf{X}$ appartenant à une facette f sera noté $\mathbf{X} \in f$. D'un point de vue topologique, chaque facette f a $\mid f \mid$ points et d'un point de vue géométrique, chaque facette est liée à la normale d'un plan ou *direction* $\mathbf{v}_i$. L'ensemble des directions $\mathcal{D}$ a pour cardinal D. La figure 6.2 montre un exemple de bâtiment 3D et sa représentation schématique.

6.2.2 Clustering des directions

À partir du modèle initial, toutes les normales aux plans sont estimées pour chaque facette du modèle (avec une condition d'horizontalité pour les normales des façades définies par les arêtes externes), conduisant ainsi à F normales appelées dans la suite du chapitre directions $\mathbf{v}$. À cet ensemble de directions, la direction verticale $\mathbf{z}$ est ajoutée en raison de son importance pour la définition d'arêtes horizontales ou de symétries verticales et pour son rôle dans la définition des façades dont la condition d'horizontalité de la normale sera imposée par orthogonalité avec la direction verticale.
Une agrégation ascendante hiérarchique est réalisée sur ces directions avec pour distance de fusion la distance entre les centres de gravité des agrégats ou clusters. Ainsi, un cluster est généré en regroupant le couple de normales dont la distance est la plus proche. La distance utilisée est la valeur absolue de l'angle aigu entre les deux normales. Une nouvelle normale est recalculée pour le nouveau cluster par la valeur moyenne des normales qu'il regroupe. Itérativement, les normales sont regroupées et le processus continue jusqu'à ce que

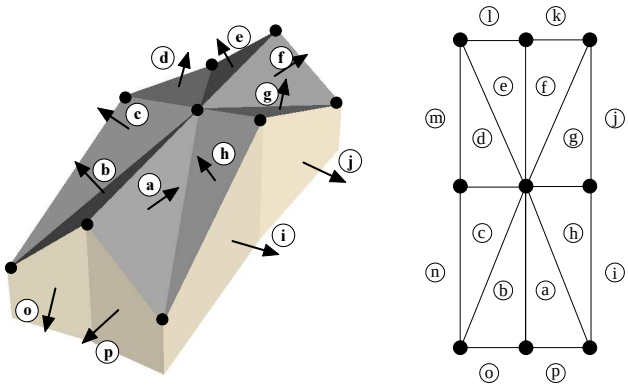


Figure 6.2 - Bâtiment 3D approximatif et sa représentation schématique annotée avec les directions (lettres) sur les facettes. Les cercles noirs symbolisent les points initiaux approchés qui ne vérifient pas les contraintes. Les plans verticaux constituant les bords du bâtiment sont déduits des points. Les directions sont représentées avec des flèches mais les orientations sont ici sans importance.

la distance minimale dépasse un seuil $\sigma_{/\!/}$.
A la fin de l'algorithme, chaque paire de clusters est distante de *plus de* $\sigma_{/\!/}$ et il n'y a plus que $D \leq F + 1$ directions. Notons que, pour ne pas altérer la direction verticale, la valeur de la direction pour ce cluster n'est pas moyennée mais gardée verticale.
Cette étape permet de gérer les propriétés de parallélisme et de verticalité des normales. Chaque facette du modèle initial est identifiée à une direction ce qui sera utile pour construire le système contraint.

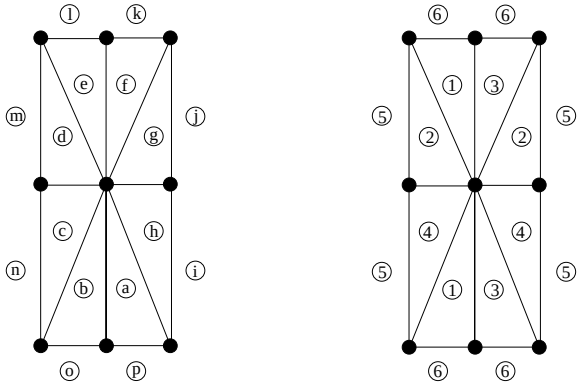


Figure 6.3 - Résultat du clustering de directions sur l'exemple précédent.

6.2.3 Graphe des contraintes

Un graphe non orienté est construit à partir des directions, chaque direction étant un noeud du graphe. Plusieurs types d'arêtes valuées sont ajoutés entre deux noeuds $\mathbf{v}_i$ et $\mathbf{v}_j$ dès qu'une de ces conditions est remplie :

- *orthogonalité* si $|\mathbf{v}_i \bullet \mathbf{v}_j| < \sin(\sigma_{\sphericalangle})$
- *symétrie verticale* si $|(\mathbf{v}_i + \mathbf{v}_j) \bullet \mathbf{z}| > \cos(\sigma_{\sphericalangle}) \cdot \|(\mathbf{v}_i + \mathbf{v}_j)\|$
- *ligne horizontale* si $|(\mathbf{v}_i \otimes \mathbf{v}_j) \bullet \mathbf{z}| < \sin(\sigma_{\sphericalangle}) \cdot \|\mathbf{v}_i \otimes \mathbf{v}_j\|$

où • représente le produit scalaire et $\otimes$ le produit vectoriel. Chaque arête c correspond à une contrainte entre deux directions $\mathbf{v}_i$ et $\mathbf{v}_j$, ce qui sera noté $c_f = \mathbf{v}_i, c_t = \mathbf{v}_j$. Dans le graphe de contraintes, l'orientation de la contrainte n'a pas d'importance (le graphe n'est pas orienté). Dans la suite en revanche, la notation précédente signifiera que c est orientée de $\mathbf{v}_i$ à $\mathbf{v}_j$ et que donc $\mathbf{v}_i$ *contraint* $\mathbf{v}_j$. Une seule arête peut exister entre deux noeuds. De fait, en cas de redondance entre la symétrie verticale et la contrainte de ligne horizontale, la symétrie est préférée car elle apporte plus de contraintes sur les directions. L'orthogonalité n'est testée qu'entre des plans verticaux (voir figure 6.4), l'orthogonalité entre une façade et la direction verticale est automatiquement intégrée[1] Mentionnons que les deux seuils $\sigma_{/\!/}$ et $\sigma_{\lhd}$ sont aisément réglés en fonction de l'erreur estimée dans le modèle approximatif ou en fonction du degré de caricature voulu pour la reconstruction.

6.2.4 Graphe de dépendances

D'après le modèle de [Bon00], les directions doivent être déduites les unes des autres dans un ordre séquentiel : chaque direction est déduite des directions précédentes dans la séquence et de règles de construction. Nous devons donc établir un ordre constructif sur ces directions ou, ce qui est équivalent, orienter le graphe de contraintes [Bon00] tout en assurant qu'aucune sur-contrainte n'apparaît. Des tentatives ont été menées pour résoudre ce problème dans des configurations générales [HV95] mais elles ne sont pas bien adaptées aux cas 3D. Une autre approche a été développée pour gérer des configurations générales (avec contraintes sur les points, les plans, les lignes...) avec des procédures heuristiques pour maximiser le nombre de contraintes appliquées au système [Bon00]. Nous réutilisons quelques-unes de ces procédures pour orienter notre graphe de contraintes.

Notations

L'algorithme intègre itérativement les directions et les contraintes qui leur sont reliées. À tout moment, l'algorithme maintient l'ensemble des directions qui ont été intégrées $\mathcal{A}$. En plus, certaines des contraintes du graphe de contraintes peuvent être confirmées (acceptées par le système à cet instant), potentielles (pas encore intégrées) ou inactives (rejetées pour des considérations de surcontraintes). On trouve ainsi les différents ensembles correspondant respectivement à chaque type $\mathcal{C}, \mathcal{P}$ et $\mathcal{I}$. Les ensembles de contraintes confirmées ou potentielles sur un noeud spécifique n (direction) seront notées respectivement $C(n)$ et $P(n)$. Toute contrainte c ajoutée dans le graphe de contraintes réduit le nombre de degrés de liberté du système contraint par $f(c)$: 2 pour une symétrie verticale, 1 pour les autres types de contraintes. Initialement, chaque noeud a un nombre de degrés de liberté de 2 et le degré courant peut être calculé à chaque instant et pour chaque noeud par l'équation 6.1 :

$$d(n) = 2 - \sum_{c \in C(n)/c_t = n} f(c) \tag{6.1}$$

Algorithme

L'algorithme d'ordre constructif (algorithme 4) place dans un premier temps le noeud correspondant à la direction verticale qui a un rôle essentiel, et rend actives toutes les contraintes qui lui sont liées en les orientant à partir de lui. En particulier, toutes les contraintes d'orthogonalité avec les façades sont ainsi introduites dans le graphe de dépendances. Ensuite, séquentiellement, il choisit dans les directions les plus

[1] On notera que le cas des toits en terrasse est géré dans le clustering des directions. En effet, à l'issue de cette étape, les plans dont la normale est proche de la direction verticale sont rattachés au cluster de cette dernière. Or, comme cela est rappelé en section 6.2.2, la valeur de la direction de ce cluster n'est pas moyennée mais reste verticale. La direction de ces plans sera donc parfaitement verticale ce qui permet la gestion des toits plats.

contraintes celle qui a le plus de contraintes potentielles qui lui sont liées. Il introduit cette direction et active ses contraintes potentielles en les orientant à partir de cette direction. Si un noeud devient surcontraint, c'est-à-dire que son degré de liberté devient négatif, des contraintes sont supprimées jusqu'à ce qu'aucune surcontrainte ne subsiste.

Algorithme 4: Algorithme d'ordre constructif sur les directions.

$\mathcal{A} = \mathcal{A} \cup \{\mathbf{z}\}$;
$\mathcal{I} = \emptyset$;
$\mathcal{C} = \mathcal{C} \cup \{P(\mathbf{z})\}$;
tant que $(\mathcal{D} \backslash \mathcal{A}) \neq \emptyset$ **faire**
  soit $m = \min_{n \in \mathcal{N}} d(n)$, $\mathcal{L} = argmin_{n \in \mathcal{N}} d(n)$;
  si $m < 0$ **alors**
    pour chaque $n \in \mathcal{L}$ **faire**
      soit $T = argmin_{c \in C(n)} f(c)$ et $t \in T$;
      répéter $\mathcal{C} = \mathcal{C} \backslash t, \mathcal{I} = \mathcal{I} \cup t$ **jusqu'à** $d(n) \geq 0$
    fin
  sinon
    soit $B = argmin_{n \in \mathcal{L}} Card(P(n))$ et $b \in B$;
    $\mathcal{A} = \mathcal{A} \cup b$;
    $\mathcal{C} = \mathcal{C} \cup P(b)$;
  fin
fin

Cas ambigus

À la fin, dans le cas de direction contrainte par deux contraintes "ligne horizontale", une des deux contraintes est supprimée. Deux configurations peuvent en effet se produire dans cette situation : soit les 2 lignes ne sont pas parallèles et le plan est horizontal (la direction est donc verticale), ce qui par définition n'est pas possible (car la direction verticale est introduite au début et n'est donc pas contrainte voir algorithme 4), soit les 2 lignes sont parallèles et une des deux contraintes n'apporte aucune information.
Quand une direction $\mathbf{v}_i$ est définie perpendiculaire à $\mathbf{v}_k$ et a une contrainte de "ligne horizontale" avec $\mathbf{v}_l$, la situation est analogue. Soit $\mathbf{v}_l \perp \mathbf{v}_k$ et la contrainte d'horizontalité n'apporte aucune information supplémentaire, soit les deux contraintes sur $\mathbf{v}_i$ sont incompatibles (ce qui ne devrait pas se produire puisque les seuils sont les mêmes pour évaluer les deux contraintes). Par conséquent, dans ce cas, nous supprimons la contrainte de ligne horizontale de $\mathcal{C}$. Ainsi, une fois que le graphe de dépendances est construit, seules 6 situations peuvent apparaître. Elles sont détaillées dans le tableau 6.1.

Tri topologique

L'algorithme assure, par construction, qu'aucune direction n'est surcontrainte (son degré de liberté n'est pas négatif) et que le graphe de dépendances obtenu est orienté et acyclique. Il est en effet orienté par construction et acyclique parce que, si une direction a été placée à une étape i de l'algorithme, par construction, *aucune contrainte orientée* ne peut exister entre cette direction et d'autres apparues à l'étape j avec $j < i$, par conséquent aucun cycle n'existe et un tri topologique peut être appliqué sur le graphe. À la fin de l'algorithme, l'ensemble des contraintes confirmées $\mathcal{C}$ peut être appliqué sur les directions. Il est subdivisé en contraintes de symétrie, d'orthogonalité, et d'arête horizontale : $\mathcal{C}_S$, $\mathcal{C}_P$, $\mathcal{C}_H$. Des illustrations du fonctionnement de l'algorithme sont données sur la figure 6.4.

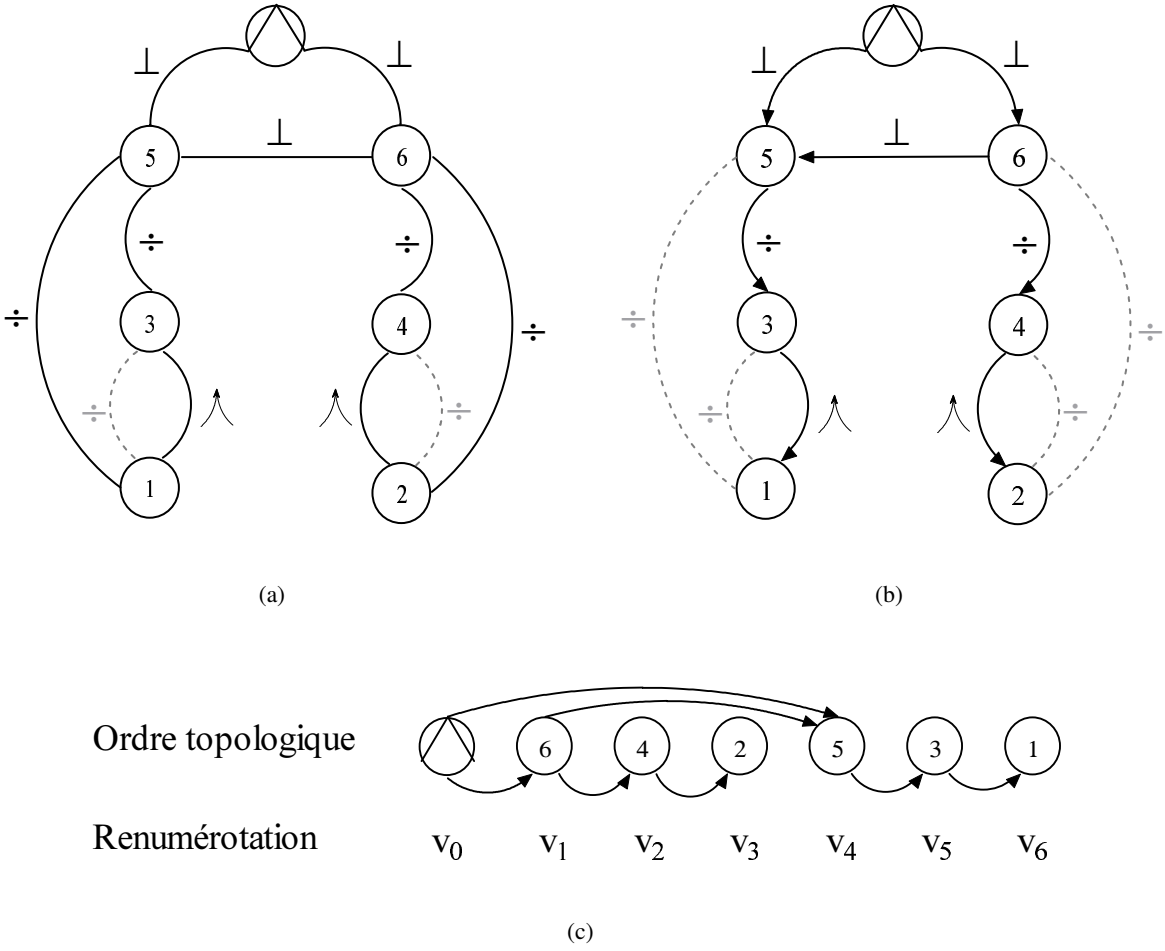


Figure 6.4 - Graphe de contraintes (a) et graphe de dépendances (b) correspondant au graphe du bâtiment présenté en figure 6.2. Chaque arête représente une contrainte entre deux directions. ⏀ définit la direction verticale, $\perp$, $\div$ et ⅄ symbolisent la contrainte d'orthogonalité, d'arête horizontale et de symétrie verticale respectivement. Les arêtes grisées sont des contraintes supprimées du graphe soit pour des raisons de redondance soit par l'algorithme d'ordre constructif 4 (voir le texte pour détails). Le tri topologique est appliqué sur le graphe de dépendances (c) et la renumérotation peut alors s'effectuer.

6.2.5 Calculs via le graphe de dépendances

D'après ce que l'on a vu, chaque direction peut être numérotée de façon à ce qu'elles puissent être calculées séquentiellement, chacune ne dépendant que des précédentes dans l'ordre et de règles définies par les contraintes données par les arêtes du graphe de dépendances. Dans la suite, nous supposerons que cet ordonnancement a été effectué et que $\forall i$, $\mathbf{v}_i$ dépend uniquement de direction(s) précédente(s) $\mathbf{v}_j, j < i$. Si $d(\mathbf{v}_i) > 0$, $\mathbf{v}_i$ n'est pas complètement contrainte par les directions précédentes (par exemple des directions arbitraires non contraintes qui ne dépendent d'aucune autre direction ou des directions perpendiculaires qui ne dépendent que d'une seule autre direction). Nous supposerons que ces directions sont reliées à des paramètres $\boldsymbol{\theta}_i$. Il est évident que certaines directions totalement contraintes (ou connues telle la direction verticale) ne dépendent d'aucuns paramètres. Cependant, pour des raisons de simplicité, nous garderons ces $\boldsymbol{\theta}_i$ inutiles dans notre notation et considérerons donc l'ensemble $\boldsymbol{\theta}_1 \cdots \boldsymbol{\theta}_D$. Les règles rappelées dans

le tableau 6.1 permettent alors de calculer séquentiellement chaque direction à partir des paramètres et des directions de dépendance, de façon à ce que *les contraintes géométriques soient vérifiées*. Les dérivées partielles $\partial \mathbf{v}_i / \partial \boldsymbol{\theta}_k \quad \forall i, k$ peuvent aisément être calculées à partir des indications fournies dans ce tableau et des règles de composition. Ces dérivées seront nécessaires dans la minimisation itérative.

6.3 Application des contraintes au modèle

Cette section rappelle essentiellement les principes de la méthode décrite dans [Gro02] avec quelques simplifications dues à la nature de notre problème (en particulier, il n'y a pas de calibration de caméras à faire).

6.3.1 Paramétrisation géométrique

Comme expliqué plus haut, chaque direction est déterminée d'après certains paramètres $\boldsymbol{\theta} = [\boldsymbol{\theta}_1^T, \ldots, \boldsymbol{\theta}_D^T]^T$. Toute paire de points appartient au même plan dont la direction de la normale est $\mathbf{v}_i$ si et seulement si leurs coordonnées $\mathbf{X}_m$, $\mathbf{X}_n$ vérifient l'équation de planarité :

$$\mathbf{v}_i^T \left(\mathbf{X}_m - \mathbf{X}_n \right) = 0 \tag{6.2}$$

Pour chaque facette du modèle initial comportant $|f|$ points, $|f| - 1$ équations peuvent donc être écrites. En concaténant ces équations, on obtient le système suivant :

$$\mathbf{B}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_D)\mathbf{X} = \mathbb{0} \tag{6.3}$$

où $\mathbf{X} = [\mathbf{X}_1^T, \ldots, \mathbf{X}_N^T]^T$ contient l'ensemble des coordonnées des points et $\mathbf{B}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_D)$, une matrice $P \times N$, contient les contraintes géométriques. Il est alors aisé de voir que la dimension M du noyau de $\mathbf{B}$ est supérieure ou égale à 3 (tous les points égaux à une valeur quelconque vérifient le système) et que, par conséquent, une *paramétrisation implicite* des points $\mathbf{X}$ est

$$\mathbf{X} = \mathbf{U}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_D)\mathbf{V} \tag{6.4}$$

avec $\mathbf{V} \in \mathbb{R}^M$ et où $\mathbf{U}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_D)$ est une matrice $3N \times M$ dont les colonnes forment une base orthonormale du noyau de $\mathbf{B}$. De fait, $\mathbf{X}$ *vérifie implicitement toutes les contraintes géométriques et* $\mathbf{V}$ *contient l'ensemble des paramètres implicites du modèle contraint*. Par la construction séquentielle, les normales des plans vérifient les contraintes géométriques et donc *toute valeur* de $\mathbf{V}$ fournit une solution du système contraint. M est le nombre de degrés de liberté de $\mathbf{X}$.
En ce qui concerne les inconnues, $\boldsymbol{\theta}$ et $\mathbf{V}$, on les regroupe dans un seul vecteur $\boldsymbol{\Theta} = [\boldsymbol{\theta}, \mathbf{V}]$. Notons quelques points importants avant d'expliquer la reconstruction suivant le Maximum de Vraisemblance.

Normalisation

La normalisation améliore grandement la stabilité des calculs numériques [Har95a]. Dans ce travail, les points 3D sont normalisés au début de façon à ce que leur centre de gravité coïncide avec l'origine et que la norme moyenne $\|\mathbf{X}\|$ vaille 1, ce qui se traduit par :

$$\left(\mathbb{1}_{1 \times N} \otimes I \right) \mathbf{X} = \mathbb{0}_{3 \times 1} \tag{6.5}$$

$$\mathbf{X}^T \mathbf{X} = N \tag{6.6}$$

L'étape de renormalisation permet de revenir au référentiel initial.

Type	Paramètre	Valeur	Derivées
connue	non	non changé	$\mathbb{O}$
arbitraire	$\boldsymbol{\theta}_i \in \mathbb{R}^3 \backslash \{\mathbb{O}\}$	$\mathbf{v}_i = \frac{\boldsymbol{\theta}_i}{\|\boldsymbol{\theta}_i\|}$	$\frac{\partial \mathbf{v}_i}{\partial \boldsymbol{\theta}_i} = \frac{1}{\|\boldsymbol{\theta}_i\|}\left(I - \frac{1}{\|\boldsymbol{\theta}_i\|^2}\boldsymbol{\theta}_i\boldsymbol{\theta}_i^T\right)$
orthogonal à $\mathbf{v}_j$	$\boldsymbol{\theta}_i$	$\mathbf{u}_1 = \left(I - \mathbf{v}_j\mathbf{v}_j^T\right)\boldsymbol{\theta}_i$, $\mathbf{u}_2 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}$ $\mathbf{v}_i = \mathbf{u}_2$	$\frac{\partial \mathbf{v}_i}{\partial \boldsymbol{\theta}_i} = \frac{1}{\|\mathbf{u}_1\|}(I - \mathbf{u}_2\mathbf{u}_2^T - \mathbf{v}_j\mathbf{v}_j^T)$ $\frac{\partial \mathbf{v}_i}{\partial \mathbf{v}_j} = \frac{1}{\|\mathbf{u}_1\|}\left(\left(\mathbf{v}_j^T\boldsymbol{\theta}_i\right)\left(I - \mathbf{u}_2\mathbf{u}_2^T\right) + \mathbf{v}_j\boldsymbol{\theta}_i^T\right)$
2 orthogonalités : $\mathbf{v}_j$, $\mathbf{v}_k$	non	$\mathbf{v}_i = \frac{\mathbf{v}_j \otimes \mathbf{v}_k}{\|\mathbf{v}_j \otimes \mathbf{v}_k\|}$	$\frac{\partial \mathbf{v}_i}{\partial \mathbf{v}_j} = -\frac{1}{\|\mathbf{v}_j \otimes \mathbf{v}_k\|}(I - \mathbf{v}_i\mathbf{v}_i^T)\mathbf{S}_{v_k}$ $\frac{\partial \mathbf{v}_i}{\partial \mathbf{v}_k} = \frac{1}{\|\mathbf{v}_j \otimes \mathbf{v}_k\|}(I - \mathbf{v}_i\mathbf{v}_i^T)\mathbf{S}_{v_j}$
ligne horizontale avec $\mathbf{v}_j$	$\boldsymbol{\theta}_i$	$\mathbf{h} = \mathbf{z} \otimes \mathbf{v}_j$, $\mathbf{u}_1 = \left(I - \mathbf{h}\mathbf{h}^T\right)\boldsymbol{\theta}_i$ $\mathbf{u}_2 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}$, $\mathbf{v}_i = \mathbf{u}_2$	$\frac{\partial \mathbf{v}_i}{\partial \boldsymbol{\theta}_i} = \frac{1}{\|\mathbf{u}_1\|}\left(I - \mathbf{u}_2\mathbf{u}_2^T - \mathbf{h}\mathbf{h}^T\right)$ $\frac{\partial \mathbf{v}_i}{\partial \mathbf{v}_j} = \frac{1}{\|\mathbf{u}_1\|}\left(\left(\mathbf{h}^T\boldsymbol{\theta}_i\right)\left(I - \mathbf{u}_2\mathbf{u}_2^T\right) + \mathbf{h}\boldsymbol{\theta}_i^T\right)\mathbf{S}_\mathbf{z}$
symétrie verticale avec $\mathbf{v}_j$	non	$\mathbf{v}_i = 2(\mathbf{v}_j \cdot \mathbf{z}) - \mathbf{v}_j$	$\frac{\partial \mathbf{v}_i}{\partial \mathbf{v}_j} = 2\mathbf{z}\mathbf{z}^T - I$

Table 6.4 - Calculs des directions $\mathbf{v}_i$ et de leurs dérivées en fonction des paramètres. $\mathbf{S}_\mathbf{v}$ est la matrice de produit vectoriel associée à un vecteur $\mathbf{v}$. Quand elle n'est pas indiquée pour un index spécifique l, la dérivée par rapport à $\boldsymbol{\theta}_l$ est 0.

Unicité de U

Toute base orthonormale du noyau $\mathbf{B}$ peut constituer les colonnes de $\mathbf{U}$. Or le noyau de $\mathbf{B}$ a une infinité de bases orthonormales. Cependant, des précautions doivent être prises pour que cette matrice soit différentiable. Nous décrivons ici une procédure permettant de définir de façon univoque la matrice $\mathbf{U}$ et qui assure en outre que cette matrice est différentiable. Supposons connue la valeur de $\mathbf{U}$ en un point $\boldsymbol{\theta}^s$. À partir de cette valeur $\mathbf{U}^s$, toute valeur de $\mathbf{U}$ peut être définie en un point quelconque par la procédure suivante. $\mathcal{U}_1$ est une matrice unitaire (quelconque) dont les colonnes forment une base de $\mathcal{N}ull(\mathbf{B}(\boldsymbol{\theta}))$,

$$\begin{aligned} \mathbf{U_1}^T\mathbf{U}^s &\overset{\text{SVD}}{=} \mathcal{U}\mathcal{D}\mathcal{V}^T \\ \mathbf{U}^i &= \mathbf{U}_1\mathcal{U}\mathcal{V}^T \end{aligned} \tag{6.7}$$

Cette procédure assure que $\mathbf{U}^i$ est unique et minimise $\|\mathbf{U}^i - \mathbf{U^s}\|_F$ dans l'ensemble des bases orthonormales de $\mathcal{N}ull(\mathbf{B}(\boldsymbol{\theta}))$.

6.3.2 Schéma de minimisation itératif

Nous estimons la solution *vérifiant les contraintes géométriques* qui approxime au mieux les données 3D initiales. La solution devrait ainsi minimiser la fonction :

$$\mathcal{Q}(\boldsymbol{\Theta}) = \|\mathbf{X}_{ini} - \mathbf{X}(\boldsymbol{\Theta})\|^2 \tag{6.8}$$

On étudiera en section 6.6 la possibilité de choisir une autre fonction $\mathcal{Q}$ plus spécifiquement reliée à l'information image. La procédure de Levendberg-Marquardt suit ensuite le schéma rappelé dans l'algorithme 5. À part les boucles externes et internes, quelques détails méritent d'être approfondis.

Initialisation

L'initialisation de $\boldsymbol{\theta}_i$ est directement liée aux valeurs initiales des directions des plans. En ce qui concerne $\mathbf{V}$, on choisit

$$\mathbf{V}^0 = \left(\mathbf{U}^0\right)^+ \mathbf{X}_{ini} \tag{6.9}$$

où $(\mathbf{U}^0)^+$ désigne la pseudo-inverse de $\mathbf{U}^0$. $\mathbf{U}^0$ est choisie quelconque comme base du noyau de $\mathbf{B}$. N'importe quelle procédure convient. Nous avons choisi le résultat de la fonction $getnull$ du langage "octave" [Gro02]. Cette solution initiale vérifie ainsi les contraintes géométriques tout en étant la plus proche des valeurs initiales au sens des moindres carrés.

Différentiabilité de X

En plus de son unicité, on peut montrer que $\mathbf{U}$ est différentiable dans un voisinage de $\boldsymbol{\theta}^s$ et que l'expression simplifiée de cette différentielle en ce point $\boldsymbol{\theta}^s$ est donnée par

$$\frac{\partial \mathbf{U}}{\partial \boldsymbol{\theta}}(\boldsymbol{\theta}^s) = -\mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \boldsymbol{\theta}}(\boldsymbol{\theta}^s)\mathbf{U}^s \tag{6.10}$$

À partir de l'équation 6.4, il s'ensuit

$$\frac{\partial \mathbf{X}}{\partial \boldsymbol{\Theta}} = \left[\frac{\partial \mathbf{U}}{\partial \boldsymbol{\theta}}\mathbf{V}; \mathbf{U}\right] \tag{6.11}$$

Par conséquent, dans un voisinage d'un point donné, $\mathbf{X}$ est différentiable et sa valeur peut aisément être déterminée.

Algorithme 5: Algorithme de Levendberg-Marquardt.

Initialisation : $\boldsymbol{\theta}^0, \mathbf{U}^0, \mathbf{V}^0 \Rightarrow \mathbf{X}^0, i = 0, \lambda = 0.001$
répéter
  Évaluer $\frac{\partial \mathbf{X}^i}{\partial \boldsymbol{\Theta}^i}$ à l'aide des équations 6.10 et 6.11
  succes=faux
  tant que *!succes* **faire**
    Évaluer $\mathcal{Q}(\boldsymbol{\Theta}^i)$ à l'aide des équations 6.4,6.7 et 6.8
    $\beta = \frac{\partial \mathbf{X}^i}{\partial \boldsymbol{\Theta}^i}^T \times (\mathbf{X}_{ini} - \mathbf{X}^i), \alpha = \frac{\partial \mathbf{X}^i}{\partial \boldsymbol{\Theta}^i}^T \frac{\partial \mathbf{X}^i}{\partial \boldsymbol{\Theta}^i}$
    $\alpha'_{jk} = \alpha_{jk}$ pour $j \neq k$ et $\alpha'_{jj} = (1 + \lambda)\alpha_{jj}$
    Résoudre $\alpha'(\delta\boldsymbol{\Theta}^i) = \beta$
    si $\mathcal{Q}(\boldsymbol{\Theta}^i + \delta\boldsymbol{\Theta}^i) \geq \mathcal{Q}(\boldsymbol{\Theta}^i)$ **alors**
      $\lambda = \lambda * 10$
    fin
    si $\mathcal{Q}(\boldsymbol{\Theta}^i + \delta\boldsymbol{\Theta}^i) < \mathcal{Q}(\boldsymbol{\Theta}^i)$ **alors**
      $\lambda = \lambda/10$, succes=vrai
      $\boldsymbol{\Theta}^s = \boldsymbol{\Theta}^i$ et $\mathbf{U}^s = \mathbf{U}^i$
      $\boldsymbol{\Theta}^{i+1} = \boldsymbol{\Theta}^i + \delta\boldsymbol{\Theta}^i, i = i + 1$
    fin
  fin
jusqu'à $\|\mathbf{X}^i - \mathbf{X}^{i-1}\|^2 < \epsilon$
solution : $\boldsymbol{\Theta}^s \Rightarrow \mathbf{X}^s$

6.4 Gestion des surcontraintes et des dégénérescences

6.4.1 Détection

La génération du graphe de dépendances assure qu'aucune surcontrainte *sur les directions* ne peut avoir lieu. Cependant, dans le système de contraintes appliqué aux points, rien ne permet d'affirmer qu'un point n'est pas surcontraint. Comme illustré par la figure 6.5, les points peuvent être surcontraints ce qui conduit à des dégénérescences du système. Un point intéressant à noter est qu'un point peut être surcontraint même si moins de contraintes sont appliquées sur les directions. Cette remarque vient du fait que notre minimisation est effectuée en deux étapes comme dans [SHS98] : tout d'abord, les directions sont calculées, ensuite les points *doivent* se conformer à ces directions. Il est tout à fait possible qu'un point ne puisse vérifier toutes les contraintes de planarité, car aucun test n'a été effectué dans le calcul des directions pour vérifier des situations que la topologie impose (voir figure 6.5). Les dégénérescences apparaissent quant à elles lorsque les valeurs sont trop éloignées de la situation contrainte imposée par l'application, ce qui conduit à la fusion de points.
Dans notre méthode, les situations de surcontraintes et de dégénérescences sont aisément mises en évidence, comme rappelé dans [Gro02] en examinant $\mathbf{U}^i$ à tout moment de l'algorithme. Si il existe deux triplets de rangées $\{3m-2, 3m-1, 3m\}$ et $\{3n-2, 3n-1, 3n\}$ $(m \neq n)$, correspondant à deux points différents qui sont égaux, cela signifie que la seule solution pour vérifier les contraintes dans la paramétrisation implicite fut de fusionner les deux points en question. C'est donc la procédure utilisée pour détecter ces situations dégénérées.

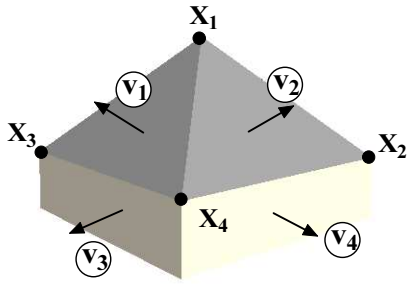


Figure 6.5 - Etude d'un cas de surcontrainte : si on suppose que les $\mathbf{v}_i$ sont arbitraires ou sont connues, une fois que $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$ sont définis, $\mathbf{X}_4$ ne peut pas vérifier les 4 contraintes de planarité. En revanche, si $\mathbf{v}_4 \div \mathbf{v}_2$ et $\mathbf{v}_1 \div \mathbf{v}_3$, $\mathbf{X}_4$ vérifie automatiquement ces contraintes.

6.4.2 Gestion

Une bonne technique pour gérer ces situations serait d'inférer des propriétés par exemple dues à des considérations topologiques pour les appliquer sur les directions [Bon00]. Dans l'exemple 6.5, en supposant que les autres directions sont fixées, la direction 4 ne peut être déterminée indépendamment des points. C'est là toute la limitation de notre méthode qui découple le problème en deux étapes bien distinctes. Dans cet exemple, supprimer les contraintes de planarité (inutiles ici) liées à une facette qui n'a que 3 points résoudrait le problème de surcontrainte. L'algorithme utilisé pour gérer ces cas de figure suit le même principe. À partir du modèle topologique, définissons $\mathcal{F}(\mathbf{v}_i)$ l'ensemble des facettes lié à une direction $\mathbf{v}_i$ et $\mathcal{F}_3(\mathbf{v}_i) = \{f \in \mathcal{F}(\mathbf{v}_i)/|f| \leq 3\}$.

À la fin de l'algorithme, en cas de succès, toutes les relations topologiques sont vérifiées dans la mesure où toutes les contraintes de planarité sur les facettes de strictement plus de 3 points sont conservées. De plus, l'algorithme utilise des heuristiques pour pénaliser les directions qui sont liées à peu de facettes dans le but de perdre le moins possible des conditions d'alignement et de parallélisme. Dans le cas où la suppression d'équations de planarité n'est pas suffisante, des contraintes sur les directions sont itérativement supprimées.

6.5 Résultats

6.5.1 Protocole d'évaluation

25 bâtiments de référence ont été modélisés à la main (sans contrainte) correspondant à 238 points. À partir de cette modélisation, les contraintes sont inférées et imposées en utilisant l'algorithme décrit dans ce chapitre, fournissant ainsi une référence contrainte qui servira de base à notre comparaison : tous les résultats donnés dans la suite le seront par rapport à cette référence contrainte. Il est en effet indispensable d'avoir une vérité qui respecte les contraintes pour pouvoir effectuer des analyses. Pour évaluer notre algorithme, un bruit gaussien est ajouté aux points avec une variance croissante et la reconstruction est réalisée en utilisant ces points bruités comme points initiaux $\mathbf{X}_{ini}$. Tous les résultats sont donnés sur une moyenne de 25 itérations. Les paramètres utilisés ont été : $\sigma_{/\!/} = 12°$ et $\sigma_{\sphericalangle} = 10°$.

Deux cas de figure ont été considérés pour l'évaluation sur des bâtiments bruités : les contraintes a priori et les contraintes a posteriori. Le premier cas est comparable dans l'esprit aux reconstructions à base de modèles : les contraintes sont connues *a priori* par une information externe (apportée ici par les contraintes du modèle de référence). On évite ainsi l'étape de clustering de directions et de construction du graphe de contraintes (bien que les valeurs initiales soient elles déduites à partir des bâtiments bruités). Dans le second cas, les contraintes sont systématiquement déduites du modèle bruité. Des contraintes vont donc être perdues

Algorithme 6: Suppression itérative de contraintes.

succes=faux
tant que *((!succes) && ($(\mathcal{C}_S \neq \emptyset)\|(\mathcal{C}_P \neq \emptyset)\|(\mathcal{C}_H \neq \emptyset)$)* **faire**
  construction du graphe de dépendances (Algorithme 4)
  soit $\mathcal{L} = \{\mathbf{v}_i / \mathcal{F}_3(\mathbf{v}_i) \neq \emptyset\}$
  $\eta(\mathbf{v}_i) = Card(\mathcal{F}(\mathbf{v}_i))$
  tant que *((!succes) && $\mathcal{L} \neq \emptyset$)* **faire**
    minimisation (algorithme 5)
    si *une dégénérescence est détectée (section 6.4.1)* **alors**
      soit $\mathbf{t} = \text{argmin}_{\mathbf{v}_i \in \mathcal{L}}(\eta(\mathbf{v}_i))$
      $f \in \mathcal{F}_3(\mathbf{t})$, $\mathcal{F}_3(\mathbf{t}) = \mathcal{F}_3(\mathbf{t}) \backslash f$
      suppression des contraintes topologiques liées à f
      si $\mathcal{F}_3(\mathbf{t}) = \emptyset$ **alors** $\mathcal{L} = \mathcal{L} \backslash \mathbf{t}$
    sinon
      succes=vrai
    fin
  fin
  si *(!succes)* **alors**
    si $\mathcal{C}_S \neq \emptyset$ **alors**
      $t = \text{argmin}_{c \in \mathcal{C}_S}(\eta(c_1) + \eta(c_2))$
    sinon
      si $\mathcal{C}_P \neq \emptyset$ **alors**
        $t = \text{argmin}_{c \in \mathcal{C}_P}(\eta(c_1) + \eta(c_2))$
      sinon
        si $\mathcal{C}_H \neq \emptyset$ **alors**
          $t = \text{argmin}_{c \in \mathcal{C}_H}(\eta(c_1) + \eta(c_2))$
        fin
      fin
    fin
    $\mathcal{C} = \mathcal{C} \backslash t$
  fin
fin

par rapport aux modèles de référence. D'un côté, nous analysons le gain en précision, c'est-à-dire la distance moyenne entre les bâtiments reconstruits et la référence (rappelons que la minimisation est faite aux moindres carrés par rapport aux points bruités !). D'un autre côté, le gain en degré de liberté est étudié pour montrer la rigidité que les contraintes apportent à la reconstruction. Comme $M = rank(\mathbf{U})$ est le degré de libertés de $\mathbf{X}$ et comme le degré de liberté des directions est réduit par les contraintes, pour chaque bâtiment B le degré de liberté total est donné par :

$$d(B) = 2 * D - \sum_{c \in \mathcal{C}} f(c) + M \tag{6.12}$$

Sans aucune contrainte, le nombre de degrés de liberté global des bâtiments est de 966.

6.5.2 Contraintes a priori

Dans ce cas, comme le montre la figure 6.6, la reconstruction est très proche de la modélisation attendue. Le lecteur trouvera en section 7.2.5 et en particulier sur la figure 7.4 d'autres résultats attestant du gain évident obtenu par les reconstructions avec contraintes. La figure 6.7 montre des résultats plus quantitatifs. On remarque que la distance est toujours plus faible sans contraintes, ce qui prouve que les contraintes naturellement rapprochent le modèle bruité de la référence. Dans le cas de contraintes a priori, la diminution du nombre de degrés de liberté est quant à elle uniquement due à des surcontraintes ou dégénérescences détectées au cours du processus et qui ont conduit à la suppression de contraintes. Une grande stabilité au bruit peut être observée pour la reconstruction a priori qui reste insensible jusqu'à $\sigma = 0.5m$. Notons qu'une évaluation visuelle des degrés de liberté des bâtiments a établi que 258 était le nombre attendu de degrés de liberté, ce qui est la valeur déterminée par l'algorithme. Ceci prouve qu'aucune contrainte n'a été perdue durant la minimisation et particulièrement durant l'étape de construction du graphe de dépendances sur la référence ce qui valide de facto nos heuristiques.

6.5.3 Contraintes a posteriori

Bien que toutes les contraintes ne soient plus détectées dès que le bruit augmente, la forme générale des bâtiments est bien plus acceptable que sans les contraintes (figure 6.6). En ce qui concerne le degré de liberté, on observe une relative stabilité jusqu'à $\sigma = 0,3m$ où seulement 10% du nombre de degrés de liberté est perdu. Notons que $\sigma = 0,3m$ est un bruit important en particulier quand des arêtes d'un mètre sont présentes dans la scène.

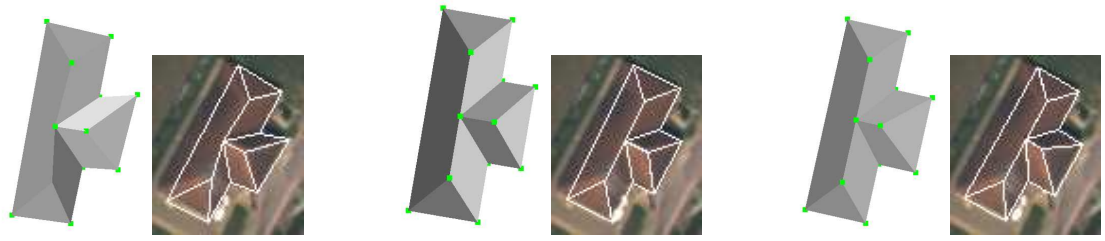
6.6 Conclusion

6.6.1 Apports

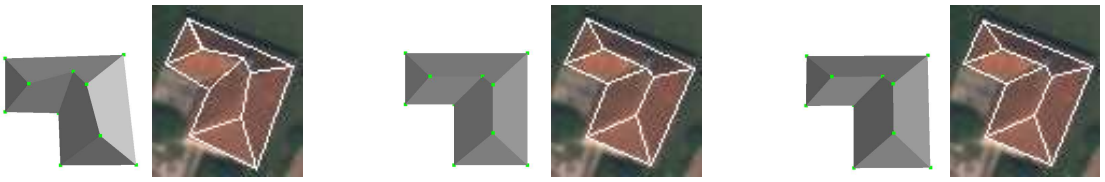
Nous avons présenté dans ce chapitre une méthode *automatique* de recalage de bâtiments avec contraintes. Cet algorithme est automatique dans le sens où les contraintes sont inférées et imposées sur le modèle sans aucune intervention manuelle, y compris pour la définition des contraintes. L'approche n'utilise que 2 seuils aisés à régler et permet par une paramétrisation implicite du modèle de contraintes de savoir exactement quelles contraintes sont réellement appliquées. Les situations de surcontraintes sont détectées puis gérées par suppression itérative de contraintes. Les résultats ont montré le gain obtenu en termes de précision et degré de liberté du modèle. Ainsi, l'algorithme a permis de réduire le nombre de degrés de liberté global de 966 à 258, soit un gain d'environ 75%. En outre, la reconstruction avec contraintes élimine un grand nombre d'artefacts visuels.

6.6.2 Extension

Une amélioration consisterait à recaler le modèle contraint *dans les images*. Une première étude a porté sur le recalage de points de jonction 3D dans les images en utilisant une généralisation 3D du détecteur développé par [FG87] (voir annexe B). Ce détecteur est toutefois fortement perturbé par le bruit dans le voisinage radiométrique ce qui limite son intérêt dans les zones urbaines où la proximité de véhicules, végétation induit nécessairement des contours forts. Les résultats décevants conduisent à considérer le problème de façon globale pour essayer de recaler le modèle filaire complet dans les images. De premiers travaux [BR97] ont montré la validité d'une telle approche. Une étude est en cours pour l'adaptation de ces techniques à notre problématique. Même si ce recalage image paraît indispensable pour une bonne précision géométrique, on peut s'interroger sur la possibilité d'un recalage d'une *caricature* de modèle dans les images qui représentent elles une information *parfaite* sans modélisation caricaturale. Nous pensons cependant que les

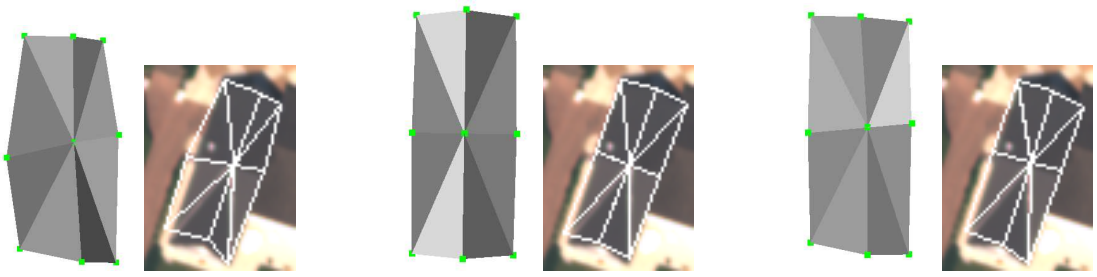


(a) 48 / 10 / 11



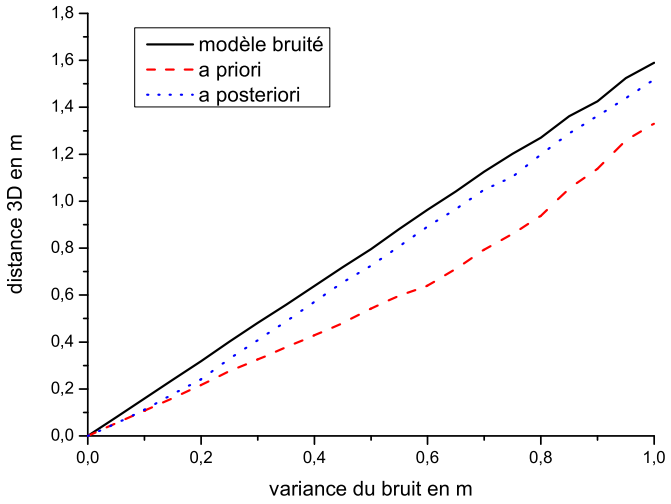
(b) 42 / 10 / 10

(c) 54 / 14 / 15

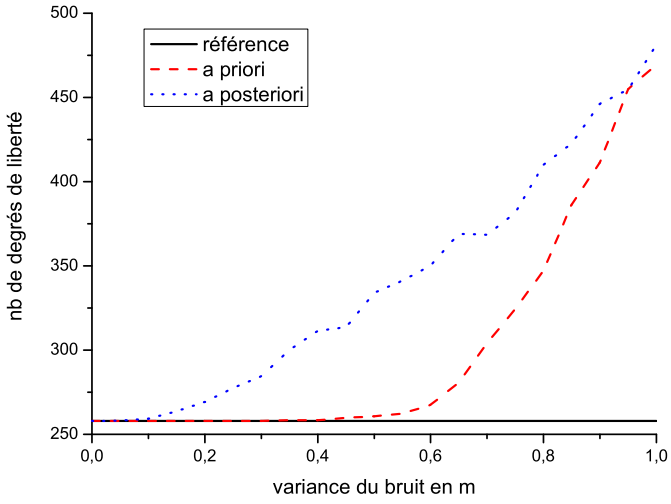


(d) 39 / 7 / 12

Figure 6.6 - Différentes vues de quelques bâtiments avec $\sigma = 0.5m$. 2 premières colonnes : vues de bâtiments bruités. 2 colonnes suivantes : vues de la mise en application de contraintes a priori. 2 dernières colonnes : vues de la mise en application de contraintes a posteriori. Les 3 nombres indiquent le nombre de degrés de liberté sans contraintes, pour le bâtiment reconstruit avec contraintes a priori et pour le bâtiment construit avec les contraintes a posteriori respectivement.



(a)



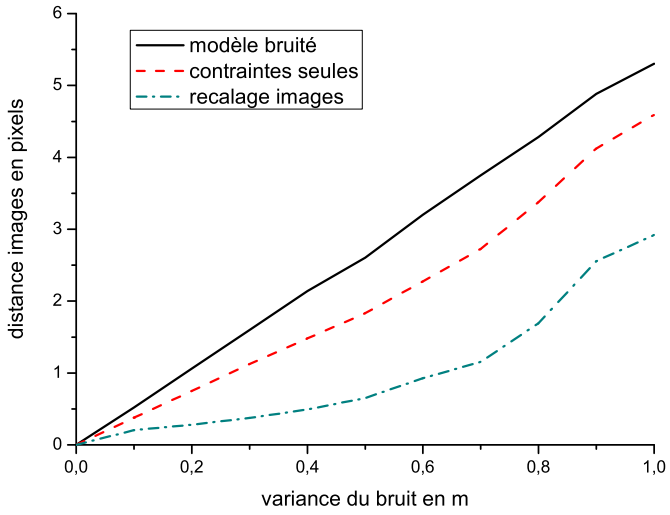
(b)

Figure 6.7 - Gain apporté par le recalage avec contraintes. (a) : distance à la référence en m. (b) : degrés de liberté.

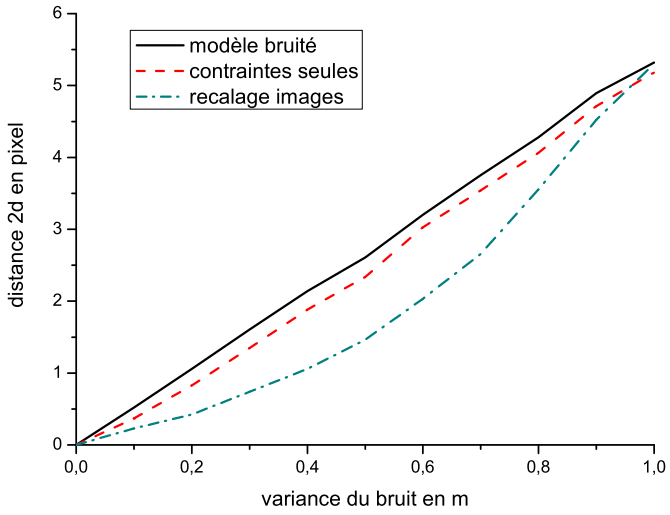
informations déterminantes (arêtes en particulier) constituant la caricature sont présentes même partiellement dans les images et que de fait le recalage a un sens. Les premiers résultats ont semblé confirmer cette conjecture [Val04]. La stratégie mise en oeuvre est une extension naturelle des travaux présentés dans ce chapitre et des techniques de mise en correspondance de modèles 3D dans les images exposées dans [BR97]. Le lecteur pourra se référer à [Val04] pour plus de détails, nous ne reproduisons sur la figure 6.8 que les premiers résultats qui dénotent une amélioration sensible de la précision géométrique avec le recalage image.

6.6.3 Perspectives

Indépendamment de son utilisation dans le cadre d'une reconstruction tout automatique de bâtiment, cet algorithme peut avantageusement être utilisé dans des systèmes semi-automatiques voire manuels de saisie de scènes urbaines. L'objectif à terme est de l'intégrer dans une plate-forme de restitution semi-automatique de bâtiments [FMJ03] pour, en aval, tirer parti des contraintes internes des bâtiments.



(a) contraintes a priori



(b) contraintes a posteriori

Figure 6.8 - Gain apporté par le recalage avec contraintes dans les images. On mesure la distance à la référence dans les images utilisées pour le recalage.

Quatrième partie

Évaluation

Chapitre 7

Résultats

7.1 Introduction

En contraste avec la très grande quantité de références portant sur des méthodes automatiques et semi-automatiques pour la reconstruction de bâtiments à partir d'images aériennes ou de données laser, il n'y a que très peu d'articles s'intéressant à l'évaluation *automatique* qualitative et quantitative de ces méthodes en dehors d'une appréciation visuelle très subjective. L'évaluation des méthodes de reconstruction est cependant une étape indispensable pour mettre en évidence les apports et les défauts de la chaîne et déterminer, le cas échéant, les aspects à améliorer ou les erreurs à signaler pour permettre l'intervention d'un opérateur. Cette évaluation est également essentielle pour comparer la méthode avec des techniques déjà existantes et ainsi convaincre un utilisateur de l'intérêt d'une approche donnée dans une chaîne de production de grande envergure.
Une partie des études se focalise sur la précision géométrique [JDA95, SSF96, HB97, HMS+01, SW03] pour laquelle l'écart moyen résiduel est systématiquement utilisé pour les coordonnées des points ou pour les valeurs des altitudes par rapport à un MNE de référence.
Une autre partie des travaux se penche sur l'évaluation de la détection de bâtiments et utilise taux de détection et autres mesures de qualité rattachées ([MS94, Wei97, SW03]). Ils s'inspirent ainsi directement du travail précurseur de [MS94]. Ces études s'intéressent aux bâtiments *en tant qu'objets* et cherchent à déterminer les bâtiments manquants par rapport à la référence ou les objets assimilés aux bâtiments dans la reconstruction et qui ne sont pas répertoriés comme tels dans la référence. Dans ces travaux, pour chaque bâtiment de la référence, la question posée est extrêmement simple : y-a-t'il un bâtiment reconstruit correspondant ? Inversement, certains s'intéressent à la question symétrique : pour chaque bâtiment reconstruit, y-a-t'il un bâtiment correspondant dans la référence. On retrouve ici les notions de "False Positive" ou "True Negative" telles que nous les avons utilisées dans le chapitre 3. Toutefois, ces techniques ne permettent pas de quantifier *l'adéquation des formes de modèles ni de prendre en compte les notions de généralisation.*
Seul [HB97] se penche sur l'aspect "respect des formes" et propose une mesure de qualité pour mesurer la dissimilarité entre la forme de référence et la forme reconstruite. Cependant, ses mesures ne concernent que des aspects de recouvrement de la référence par la reconstruction et non pas à proprement parler de respect de la forme globale (préservation des angles, topologie cohérente) pour lequel les auteurs restent très évasifs :

> *all reconstruted buildings actually look like buildings and their shape is close to the reference data, i.e. the human interpretation.*

En dehors de cet article, aucune étude y compris dans les plus récentes et les plus exhaustives [SW03] ne s'intéresse au *respect des formes* en dehors de tout aspect de détection proprement dit tel qu'il a été évoqué. L'évaluation de ce respect des formes reste extrêmement difficile en raison du phénomène de généralisation qui autorise certaines caricatures de représentation. De fait, un modèle peut être accepté même s'il ne cor-

respond pas exactement à la réalité. Cependant, toutes les caricatures ne sont pas acceptables comme nous l'avons déjà évoqué en section 2.1.1 et l'évaluation automatique est donc extrêmement délicate. La complexité de cette évaluation automatique explique qu'aucune étude quantitative n'a été faite pour quantifier ce *respect des formes*. Seul [Fuc01] a proposé une approche fondée sur une évaluation visuelle permettant de prendre en compte la généralisation dans le processus d'évaluation.
La suite de ce chapitre sera consacrée dans un premier temps (section 7.2) à une analyse qualitative, essentiellement pour mettre en évidence sur exemples réels, les possibilités de l'algorithme (bâtiments complexes, bâtiments avec discontinuités internes, apports des contraintes, possibilité de généralisation). Dans la suite (section 7.3), nous ferons une étude plus approfondie du respect structurel de notre méthode de reconstruction, c'est-à-dire du respect des formes 3D par rapport à la réalité en tenant compte des problèmes de généralisation. Compte tenu de la complexité de cette phase d'évaluation, à l'instar de ce qui est fait dans [Fuc01], nous fournirons pour différentes valeurs de β une évaluation basée sur une inspection visuelle d'une centaine de bâtiments. L'annexe E permettra au lecteur d'avoir un aperçu des différents résultats ainsi que des références utilisées pour cette comparaison. Enfin, une évaluation géométrique sera faite en section 7.4. Assez classiquement, elle est basée sur une évaluation de la précision de coordonnées ponctuelles. Pour les deux derniers aspects qui fournissent une évaluation plus quantitative, les protocoles d'évaluation sont dûment expliqués en début de chaque section.

7.2 Quelques résultats qualitatifs

Cette section cherche à mettre en évidence les avantages de la méthode décrite pour un certain nombre de configurations courantes en milieu urbain. Elle s'efforce aussi de mettre en avant les apports d'une telle méthode sur des méthodes à base de modèles logiquement plus robustes mais moins générales. On montrera également, toutefois, les analogies que l'on peut voir apparaître entre ces deux types de stratégies, en jouant sur le terme déterminant le niveau de généralisation souhaité.

7.2.1 Quelques bâtiments à la loupe

Le lecteur trouvera dans l'annexe D une description des différentes étapes de l'algorithme sur quelques bâtiments typiques de zones péri-urbaines. Sur des zones plus étendues, les résultats sont présentés en annexe E pour différents réglages du paramètre β. Dans la suite de ce chapitre, nous reviendrons sur ces différents résultats.

7.2.2 Bâtiments avec discontinuités altimétriques internes

La méthode décrite permet de traiter des discontinuités internes aux bâtiments. Ainsi, l'algorithme a correctement reconstruit les bâtiments accolés présentés en figures D.5 et D.6. Ce type de discontinuités n'est pas rare en milieu péri-urbain et met généralement en échec toutes les méthodes "bottom-up" qui présupposent très souvent un polyèdre continue de la forme $z = f(x, y)$ ([FKL$^+$98, Jib02, EBM02] par exemple). De même, comme pour la méthode présentée par [Fuc01], notre algorithme peut gérer des défauts de focalisation et représenter des bâtiments distincts sur une seule et même zone comme l'atteste la figure D.7. Théoriquement donc, on pourrait reconstruire des ensembles de bâtiments importants tels qu'on peut en trouver en zones urbaines denses. Cependant, l'augmentation substantielle de combinatoire que génèrent ces défauts de focalisation rend pratiquement ce genre de tentative utopique.

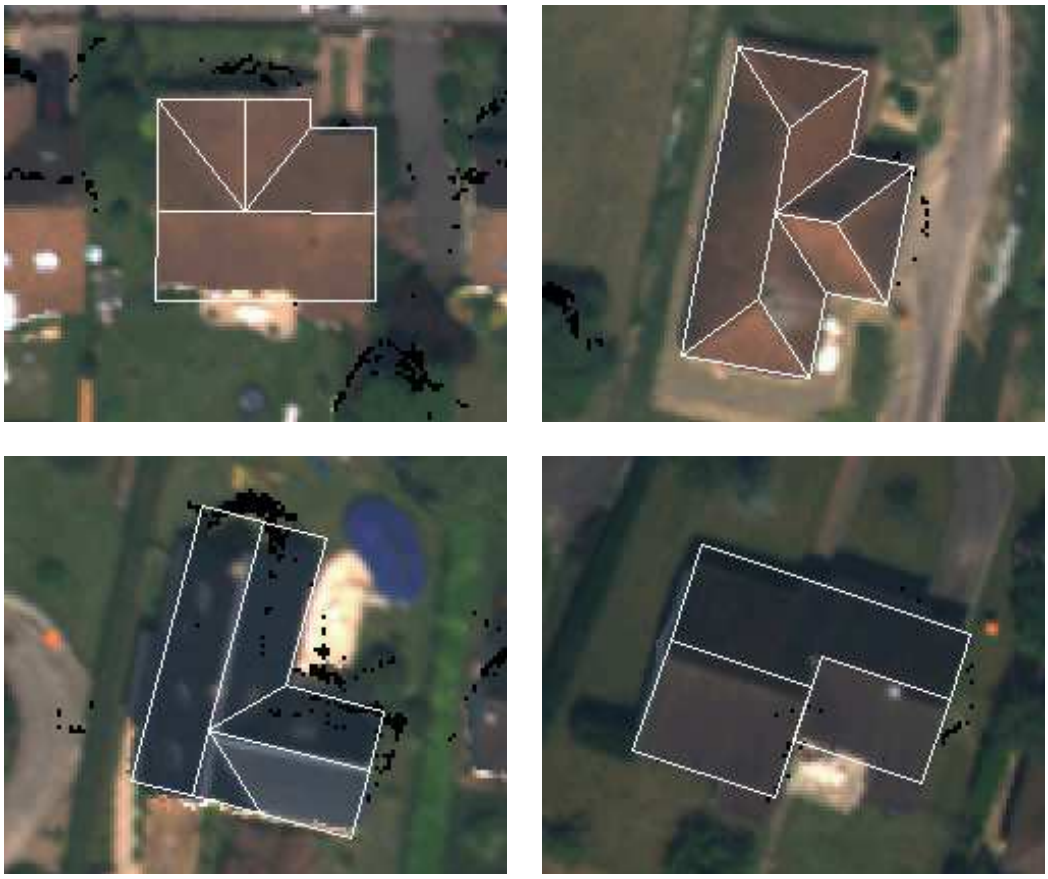


Figure 7.1 - Exemples de bâtiments complexes reconstruits avec succès par l'algorithme.

7.2.3 Bâtiments complexes

Contrairement aux approches à base de modèles, la méthode n'est pas limitée à un ensemble de formes, mais peut reconstruire des structures arbitrairement complexes. La figure 7.1 montre quelques modèles complexes de bâtiments que l'on peut trouver en milieu péri-urbain. On remarquera que les premiers modèles ne sont pas aisément modélisables par une approche modèles telle que celle de [Fuc01]. Et quoi qu'il en soit, dans ce type d'approches, accroître la base de modèles altère à la fois la robustesse de l'approche et le temps nécessaire pour la reconstruction de bâtiments.

7.2.4 Niveau de généralisation

Nous présentons ici qualitativement avant d'en faire une étude plus approfondie dans la suite (section 7.3.2) l'influence du paramètre β sur le niveau de généralisation de la représentation. Un β élevé permet une description plus fine des détails, mais rend le processus moins robuste aux erreurs des détecteurs de primitives (en particulier le détecteur de segments dont l'influence dans le choix du modèle est importante). En revanche, un β faible donne une représentation très caricaturée de la réalité au détriment du détail (figure 7.2). Ce choix dépend bien entendu de l'application envisagée et des buts fixés : niveau de détail, compacité de la représentation...

(a) $\beta = 10$

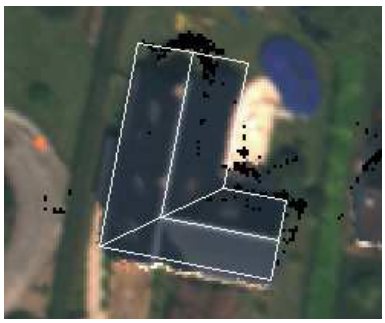
(b) $\beta = 50$

(c) $\beta = 75$

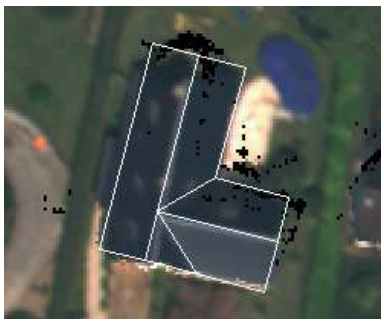
(d) $\beta = 10$

(e) $\beta = 75$

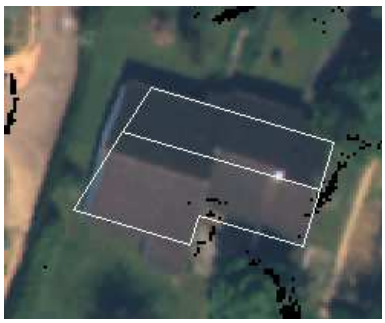
(f) $\beta = 150$



(g) $\beta = 10$



(h) $\beta = 50$



(i) $\beta = 10$

(j) $\beta = 75$

Figure 7.2 - Rôle du paramètre β dans le niveau de généralisation.

7.2.5 Contraintes

Les contraintes sont utilisées à la fois pour choisir le meilleur modèle en regard d'un certain nombre d'hypothèses architecturales rudimentaires (arêtes horizontales, symétrie verticale, parallélisme, orthogonalité) et pour le recaler dans les images en utilisant cette information afin d'obtenir une représentation plus conforme à l'intuition. Les figures 7.3 et 7.4 illustrent l'intérêt des contraintes dans ces deux étapes. Il est évident, au vu des résultats, que l'algorithme manque d'une fonction permettant d'intégrer une information image dans le recalage. Cette étude est en cours et devrait sensiblement améliorer les résultats en ce qui concerne la précision.

7.3 Évaluation structurelle

7.3.1 Protocole d'évaluation

À l'instar de ce qui est fait dans [Fuc01], nous cherchons ici à déterminer les capacités de l'algorithme à reconstruire la forme correcte du bâtiment ou, le cas échéant, une généralisation acceptable de celui-ci. Nous définissons dans ce qui suit les critères d'évaluation utilisés ainsi que les critères de généralisation tolérée.
Toutes les zones étudiées, au nombre de 113, comportent un seul bâtiment et se situent en milieu périurbain. Du fait de la combinatoire de l'exploration de toutes les représentations possibles et de l'absence d'heuristiques pour la contrer, le système n'est en effet pas prêt pour le milieu urbain dense compte tenu de la complexité des formes et des erreurs inévitables de focalisation dans ce milieu.
Sur chaque zone, chaque bâtiment reconstruit est classé dans une des 4 catégories suivantes :

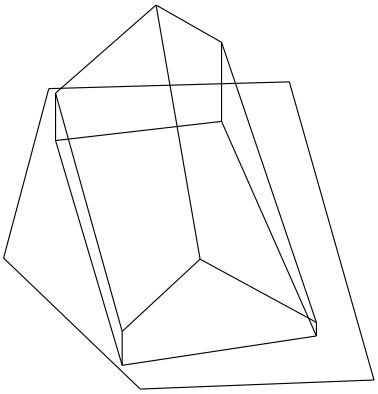
Correct Il s'agit d'un bâtiment dont la reconstruction est rigoureusement conforme à la réalité, incluant donc tous les décrochements, chiens-assis, même peu importants.

Généralisé Il s'agit d'un bâtiment correspondant à une caricature acceptable de la réalité. Nous tolérons, comme annoncé précédemment un niveau de généralisation de 1,5m. Une représentation correspondant à une caricature de la réalité dans la limite de ce niveau de généralisation sera donc classée dans la catégorie "généralisé".

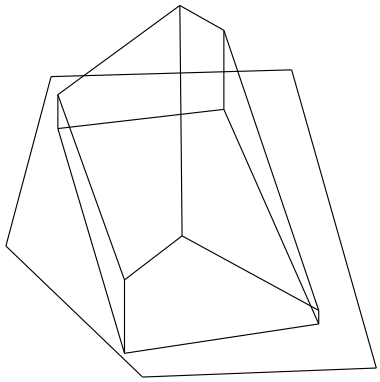
Surgénéralisé Cette catégorie un peu particulière regroupe toutes les représentations qui correspondent à une caricature de la réalité "acceptable" (au sens du respect de la forme générale) mais dans un niveau de généralisation supérieur à 1,5m. En fonction de l'application et du degré souhaité de fidélité à la réalité, ces représentations peuvent être acceptées ou refusées. Nous utilisons cette catégorie pour mettre en exergue les capacités de l'algorithme à fournir une caricature de la zone.

Faux Il s'agit de représentations qui ne correspondent à aucune caricature de la réalité quel que soit le niveau de généralisation.

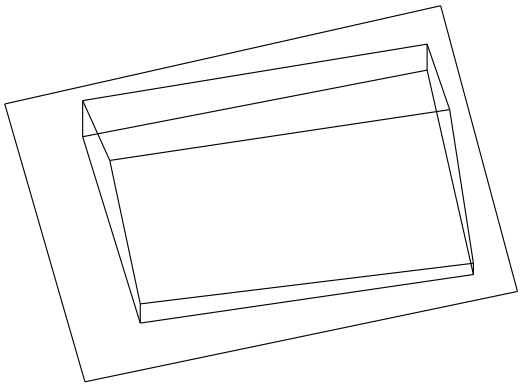
Tous ces éléments fixent les objectifs de reconstruction du système et nous chercherons à montrer par ces catégories que l'application, si elle n'est pas toujours capable de produire une représentation strictement conforme à la réalité, pour diverses raisons qui seront détaillées en section 7.3.3, peut fournir une caricature acceptable de cette réalité.
Il est entendu dans cette section que nous ne nous intéressons qu'à la *forme générale* du bâtiment en dehors de toute considération de précision ou de contraintes sur les primitives. Nous calquons ainsi notre analyse sur celle faite par [Fuc01] qui s'intéressait dans une première étude uniquement à la forme des bâtiments.
Dans la section suivante, une étude détaillée de l'influence du paramètre β qui règle le compromis entre adéquation aux données et simplicité du modèle est effectuée, afin, entre autres, de régler ce paramètre primordial.



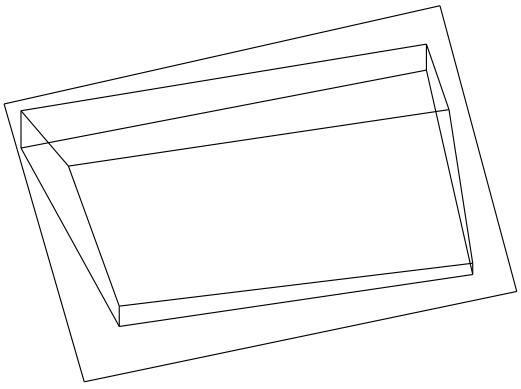
(a) solution n°1 - $\mathcal{L}_M = 336.35$



(b) solution n°2 - $\mathcal{L}_M = 348.35$



(c) solution n°1 - $\mathcal{L}_M = 324.35$



(d) solution n°2 - $\mathcal{L}_M = 374.301$

Figure 7.3 - Rôle des contraintes dans le choix du meilleur modèle. Dans le premier exemple, c'est la contrainte de symétrie verticale qui assure le choix de la solution 1, dans le deuxième, le parallélisme et l'orthogonalité expliquent la différence de longueur de description entre les deux premières solutions.

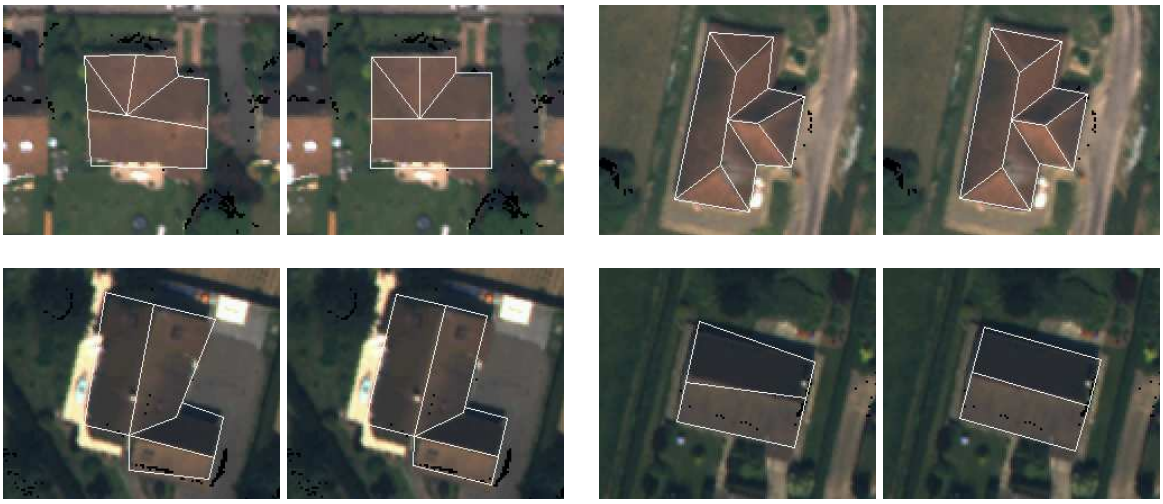


Figure 7.4 - Importance des contraintes dans le recalage.

7.3.2 Influence du paramètre beta

Le graphique 7.5 présente les évaluations en fonction de différentes valeurs de β. On constate que pour des β relativement faibles, le taux de reconstructions acceptables ("correct" + "généralisé") oscille entre 69% et 72,5% avec un score maximum pour $\beta = 25$. Si on accepte les formes surgénéralisées, on obtient un score maximum de 83%. Au-delà de la simple estimation de paramètre, ce graphique montre également toute l'importance du facteur modèle dans l'algorithme. Il est en effet intéressant de constater une rapide décroissance du taux de reconstructions acceptables voire surgénéralisées avec l'accroissement de β ce qui démontre l'importance du terme $P(M)$ dans le choix du bon modèle. On remarquera en particulier que lorsque l'algorithme ne repose que sur le terme d'adéquation aux données ($\beta \rightarrow +\infty$), les résultats acceptables sont inférieurs à 35%. L'analyse du graphique montre une décroissance très forte du taux de formes surgénéralisées avec l'augmentation de β. Cette évolution est relativement logique : les formes surgénéralisées et, dans une moindre mesure les formes généralisées, par définition, représentent une caricature et ne s'appuient donc pas sur des éléments réels présents dans l'information image ou les segments 3D. Par conséquent, ces représentations dépendent essentiellement du terme $P(M)$ dont l'importance est donc une fois de plus soulignée pour le choix de ces formes caricaturées.

7.3.3 Sources d'erreurs

Les sources d'erreurs sont directement liées aux primitives. Dans la majorité des cas, il s'agit de cas de figure où l'on ne respecte pas les hypothèses de base de l'algorithme de reconstruction de bâtiments : primitives importantes non détectées, en particulier les façades (zone2, bâtiments 8 et 39 par exemple), primitives mal localisées en géométrie (zone1, bâtiments 1 et 5). Quelques rares erreurs sont également dues à la présence de segments 3D en correspondance par hasard avec des intersections de plans.
Dans l'ensemble, le principal détecteur incriminé dans les sources d'erreurs est le détecteur de façades. L'introduction de limites cadastrales devrait sensiblement atténuer les problèmes dus à ce détecteur.

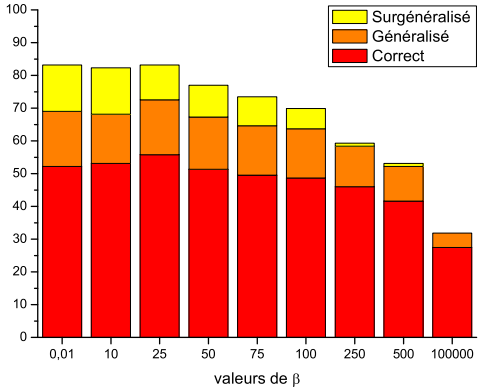


Figure 7.5 - Influence du paramètre β sur la classification des bâtiments, en $\%$.

7.4 Évaluation de la précision

7.4.1 Protocole d'évaluation

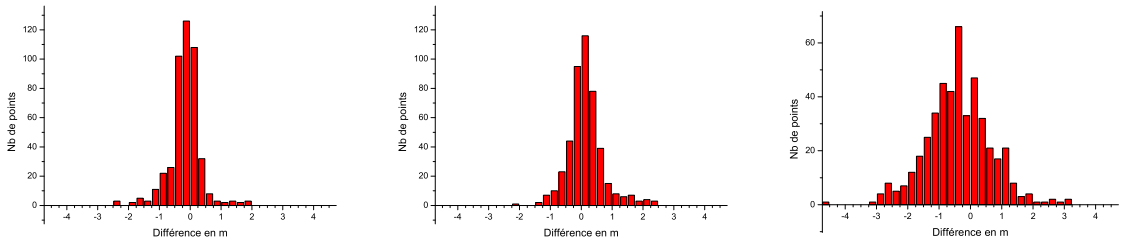
L'évaluation de la précision est faite pour une seule échelle, à savoir $\beta = 25$ qui fournit les meilleurs résultats. Pour chaque bâtiment "correct" ou "généralisé", la position exacte des points reconstruits *correspondant à une réalité physique*[1] est évaluée, par un opérateur, par pointé manuel dans $d > 2$ images.
Concrètement, les modèles reconstruits sont projetés sur une orthophotographie. L'opérateur sélectionne par une sélection dans cette orthophotographie les points dont il souhaite estimer la localisation réelle. Les n images de la zone sont alors centrées sur ce point *sans qu'aucune projection du point reconstruit ne soit faite* pour ne pas biaiser la saisie. Par un clic dans d images dans lesquelles le détail est visible, l'opérateur détermine le point de référence.

7.4.2 Résultats

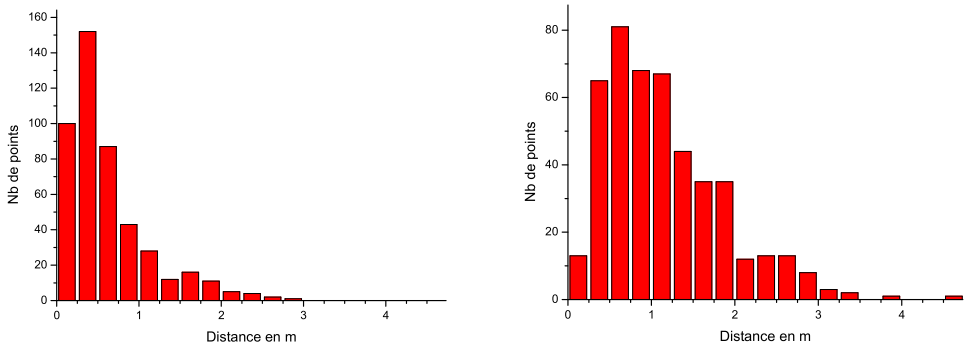
Les résultats sont présentés dans les figures 7.6 qui illustrent les histogrammes des écarts en X, Y, Z, en distance planimétrique et en distance 3D sur 468 points. Ces résultats montrent une précision satisfaisante avec un écart type sur les coordonnées planimétriques de 50cm en X et de 57cm en Y soit moins de 2,2 fois la résolution planimétrique des images. En altimétrie, on obtient une précision de 1,03 m ce qui correspond à 1,65 fois la précision altimétrique nominale, compte tenu de la résolution des clichés (25cm) et du B/H (0,4). Nous estimons que ces résultats sont corrects et témoignent d'un accroissement sensible de la précision par rapport à la stéréoscopie simple (pour laquelle [Fuc01] obtenait 9 fois la précision planimétrique[2]).
Toutefois, certains points sont à plus de 3m de la position physique du détail qu'ils sont censés matérialiser. Indiscutablement, si le recalage avec contraintes apporte un gain en précision, la chaîne de traitement souffre encore de ne pas utiliser dans cette étape une information image pour améliorer encore cette précision.
On notera également un biais sur les écarts en X, Y, Z : -16cm en X, 15cm en Y et -38cm en Z. Un rapide test de Student permet de vérifier que ces valeurs, bien qu'inférieures aux résolutions planimétriques et altimétriques des images, représentent bien un biais. Ce biais s'explique probablement par le fonctionnement

[1] nous excluons de fait tous les points reconstruits correspondant à une généralisation de la forme et qui ne sont donc que des points virtuels, sans homologues dans les images.

[2] La comparaison entre les résultats doit se faire avec prudence, car la philosophie des deux systèmes est totalement différente : l'approche modèle utilisée par [Fuc01] sera inévitablement moins précise que la nôtre qui repose directement sur les primitives.



(a) Histogramme des écarts en X. (b) Histogramme des écarts en Y. (c) Histogramme des écarts en Z.



(d) Histogramme des distances planimétriques. (e) Histogramme des distances 3D.

Figure 7.6 - Évaluation de la précision géométrique.

de l'algorithme et les ombres portées par les bâtiments. Ainsi, les primitives planes et en particulier les façades sont déduites des transitions radiométriques présentes dans les images par le biais de l'orthoimage de gradient et de la projection des segments 3D. La présence de limites d'ombre proches des façades réelles de bâtiments dans les images peut biaiser le positionnement des primitives verticales extraites qui sont, avec les facettes non verticales, à la base de la construction des modèles de bâtiment. Pour les zones qui nous intéressent, la position du soleil indique que le vecteur solaire, c'est à dire la direction des rayons du soleil est $\mathbf{V}_s = (-0,264783; 0,459127; -0,847993)$ soit, en projection verticale après normalisation $\mathbf{v}_s = (-0,499; 0.866)$. Ces chiffres confirment bien la tendance à un biais négatif en X, positif en Y. Le biais négatif en Z s'expliquant par le fait qu'un mauvais appariement avec une limite d'ombre a naturellement tendance à écarter la primitive verticale de sa position réelle par rapport aux bâtiments et de fait, pour la majorité des bâtiments aux toits en pente, à baisser l'altitude de la gouttière.

7.5 Conclusion

Ce chapitre a étudié les caractéristiques de l'algorithme, aussi bien qualitativement en mettant en exergue les possibilités de la méthode, que quantitativement en validant le système sur des images réelles représentant une centaine de bâtiments. Une étude a montré l'influence du paramètre réglant le compromis entre

adéquation aux données et simplicité du modèle et l'importance de l'introduction de données externes pour le choix de la meilleure représentation.
L'algorithme a montré ses possibilités, sur données réelles, en matière de modélisation de bâtiments complexes, présentant ou non des discontinuités internes. L'apport des contraintes a également été noté aussi bien dans la phase de choix du modèle que dans celle du recalage a posteriori. Nous souhaitons insister sur l'intérêt de ces dernières qui permettent, sans perte de généralité, l'introduction de modèles architecturaux simples dans la modélisation de bâtiments.
Quantitativement, environ 72,5% des représentations sont acceptables, ce qui est un résultat satisfaisant même si de nombreux aspects sont perfectibles. Dans bien des cas (environ 17%), la représentation n'est pas une exacte modélisation de la réalité mais une caricature acceptable, ce qui met en avant les capacités de généralisation de l'algorithme, *en dehors de toute base de modèles*. Le principe "la géométrie induit la topologie" est ici utilisé pour parvenir à ces possibilités de généralisation.
Sur le plan géométrique, le système atteint une précision de l'ordre de 2,2 fois la résolution planimétrique des images et 1,7 fois la résolution altimétrique, ce qui est correct même si le système manque d'un retour image pour améliorer encore cette précision.
Globalement, ce système possède donc des possibilités intéressantes, essentiellement parce qu'il ne repose sur aucune modélisation déclarative d'un bâtiment et peut donc virtuellement modéliser n'importe quel bâtiment voire ensemble de bâtiments et parce que, cependant il permet, le cas échéant, de proposer des modélisations généralisées de la réalité.

Chapitre 8

Conclusion

Nous pensons que la méthode proposée répond à un certain nombre des objectifs fixés dans le cadre de la reconstruction automatique de bâtiments en milieu urbain. Pour autant, elle présente certaines lacunes qui empêchent sa mise en production immédiate. Dans un premier temps, nous analyserons donc notre contribution, en mettant en évidence les faiblesses et limites de l'approche proposée, mais aussi en soulignant ses points forts qui constituent, nous l'espérons, autant d'arguments pour la poursuite de l'exploration de la stratégie décrite dans ce mémoire. Nous verrons également que les perspectives d'évolution sont nombreuses et pourraient combler tout ou partie des problèmes importants pour éventuellement une utilisation dans le cadre de la constitution d'une base de données 3D urbaines.

8.1 Analyse

8.1.1 Inconvénients

Focalisation manuelle

Ceci constitue certainement le point le plus critique pour le développement d'une chaîne de production basée sur le système actuel. Volontairement, nous avons écarté la focalisation automatique pour nous consacrer au problème de la reconstruction complète d'un bâtiment générique. Néanmoins, des outils de focalisation automatique sur des zones de sursol [BSZF99, CJC01] devraient être testés et évalués en conjonction avec le système proposé. Nous soulignerons à cet effet que la possibilité pour la méthode de reconstruire plusieurs bâtiments sur une même zone de focalisation devrait permettre de pallier les erreurs de focalisation inévitables dans les approches automatiques.

Dépendance aux primitives extraites

C'est un point prévisible et prévu de l'approche. La généralité de la méthode s'obtient inévitablement au détriment d'une dépendance aux données initiales pour la construction des hypothèses. C'est indéniablement un inconvénient par rapport aux approches strictement basées modèles, son pendant étant la généralité. De fait, cette dépendance nécessite de travailler avec des résolutions relativement élevées pour pouvoir extraire les primitives pertinentes. Nous pensons que 25cm de résolution image est un minimum et qu'une résolution de 10-15cm serait idéale. Toutefois, soulignons que cette dépendance reste limitée à la géométrie des primitives extraites et non pas à leur délimitation, ce qui limite l'influence des détecteurs de primitives dans la chaîne.

Gestion de la combinatoire

À l'heure de la rédaction de ce manuscrit, des méthodes pour contrer la combinatoire inhérente à la recherche des cliques maximales sont à l'étude. La solution à ce problème est indispensable pour pouvoir gérer des bâtiments ou agrégats de bâtiments complexes dans des zones urbaines denses.

Absence de système d'alertes

Cette absence pénalise grandement la stratégie tout automatique dans la mesure où elle impose un contrôle visuel systématique de toutes les reconstructions par un opérateur. L'hypothèse d'une industrialisation du processus passerait nécessairement par le développement d'un tel outil. Il pourrait s'appuyer sur la longueur de description du modèle $\mathcal{L}(M)$ dont les valeurs importantes traduisent souvent des modèles complexes et faux.

8.1.2 Apports

Reconstruction de segments 3D en multi-vues

Nous avons présenté une méthode de reconstruction de segments 3D répondant aux critères du multi-vues. Cette méthode a déjà porté ses fruits dans le cadre d'une caricature de surface par triangulation contrainte [MPT02]. Nous avons également montré l'apport du multi-vues dans cette étape.

Généralité

Nous avons proposé une méthode originale de détection de polyèdres sans surplomb à partir d'un ensemble de primitives planes orientées, basée sur une extension des travaux de [Jib02]. La démonstration théorique de la validité de l'algorithme d'extraction d'hypothèses de polyèdres sans surplomb constitue un des aspects importants de cette thèse. Cette modélisation assure aussi la généralité de l'approche qui permet, virtuellement, de reconstruire la quasi-totalité des bâtiments en milieu urbain, en opposition aux approches à partir de librairies de modèles. Nous soulignons ici que la méthode permet théoriquement de reconstruire un *ensemble* de bâtiments sur une même zone comme l'atteste l'exemple D.7 présenté en annexe D.

Système ouvert et contraint

Nous pensons que la modélisation bayésienne du choix du modèle de bâtiment est une autre contribution importante de l'approche proposée pour deux raisons essentielles. La première réside dans le fait que le système reste ouvert et il est ainsi aisé d'introduire de nouvelles primitives dans la formulation. La deuxième est la possibilité d'utiliser des contraintes pour guider le choix du modèle. Cet aspect est particulièrement novateur et permet l'introduction d'informations architecturales externes dans le processus de décision sans toutefois nuire à la généralité du système.

Capacités de généralisation

L'introduction d'informations externes permet aussi au système d'avoir des capacités de généralisation. Ces capacités reposent également sur la stratégie choisie d'induire la topologie à partir de la géométrie, comme illustré dans le chapitre 2. Cette faculté de proposer, le cas échéant, une caricature acceptable de la réalité permet au système de s'affranchir de certaines occultations telles que la végétation ou d'autres bâtiments, d'éliminer certains détails inutiles dans la modélisation (chiens-assis, cheminées, petits décrochements...). Cet aspect représente une analogie avec les systèmes reposant sur une librairie de modèles qui

étaient, jusque-là, les seuls à pouvoir proposer une forme généralisée de la réalité. Ces capacités de généralisation représentent l'une des contributions essentielles du système mis en oeuvre. Soulignons aussi que le système peut fournir plusieurs solutions à différentes échelles comme cela a été montré (section 7.2.4). Cet aspect peut revêtir une importance particulière dans un processus interactif en laissant à un opérateur externe le choix du modèle final.

Recalage avec contraintes

La technique de recalage avec contraintes proposée est générique et ne nécessite pas l'intervention de l'opérateur pour le choix des contraintes. Elle peut s'utiliser dans un module de recalage a posteriori (les contraintes ne sont pas connues initialement : c'est le cas du système décrit dans ce manuscrit) ou a priori (les contraintes sont connues a priori ce qui est le cas des approches à base de modèles).
En outre, nous avons présenté de premiers résultats très prometteurs pour le recalage avec contraintes dans les images. Cette stratégie permet de revenir à l'information initiale et devrait améliorer sensiblement les résultats présentés en terme de précision.

8.2 Perspectives

8.2.1 Optimisation

L'optimisation représente probablement la première chose à étudier pour élargir les possibilités de l'algorithme. Beaucoup de travaux ont déjà abordé l'utilisation d'heuristiques dans la recherche des meilleures cliques maximales. On pourra se référer à [BBPP99] pour plus de détails concernant les stratégies employées. Toutes ces techniques utilisent des choix successifs plus ou moins judicieux de noeuds du graphe pour trouver *la* meilleure clique maximale possible sans explorer l'ensemble des solutions. Ici, la décomposition des facteurs d'évaluation du modèle en termes additifs (section 5.5.1, au moins en ce qui concerne l'évaluation de son adéquation aux données) simplifie grandement cette étape : le choix itératif pouvant s'effectuer en fonction du poids de chaque facette.
Il serait pertinent, à ce niveau de ne pas se contenter d'une seule solution et donc de l'exploration d'une seule branche dans l'arbre des possibles, mais de pouvoir proposer des modèles *suffisamment distincts* pour le choix de la représentation finale. A ce titre, la stratégie développée par [BP01] nous paraît prometteuse.

8.2.2 Introduction d'une librairie de modèles

Les librairies de modèles fournissent une information sémantique que les systèmes à base de primitives ne peuvent donner (bâtiment en "L", bipan symétrique...). Sans perdre la généralité de notre approche, il serait intéressant de confronter la ou les premières solutions proposées à une librairie de modèles qui, le cas échéant, attribuerait à la solution une information sémantique supplémentaire.

8.2.3 Introduction des emprises planimétriques

L'introduction des emprises planimétriques des bâtiments telles que celles fournies par une vectorisation automatique du cadastre est une perspective très intéressante pour l'approche proposée, compte tenu de la généralisation du RGE sur le territoire français. D'une part, ces emprises régleront le problème de focalisation automatique évoqué plus haut, d'autre part les limites planimétriques fournissent autant d'hypothèses de façades et de contraintes sur les normales aux plans. Compte tenu de la méthodologie générale et de l'importance accordée aux plans verticaux, ces limites doivent pouvoir s'intégrer rapidement (c'était le but avoué) à notre approche. Par rapport au système mis en place par [Jib02], le nôtre pourra modéliser des bâtiments ou agrégats de bâtiments présentant des discontinuités internes voire plusieurs bâtiments sur une même zone et

donc gérer des défauts de focalisation. De la même façon, notre approche peut remettre en cause les emprises cadastrales qui présentent parfois quelques défauts.

8.3 Bilan

Dans cette thèse, nous avons présenté une méthode originale et générale de reconstruction de bâtiments à partir de multiples images aériennes géoréférencées. Nous avons également proposé des outils pour la détection de primitives en nous efforçant en permanence de contrer le problème combinatoire lié au contexte de multiscopie, une analyse approfondie d'un des détecteurs a par ailleurs montré l'apport du multi-vues, ce qui était un des objectifs de ces travaux. Enfin, l'introduction de contraintes géométriques à différents niveaux dans l'algorithme représente un aspect nouveau par rapport aux travaux antérieurs en reconstruction de bâti. Nous espérons avoir convaincu le lecteur de l'intérêt de ces contraintes pour apporter une information externe tout en préservant la généralité de la modélisation.
En regard de la problématique globale, nous pensons qu'une modélisation totalement automatique de l'ensemble des bâtiments d'une zone urbaine dense sous la forme de modèles simples et uniquement à partir d'images aériennes (sans informations cadastrales) est actuellement hors de portée compte tenu de la complexité des structures présentes et de l'absence de focalisation a priori. Si l'application vise le réalisme visuel sans accorder d'importance à la modélisation des bâtiments alors une caricature de surface telle que celle proposée dans [MPT02] nous paraît une excellente alternative. Cette modélisation n'est évidemment pas envisageable pour la constitution de bases de données 3D pour lesquelles les travaux les plus récents semblent confirmer que seule l'utilisation de données externes (comme les limites cadastrales) permettrait, à court ou moyen terme d'espérer obtenir une modélisation 3D automatique d'une ville.

Annexes

Annexe A

Reconstruction avec la représentation de Plücker

On rappelle ici une méthode de reconstruction basée sur la représentation d'une droite par ses coordonnées de Plücker. Cette méthode s'appuie sur les derniers résultats de [För01] et a l'avantage de s'intégrer naturellement dans une formalisation géométrique qui permet des tests aisés tenant compte de l'incertitude des primitives géométrique [FBH00]. Nous rappelons ici brièvement le principe de la représentation d'une droite par ses coordonnées de Plücker et la méthode d'estimation optimale de droite définie comme intersection de plans $\mathbf{A}_i$ en utilisant cette représentation. Nous invitons le lecteur à se référer à [För01] pour de plus amples détails.

A.1 Coordonnées homogènes et notations

Dans cette partie, nous représentons toutes les entités géométriques à l'aide de leurs coordonnées homogènes qui ont prouvé leur intérêt dans le cadre de la géométrie projective dont l'usage a simplifié nombre de problèmes en vision par ordinateur [FLP01, HZ02]. Ainsi, un point $\mathbf{X} = (U, V, W, T)^t = (\boldsymbol{X}_0^t, X_h)$ et un plan $\mathbf{A} = (A, B, C, D)^t = (\boldsymbol{A}_h^t, A_0)$ sont définis par des vecteurs à 4 paramètres composés d'une partie homogène indexée par h et d'une partie euclidienne indexée par 0. Pour une droite 3D orientée $\mathbf{L}$, la représentation en coordonnées homogènes utilise les 6 paramètres de Plücker :

$$\mathbf{L} = \begin{pmatrix} L_1 \\ L_2 \\ L_3 \\ L_4 \\ L_5 \\ L_6 \end{pmatrix} = \begin{pmatrix} \boldsymbol{L}_h \\ \boldsymbol{L}_0 \end{pmatrix} \tag{A.1}$$

Conformément aux principes de la représentation en coordonnées homogènes, le vecteur est divisé en une partie homogène $\mathbf{L}_h$, vecteur à 3 paramètres et une partie euclidienne $\mathbf{L}_0$ à 3 paramètres également. Une droite orientée 3D peut être obtenue à partir de 2 points 3D de coordonnées euclidiennes $\boldsymbol{X}$ et $\boldsymbol{Y}$ par :

$$\mathbf{L}(\mathbf{X} \wedge \mathbf{Y}) = \begin{pmatrix} \boldsymbol{L}_h \\ \boldsymbol{L}_0 \end{pmatrix} = \begin{pmatrix} \boldsymbol{Y} - \boldsymbol{X} \\ \boldsymbol{X} \otimes \boldsymbol{Y} \end{pmatrix} \tag{A.2}$$

$\mathbf{L}_h$ représente donc la direction de la droite 3D et $\mathbf{L}_0$ la normale du plan passant par la droite et l'origine (figure A.1).

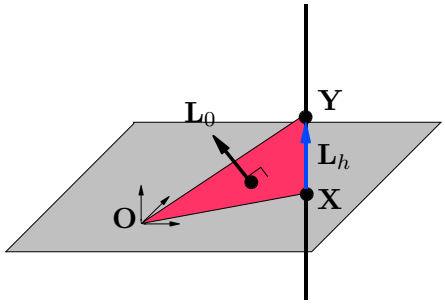


Figure A.1 - Représentation de Plücker d'une droite 3D.

En utilisant ces représentations, par exemple, la distance d'une entité à l'origine est donnée par :

$$d_{\mathbf{X},\mathbf{O}} = \frac{\|\boldsymbol{X}_0\|}{|\ X_h\ |} \qquad d_{\mathbf{A},\mathbf{O}} = \frac{|\ A_0\ |}{\|\boldsymbol{A}_h\|} \qquad d_{\mathbf{L},\mathbf{O}} = \frac{\|\boldsymbol{L}_0\|}{\|\boldsymbol{L}_h\|}$$

Pour la droite 3D, on notera la droite 3D duale

$$\bar{\mathbf{L}} = \begin{pmatrix} \boldsymbol{L}_0^t \\ \boldsymbol{L}_h^t \end{pmatrix} = \mathsf{C}\mathbf{L} \qquad \text{avec} \quad \mathsf{C} = \begin{pmatrix} \mathbf{O} & \boldsymbol{I} \\ \boldsymbol{I} & \mathbf{O} \end{pmatrix} \tag{A.3}$$

Les paramètres de la droite 3D doivent vérifier la contrainte de Plücker qui traduit l'orthogonalité entre $\boldsymbol{L}_h$ et $\boldsymbol{L}_0$:

$$L_1L_4 + L_2L_5 + L_3L - 6 = \boldsymbol{L}_h^t\boldsymbol{L}_0 = \frac{1}{2}\mathbf{L}^t\bar{\mathbf{L}} = 0 \tag{A.4}$$

Par rapport à la représentation $\mathbf{abpq}$ complète, bien que non minimale, cette représentation est élégante et en pratique très facile à utiliser. Elle ne présente pas, en outre, le problème de choix de carte.
Dans la suite, nous aurons besoin des matrices $\boldsymbol{\Gamma}(\mathbf{L})$ et $\bar{\boldsymbol{\Pi}}(\mathbf{A})$ définies par

$$\bar{\boldsymbol{\Pi}}(\mathbf{A})^t = \begin{pmatrix} -\mathsf{S}(\boldsymbol{A}_h) & A_0\boldsymbol{I} \\ -\boldsymbol{A}_h & \mathbf{0}_t \end{pmatrix} \qquad \boldsymbol{\Gamma}(\mathbf{L}) = \begin{pmatrix} -\mathsf{S}(\boldsymbol{L}_h) & -\boldsymbol{L}_0 \\ -\boldsymbol{L}_0^t & 0 \end{pmatrix} \tag{A.5}$$

où $\mathsf{S}(\mathbf{v})$ désigne la matrice de produit vectoriel relative au vecteur à 3 paramètres $\mathbf{v}$. Pour ces matrices, nous avons aussi leurs versions duales :

$$\boldsymbol{\Pi}(\mathbf{A}) = \mathsf{C}\bar{\boldsymbol{\Pi}}(\mathbf{A}) \qquad \bar{\boldsymbol{\Gamma}}(\mathbf{L}) = \boldsymbol{\Gamma}(\mathsf{C}\mathbf{L}) = \boldsymbol{\Gamma}(\bar{\mathbf{L}}) \tag{A.6}$$

A.2 Reconstruction algébrique

La contrainte d'incidence droite/plan s'écrit $\mathbf{w_i} = \boldsymbol{\Pi}(\mathbf{A_i})\mathbf{L} = \bar{\boldsymbol{\Gamma}}(\mathbf{L})\mathbf{A_i} = \mathbf{0}$. Chercher la droite résultant de l'intersection de plans revient donc à trouver $\mathbf{L}$ qui minimise la distance algébrique

$$\boldsymbol{\Omega} = \sum {\mathbf{w_i}}^t\mathbf{w_i} \tag{A.7}$$

Considérant que les coordonnées homogènes définissent les entités à un facteur d'échelle près, cette minimisation se fait sous la contrainte $\mathbf{L}^t\mathbf{L} = 1$. En outre, pour les droites 3D, on a la contrainte de Plücker $\mathbf{L}^t\bar{\mathbf{L}} = 0$. La résolution de ce système est immédiate et ne nécessite aucune itération. Elle est détaillée dans [För01] et ne sera pas reprise ici.

A.3 Reconstruction optimale

Pour avoir une minimisation optimale, prenant en compte les incertitudes des plans et donc pondérant chaque contrainte d'incidence droite/plan par sa matrice de variance covariance, nous sommes ramenés à la minimisation de

$$\mathbf{\Omega} = \sum_i \mathbf{w_i}^t \mathbf{\Lambda}^+_{\mathbf{w_i}} \mathbf{w_j} \tag{A.8}$$

où $\Lambda^+_{\mathbf{w_i}}$ dénote la pseudo-inverse de la matrice de variance covariance de $\mathbf{w_i}$. En raison de la dégénérescence, cette matrice est en effet de rang 2 et l'on est de fait contraint d'utiliser la pseudo-inverse. On peut montrer alors que, prenant en compte la contrainte de normalisation $\mathbf{L}^t\mathbf{L} = 1$, ce système admet comme solution $\hat{\mathbf{L}}$ vecteur propre correspondant à la plus petite valeur propre de

$$\mathbf{N} = \sum_i \mathbf{\Pi}(\hat{\mathbf{A}_i})\big(\bar{\mathbf{\Gamma}}(\hat{\mathbf{L}})\Lambda_{\mathbf{A_i}}\bar{\mathbf{\Gamma}^t}(\hat{\mathbf{L}})\big)^+\mathbf{\Pi}^t(\mathbf{A_i}) \tag{A.9}$$

En prenant en compte la contrainte de Plücker, on obtient finalement

$$\hat{\hat{\mathbf{L}}} = \hat{\mathbf{L}} - \frac{1}{2}\frac{\hat{\bar{\mathbf{L}}}^t\hat{\mathbf{L}}}{\hat{\bar{\mathbf{L}}}^t\mathbf{N}^+\hat{\mathbf{L}}}\mathbf{N}^+\hat{\bar{\mathbf{L}}} \tag{A.10}$$

La matrice de variance-covariance de l'estimée finale est alors :

$$\Gamma_{\hat{\hat{\mathbf{L}}}} = \mathbf{N}^+ - \frac{1}{\hat{\bar{\mathbf{L}}}^t\mathbf{N}^+\hat{\mathbf{L}}}\mathbf{N}^+\hat{\bar{\mathbf{L}}}\hat{\bar{\mathbf{L}}}^t\mathbf{N}^+ \tag{A.11}$$

On remarquera que ce système (calcul de $\mathbf{N}$) dépend des valeurs estimées $\hat{\mathbf{L}}$ et $\hat{\mathbf{A}_i}$, ce qui induit un système itératif analogue à celui que nous proposons dans lequel, partant d'une évaluation de $\hat{\mathbf{L}}$ par minimisation algébrique, solution $\hat{\mathbf{L}}$ et observations $\hat{\mathbf{A}_i}$ sont rééstimées itérativement en tenant compte des valeurs à l'itération précédente. Une description précise de l'algorithme est faite dans [För01].

Annexe B

Recalage de jonctions 3D dans les images

Nous proposons dans cette annexe une méthodologie de recalage de jonctions en 3D généralisant dans l'espace le recalage de coins de Förstner [FG87]. Dans un premier temps, pour une meilleure compréhension, nous récapitulons la technique de recalage de coins de Förstner avant de présenter la généralisation. Quelques résultats démontrent la viabilité de l'approche et la difficulté de l'utiliser en tout automatique sur des images réelles.

B.1 Recalage de coins

Förstner définit les coins ou jonctions dans l'image comme les points d'intersection de plusieurs contours supposés rectilignes. Si on connaît une localisation approximative de ces points d'intérêt et donc des contours, il suffit de calculer le point d'intersection des droites porteuses de ces contours pour obtenir un recalage subpixellique de la position de la jonction. Pour éviter les étapes d'extraction de contours et d'approximation polygonale, Förstner utilise le gradient de l'image comme information.
L'équation d'une droite $l_{\mathbf{p}'}$ passant par un pixel $\mathbf{p}' = (x', y')$ et perpendiculaire au gradient de l'image en ce point $\nabla I(\mathbf{p}') = (I_x(\mathbf{p}'), I_y(\mathbf{p}'))$ est donnée par :

$$D_{\mathbf{p}'}(\mathbf{p}) = (\nabla I(\mathbf{p}'))^t(\mathbf{p} - \mathbf{p}') = I_x(\mathbf{p}')(x - x') + I_y(\mathbf{p}')(y - y') = 0 \tag{B.1}$$

pour $\mathbf{p}(x, y) \in \mathbb{R}^2$. La position de la jonction 2D est alors définie comme le point $\mathbf{p}$ qui minimise la distance orthogonale aux droites $l_{\mathbf{p}'}(\mathbf{p})$ situées dans un voisinage de la jonction soit

$$\min_{\mathbf{p}\in\mathbb{R}^2} \sum_{\mathbf{p}'\in\mathbb{R}^2} (D_{\mathbf{p}'}(\mathbf{p}))\omega(\mathbf{p}') \tag{B.2}$$

où ω est une fonction de fenêtrage (fenêtre rectangulaire classique ou gaussienne par exemple). En développant ce système, on peut le réécrire sous la forme d'un moindre carré classique dont la solution est donnée par :

$$\mathbf{p} = (\mathbf{A}^t\mathbf{A})^{-1}\mathbf{A}^t\mathbf{b} \tag{B.3}$$

avec

$$\mathbf{A}^t\mathbf{A} = \sum_{\mathbf{p}'\in\mathbb{R}^2} \begin{pmatrix} I_x^2(\mathbf{p}') & I_xI_y(\mathbf{p}') \\ I_yI_x(\mathbf{p}') & I_y^2(\mathbf{p}') \end{pmatrix} \omega(\mathbf{p}') \tag{B.4}$$

$$\mathbf{A}^t\mathbf{b} = \sum_{\mathbf{p}'\in\mathbb{R}^2} \begin{pmatrix} I_x^2(\mathbf{p}')x' & I_xI_y(\mathbf{p}')y' \\ I_yI_x(\mathbf{p}')x' & I_y^2(\mathbf{p}')y' \end{pmatrix} \omega(\mathbf{p}') \tag{B.5}$$

Le lecteur intéressé pourra se référer à [Bla97] pour une étude approfondie des caractéristiques de ce détecteur en comparaison avec d'autres détecteurs de la littérature.

B.2 Recalage de jonctions 3D

En 3D, en supposant connues les matrices de transformation perspective 3x4 $\mathcal{T}_i$ correspondant aux n images, on rappelle que pour une image i et un point 3D $P(X,Y,Z)$, nous avons :

$$\tilde{\mathbf{p}} = s\mathcal{T}_i\tilde{\mathbf{P}} \tag{B.6}$$

En notant $\mathbf{l}^i_{\mathbf{p}'} = [I^i_x(\mathbf{p}'), I^i_y(\mathbf{p}'), -(I^i_x(\mathbf{p}')x' + I^i_y(\mathbf{p}')y'])^t$, on peut écrire, pour une droite $l^i_{\mathbf{p}'}$ dans une image i passant par $\mathbf{p}'$

$$D^i_{\mathbf{p}'}(\mathbf{p}) = \tilde{\mathbf{x}}^t\mathbf{l}^i_{\mathbf{p}'} = \tilde{\mathbf{P}}^t\mathcal{T}_i^t\mathbf{l}^i_{\mathbf{p}'} = 0 \tag{B.7}$$

En introduisant la matrice

$$\mathbf{v}_i(\mathbf{p}') = \begin{bmatrix} a_i(\mathbf{p}') \\ b_i(\mathbf{p}') \\ c_i(\mathbf{p}') \\ d_i(\mathbf{p}') \end{bmatrix} = \mathcal{T}_i^t\mathbf{l}^i_{\mathbf{p}'} \tag{B.8}$$

on obtient l'équation B.9 :

$$D^i_{\mathbf{p}'}(\mathbf{p}) = a_i(\mathbf{p}')X + b_i(\mathbf{p}')Y + c_i(\mathbf{p}')Z + d_i(\mathbf{p}') = 0 \tag{B.9}$$

La recherche de la meilleure jonction 3D se traduit par la minimisation des distances orthogonales à *toutes* les droites $l^i_{\mathbf{p}'}$ proches de la projection de $\mathbf{P}$ dans *toutes* les images soit :

$$\min_{\mathbf{P}\in\mathbb{R}^2} \sum_{i=1}^{i=n} \sum_{\mathbf{p}'\in\mathbb{R}^2} (D^i_{\mathbf{p}'}(\mathbf{p}))\omega(\mathbf{p}') \tag{B.10}$$

On est donc ramené à la résolution d'un système aux moindres carrés dont la solution est donnée par

$$\mathbf{P} = \mathbf{A^t A}^{-1}\mathbf{A^t b} \tag{B.11}$$

avec

$$\begin{aligned} \mathbf{A^t A} &= \sum_{i=1}^{i=n} \sum_{\mathbf{p}'\in\mathbb{R}^2} \begin{pmatrix} a_i(\mathbf{p}')^2 & a_i(\mathbf{p}')b_i(\mathbf{p}') & a_i(\mathbf{p}')c_i(\mathbf{p}') \\ b_i(\mathbf{p}')a_i(\mathbf{p}') & b_i(\mathbf{p}')^2 & b_i(\mathbf{p}')c_i(\mathbf{p}') \\ c_i(\mathbf{p}')a_i(\mathbf{p}') & c_i(\mathbf{p}')b_i(\mathbf{p}') & c_i(\mathbf{p}')^2 \end{pmatrix} \omega(\mathbf{p}') \\ \mathbf{A^t b} &= -\sum_{i=1}^{i=n} \sum_{\mathbf{p}'\in\mathbb{R}^2} \begin{pmatrix} a_i(\mathbf{p}')d_i(\mathbf{p}') \\ b_i(\mathbf{p}')d_i(\mathbf{p}') \\ c_i(\mathbf{p}')d_i(\mathbf{p}') \end{pmatrix} \omega(\mathbf{p}') \end{aligned} \tag{B.12}$$

B.3 Résultats

Nous ne présenterons que peu de résultats sur la technique de recalage de coins 2D qui a déjà été étudiée en détail dans [FG87, Bla97] et sur laquelle la figure B.1 fournit quelques illustrations.

Dans la figure B.2, nous présentons quelques exemples de recalage de jonctions 3D pour lesquels, qualitativement, le recalage a fourni un gain indéniable de précision. La fonction de fenêtrage utilisée est une fenêtre rectangulaire classique de taille 10x10 pixels et compte tenu de la qualité des images, nous avons effectué un suréchantillonnage d'un facteur 2 (ce qui correspond donc à 400 pixels par fenêtre).

En revanche, dès que la fenêtre grandit (ce qui permet de gagner en précision [Bla97]) ou que des détails linéaires ne passant pas par la jonction recherchée sont présents dans le voisinage, la reconstruction est très fortement perturbée (figure B.3).

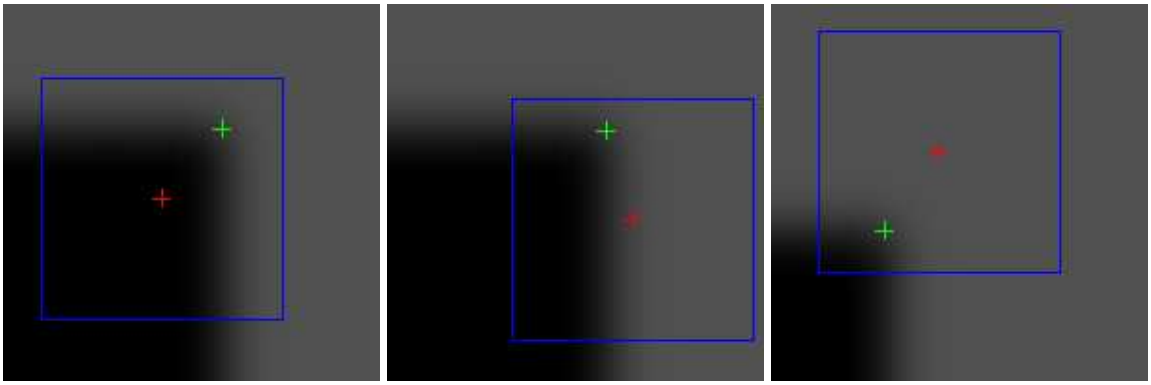


Figure B.1 - Recalage de jonctions en 2D dans différentes configurations simulées démontrant la robustesse de l'approche, le bruit gaussien ajouté sur l'image initiale de coin a un écart type $\sigma = 20$pixels. Le rectangle bleu délimite la fenêtre utilisée pour le voisinage. La croix rouge montre l'emplacement de la position initiale de la jonction, la croix verte le résultat du recalage.

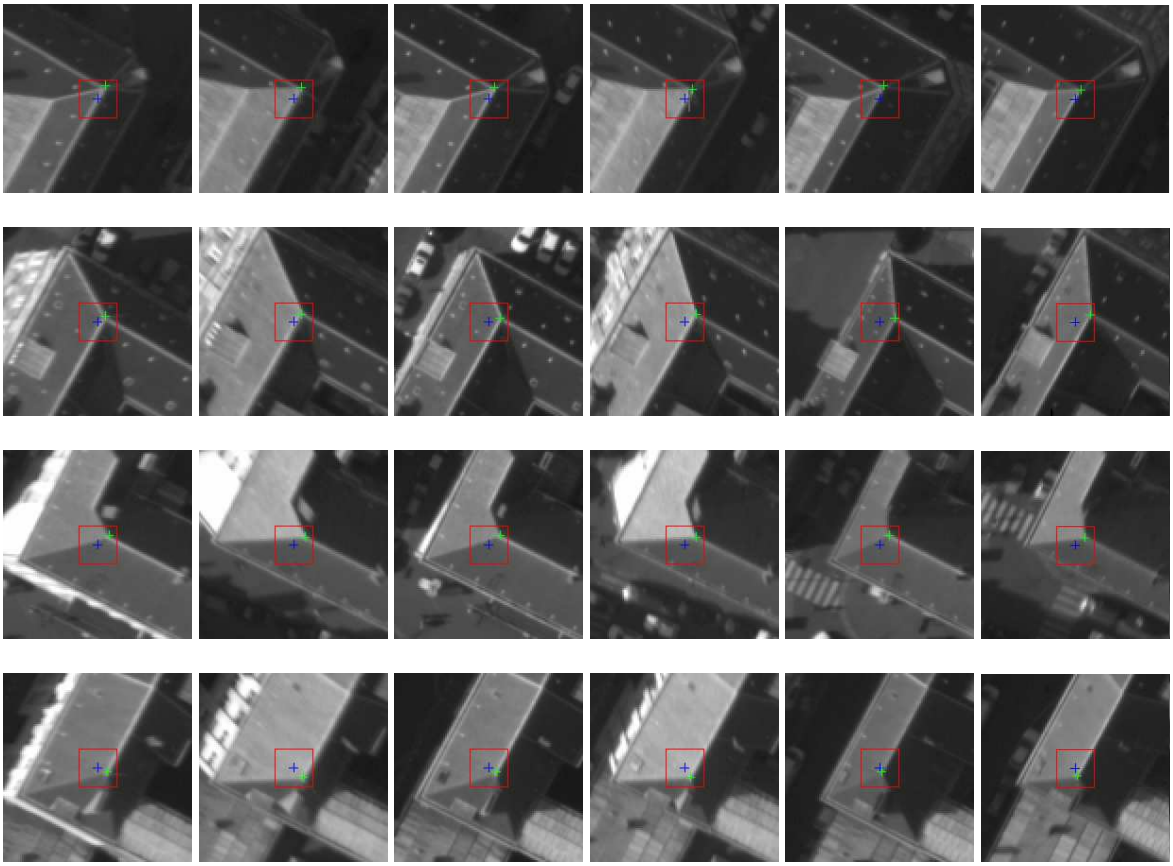


Figure B.2 - Recalage de jonctions 3D sur images réelles. Chaque rangée représente un nouvel exemple de jonction, les différentes vignettes dans une rangée représentent les images utilisées et sur lesquelles a été projeté le résultat. La croix bleue symbolise la jonction initiale, la verte le résultat du recalage. Le rectangle délimite la fonction de fenêtrage.

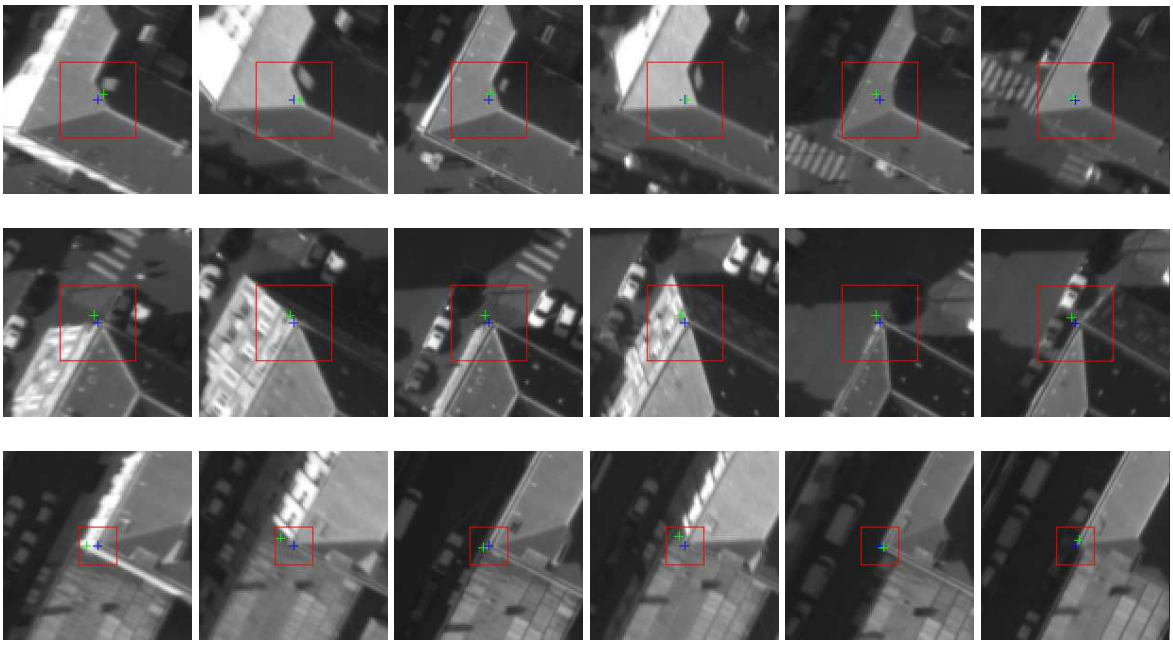


Figure B.3 - Autres exemples de recalage.

Dans un cadre automatique et dans un environnement urbain où les objets parasites (véhicules, appentis, végétations...) sont nombreux, cet outil semble donc difficile à utiliser. Nous avons ainsi bruité un ensemble de points 3D de référence (issus de la saisie manuelle de la section 6.5) avec un bruit gaussien faible d'écart type 10cm. Après recalage des jonctions, l'écart type est de 31cm en moyenne en X et Y et de 60cm en Z ce qui représente sur les images un écart-type de 1,36 pixel (0,43 pour les points bruités - ce qui est logique compte tenu de la résolution de 25 cm des images) et traduit la dégradation de la précision.
Nous pensons donc que cet outil est une bonne technique de recalage de jonction pour une utilisation manuelle, c'est-à-dire où l'opérateur juge, en fonction du voisinage radiométrique de la pertinence d'effectuer un recalage. En revanche, l'utilisation automatique dans un environnement urbain paraît illusoire.

Annexe C

Incertitude géométrique de facettes planes

L'objectif dans cette partie est de munir les plans déterminés dans le chapitre 4 d'une mesure d'incertitude géométrique. Cette incertitude est importante pour faire des tests statistiques robustes avec des primitives issues d'autres détecteurs [För01].

C.1 Points indépendants

Dans ce qui suit, on supposera que la définition du meilleur plan se fait aux moindres carrés. L'utilisation d'un M-estimateur ne change pas le raisonnement puisqu'il s'agit d'un processus de moindres carrés pondérés itératifs [ZF92] et qu'en utilisant les valeurs de la dernière itération, on obtient le même résultat.
L'équation du plan recherché s'écrit avec les 3 paramètres a, b et d et l'équation :

$$ax + by + d = z \tag{C.1}$$

La résolution aux moindres carrés donnant la solution $\mathbf{X} = [a, b, d]^t$ conduit à

$$\hat{\mathbf{X}} = (\mathbf{A^t A})^{-1}\mathbf{A^t B} = \mathbf{\Phi} \tag{C.2}$$

avec les expressions suivantes pour $\mathbf{A^t A}$ et $\mathbf{A^t B}$

$$\mathbf{A^t A} = \begin{bmatrix} \sum x^2 & \sum xy & \sum x \\ \sum xy & \sum y^2 & \sum y \\ \sum x & \sum y & \sum 1 \end{bmatrix} \qquad \mathbf{A^t B} = \begin{bmatrix} -\sum xz \\ -\sum yz \\ -\sum z \end{bmatrix} \tag{C.3}$$

où les sommations sont faites sur un ensemble de points $\mathbf{P} = \{\mathbf{p}_i(x_i, y_i, z_i)\}$. En considérant tous les points *indépendants* et ayant la même variance/covariance $\mathbf{\Lambda}$, on sait que [ZF92]

$$\mathbf{\Lambda}_{\hat{\mathbf{X}}} = \frac{\partial \mathbf{\Phi}}{\partial \mathbf{P}} \mathbf{\Lambda}_{\mathbf{P}} \frac{\partial \mathbf{\Phi}}{\partial \mathbf{P}}^t \tag{C.4}$$

avec

$$\mathbf{\Lambda}_{\mathbf{P}} = \begin{bmatrix} \mathbf{\Lambda} & \mathbf{0} & \mathbf{0} & \dots \\ \mathbf{0} & \mathbf{\Lambda} & \mathbf{0} & \dots \\ \mathbf{0} & \mathbf{0} & \mathbf{\Lambda} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{C.5}$$

On obtient ainsi

$$\mathbf{\Lambda}_{\hat{\mathbf{X}}} = \sum_i \frac{\partial \mathbf{\Phi}}{\partial \mathbf{p_i}} \mathbf{\Lambda} \frac{\partial \mathbf{\Phi}}{\partial \mathbf{p_i}}^t \tag{C.6}$$

En utilisant des expressions analogues à la suivante pour les dérivées partielles :

$$\frac{\partial \mathbf{\Phi}}{\partial x_i} = -(\mathbf{A^t A})^{-1} \frac{\partial \mathbf{A^t A}}{\partial x_i} (\mathbf{A^t A})^{-1} \mathbf{A^t B} + (\mathbf{A^t A})^{-1} \frac{\partial \mathbf{A^t B}}{\partial x_i} \tag{C.7}$$

on peut aisément calculer $\mathbf{\Lambda}_{\hat{\mathbf{X}}}$.
Cette expression est directement utilisable sur des MNE lasers par exemple où chaque point est mesuré indépendamment de ses voisins. En revanche, cette expression n'est pas utilisable telle quelle pour des MNE issus de corrélation.

C.2 Points issus de corrélation

Dans le cas d'un MNE issu de corrélation, la propriété d'indépendance des points *n'est pas vérifiée* compte tenu du recouvrement des fenêtres de corrélation entre points successifs. Si le calcul de la corrélation s'est fait sur une fenêtre de taille T^2, on peut se ramener à l'indépendance des points en supposant que l'on prend seulement un point sur T^2 points. C'est une modélisation approximative, mais suffisante, pour notre application, des phénomènes d'autocorrélation spatial entre points calculés par corrélation. On a alors :

$$\mathbf{\Lambda}_{\hat{\mathbf{X}}} = T^2 \sum_i \frac{\partial \mathbf{\Phi}}{\partial \mathbf{p_i}} \mathbf{\Lambda} \frac{\partial \mathbf{\Phi}}{\partial \mathbf{p_i}}^t \tag{C.8}$$

C.3 Résultats

La figure C.1 montre un exemple visuel de points intersection de 3 plans projetés sur le MNE de corrélation avec leurs ellipses d'incertitude. On remarquera l'influence de l'orientation des plans, de la taille des régions définissant ses plans.

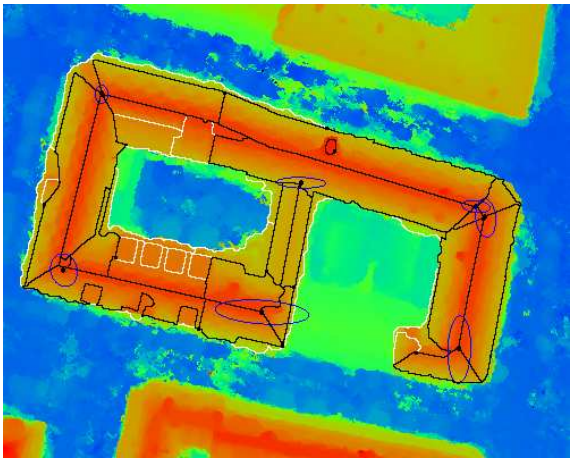
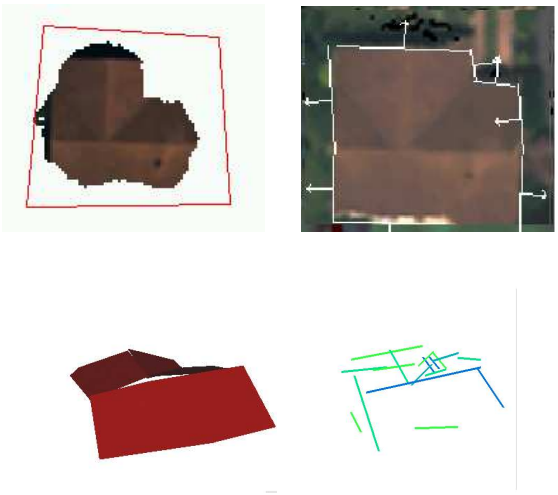


Figure C.1 - Points projetés avec incertitude.

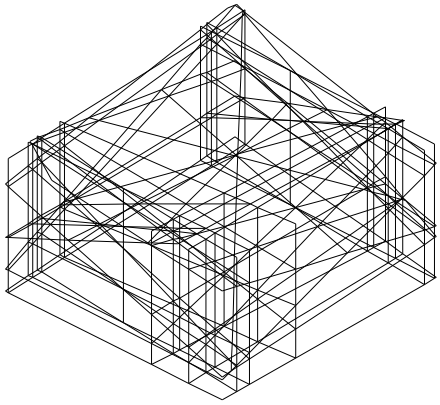
Annexe D

L'algorithme en quelques planches

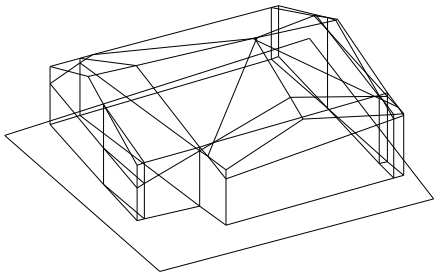
Cette annexe récapitule visuellement sur quelques exemples réels le fonctionnement de l'algorithme et ses différentes étapes. Le lecteur est invité à se référer aux chapitres principaux pour une description plus en profondeur de ces étapes et les justifications théoriques afférentes.



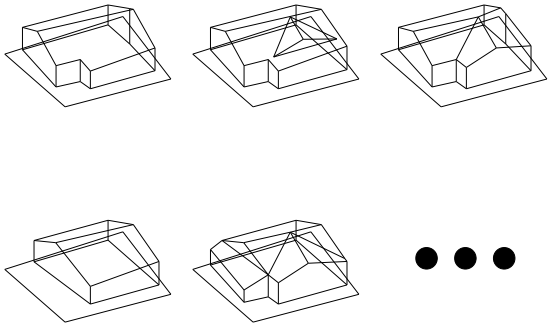
(a) Masque de focalisation, façades, facettes et segments.



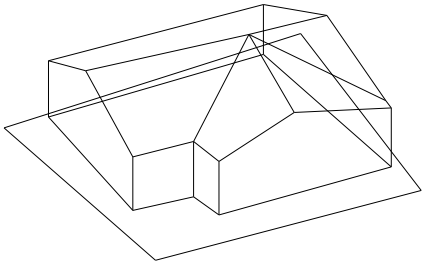
(b) Arrangement de plans.



(c) Après filtrages.



(d) Recherche de l'ensemble des solutions.

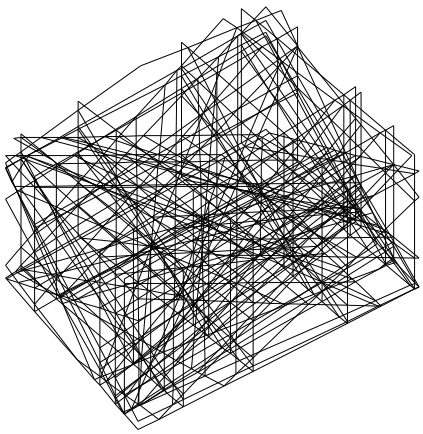


(e) Choix de la meilleure solution.

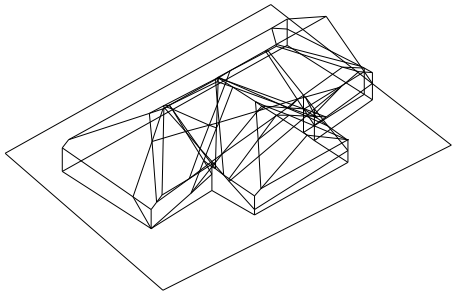
(f) La solution avant et après recalage avec contraintes.

Figure D.1 - Les différentes étapes de l'algorithme appliqué à un exemple.

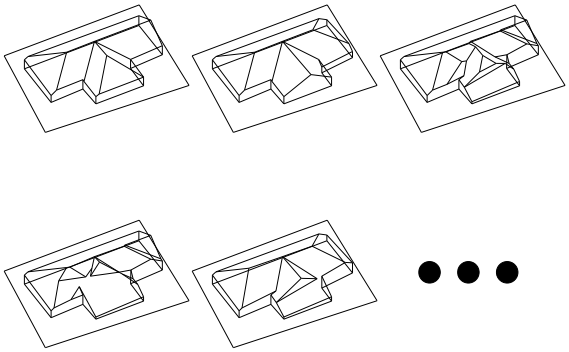
(a) Masque de focalisation, façades, facettes et segments.



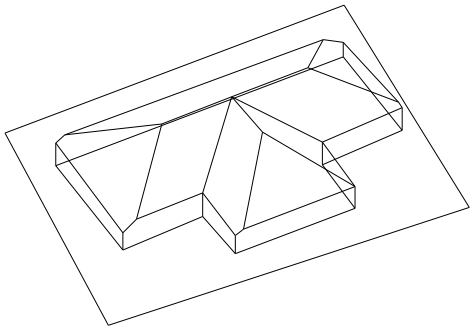
(b) Arrangement de plans.



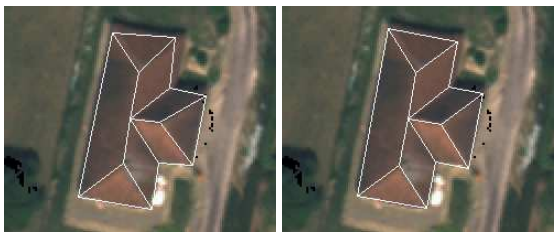
(c) Après filtrages.



(d) Recherche de l'ensemble des solutions.

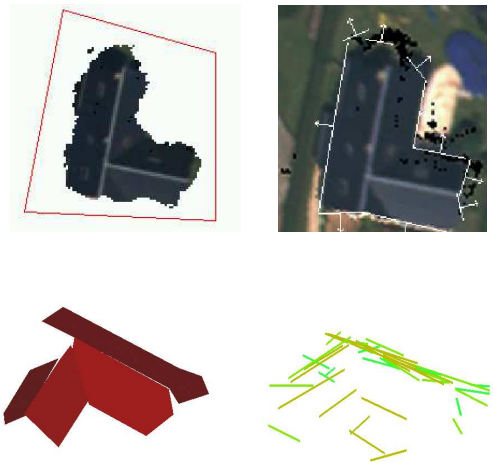


(e) Choix de la meilleure solution.

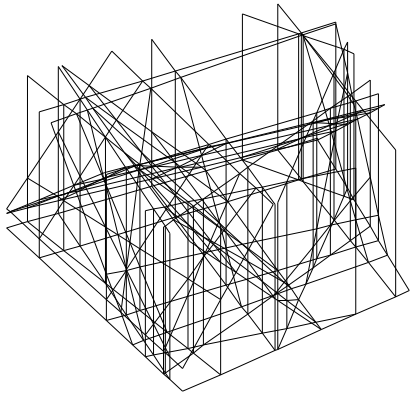


(f) La solution avant et après recalage avec contraintes.

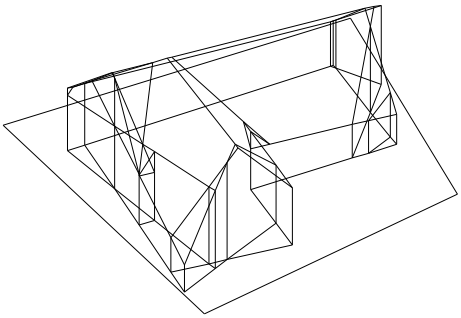
Figure D.2 - Les différentes étapes de l'algorithme appliqué à un exemple.



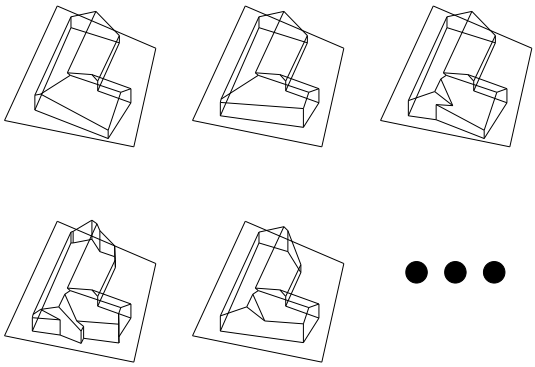
(a) Masque de focalisation, façades, facettes et segments.



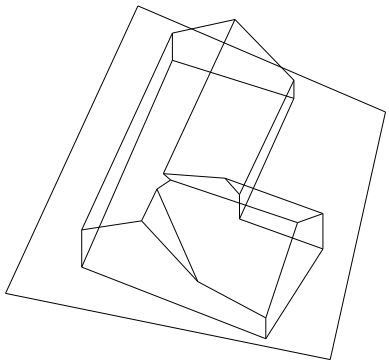
(b) Arrangement de plans.



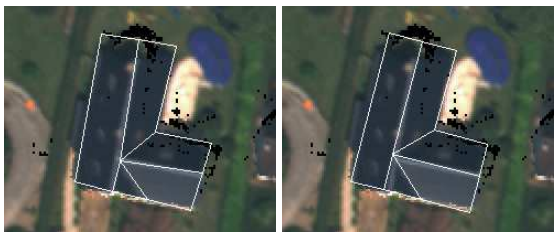
(c) Après filtrages.



(d) Recherche de l'ensemble des solutions.

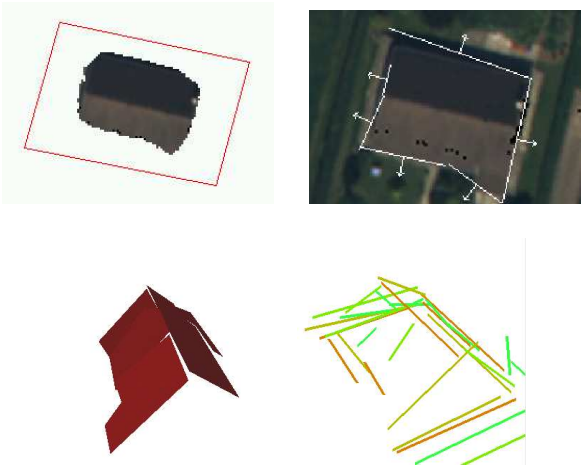


(e) Choix de la meilleure solution.

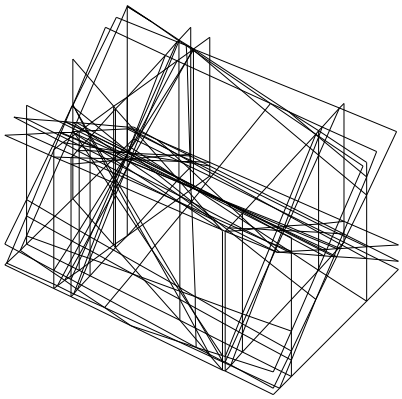


(f) La solution avant et après recalage avec contraintes.

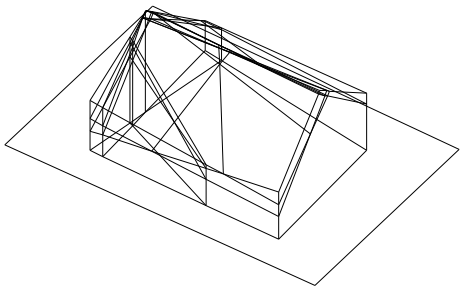
Figure D.3 - Les différentes étapes de l'algorithme appliqué à un exemple.



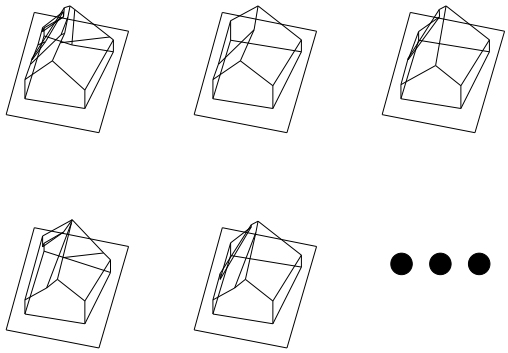
(a) Masque de focalisation, façades, facettes et segments.



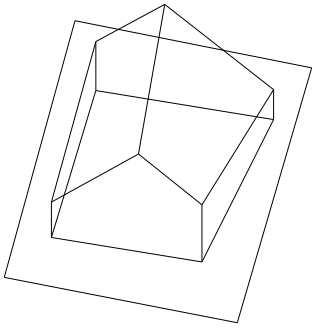
(b) Arrangement de plans.



(c) Après filtrages.



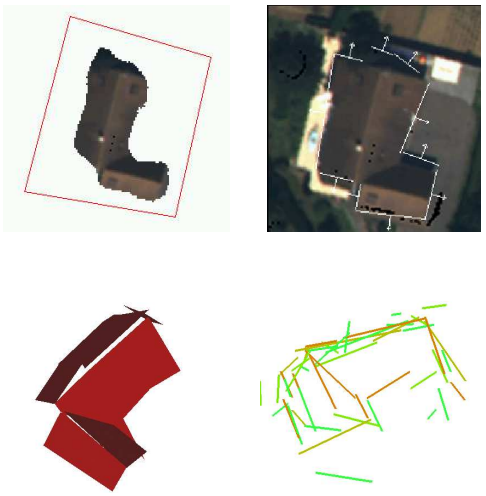
(d) Recherche de l'ensemble des solutions.



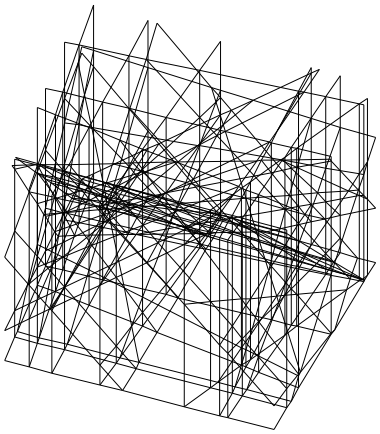
(e) Choix de la meilleure solution.

(f) La solution avant et après recalage avec contraintes.

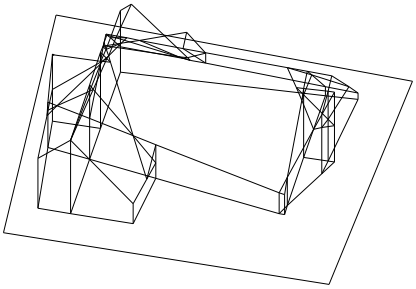
Figure D.4 - Les différentes étapes de l'algorithme appliqué à un exemple.



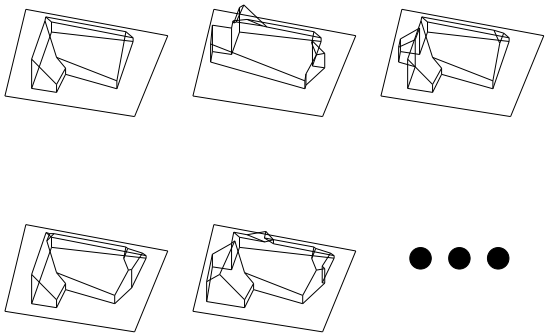
(a) Masque de focalisation, façades, facettes et segments.



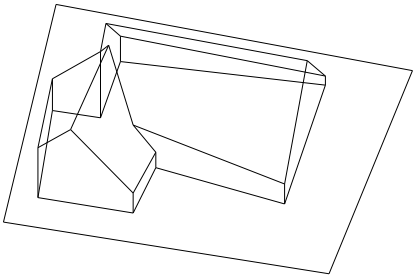
(b) Arrangement de plans.



(c) Après filtrages.



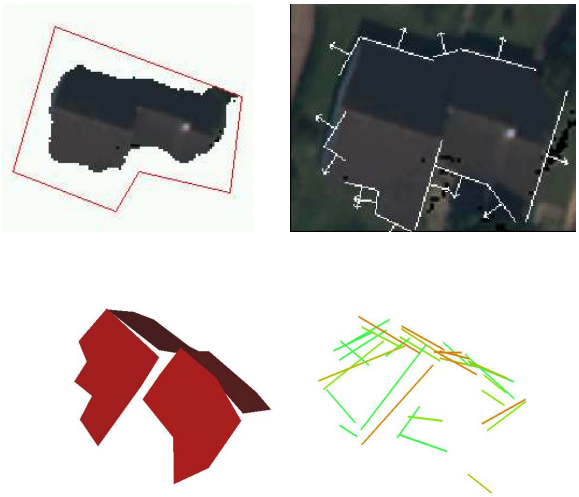
(d) Recherche de l'ensemble des solutions.



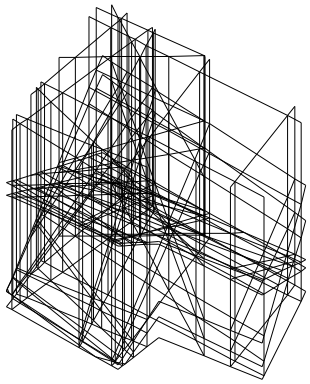
(e) Choix de la meilleure solution.

(f) La solution avant et après recalage avec contraintes.

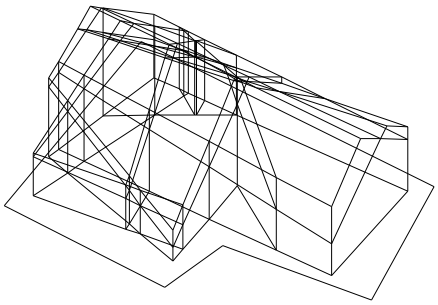
Figure D.5 - Les différentes étapes de l'algorithme appliqué à un exemple.



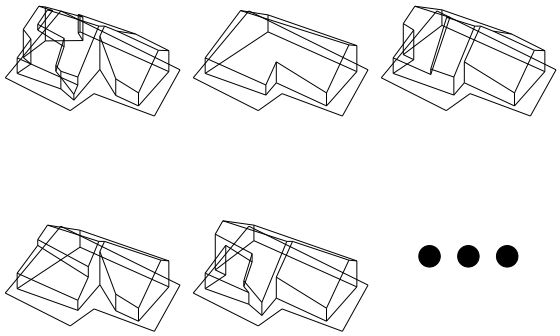
(a) Masque de focalisation, façades, facettes et segments.



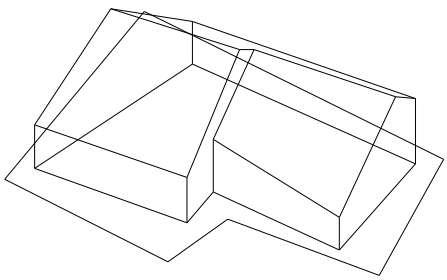
(b) Arrangement de plans.



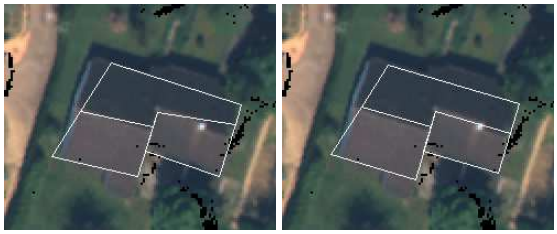
(c) Après filtrages.



(d) Recherche de l'ensemble des solutions.

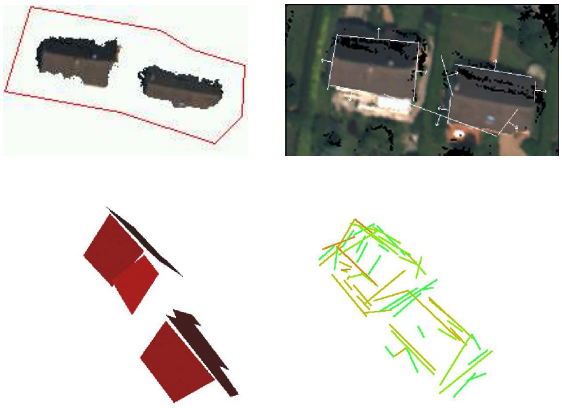


(e) Choix de la meilleure solution.

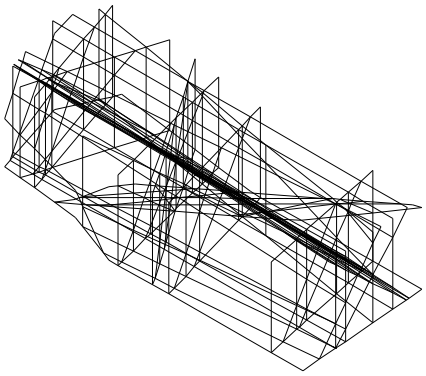


(f) La solution avant et après recalage avec contraintes.

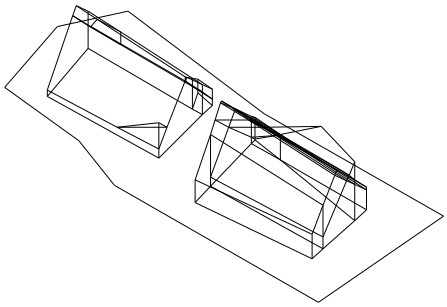
Figure D.6 - Les différentes étapes de l'algorithme appliqué à un exemple.



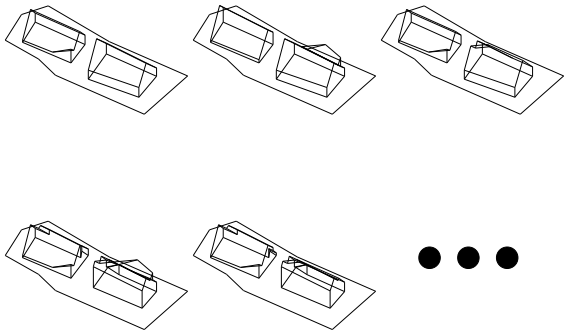
(a) Masque de focalisation, façades, facettes et segments.



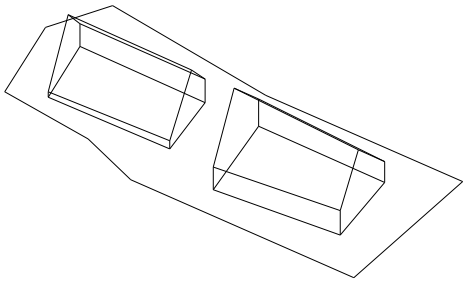
(b) Arrangement de plans.



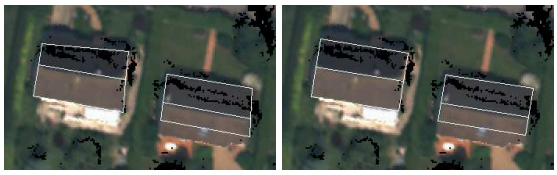
(c) Après filtrages.



(d) Recherche de l'ensemble des solutions.



(e) Choix de la meilleure solution.



(f) La solution avant et après recalage avec contraintes.

Figure D.7 - Les différentes étapes de l'algorithme appliqué à un exemple. Cet exemple illustre en particulier les capcaités de l'algorithme à pallier les défauts de focalisation.

Annexe E

Zones de travail

Cette annexe présente les 2 zones de travail qui ont servi de support à l'évaluation de l'algorithme. On fournit en particulier les résultats aux différents β.

E.1 Zone 1

E.1.1 Référence

Figure E.1 - Zone 1 avec bâtiments numérotés.

Figure E.2 - Référence pour la zone 1.

E.1.2 Résultats aux différents β

Figure E.3 - Résultats zone 1, $\beta = 0.01$.

Figure E.4 - Résultats zone 1, $\beta = 10$.

Figure E.5 - Résultats zone 1, $\beta = 25$.

Figure E.6 - Résultats zone 1, $\beta = 50$.

Figure E.7 - Résultats zone 1, $\beta = 75$.

Figure E.8 - Résultats zone 1, $\beta = 100$.

Figure E.9 - Résultats zone 1, $\beta = 250$.

Figure E.10 - Résultats zone 1, $\beta = 500$.

Figure E.11 - Résultats zone 1, $\beta = 100000$.

E.2 Zone 2

E.2.1 Référence

Figure E.12 - Zone 2 avec bâtiments numérotés.

Figure E.13 - Référence pour la zone 2.

E.2.2 Résultats aux différents β

Figure E.14 - Résultats zone 2, $\beta = 0.01$.

Figure E.15 - Résultats zone 2, $\beta = 10$.

Figure E.16 - Résultats zone 2, $\beta = 25$.

Figure E.17 - Résultats zone 2, $\beta = 50$.

Figure E.18 - Résultats zone 2, $\beta = 75$.

Figure E.19 - Résultats zone 2, $\beta = 100$.

Figure E.20 - Résultats zone 2, $\beta = 250$.

Figure E.21 - Résultats zone 2, $\beta = 500$.

Figure E.22 - Résultats zone 2, $\beta = 100000$.

Publications

Publications en rapport avec les travaux de thèse

N. PAPARODITIS, G. MAILLET, F. TAILLANDIER, H. JIBRINI, F. JUNG, L. GUIGUES et D. BOLDO. « Multi-image 3D feature and DSM extraction for change detection and building reconstruction ». Dans E.P. BALTSAVIAS, A. GRUEN et L.V. GOOL, éditeurs, *Automatic Extraction of Man-Made Objects from Aerial and Space Images (III)*, Centro Stefano Franscini, Monte Verità, Ascona, 2001. A.A. Balkema Publishers.

F. TAILLANDIER et R. DERICHE. « Reconstruction de primitives linéaires 3D en multi-vues pour la modélisation de scènes urbaines ». Dans *Actes du 13ème Congrès Francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle*, volume I, pages 115–124, Angers, France, janvier 2002. AFRIF-AFIA.

F. FUCHS, H. JIBRINI, G. MAILLET, M. PIERROT-DESEILLIGNY et F. TAILLANDIER. « Trois approches pour la construction automatique de modèles 3D de batiments en imagerie aerienne haute résolution ». *Bulletin de la Société Française de Photogrammétrie et de Télédétection*, 166 :10–18, mai 2002.

F. TAILLANDIER et R. DERICHE. « Reconstruction of 3D Linear Primitives from Multiple Views for Urban Areas Modelisation ». Dans R.KALLIANY et F. LEBERL, éditeurs, *Proceedings of the ISPRS Commission 3 Symposium on Photogrammetric Computer Vision (PCV'02)*, volume 34 de *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages B267–272, Graz, Austria, septembre 2002. Institute for Computer Graphics and Vision, Graz University of Technology.

G. MAILLET, N. PAPARODITIS et F. TAILLANDIER. « TIN Surface Reconstruction From Multiple Calibrated Aerial Images in Urban Areas. ». Dans *International Conference on Image Processing (ICIP 02)*, volume 3, pages 521–524, Rochester, septembre 2002. IEEE Computer Society.

F. TAILLANDIER, L. GUIGUES et R. DERICHE. « A Framework for Constrained Multi-Scale Range Image Segmentation ». Dans *Proceedings of the International Conference on Image Processing (ICIP'03)*, Barcelona, Spain, septembre 2003. IEEE Signal Processing Society.

F. TAILLANDIER et R. DERICHE. « Reconstruction automatique de bâtiments à partir d'images aériennes : une approche probabiliste générique. ». Dans *Actes du 14ème Congrès Francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle*, Toulouse, France, janvier 2004. AFRIF-AFIA.

F. TAILLANDIER et R. DERICHE. « Automatic Buildings Reconstruction from Aerial Images : a Generic Bayesian Framework. ». Dans *Proceedings of the XXth ISPRS Congress*, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Istanbul, Turkey, juillet 2004. ISPRS. to be published.

Publications au cours de la thèse, en rapport avec les travaux de DEA

X. WANG, S. TOTARO, F. TAILLANDIER, A.R. HANSON et S. TELLER. « Recovering Facade Texture and Microstructure from Real-World Images ». Dans *Proceedings of the 2nd International Workshop on Texture Analysis and Synthesis in conjunction with ECCV'02*, pages 145–149, Copenhague, Denmark, juin 2002.

X. WANG, S. TOTARO, F. TAILLANDIER, A.R. HANSON et S. TELLER. « Recovering Facade Texture and Microstructure from Real-World Images ». Dans R.KALLIANY et F. LEBERL, éditeurs, *Proceedings of the ISPRS Commission 3 Symposium on Photogrammetric Computer Vision (PCV'02)*, volume 34 de *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages A381–386, Graz, Austria, septembre 2002. Institute for Computer Graphics and Vision, Graz University of Technology.

Glossaire

A

aérotriangulation : processus qui permet de déterminer les paramètres extrinsèques des clichés d'une mission aérienne.

algorithme de partage des eaux : si l'on compare une image en niveaux de gris à un relief terrestre, cette image peut se diviser en plusieurs zones : les niveaux sombres correspondent aux fonds des bassins versants et les niveaux clairs correspondent aux crêtes des bassins versants. Un bassin versant correspond à un espace géographique dans lequel toutes les eaux convergent vers le même minimum. Les différents bassins versant sont délimités par une ligne de crête appelée ligne de partage des eaux. L'algorithme du même nom extrait ces lignes. Cet algorithme est fréquemment utilisé pour obtenir une sursegmentation soit un partitionnement de l'image en régions élémentaires de petite taille.

attaque droite : désigne une extrémité de toit en forme de bipan classique.

attaque en V : désigne une extrémité de toit en forme de bipan "étété" par un plan oblique. L'extrémité est donc un triangle en vue verticale d'où le nom donné.

B

bâti : thème cartographique désignant l'ensemble des bâtiments.

bayésienne (modélisation) : approche fondée sur le conditionnement et la formule d'inversion de Bayes. La formule d'inversion de Bayes ou théorème de Bayes permet de relier les probabilités a posteriori et les probabilités a priori : $P(A \mid B) = \frac{P(B|A)P(A)}{P(B)}$.

bottom-up (approches) : approches qui partent des données pour inférer la reconstruction sans a priori sur le modèle à reconstruire.

C

caméra sténopé : situation dans laquelle tous les rayons perspectifs convergent en un point unique : le sommet de prise de vue. C'est le modèle classique utilisé dans la quasi-totalité des travaux de Vision par Ordinateur pour modéliser une projection sur un plan image.

chiens-assis : superstructure d'un toit composée d'un pan principal et de deux pans latéraux obliques [Fuc01]

clique maximale : une clique est un ensemble de noeuds d'un graphe tel que 2 noeuds quelconques sont reliés par une arête. Elle est maximale si on ne peut rajouter de noeuds sans violer la définition de clique.

D

degrés de liberté : nombre de variables indépendantes dans la loi du χ^2 et plus généralement dans un ensemble d'équations.

E

ensemble stable maximal : un ensemble stable est un ensemble de noeuds d'un graphe qui ne comprend que des noeuds non adjacents 2 à 2. Il est stable si on ne peut rajouter de noeuds sans violer ce principe.

Espace-Objet : désigne l'espace de reconstruction, c'est-à-dire l'espace 3D dans lequel sont reconstruits les objets.

estimations robustes : techniques de régression permettant, au contraire de la minimisation aux moindres carrés, d'éliminer les erreurs les plus importantes. Ces techniques ont fait l'objet de nombreuses études, on peut citer les classes des M-estimateurs, Least Median of Squares, RANSAC...

F

focalisation (zone de) : zone de la scène sur laquelle se focalise l'algorithme ou l'opérateur en supposant la présence d'un bâtiment. Idéalement, dans notre cas, chaque zone de focalisation devrait isoler un *seul* bâtiment.

Freeman (séquence de) : un contour dans une image peut être décrit par la position absolue d'un de ses points terminaux suivie par la position *relative* des points restants par rapport à leur prédecesseur. En 8-connexité, on obtient ainsi pour un contour fermé de n pixels dans une image NxM une longueur de description de $\log N + \log M + n \log 8$.

G

généralisation : simplifications de formes ou caricatures qui interviennent couramment dans les processus de production cartographique pour adapter la représentation à l'échelle de travail.

généricité (ou généralité) : capacité de l'algorithme à pouvoir traiter des formes arbitraires de bâtiments.

graphe complet : graphe dans lequel chaque noeud est adjacent à tous les autres.

H

Hough (transformée de) : technique classique pour extraire des objets définis sous forme paramétrique $f(a, b, \cdots) = 0$. La transformée de Hough définit deux espaces : l'espace objet et l'espace des paramètres. Sommairement, les éléments de l'espace objet vont voter dans l'espace des paramètres. Une détection des maxima locaux dans l'espace des paramètres permet de déterminer les objets principaux via leur paramétrisation.

hypothesize-and-verify (approches) : approches combinant une stratégie bottom-up suivie d'une stratégie top-down.

I

image de profondeur (ou de distances) : la valeur de chaque pixel encode en fait une distance ou une profondeur ou une altitude. Un MNE est une image de profondeur.

image géoréférencée : se dit lorsque les paramètres de projection d'une image sont parfaitement connus.

L

loi du χ^2 : loi suivie par une variable aléatoire positive qui s'écrit comme la somme $\sum_1^d \epsilon_i^2$ de d carrés de variables aléatoires gaussiennes réduites *indépendantes* ϵ_i. Cette variable suit une loi de densité $c_d \exp(-y/2) y^{(d/2)-1}$ où c_d est une constante de normalisation .

loi gaussienne (ou loi normale) : les variables aléatoires suivant une telle loi ont une densité $p(x) = \frac{1}{\sqrt{2\pi}} \exp -x^2/2$. Il s'agit d'une densité gaussienne centrée d'écart type normalisé à 1.

M

matrice de transformation perspective : matrice $\mathcal{T}$ permettant de relier un point 3D $\mathbf{M}$ et sa projection $\mathbf{m}$ par la relation $\tilde{\mathbf{m}} = \mathcal{T}\tilde{\mathbf{M}}$ dans laquelle apparaissent les coordonnées homogènes.

Minimum Descriptor Length (MDL) : technique reprise de la théorie du codage du signal et qui consiste à chercher à coder l'information avec un nombre minimum de bits donc une longueur de codage minimum. On utilise énormément le MDL pour faire de la sélection de modèles : avoir un modèle simple d'un phénomène permet un codage plus court de ce modèle, mais plus long des erreurs entre le modèle et le phénomène et inversement pour un modèle plus complexe. La technique MDL permet de choisir la modélisation la plus adaptée en utilisant un langage de description commun pour le modèle et les erreurs.

Modèle Numérique d'Élévation (MNE) ou Modèle Numérique de Surface (MNS) : représentation maillée de la surface topographique, c'est-à-dire de l'ensemble des objets vus depuis les clichés (sol, bâtiments, végétation, ponts, etc compris). Cette représentation est la plupart du temps une représentation 2D1/2 sous forme d'une grille d'altitude fournissant pour chaque (u, v) une altitude z. Le géoréférencement de la grille intrinsèque au MNE permet alors de définir le point 3D (X, Y, Z) correspondant.

Modèle Numérique de Terrain (MNT) : contrairement au MNE, un MNT ne représente que la surface du sol, à l'exclusion des objets constituant le sursol que sont la végétation, les bâtiments en particulier.

monoscopie : se dit d'une situation dans laquelle l'objet n'est vu que d'une seule image. Dans ce cas, les informations d'ombres projetées sont en général utilisées pour déterminer la forme tridimensionnelle de l'objet.

multiscopie : se dit d'une situation dans laquelle l'objet est vu sous $n > 2$ points de vue. Par rapport à la stéréoscopie simple, ce contexte apporte une redondance d'informations qui permet d'améliorer la précision et l'exhaustivité des renseignements en particulier en limitant le nombre de parties cachées.

N

niveau de généralisation : taille maximale du détail qui peut être omis ou caricaturé dans une modélisation représentative de la réalité.

NP-complet (problème) : un problème qui n'est pas soluble par un algorithme déterministe en temps polynomial en fonction de la taille des données est dit NP. On dit qu'un problème est NP-complet quand tous les problèmes de la classe NP lui sont réductibles. Nous ne faisons qu'effleurer ici les notions de complexité des algorithmes. Nous renvoyons le lecteur intéressé à plus de détail dans [Pap94].

O

occultation : ou partie cachée. Ce phénomène est très courant en stéréoscopie où certaines parties ne sont pas visibles en raison de l'inclinaison de la prise de vue et du relief qui masque des informations.

ordre constructif : ordre permettant de déduire les éléments les uns à la suite des autres.

orthophotographie : photographie supersposable à une carte. C'est donc une image recréée de façon à avoir une transformation axonométrique et non plus perspective. Ceci implique de corriger tous les effets de relief et de perspective.

P

paramètres de projection : pour un cliché aérien, regroupent les paramètres intrinsèques et extrinsèques de la caméra liée au cliché. En supposant une caméra sténopé, la connaissance de ces paramètres permet le calcul de la projection $\mathbf{m}$ d'un point 3D $\mathbf{M}$ par une relation faisant intervenir une matrice de transformation perspective.

paramètres extrinsèques : sont liés à la position de la caméra dans l'espace au moment de la prise de vue et regroupent donc les paramètres de translation et de rotation définissant la position du sommet de prise de vue ($\mathbf{C}_i$) et l'orientation de la caméra (roulis, tangage, lacet) dans le référentiel global. Les paramètres extrinsèques constituent la calibration externe de la caméra et sont déterminés par des techniques d'aérotriangulation.

paramètres intrinsèques : regroupent l'ensemble des données issues de la calibration interne de la caméra (focale, position du point principal dans l'image, taille des pixels, distorsion éventuelle).

paramétrique (modèle) : modèle de bâtiment décrit par un ensemble de paramètres géométriques : la structure topologique est définie au départ, les paramètres définissent les dimensions du bâtiment.

partition (d'une image) : découpage d'une image en régions disjointes dont l'union recouvre la totalité de l'image.

polyédrique (modèle) : modèle de bâtiment générique décrit comme juxtaposition de facettes planes adjacentes. Seuls les bâtiments aux formes arrondies ne peuvent être représentés par cette modélisation.

prismatique (modèle) : modèle de bâtiment dont la forme planimétrique est un polygone quelconque et qui a un toit plat situé à une altitude z.

R

raster (information) : une information raster est une information discrète décrite sous forme d'une grille régulière de valeurs. Typiquement, une image est une information raster puisque chaque pixel fournit une valeur discrète. Une information raster n'est pas structurée et s'oppose ainsi à une information vecteur.

représentation 2D1/2 : représentation dans laquelle chaque point planimétrique (x, y) est affecté d'une seule altitude.

S

segmentation : processus de partitionnement d'une image en régions.

sommet de prise de vue : désigne, dans un modèle sténopé de caméra, le point vers lequel convergent tous les rayons perspectifs.

stéréoscopie : se dit d'une situation dans laquelle deux images permettent de voir l'objet depuis deux points de vue différents, ce qui permet d'en percevoir la forme tridimensionnelle.

structurel (modèle) : modèle de bâtiment décrit comme assemblage de modules élémentaires simples.

sursegmentation : partionnement extrêmement fin de l'image en régions de petite taille.

T

top-down (approches) : approches mises en oeuvre dans les travaux faisant intervenir des bibliothèques de modèles. Ces approches recherchent un modèle spécifié dans les données.

tri topologique : le tri topologique d'un graphe orienté acyclique G consiste à ordonner linéairement tous ses sommets de sorte que si G contient un arc (u, v), u apparaisse avant v.

V

vecteur (information) : une information vecteur est une information décrite sous forme d'un ensemble de primitives (plans, arêtes, points, surfaces plus complexes) et éventuellement de leurs relations topologiques. Une information vecteur structure donc la scène et apporte une description continue de l'environnement. Elle s'oppose à une information raster.

vectorisation : processus de passage d'une information raster à une information vectorielle.

voxels : volume élémentaire de l'espace 3D discrétisé.

Bibliographie

[ADWS02] L. ALVAREZ, R. DERICHE, J. WEICKERT et J. SANCHEZ. « Dense Disparity Map Estimation Respecting Image Discontinuities : A PDE and Scale-Space Based Approach ». *International Journal of Visual Communication and Image Representation*, 13(1/2) :3–21, mars 2002.

[AF87] N. AYACHE et B. FAVERJON. « Efficient registration of stereo images by matching graph of edge segments ». *International Journal on Computer Vision*, 1(2) :107–131, avril 1987.

[AF00] B. AMERI et D. FRITSCH. « Automatic 3D Building Reconstruction Using Plane-Roof Structures ». Dans *ASPRS*, Washington DC, 2000.

[Aya89] N. AYACHE. *Vision Stéréoscopique et Perception Multisensorielle : Application à la Robotique Mobile*. Collection Science Informatique, G. Huet. Inter-Editions, Paris, mai 1989. 341 pages.

[Bai97] C. BAILLARD. « *Analyse d'images aériennes stéréoscopiques pour la restitution 3D des milieux urbains* ». Thèse de doctorat, Ecole Nationale Supérieure des Télécommunications, Paris, octobre 1997.

[BBPP99] I.M. BOMZE, M. BUDINISH, P.M. PARDALOS et M. PELLILO. The Maximum Clique Problem. Dans D.-Z. DU et P.M PARDALOS, éditeurs, *Handbook of Combinatorial Optimization (supp Vol. A)*, pages 1–74. Kluwer, Dordrecht, 1999.

[BD00] C BAILLARD et O. DISSARD. « A Stereo Matching Algorithm for Urban Digital Elevation Models ». *Photogrammetric Engineering and Remote Sensing*, 66(9) :1119–1128, septembre 2000.

[BFJW01] D. BOLDO, F. FUCHS, F. JUNG et Y. WEHRSTED. « Automating the aerial triangulation of digital aerial frame camera images ». Dans N. EL-SHEIMY, éditeur, *Proceedings of the 3rd International Symposium on Mobile Mapping Technology (MMS)*, Cairo, Egypt, janvier 2001. (CD-ROM).

[BHF01] C. BRENNER, N. HAALA et D. FRITSCH. « Towards Fully Automated 3D City Model Generation ». Dans E.P. BALTSAVIAS, A. GRUEN et L.V. GOOL, éditeurs, *Automatic Extraction of Man-Made Objects from Aerial and Space Images (III)*, pages 47–57, Centro Stefano Franscini, Monte Verità, Ascona, 2001. A.A. Balkema Publishers.

[BK73] C. BRON et J. KERBOSCH. « Algorithm 457 : Finding all cliques of an undirected graph ». *Communications of the ACM*, 16(9) :575–577, septembre 1973. ISSN :0001-0782.

[Bla97] T. BLASZKA. « *Approches par modèles en vision précoce* ». Thèse de doctorat, Université de Nice Sophia-Antipolis, Sophia-Antipolis, février 1997.

[BM99] C. BAILLARD et H. MAITRE. « 3-D Reconstruction of Urban Scenes from Aerial Stereo Imagery : A Focusing Strategy ». *Computer Vision and Image Understanding*, 76(3) :244–258, décembre 1999.

[Bol02] D. BOLDO. « *Fabrication conjointe de MNS et d'ortho-images pour la visualisation perspective de scènes urbaines* ». Thèse de doctorat, Université René Descartes, Paris, juin 2002.

[Bon00] D. BONDYFALAT. « *Interaction entre Symbolique et Numérique ; Application à la Vision Artificielle* ». Thèse de doctorat, Univ de Nice Sophia-Antipolis, Sophia-Antipolis, septembre 2000.

[BP01] R. BATTITI et M. PROTASI. « Reactive Local Search for the Maximum Clique Problem ». *Algorithmica*, 29(4) :610–637, avril 2001.

[BPDR03] F. BRETAR, M. PIERROT-DESSEILLIGNY et M. ROUX. « Estimating Intrinsic Accuracy of Airborne Laser Data with local 3D-offsets ». Dans H-G. MAAS, G. VOSSELMAN et A. STREILEIN, éditeurs, *Proceedings of the ISPRS working group III/3 workshop on 3-D reconstruction from airborne laserscanner and InSAR data*, volume 34/3/W13, pages 20–26, Dresden, Germany, octobre 2003. ISSN :1682-1750.

[BPPDH03] O. BENTRAH, N. PAPARODITIS, M. PIERROT-DESEILLIGNY et R. HORAUD. « Modeling Urban 3D Environments : the Stereopolis System ». Dans *International Advanced Workshop on Virtual Geographic Environments and Geocollaboration*, Hong-Kong, décembre 2003.

[BR97] J.R. BEVERIDGE et E.M. RISEMAN. « How Easy is Matching 2D Line Models Using Local Search ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(6) :564–579, juin 1997. ISSN :0162-8828.

[BSZF99] C. BAILLARD, C. SCHMID, A. ZISSERMAN et A. FITZGIBBON. « Automatic Line matching and 3D Reconstruction of Buildings from Multiple Views ». Dans *Proceedings of ISPRS Conference on Automatic Extraction of GIS Objects from Digital Imagery, IAPRS*, volume 32(3), pages 69–80, septembre 1999.

[BY95] J-D. BOISSONNAT et M. YVINEC. *Géométrie Algorithmique*. Collection Informatique, C. Puech and J-M. Rifflet. Ediscience International, Paris, 1995.

[BZ99] C. BAILLARD et A. ZISSERMAN. « Automatic Reconstruction of piecewise planar models from multiple views ». Dans *Proceedings of 18th the Conference on Computer Vision and Pattern Recognition (CVPR'99)*, Fort Collins, CO, juin 1999. IEEE Computer Society.

[Cho00] G. CHOU. « *Large-Scale reconstruction : a triangulation-based approach* ». PhD Thesis, Massachusetts Institute of Technology, Cambridge, MA, août 2000.

[CJC+98] R. COLLINS, C. JAYNES, Y-Q. CHENG, X. WANG, F. STOLLE, E. RISEMAN et A. HANSON. « The Ascender System : Automated Site Modeling from Multiple Aerial Images ». *Computer Vision and Image Understanding*, 72(2) :143–162, novembre 1998.

[CJC01] M. CORD, M. JORDAN et JP. COQUEREZ. « Accurate Building Structure Recovery from High Resolution Aerial Imagery ». *Computer Vision and Image Understanding*, 82(2) :138–173, mai 2001.

[CJPDS03] N. CHEHATA, F. JUNG, M. PIERROT-DESEILLIGNY et G. STAMON. « A Region-Based Matching Approach for 3D-Roof Reconstruction from High Resolution satellite Stereo Pairs ». Dans Changming SUN, Hugues TALBOT, Sébastien OURSELIN et Tony ADRIAANSEN, éditeurs, *Proceedings of the Seventh International Conference on Digital Image Computing : Techniques and Applications, DICTA 2003*, pages 889–898, Macquarie University, Sydney, Australia, décembre 2003. APRS, CSIRO Publishing.

[Col96] R.T. COLLINS. « A Space-Sweep Approach to True Multi-Image Matching ». Dans *Proceedings of the 15th Conference on Computer Vision and Pattern Recognition (CVPR'96)*, pages 358–365, San Francisco, CA, juin 1996. IEEE Computer Society.

[COM03] « Contrat d'Objectifs et de Moyens pour les années 2003 à 2006 entre l'État et l'IGN ». http ://www.ign.fr/telechargement/Pi/publication/Plaquette-COM.pdf, 2003.

[CPJ98] M. CORD, N. PAPARODITIS et M. JORDAN. « Dense, Reliable and Depth Discontinuity Preserving DEM Computation from H.R.V Urban Stereopairs ». Dans *Proceedings of ISPRS Commission II Symposium*, volume 32(2), Cambridge, UK, juillet 1998.

[Dan94] T. DANG. « *Interprétation et restitution automatique des bâtiments isolés à partir d'un couple stéréoscopique d'images aériennes* ». Thèse de doctorat, ENST, Paris, décembre 1994.

[Der87] R. DERICHE. « Using Canny's Criteria to Derive a Recursively Implemented Optimal Edge Detector ». *International Journal of Computer Vision*, 1(2) :167–187, mai 1987.

[DTM96] P.E. DEBEVEC, C.J. TAYLOR et J. MALIK. « Modeling and Rendering Architecture from Photographs : a Hybrid Geometry-and Image-Based Approach ». Dans *Proceedings of SIGGRAPH 96*, Computer Graphics Proceedings, Annual Conference Series, pages 11–20, août 1996.

[DVF92] R. DERICHE, R. VAILLANT et O. FAUGERAS. From Noisy Edges Points to 3D Reconstruction of a Scene : A Robust Approach and Its Uncertainty Analysis. Dans *Theory and Applications of Image Analysis*, volume 2 de *Series in Machine Perception and Artificial Intelligence*, pages 71–79. World Scientific, 1992.

[EBM02] A.F. ELAKSHER, J. BETHEL et E.M. MIKHAIL. « Reconstructing 3D buildings wireframes from multiple images ». Dans R.KALLIANY et F. LEBERL, éditeurs, *Proceedings of the ISPRS Commission 3 Symposium on Photogrammetric Computer Vision (PCV'02)*, volume 34 de *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages A91–96, Graz, Austria, septembre 2002. Institute for Computer Graphics and Vision, Graz University of Technology.

[Ede87] H. EDELSBRUNNER. *Algorithms in Combinatorial Geometry*. EATCS Monographs on theoretical Computer Science. Springer-Verlag, New York, Inc, 1987.

[FBH00] W. FÖRSTNER, A. BRUNN et S. HEUEL. « Statistically Testing Uncertain Geometric Relations ». Dans G. SOMMER, N. KRÜGER et Ch. PERWASS, éditeurs, *Mustererkennung 2000*, pages 17–26. DAGM, Springer, septembre 2000.

[FG87] W. FÖRSTNER et E. GÜLCH. « A fast operator for detection and precise location of distinct points, corners and centers of circular features ». Dans *Proceedings of Intercommission Conference on Fast Processing of Photogrammetric Data*, pages 281–305, Interlaken, Suisse, juin 1987. ISPRS.

[FJM+02] F. FUCHS, H. JIBRINI, G. MAILLET, M. PIERROT-DESEILLIGNY et F. TAILLANDIER. « Trois approches pour la construction automatique de modèles 3D de batiments en imagerie aerienne haute résolution ». *Bulletin de la Société Française de Photogrammétrie et de Télédétection*, 166 :10–18, mai 2002.

[FKL+98] A. FISCHER, T. KOLBE, F. LANG, A. CREMERS, W. FÖRSTNER, L. PLÜMER et V. STEINHAGE. « Extracting Buildings from Aerial Images Using Hierarchical Aggregation in 2D and 3D ». *Computer Vision and Image Understanding*, 72(2) :163–185, novembre 1998.

[FLM99] F. FUCHS et H. LE-MEN. « Building Reconstruction on Aerial Images Through Multi-Primitive Graph Matching ». Dans W.G. KROPATSCH et J-M. JOLION, éditeurs, *Proceedings of the 2nd IAPR-TC-15 Workshop on Graph-based Representations*, pages 21–30, Vienna, Austria, mai 1999. Österreiche Computer Gesellschaft.

[FLP01] O. FAUGERAS, Q.T. LUONG et T. PAPADOPOULO. *The Geometry of Multiple Images*. The MIT Press, 2001.

[FMJ03] D. FLAMANC, G. MAILLET et H. JIBRINI. « 3D City Models : an Operational Approach using Aerial Images and Cadastral Maps ». Dans H. EBNER, C. HEIPKE, H. MAYER et K. PAKZAD, éditeurs, *Proceedings of the ISPRS Conference Photogrammetric on Image Analysis (PIA'03)*,

volume 34 de *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 53–58, Münich, Germany, septembre 2003. Institute for Photogrammetry and GeoInformation University of Hannover, Germany. ISSN : 1682-1750.

[För01] W. FÖRSTNER. Direct optimal estimation of geometric entities using algebraic projective geometry. Dans *Festschrift anläßlich des 60. Geburtstages von Prof. Dr.-Ing. Bernhard Wrobel*, pages 69–87. Technische Universität Darmstadt, 2001.

[Fuc95] C. FUCHS. « Polymorphic Grouping for Image Segmentation ». Dans *Proceedings of the 5th International Conference on Computer Vision (ICCV'95)*, pages 175–182, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA, juin 1995. IEEE Computer Society.

[Fuc01] F. FUCHS. « *Contribution à la reconstruction du bâti en milieu urbain, à l'aide d'images aériennes stéréoscopiques à grande échelle. Étude d'une approche structurelle* ». Thèse de doctorat, Université de René Descartes, Paris, avril 2001.

[GBB98] P. GROS, O. BOURNEZ et E. BOYER. « Using Local Planar Geometric Invariant to Match and Model Images of Line Segments ». *Computer Vision and Image Understanding*, 69(2) :135–155, février 1998.

[GLMC03] L. GUIGUES, H. LE-MEN et J-P. COQUEREZ. « Scale-Sets Image Analysis ». Dans *Proceedings of the International Conference on Image Processing (ICIP'03)*, volume 2, pages 45–48, Barcelona, Spain, septembre 2003. IEEE Signal Processing Society.

[Gro02] E. GROSSMANN. « *Maximum Likelihood 3D Reconstruction From One or More Uncalibrated Views Under Geometric Constraints* ». PhD Thesis, Universidade Tecnica de Lisboa, Lisboa, Portugal, octobre 2002.

[Gru98] A. GRUEN. « TOBAGO - a Semi-Automated Approach for the Generation of 3-D Building Models. ». *Journal of Photogrammetry and Remote Sensing*, 53(2) :108–118, avril 1998.

[GSV02] E. GROSSMANN et J. SANTOS-VICTOR. « Maximum Likelihood 3D Reconstruction from One or More Images Under Geometric Constraints ». Dans *British Machine vision Conference*, Cardiff, UK, septembre 2002.

[Gui03] L. GUIGUES. « *Modèles multi-échelles pour la segmentation d'images* ». Thèse de doctorat, Université de Cergy-Pontoise, Cergy-Pontoise, décembre 2003.

[GW98] A. GRUEN et X. WANG. « CC-Modeler : a topology generator for 3D city models ». *Journal of Photogrammetry and Remote Sensing*, 53(5) :286–295, octobre 1998.

[Har95a] R. HARTLEY. « In defense of the 8-point algorithm ». Dans *Proceedings of the 5th International Conference on Computer Vision (ICCV'95)*, pages 1064–1070, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA, juin 1995. IEEE Computer Society.

[Har95b] R. HARTLEY. « A linear method for reconstruction from lines and points ». Dans *Proceedings of the 5th International Conference on Computer Vision (ICCV'95)*, pages 882–887, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA, juin 1995. IEEE Computer Society.

[HB97] O. HENRICSSON et E.P. BALTSAVIAS. « 3-D building reconstruction with ARUBA : a qualitative and quantitative evaluation ». Dans A. GRUEN, E.P. BALTSAVIAS et O. HENRICSSON, éditeurs, *Automatic Extraction of Man-Made Objects from Aerial and Space Images (II)*, pages 65–76, Centro Stefano Franscini, Monte Verità, Ascona, 1997. Birkhaeuser Verlag, Basel.

[HB98] N. HAALA et C. BRENNER. « Interpretation of urban surface models using 2D building information ». *Computer Vision and Image Understanding*, 72(2) :204–214, novembre 1998.

[Hen98] O. HENRICSSON. « The role of Color Attributes and Similarity Grouping in 3-D Building Reconstruction ». *Computer Vision and Image Understanding*, 72(2) :163–184, novembre 1998.

[HJBJ+96] A. HOOVER, G. JEAN-BAPTISTE, X. JIANG, P.J. FLYNN, H. BUNKE, D. GOLDGOF, K. BOWYER, D. EGGERT, A. FITZGIBBON et R. FISHER. « An experimental Comparison of Range Segmentation Algorithms ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(7) :673–689, juillet 1996.

[HLF00] S. HEUEL, F. LANG et W. FÖRSTNER. « Topological and Geometrical Reasoning in 3D Grouping for Reconstructing Polyhedral Surfaces ». Dans *Proceedings of the XIXth ISPRS Congress*, volume 33 de *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 397–404, Amsterdam, juillet 2000. ISPRS.

[HMS+01] A. HANSON, M. MARENGONI, H. SCHULTZ, F. STOLLE, E. RISEMAN et C. JAYNES. « Ascender II : a Framework for Reconstruction of Scenes from Aerial Images ». Dans E.P. BALTSAVIAS, A. GRUEN et L.V. GOOL, éditeurs, *Automatic Extraction of Man-Made Objects from Aerial and Space Images (III)*, pages 25–34, Centro Stefano Franscini, Monte Verità, Ascona, 2001. A.A. Balkema Publishers.

[HP76] S.L. HOROWITZ et T. PAVLIDIS. « Picture Segmentation by a Tree Traversal Algorithm ». *ournal of the Association for Computing Machinery*, 23(2) :368–388, avril 1976.

[HS89] R. HORAUD et T. SKORDAS. « Stereo Correspondence through Feature Groupings and Maximal Cliques ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(11) :1168–1180, novembre 1989.

[HV95] C. HOFFMANN et P. VERMEER. Geometric Constraint Solving in R^2 and R^3. Dans D.Z. DU et F. HWANG, éditeurs, *Computing in Euclidean Geometry*, pages 266–298. World Scientific Publishing, Singapore, 2^{nd} édition, 1995.

[HZ02] R. HARTLEY et A. ZISSERMAN. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2002.

[IJ01] H. ISHIKAWA et I.H. JERMYN. « Region Extraction from Multiple Images ». Dans *Proceedings of the 8th IEEE International Conference on Computer Vision (ICCV 01)*, volume 1, pages 509–516. IEEE Computer Society, juillet 2001.

[JDA95] O. JAMET, O. DISSARD et S. AIRAULT. « Building Extraction from Stereo Pairs of Aerial Images : Accuracy and Productivity Constraint of a Topographic Production Line ». Dans A. GRUEN, O. KUEBLER et P. AGOURIS, éditeurs, *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, pages 231–240, Centro Stefano Franscini, Monte Verità, Ascona, 1995. Birkhaeuser Verlag, Basel.

[JHRS97] C. O. JAYNES, A. HANSON, E. RISEMAN et H. SCHULTZ. « Building reconstruction from Optical and Range Images ». Dans *Proceedings of the 16th Conference on Computer Vision and Pattern Recognition (CVPR'97)*, pages 380–386, San Juan, Puerto Rico, juin 1997. IEEE Computer Society.

[Jib02] H. JIBRINI. « *Reconstruction automatique des bâtiments en modèles polyèdriques 3D à partir de données cadastrales vectorisées 2D et d'un couple d'images aériennes à haute résolution* ». Thèse de doctorat, Ecole Nationale Supérieure des Télécommunications, Paris, avril 2002.

[JPPDM00] H. JIBRINI, N. PAPARODITIS, M. PIERROT-DESEILLIGNY et H. MAITRE. « Automatic Building Reconstruction from very high resolution aerial stereopairs using cadastral ground plans ». Dans *Proceedings of the XIXth ISPRS Congress*, volume 33 de *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Amsterdam, juillet 2000. ISPRS.

[JRH03] C. JAYNES, E. RISEMAN et A. HANSON. « Recognition and Reconstruction of Buldings from Multiple Aerial Images ». *Computer Vision and Image Understanding*, 90(1) :68–98, avril 2003.

[JTP02] F. JUNG, V. TOLLU et N. PAPARODITIS. « Extracting 3D Edgel Hypotheses From Mutliple Calibrated Images : a Step Towards The Reconstruction of Curved and Straight Oject Boundary Lines ». Dans R.KALLIANY et F. LEBERL, éditeurs, *Proceedings of the ISPRS Commission 3 Symposium on Photogrammetric Computer Vision (PCV'02)*, volume 34 de *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages B100–104, Graz, Austria, septembre 2002. Institute for Computer Graphics and Vision, Graz University of Technology.

[KDN+95] T. KANUNGO, B. DOM, W. NIBLACK, D. STEELE et J. SHEINVALD. « MDL-Based Multi-Band Image Segmentation Using a Fast Region Merging Scheme ». Rapport Technique RJ 9960, IBM Research Division, Almaden Research Center, San Jose, Ca 95120-6099, mai 1995.

[KHN01] Z. KIM, A. HUERTAS et R. NEVATIA. « Automatic Description of Buildings with Complex Rooftops from Multiple Images ». Dans *Proceedings of 20th the Conference on Computer Vision and Pattern Recognition (CVPR'01)*, volume 2, pages 272–279, Hauai, Hawaii, décembre 2001. IEEE Computer Society.

[Koe02] M. KOEHL. « Modélisation géométrique et sémantique dans un système d'information topographique tridimensionnel ». Rapport Technique, Séminaire interne de l'UMR MAP CNRS/MCC 694, Sousse (Tunisie), avril 2002. 6 pages.

[Kol99] T.H. KOLBE. « *Identifikation und Rekonstruktion von Gebäuden in Luftbildern mittels unsharfer Constraints* ». PhD Thesis, Institute for Environmental Sciences, University of Vechta, 1999.

[Kra93] K. KRAUS. *Photogrammetry*, volume 1. Dümmler/Bonn, 4nd édition, 1993. with contributions by P. Waldhäusl.

[LN98] C. LIN et R. NEVATIA. « Building Detection and Description from a Single Intensity Image ». *Computer Vision and Image Understanding*, 72(2) :101–121, novembre 1998.

[May99] H. MAYER. « Automatic Object Extraction from Aerial Imagery - A Survey Focusing on Buildings ». *Computer Vision and Image Understanding*, 74(2) :138–149, mai 1999.

[ML92] H. MAITRE et W. LUO. « Using Models to Improve Stereo Reconstruction ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2) :269–277, 1992.

[MN85] G. MEDIONI et R. NEVATIA. « Segment-based Stereo Matching ». *Computer Vision, Graphics and Image Processing*, 31(3) :2–18, juillet 1985.

[Moo96] T. MOONS. « Epipolar, trifocal and quadrifocal geometry in a nutshell ». Rapport Technique KUL/ESAT/MI2/9611, K.U. Leuven, septembre 1996.

[MPT02] G. MAILLET, N. PAPARODITIS et F. TAILLANDIER. « TIN Surface Reconstruction From Multiple Calibrated Aerial Images in Urban Areas. ». Dans *International Conference on Image Processing (ICIP 02)*, volume 3, pages 521–524, Rochester, septembre 2002. IEEE Computer Society.

[MS94] J.C. MCGLONE et J.A. SHUFELT. « Projective and Object Space Geometry for Monocular Building Extraction ». Dans *Proceedings of the 13th Conference on Computer Vision and Pattern Recognition (CVPR'97)*, pages 54–61, Seatle, USA, juin 1994. IEEE Computer Society.

[Nev96] J. NEVEU. *Probabilités*. École Polytechnique, 1996.

[NN01] S. NORONHA et R. NEVATIA. « Detection and Modeling of Buildings from Multiple Aerial Images ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(5) :501–518, mai 2001.

[ODZ03] M. ORTNER, X. DESCOMBES et J. ZERUBIA. « Un nouveau modèle pour l'extraction de caricatures de bâtiments sur des Modèles Numériques d'Elévation ». Dans *Proceedings of Traitement et Analyse de l'Information Méthodes et Applications Conference*, Hammamet, Tunisie, octobre 2003.

[Pap94] C.H. PAPADIMITRIOU. *Computational Complexity*. Addison-Wesley Pub Co, 1994.

[PCJC98] N. PAPARODITIS, M. CORD, M. JORDAN et J.-P. COQUEREZ. « Building detection and reconstruction from mid-and high resolution aerial imagery ». *Computer Vision and Image Understanding*, 72(2) :122–142, novembre 1998.

[Pen06] L. PENARD. « *Reconstruction et visualisation de scènes urbaines à partir d'images terrestres et aériennes géoréférencées haute résolution* ». Thèse de doctorat, Paris V, Paris, 2006. en cours.

[PF98] T. PAPADOPOULO et O. FAUGERAS. « A new characterization of the trifocal tensor ». Dans H. BURKHARDT et B. NEUMANN, éditeurs, *Proceedings of the 5th European Conference on Computer Vision (ECCV'98)*, volume 1406 de *Lecture Notes in Computer Science*, pages 109–123, Freiburg, Germany, juin 1998. Springer.

[PMT+01] N. PAPARODITIS, G. MAILLET, F. TAILLANDIER, H. JIBRINI, F. JUNG, L. GUIGUES et D. BOLDO. « Multi-image 3D feature and DSM extraction for change detection and building reconstruction ». Dans E.P. BALTSAVIAS, A. GRUEN et L.V. GOOL, éditeurs, *Automatic Extraction of Man-Made Objects from Aerial and Space Images (III)*, Centro Stefano Franscini, Monte Verità, Ascona, 2001. A.A. Balkema Publishers.

[PTJ00] N. PAPARODITIS, C. THOM et H. JIBRINI. « Surface reconstruction in urban areas from multiple views of aerial digital frames ». Dans *Proceedings of the XIXth ISPRS Congress*, volume 33 de *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Amsterdam, juillet 2000. ISPRS.

[RB02] F. ROTTENSTEINER et C. BRIESE. « A New Method for Building Extraction in Urban Areas from High-Resolution LIDAR Data ». Dans R.KALLIANY et F. LEBERL, éditeurs, *Proceedings of the ISPRS Commission 3 Symposium on Photogrammetric Computer Vision (PCV'02)*, volume 34 de *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages B295–301, Graz, Austria, septembre 2002. Institute for Computer Graphics and Vision, Graz University of Technology.

[RC98] S. ROY et I.J COX. « A Maximum-Flow Formulation of the N-camera Stereo Correspondence Problem ». Dans *Proceedings of the 7th International Conference on Computer Vision (ICCV 98)*, pages 492–499, Bombay, India, janvier 1998. IEEE Computer Society.

[Ris89] J. RISSANEN. *Stochastic Complexity in Statistical Inquiry*, volume 15 de *World Scientific Series in Computer Science*. World Scientific Publishing Co., Inc., River Edge, NJ, USA, 1989.

[Sch99] C. SCHMID. « A Structured Probabilistic Model for Recognition ». Dans *Proceedings of the Conference on Computer Vision and Pattern Recognition*, volume 2, pages 485–490, Fort Collins, CO, juin 1999.

[Sha94] A. SHASHUA. « Trilinearity in Visual Recognition by Alignment ». Dans J-O. EKLUNDH, éditeur, *Proceedings of the 3th European Conference on Computer Vision (ECCV'94)*, volume 800 de *Lecture Notes in Computer Science*, pages 479–484, Stockholm, Sweden, mai 1994. Springer.

[SHS98] H-Y. SHUM, M. HAN et R. SZELISKI. « Interactive Construction of 3D Models from Panoramic Mosaics ». Dans *Proceedings of 17th the Conference on Computer Vision and Pattern Recognition (CVPR'98)*, pages 427–433, Santa Barbara, CA, juin 1998. IEEE Computer Society.

[SMVG02] S. SCHOLZE, T. MOONS et L. VAN-GOOL. « A probabilistic Approach to Building Roof Reconstruction using Semantic Labelling ». Dans Luc J. Van GOOL, éditeur, *Proceedings of the 24th DAGM Symposium*, volume 2449 de *Lecture Notes in Computer Science*, pages 257–264, Zürich, Switzerland, septembre 2002. Springer.

[SSF96] M. SESTER, W. SCHNEIDER et D. FRITSCH. « Results of the Test on Image Understanding of ISPRS Working Group III/3 ». Dans *Proceedings of the XVIIIth ISPRS Congress*, volume 31 de *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages B3 :768–773, Vienna, juillet 1996. ISPRS.

[SV00] I. SUVEG et G. VOSSELMAN. « 3D Reconstruction of Building Models ». Dans *Proceedings of the XIXth ISPRS Congress*, volume 33 de *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages B3 :538–545, Amsterdam, juillet 2000. ISPRS.

[SVP+96] M. SCHOLL, A. VOISARD, J-P. PELOUX, L. RAYNAL et P. RIGAUX. *SGBD Géographiques. Spécificités.* International Thomson Publishing, France, février 1996.

[SW03] H-F. SCHUSTER et U. WEIDNER. « A new Approach Towards Quantitative Quality Evaluation of 3D Building Models ». Dans J. SCHIEWE, L. HAHN, M. MADDEN et M. SESTER, éditeurs, *Proceedings of the ISPRS Commission IV Joint Workshop "Challenges in Geospatial Analysis, Integration and Visualization II"*, pages 156–163, Stuttgart, Germany, septembre 2003.

[SZ97] C. SCHMID et A. ZISSERMAN. « Automatic line matching across views ». Dans *Proceedings of the 16th Conference on Computer Vision and Pattern Recognition (CVPR'97)*, pages 666–671, San Juan, Puerto Rico, juin 1997. IEEE Computer Society.

[Tar96] J.P. TAREL. « Reconstruction globale et robuste de facettes 3D ». Rapport Technique RR-2813, INRIA, février 1996.

[TD02a] F. TAILLANDIER et R. DERICHE. « Reconstruction de primitives linéaires 3D en multi-vues pour la modélisation de scènes urbaines ». Dans *Actes du 13ème Congrès Francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle*, volume I, pages 115–124, Angers, France, janvier 2002. AFRIF-AFIA.

[TD02b] F. TAILLANDIER et R. DERICHE. « Reconstruction of 3D Linear Primitives from Multiple Views for Urban Areas Modelisation ». Dans R.KALLIANY et F. LEBERL, éditeurs, *Proceedings of the ISPRS Commission 3 Symposium on Photogrammetric Computer Vision (PCV'02)*, volume 34 de *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages B267–272, Graz, Austria, septembre 2002. Institute for Computer Graphics and Vision, Graz University of Technology.

[TD04a] F. TAILLANDIER et R. DERICHE. « Automatic Buildings Reconstruction from Aerial Images : a Generic Bayesian Framework. ». Dans *Proceedings of the XXth ISPRS Congress*, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Istanbul, Turkey, juillet 2004. ISPRS. to be published.

[TD04b] F. TAILLANDIER et R. DERICHE. « Reconstruction automatique de bâtiments à partir d'images aériennes : une approche probabiliste générique. ». Dans *Actes du 14ème Congrès Francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle*, Toulouse, France, janvier 2004. AFRIF-AFIA.

[TGD03] F. TAILLANDIER, L. GUIGUES et R. DERICHE. « A Framework for Constrained Multi-Scale Range Image Segmentation ». Dans *Proceedings of the International Conference on Image Processing (ICIP'03)*, Barcelona, Spain, septembre 2003. IEEE Signal Processing Society.

[TMHF00] Bill TRIGGS, Philip MCLAUCHLAN, Richard HARTLEY et Andrew FITZGIBBON. Bundle Adjustment – A Modern Synthesis. Dans W. TRIGGS, A. ZISSERMAN et R. SZELISKI, éditeurs, *Vision Algorithms : Theory and Practice*, LNCS, pages 298–375. Springer Verlag, 2000.

[TZ97] P.H.S. TORR et A. ZISSERMAN. « Robust parameterization and computation of the trifocal tensor ». *Image and Vision Computing*, 15(8) :591–605, 1997.

[Val04] B. VALLET. « recalage de modèles 3D géométriques contraints dans les images ». Rapport de dea, DEA Intelligence Artificielle, Reconnaissance des Formes et Applications, Paris VI, juillet 2004.

[VC02] S. VINSON et L.D. COHEN. « Extraction de bâtiments complexes à partir d'images aériennes et de MNE ». Dans *Actes du 13ème Congrès Francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle*, volume 1, pages 125–134, Angers, France, janvier 2002. AFRIF-AFIA.

[vdH99] F-A. van den HEUVEL. « A Line-photogrammetric Mathematical Model for the Reconstruction of Polyhedral Objects ». Dans Sabry F. EL-HAKIM, éditeur, *Videometrics VI, SPIE*, volume 3641, pages 60–71, 1999.

[vdHV97] F.A van den HEUVEL et M.G. VOSSELMAN. « Efficient 3D-modelling of buildings using a priori geometric object information ». Dans Sabry F. EL-HAKIM, éditeur, *SPIE*, volume 3174, pages 38–49, 1997.

[Ves00] C. VESTRI. « *Outils pour la reconstruction automatique de bâtiments à partir d'imagerie aérienne* ». Thèse de doctorat, INRIA, Sophia-Antipolis, septembre 2000.

[Vez95] J-M VEZIEN. « *Techniques de reconstruction globale par analyse de paires d'images stéréoscopiques* ». Thèse de doctorat, Université Paris VII, Paris, avril 1995.

[VS01] G. VOSSELMAN et I. SUVEG. « Map based building reconstruction from laser data and images ». Dans E.P. BALTSAVIAS, A. GRUEN et L.V. GOOL, éditeurs, *Automatic Extraction of Man-Made Objects from Aerial and Space Images (III)*, pages 231–239, Centro Stefano Franscini, Monte Verità, Ascona, 2001. A.A. Balkema Publishers.

[WA96] W. WILLUHN et F. ADE. « A Rule-Based System for House Reconstruction from Aerial Images ». Dans *13th International Conference on Pattern Recognition (ICPR 96)*, volume 3, pages 885–889, Vienna, Austria, août 1996. IEEE Computer Society.

[Wei97] U. WEIDNER. « *Gebäudeerfassung aus Digitalen Oberflächenmodellen* ». PhD Thesis, institut für Photogrammetrie, Universität Bonn, 1997.

[WF95] U. WEIDNER et W. FÖRSTNER. « Towards Automatic Building Reconstruction from High-Resolution Digital Elevation Models ». *Journal of Photogrammetry and Remote Sensing*, 50(4) :38–49, août 1995.

[WTT+02a] X. WANG, S. TOTARO, F. TAILLANDIER, A.R. HANSON et S. TELLER. « Recovering Facade Texture and Microstructure from Real-World Images ». Dans *Proceedings of the 2nd International Workshop on Texture Analysis and Synthesis in conjunction with ECCV'02*, pages 145–149, Copenhague, Denmark, juin 2002.

[WTT+02b] X. WANG, S. TOTARO, F. TAILLANDIER, A.R. HANSON et S. TELLER. « Recovering Facade Texture and Microstructure from Real-World Images ». Dans R.KALLIANY et F. LEBERL, éditeurs, *Proceedings of the ISPRS Commission 3 Symposium on Photogrammetric Computer*

Vision (PCV'02), volume 34 de *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages A381–386, Graz, Austria, septembre 2002. Institute for Computer Graphics and Vision, Graz University of Technology.

[XZ96] G. XU et Z. ZHANG. *Epipolar Geometry in stereo, motion and object recognition*. Kluwer Academic Publishers, 1996.

[ZF92] Z. ZHANG et O. FAUGERAS. *3D scene analysis : a stereo based approach*. Springer, Berlin, Heidelberg, 1992.