



HAL
open science

Contribution à la reconstruction du bâti en milieu urbain, à l'aide d'images aériennes stéréoscopiques à grande échelle : étude d'une approche structurale

Frank Fuchs

► To cite this version:

Frank Fuchs. Contribution à la reconstruction du bâti en milieu urbain, à l'aide d'images aériennes stéréoscopiques à grande échelle : étude d'une approche structurale. Vision par ordinateur et reconnaissance de formes [cs.CV]. Paris 5, 2001. Français. NNT : . tel-02155360

HAL Id: tel-02155360

<https://hal.science/tel-02155360>

Submitted on 13 Jun 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITE RENE DESCARTES - PARIS V
Centre Universitaire des Saints-Pères
UFR DE MATHEMATIQUES ET INFORMATIQUE

**Thèse présentée en vue de l'obtention du grade de Docteur de
l'Université RENE DESCARTES - PARIS V**

Discipline : Sciences de la Vie et de la Matière
Spécialité : Informatique

Par Monsieur Frank Fuchs

Sujet de la thèse :
Contribution à la reconstruction du bâti en milieu urbain, à l'aide
d'images aériennes stéréoscopiques à grande échelle.
Etude d'une approche structurelle.

Soutenue le 27 avril 2001, devant le jury composé de :

Monsieur Jean-Michel Jolion	Président
Monsieur Rachid Deriche	Rapporteur
Monsieur Henri Maître	Rapporteur
Monsieur Georges Stamon	Directeur de thèse
Monsieur Hervé Le Men	Directeur de thèse
Monsieur Michel Gondran	Examineur
Monsieur Horst Bunke	Examineur

A ma femme, à mes filles.

Remerciements

Je remercie en premier lieu - et sans ironie - le contribuable français qui finance par l'impôt une partie des activités de l'Institut Géographique National. Il a financé ces travaux.

Je remercie Georges Stamon qui m'a encouragé à m'engager dans cette voie dès le premier jour et a dirigé cette thèse et supporté mes piques. Je remercie Hervé Le Men qui a dirigé cette thèse et donné des conseils toujours pertinents.

Je remercie Grégoire qui a contribué à l'évaluation des travaux, ainsi qu'à la rédaction de la partie de mémoire correspondant.

Merci également au Directeur du Service de la Recherche, Serge, ainsi qu'à Daniel et Marc, Chefs consécutifs du laboratoire MATIS.

Merci aussi à mes collègues, ceux de MATIS bien entendu (Olivier, Olivier, Olivier et Olivier, Franck et Franck, Laurent et Laurent, histoire de procéder par la partie gauche de l'histogramme, et tous les autres), mais aussi à mes collègues du laboratoire SIP qui ont bien voulu supporter cette laborieuse répétition de soutenance et d'autres choses.

Je remercie les développeurs ayant contribué au 'noyau', sans oublier Jérôme Hummel et Gilles Roux.

Merci à Laurent, Patrick et Franck pour leur participation au club des 4.

Merci à ma famille.

Merci à mon épouse et mes deux filles.

Mots-clés :

Stéroroscopie, reconstruction du bâti, graphes relationnels attribués, appariement de graphes, tolérance d'erreur, grammaire de graphes

Résumé :

L'Institut Géographique National, producteur de données géographiques, poursuit ses efforts de recherche visant à automatiser la saisie de bases de données localisées. Ainsi, après des études dont le contexte applicatif visait la Base de Données Topographique (BD Topo (R)), nous abordons ici le thème du bâti dont nous cherchons une reconstruction (obtention d'une description en mode *vecteur*) dite 'grande échelle', c'est-à-dire comprenant la structure interne des toits.

Ces travaux se placent dans la suite des travaux menés au MATIS par C. Baillard, concernant l'extraction de modèles numériques d'élévation (MNE), et la caractérisation du sursol. Même si ces travaux se placent dans le cadre d'images en niveaux de gris de résolution moyenne (50 cm à 1m), nous les utilisons aux résolutions dites fortes : 10 à 20 cm. Les images utilisées sont en niveaux de gris. Le modèle numérique d'élévation permet de cerner la forme des toits, et les zones de sursol permettent de focaliser les travaux de reconstruction sur des surfaces de taille raisonnable (au maximum quelques bâtiments). La méthode proposée traite ces zones de sursol indépendamment les unes des autres.

L'approche proposée est structurelle, c'est-à-dire qu'elle s'appuie sur une base de modèles. Ceci place donc la méthode à mi-chemin entre les approches génériques (dans lesquelles on exprime des règles portants sur des objets simples, de façon à pouvoir modéliser, par exemple, *tous* les polyèdres), et les approches spécifiques (par exemple les méthodes basées sur des modèles paramétriques déformables). Les modèles sont donnés de façon déclarative, ce qui confère une certaine souplesse et une certaine extensibilité à la méthode.

L'existence de modèles implique deux problèmes bien particuliers. Le premier est la reconstruction d'une scène à l'aide d'un modèle dont on sait qu'il est adapté à cette scène. Le second est précisément le choix d'un modèle adapté. Ces deux points sont abordés dans la stratégie générale qui suit.

La stratégie générale se compose de trois étapes majeures. La première consiste à détecter dans les données initiales, des objets tri-dimensionnels simples, que l'on appellera primitives. Les primitives ont des natures différentes (segments de droites, portions de plan, façades). Les trois types de primitives sont détectés indépendamment les uns des autres.

La seconde étape majeure consiste à reconstruire effectivement un bâtiment. Les modèles proposés sont donnés sous forme de graphes relationnels attribués. Les graphes ont en effet montré leur faculté de modélisation. Toutefois peu de méthodes ont réellement utilisé les graphes pour des travaux de reconstruction sur des données complexes. Il est vrai que la combinatoire du problème le rend ardu. Des avancées récentes dans le domaine du calcul d'isomorphisme de graphes avec tolérance d'erreur permettent toutefois d'aborder concrètement le problème de l'appariement entre un modèle et des données par cette voie. Par ailleurs, ces techniques récentes se montrent adaptées, pour des raisons théoriques, au cas du bâti.

Le problème principal est la mise en correspondance entre un modèle et des données, c'est-à-dire l'identification, dans les primitives, de celles qui permettent de définir la géométrie du bâtiment à reconstruire. Nous posons cette question en terme d'appariement de graphes, et plus précisément,

nous nous plaçons dans le formalisme des distances d'édition. Ceci permet, pour un modèle donné, de calculer un sous ensemble des données qui ressemble au mieux à ce modèle. La fonction de ressemblance implantée est intégrée dans la distance d'édition. Elle correspond, dans la pratique, à la compatibilité entre les objets géométriques détectés et les propriétés géométriques qu'ils doivent posséder pour être expliqués par le modèle. La combinatoire du problème d'appariement reste toutefois une limite importante de la méthode, c'est pourquoi nous proposons l'utilisation d'une grammaire de graphes, qui permet à la fois de produire des modèles homogènes entre eux (car ils sont générés avec un ensemble de règles, et non séparément), mais surtout de trouver des appariements partiels, que l'on peut exploiter lors d'un appariement global. Cet aspect est traité à la fois d'un point de vue théorique, dans le cadre des grammaires de graphes par remplacement d'hypergraphes, et d'un point de vue pratique.

Après mise en correspondance, une reconstruction finale est réalisée. Cette reconstruction s'appuie exclusivement sur les informations contenues dans le modèle, et, à nouveau, sur la distance d'édition.

Enfin, la troisième étape consiste à choisir, parmi les reconstructions effectuées, celle qui est la plus adaptée aux données, et à décider si elle est pertinente ou non (on sait en effet a priori qu'en raison du nombre fini de modèles engendrés par la grammaire, la méthode ne peut pas traiter toutes les scènes).

Le mémoire présente des résultats obtenus en exploitant la méthode sur des scènes aériennes réelles. Il propose également une évaluation de la méthode, dans un cadre automatique, mais jette aussi des ponts vers la problématique semi-automatique, qu'il faut envisager dès que l'on considère le contexte réel de production que représente l'IGN.

L'extension majeure à apporter à ces travaux concerne le pilotage du choix du modèle a priori à l'aide des données, et non a posteriori, c'est-à-dire après avoir tenté plusieurs reconstructions, dont la très grande majorité est vaine. Il faudra également étudier en profondeur l'impact de nouveaux capteurs et de prises de vues à fort recouvrement, qui devraient tous deux avoir des influences à tous les niveaux de la méthode.

Mise en garde pour le lecteur

Guide de lecture

Je tenais à proposer au lecteur quelques astuces de lecture afin de rendre le travail de lecture un peu moins difficile :

- Dans ce mémoire, j’emploierai, comme le veut la coutume, le pronom pluriel «nous» plutôt que le pronom singulier «je» afin de rendre l’exposé des travaux moins personnel.
- Le pronom indéfini «on» sera également largement utilisé afin d’impliquer le lecteur dans la lecture du document. Ainsi, les termes «on remarque, on notera», ne signifieront pas «je remarque, je noterai», mais ils signifieront que le lecteur est invité à remarquer que, à noter que ...
- On trouvera (comprenez : le lecteur trouvera) en page 145 un glossaire. Ce glossaire doit servir tant au lecteur initié qu’au lecteur non initié. Pour le non initié, il donnera certaines explications sur des termes habituels pour l’initié, ce qui lui permettra de rentrer dans le discours. Pour l’initié, ce glossaire précise dans quel sens certains termes sont utilisés. En effet, ce mémoire nécessite une terminologie précise car certains objets manipulés dans ces travaux seront dénommés par des mots de sens voisins, tout en étant très différents, d’où la nécessité de bien préciser les sens attribués aux mots dans ce document.
- Dans le texte, les mots soulignés sont présents dans le glossaire¹. Nous ne soulignerons toutefois pas toutes les occurrences des mots du glossaire, mais seulement les quelques premières. Les mots en *italique* sont ceux sur lesquels nous souhaitons insister dans la phrase.
- Par ailleurs, de nombreuses images sont reportées à la fin du mémoire, afin de rendre la partie textuelle du mémoire un peu plus fluide.
- Pour les images de Colombes que l’on rencontrera dans ce mémoire, la résolution est de 8 cm, mais les Modèles Numériques d’Élévation sont calculés à une résolution de 10cm. Ceci expliquera qu’en apparence les aires de surfaces censées être égales sembleront différer quelque peu.
- Lorsque nous ferons référence à des numéros de zones, le lecteur pourra se reporter à l’annexe H pour voir les zones correspondantes.

¹Sauf le mot «souligné» lui-même, qui est présent dans cette phrase, mais pas dans le glossaire, car personne n’en ignore le sens.

Table des matières

1	Introduction	1
1	Cadre de la thèse	1
2	Contexte et objectifs	1
2.1	Contexte et intérêt à l'IGN	1
2.2	Intérêt du sujet	3
2.3	Echelle de travail, résolution des images	3
2.4	Objectifs	5
2.5	Les difficultés de la reconstruction du bâti	10
2.6	Le contexte spécifique du bâti en milieu urbain.	11
2.7	Rappels concernant la stéréoscopie.	11
3	Etat de l'art en reconstruction du bâti.	13
3.1	Points communs	13
3.2	Monoscopie, stéréoscopie, multiscopie	14
3.3	Degré d'interactivité	15
3.4	Niveau des Informations Externes	16
3.5	Généralité des Formes Reconstituées	17
3.6	Etats de l'art	18
4	Présentation de l'approche	18
4.1	Justification	18
4.2	Les données initiales	19
4.3	Présentation	20
4.4	Cas de la végétation	20
4.5	Diviser pour régner	21
4.6	De la zone aux objets simples	21
4.7	Les graphes entrent en scène	22
4.8	Les modèles grandissent.	23
4.9	Et la reconstruction grandit.	24
5	Plan du mémoire	24
6	Conclusion	24
2	Détection de Primitives	29
1	Introduction	29
2	Nettoyage du MNE	29

2.1	Motivation	29
2.2	Méthode de Nettoyage du MNE.	30
3	Les Primitives Planes	32
3.1	Définition	32
3.2	Etat de l'Art en Segmentation d'Images de Profondeur.	33
3.3	Stratégie	33
3.4	Méthode	34
3.5	Exploitation de la méthode	39
3.6	Résultats	42
4	Les primitives linéaires	43
4.1	Définition	43
4.2	Calcul	44
4.3	Résultats	48
5	Les discontinuités du MNE	49
5.1	Définition	49
5.2	Pourquoi pas une polygonisation du contour?	49
5.3	Calcul	50
6	Les primitives ponctuelles	50
6.1	Définition	50
6.2	Utilisation	50
7	Conclusion	51
3	Appariement de graphes	53
1	Notations	53
1.1	Graphe, ou plutôt, hyper-graphe.	53
1.2	Sous-graphe	55
1.3	Différence	56
1.4	Union	56
1.5	Isomorphisme	56
1.6	Isomorphisme de sous-graphe	57
2	Généralités.	57
2.1	Isomorphisme de graphes et de sous-graphes	57
2.2	Isomorphisme de sous-graphe avec tolérance d'erreur	57
3	Le cas du bâti.	59
4	Méthode standard	59
4.1	Formalisme	59
4.2	Décomposition d'un graphe	61
4.3	Appariement	63
5	Notre contribution	68
5.1	Sommets et labels utilisés	68
5.2	Expérimentation	71
5.3	Reconstruction effective du modèle	74
5.4	Le cas des primitives surfaciques	79
5.5	Redondance des types de primitives	80

5.6	Le cas des automorphismes	80
6	Heuristique stochastique	80
6.1	Heuristique	81
6.2	Intégration	81
7	Appariement avec connaissance externe	83
7.1	Modélisation de l'information externe.	83
7.2	Exemple	84
7.3	Première approche.	84
7.4	Approche efficace.	84
7.5	Appariement de graphes AP	87
7.6	Analyse de la complexité	88
7.7	Expérimentation	89
8	Conclusion	89
4	Exploitation d'une grammaire de graphes.	91
1	Tour d'horizon des grammaires de graphes	91
2	Formalisme	92
2.1	Règle	92
2.2	Application d'une règle	93
2.3	Grammaire d'hyper-graphes	94
3	La grammaire du système	95
3.1	Engendrer les modèles de bâtiments	95
3.2	Un cas particulier : les relations ternaires	96
4	Réseau grammatical	99
4.1	Généralités	99
4.2	Exemple	100
4.3	Applications équivalentes	100
5	Décomposition	101
6	Appariement progressif	102
6.1	Réseau et modèles	102
6.2	Appariement multi-modèles	102
7	Conclusion	105
5	Vers une reconstruction	107
1	Introduction	107
2	Le cas des repliements	107
3	Analyse	108
3.1	Rappels de la stratégie	108
3.2	Contexte	108
3.3	Les 2 MNEs considérés	109
3.4	Evaluation d'une combinaison : analyse	109
3.5	Production d'alertes	110
3.6	Evaluation d'une combinaison	111
4	Vers une stratégie de reconstruction : approche arborescente	112

5	Vers une reconstruction : approche linéaire	113
5.1	Génération d'hypothèses	113
5.2	Combinaison d'hypothèses	116
5.3	Minimisation de la fonction d'évaluation	117
6	Conclusion	118
6	Résultats et évaluation	119
1	Evaluation a priori	119
2	Evaluation réelle	124
2.1	Quelques résultats clés.	124
2.2	Protocole d'évaluation	124
2.3	Evaluation générale des formes	126
2.4	Précision géométrique des sommets 3-D	126
2.5	Evaluation du mécanisme d'alerte	127
3	Extensions	128
4	Conclusion	128
7	Conclusion	137
1	Analyse de la contribution	137
1.1	Inconvénients	137
1.2	Les contributions	138
2	Extensions possibles	139
3	Conclusion	140
A	Glossaire	145
B	Résumé des notations	151
C	Figures pour la zone 172.	153
D	Figures pour la zone 22.	157
E	Figures pour la zone 19.	161
F	Figures de la zone 17	167
G	Annexe : fonctions géométriques utilisées	171
1	Relations Ponctuel / Ponctuel	171
2	Relations Ponctuel / Linéaire	171
3	Relations Ponctuel / Surfactive	172
4	Relations Ponctuel / Discontinuité	172
5	Relations Linéaire / Linéaire	172
6	Relations Linéaire / Surfactive	173
7	Relations Linéaire / Discontinuité	174
8	Relations Surfactive / Surfactive	175

9	Relations Surfaique / Discontinuité	175
10	Relations Discontinuité / Discontinuité	176
11	Commentaires	176
H	Les reconstructions	177

Chapitre 1

Introduction

1 Cadre de la thèse

Ce mémoire présente les travaux de thèse qui ont été menés sous la direction de Monsieur Hervé Le Men, Directeur Technique Adjoint de L'Institut Géographique National (IGN) ¹, et de Monsieur le Professeur Georges Stamon, Directeur du Laboratoire «Systèmes Intelligents de Perception» ² (SIP) de l'Université René Descartes, Paris 5.

Physiquement, les travaux ont été menés au laboratoire «Méthodes d'Analyse et de Traitement d'Images pour la Stéréorestitution» (MATIS) de l'IGN. Ils ont été financés par l'IGN, dont je suis membre. La thèse a débuté en décembre 1997.

Nous disposons d'outils informatiques récents. La station utilisée est un ordinateur PC, 650 MHz, sous Linux, équipé de 256 Mo de RAM. Les images utilisées (argentiques ou numériques) sont toutes produites par l'IGN. Nous disposons d'une bibliothèque c++ développée au laboratoire, permettant des manipulations aisées de structures de données de haut niveau, telles que les images, les géoréférences de modèles numériques d'élévation, les scènes tri-dimensionnelles, ainsi que d'outils de visualisation 2-D et 3-D.

2 Contexte et objectifs

2.1 Contexte et intérêt à l'IGN

L'IGN, Institut Géographique National, voit une grande part de son activité consacrée à la cartographie, qui, littéralement, se limite à la représentation bidimensionnelle d'une information tridimensionnelle. Mais sa vocation géographique ne se limite pas à la cartographie. C'est pourquoi l'introduction de la troisième dimension dans les cartes informatiques que sont les bases de données géographiques, permet de se rapprocher de la notion de Géographie, qui est l'étude et pas seulement la représentation de l'espace. En effet, si l'on représente l'espace 3-D dans lequel nous vivons, par des objets eux-mêmes tri-dimensionnels, la perception de cet espace ne peut qu'en être meilleure.

¹Site internet : <http://www.ign.fr> Site recherche : <http://recherche.ign.fr>

²Site internet : <http://www.math-info.univ-paris5.fr/sip-lab/>

L'IGN, comme d'autres institutions dans le monde, mène donc des recherches visant à restituer non seulement le relief, mais aussi à interpréter certains thèmes géographiques. Le bâti est un des thèmes majeurs de ces recherches, et ce mémoire développe des travaux inscrits dans cet axe. Ces travaux entrent donc pleinement dans la vocation géographique de l'Institut.

Ces travaux sont donc motivés par le besoin d'une meilleure appréhension de l'environnement, notamment urbain. Mais cette optique n'est pas la seule fin de cet axe de recherche. Citons par exemple d'autres contextes dans lesquels la représentation du bâti possède une grande importance :

- la modélisation radiométrique des images. Il s'agit de la recherche, à partir des images «brutes», de modèles atmosphériques, du modèle du capteur, et enfin d'un modèle 3-D de la scène observée, des propriétés radiométriques «vraies» des objets présents dans la scène. Dans ce cadre, une représentation correcte des pentes des toits est importante, car l'apparence d'une surface dépend sensiblement de l'incidence avec laquelle elle est observée, et donc du point de vue *et* de sa pente.
- la constitution d'ortho-images avec redressement des bâtiments. Il s'agit ici d'obtenir une image superposable à une carte. Ceci est possible à l'aide d'une représentation de la surface topographique. La figure 1.1 présente une ortho-image dans laquelle la surface topographique est supposée plane et horizontale. On y voit les façades des bâtiments. Lorsque la surface topographique tient compte des bâtiments, l'ortho-image correspondante ne présente plus les façades des bâtiments. Une telle image se trouve à droite dans la figure 1.1. Cette ortho-image a été réalisée à l'aide d'une seule image aérienne. Dans ce cas, en raison des occultations liées aux bâtiments, des informations manquent pour réaliser une ortho-image complète. Ces informations manquantes correspondent aux parties noires de cette image. Dans la pratique, en considérant plusieurs images prises de différents points de vue, il est possible de combler ces trous. L'intérêt d'avoir une reconstruction précise du bâti dans ce cas est de pouvoir produire des ortho-images propres au voisinage des discontinuités de la surface topographique, particulièrement au voisinage des gouttières des bâtiments.

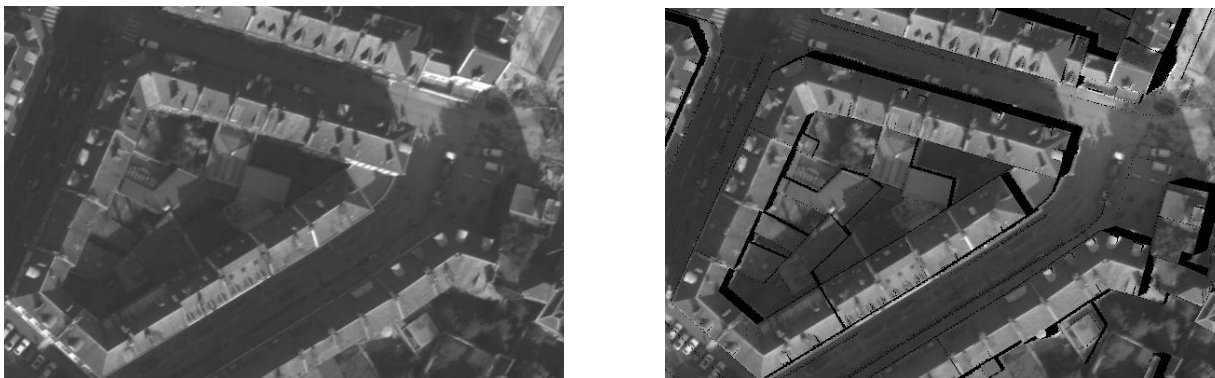


FIG. 1.1 – Deux ortho-images d'un îlot. A gauche : bâtiments non redressés, à droite : bâtiments redressés.

2.2 Intérêt du sujet

La reconstruction du bâti, notamment en milieu urbain, est un sujet de grand intérêt dans la communauté scientifique et pas seulement à l'IGN (on trouvera en section 3 un état de l'art en matière de reconstruction du bâti), car il est porteur de nombreuses applications différentes de celles citées dans la section précédente. On trouvera notamment dans [Bre99, Fri99] une revue des champs d'application possibles. Ainsi, les thèmes tels que l'étude de la propagation des ondes électromagnétiques (une application majeure ces dernières années est l'optimisation du placement des antennes des réseaux de téléphonie mobile), les études environnementales, la gestion du milieu, les jeux vidéos, l'architecture, sont des domaines d'emploi où l'information tri-dimensionnelle qu'apporte la reconstruction fine du bâti est importante.

Par ailleurs on notera que les modélisations 3-D issues de reconstructions à partir d'images aériennes peuvent être un point de départ pour des reconstructions plus fines, de type architectural. Ainsi, P. Debevec mène des travaux de rendu réaliste d'ensembles architecturaux ([DTM00]), de même que J. Marchadier ([MME00]) mène une thèse en aval de la problématique de reconstruction du bâti, concernant l'habillage de façades.

2.3 Echelle de travail, résolution des images

La cartographie ne peut être conçue sans la notion d'échelle d'utilisation : la façon de représenter une portion de terrain dépend de l'échelle de la carte finale - et donc de l'application à laquelle la carte se destine - ou de l'échelle d'utilisation de la base de données, le cas échéant. Ainsi, les objets présents dans la base de données ou la carte possèdent plus ou moins de détails. On trouve typiquement quasiment tous les bâtiments isolés sur la carte au 1 / 25000 de l'IGN, ou dans la BD Topo (R), alors qu'aux échelles plus faibles (typiquement 1/100000) on applique une généralisation (voir la thèse [Rua99] qui traite de généralisation) afin de ne laisser transparaître dans la carte que les traits caractéristiques pertinents à ces échelles : les axes principaux des villes et les quartiers. A titre d'exemple, la figure 1.2 présente un extrait de carte d'échelle 1/25000 (à gauche) où l'on distingue des bâtiments individuels, un extrait de carte d'échelle 1/100000 (au centre) où les axes principaux sont visibles en milieu urbain, et un extrait de carte d'échelle 1/1000000, où seuls les axes entre communes subsistent ³.

De même, les techniques d'imagerie (aérienne, satellitale, robotique) dépendent des résolutions des images utilisées, et, bien entendu, les résultats que l'on peut attendre d'une exploitation des images dépendent sensiblement de la résolution de ces images. On notera qu'une échelle n'a pas d'unité (c'est un rapport de longueurs) alors qu'une résolution s'exprime en unité métrique. Par ailleurs, ces deux notions sont inversement proportionnelles. Ainsi, une petite résolution (quelques centimètres) correspond à une prise de vue à grande échelle. Considérons le croquis 1.3, sur lequel on peut voir un dispositif de prise de vue de type sténopé se trouvant à une hauteur H , de focale f . Soit d une distance mesurée entre deux points de l'image et D la distance entre les deux points correspondants au sol, alors le rapport d/D est l'échelle de prise de vue ⁴. Si d est la taille d'un pixel

³Nous n'avons discuté ici que des thèmes «bâti» et «routier», le lecteur verra que les mêmes phénomènes s'appliquent aux autres thèmes, comme par exemple les cours d'eau.

⁴On voit également que d/D est égal à f/H , qui est le rapport entre la distance focale de la caméra et l'altitude au-dessus du sol.



FIG. 1.2 – Extraits de cartes à des échelles différentes

⁵, alors D est la résolution de l'image. On voit donc bel et bien que $e = d/r$, où r la résolution de l'image, et e l'échelle de prise de vue.

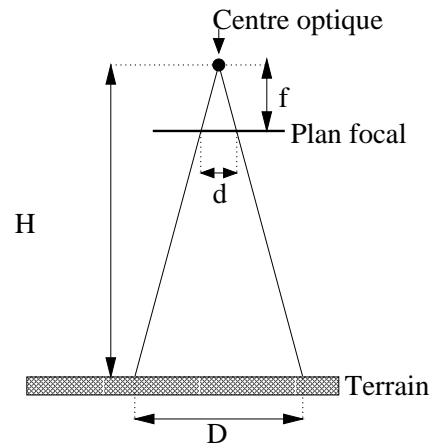


FIG. 1.3 – Représentation des distances caractéristiques d'une prise de vue : la hauteur de vol au dessus du sol (H), la distance focale (f), une distance image (d) et sa distance correspondante au sol (D).

A titre d'exemple, la figure 1.4 présente des images aériennes possédant des résolutions différentes. Les résolutions respectives de ces images sont : 20cm, 40cm, 80cm, 2m, 4m. Les images sont carrées, de côté 200 pixels. Leur taille sur la feuille est de 3cm, ce qui leur donne donc des échelles respectives de : $3/4000$, $3/8000$, $3/16000$, $3/40000$, $3/80000$. Les zones couvertes sur le terrain sont des carrés de côtés respectifs 40m, 80m, 160m, 400m, 800m.

Dans la gamme des échelles de prise de vue aériennes, ces travaux se situent dans les grandes échelles, puisque la résolution typique des images utilisées est décimétrique.

Bien entendu, du point de vue de l'architecte, le domaine d'échelles dans lequel nous nous si-

⁵Dans le cas d'une image argentique, ceci correspond au pas de scannage. Dans le cas d'une caméra numérique, ceci correspond à la taille physique des éléments du capteur (CCD, CMOS ...).



FIG. 1.4 – Images de résolutions différentes sur la ville d’Amiens

tuons est celui des faibles échelles, puisque nous excluons (voir sections suivantes) dès à présent les superstructures des toits des bâtiments telles que lucarnes et cheminées.

2.4 Objectifs

2.4.1 Qu’est-ce qu’une reconstruction ?

Nous entendons par reconstruction la représentation vectorielle d’une scène. Plus précisément, nous rechercherons un ensemble de sommets 3-D, d’arêtes 3-D, et de faces 3-D, représentant le polyèdre formé par les pans de toit.

Il y a donc une correspondance forte entre les objets recherchés et les objets réels, puisque

- les faces, qui sont des portions de plans, correspondent aux pans de toits.
- les arcs, qui sont des segments de droites, délimitent ces portions de pans. Ils correspondent aux gouttières, faîtes, et arêtes de rive des toits.
- enfin, les points 3-D, qui sont eux-mêmes les extrémités des segments, correspondent aux intersections entre arêtes faîtières, entre arêtes faîtières et arêtes de rives, entre gouttières et arêtes de rives, etc.

Ces objets sont reliés entre eux par des informations de topologie, que nous venons d’exprimer :

- Un point 3-D est l’extrémité de tel ou tel segment.
- Un segment borde une ou deux portions de pans.
- Un point appartient à tel ou tel pan.

On notera que l’information cherchée est plus riche qu’une juxtaposition d’objets tri-dimensionnels, puisque les relations qui relient les objets sont elles aussi recherchées : pour un pan de toit donné, la représentation doit savoir dire quels sont les segments 3-D qui en forment le pourtour.

Nous recherchons donc un ensemble d’objets 3-D et sa topologie sommet / arc / face. Ce modèle est très proche des modèles topologiques au sens des bases de données géographiques (voir [SVP⁺96], pages 28-31).

On remarquera que le passage des données initiales à cette reconstruction est une vectorisation, qui a pour propriété importante de représenter l’information de façon très compacte, ce qui est toujours le cas lors de passage du mode maillé (ou raster), au mode vecteur, comme cela est explicité dans la figure 1.5. Dans le cas d’images en niveau de gris ou en couleur, la représentation vectorielle perd toutefois l’aspect radiométrique des images.

Une reconstruction peut être obtenue de manière manuelle, comme c’est le cas à l’IGN pour la

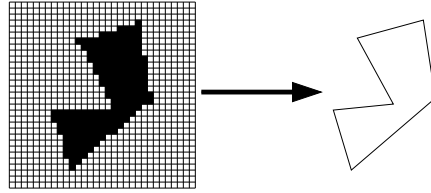


FIG. 1.5 – Passage du mode maillé (raster) au mode vecteur : la vectorisation de ce polygone en 6 points permet de synthétiser l'information.

saisie de la BD Topo (R) ⁶, où les bâtiments sont saisis à l'aide de stations de restitution photogrammétrique, par un opérateur humain. Notons que l'IGN a produit des données urbaines précises, suivant un protocole de saisie spécifique, ainsi qu'un logiciel de tracé de perspectives urbaines (appelé TRAPU). Une vue perspective de données de type TRAPU est présentée par la figure 1.6. On y voit le niveau de détail de ce type de données, puisque de nombreux détails sont présents dans ces données.

Des méthodes semi-interactives peuvent également être envisagées (voir l'état de l'art), typiquement dans le cas de bâtiments de même forme ([Gui97]).

Enfin, des méthodes automatiques sont bien entendu l'objet de beaucoup de travaux, dont celui-ci.

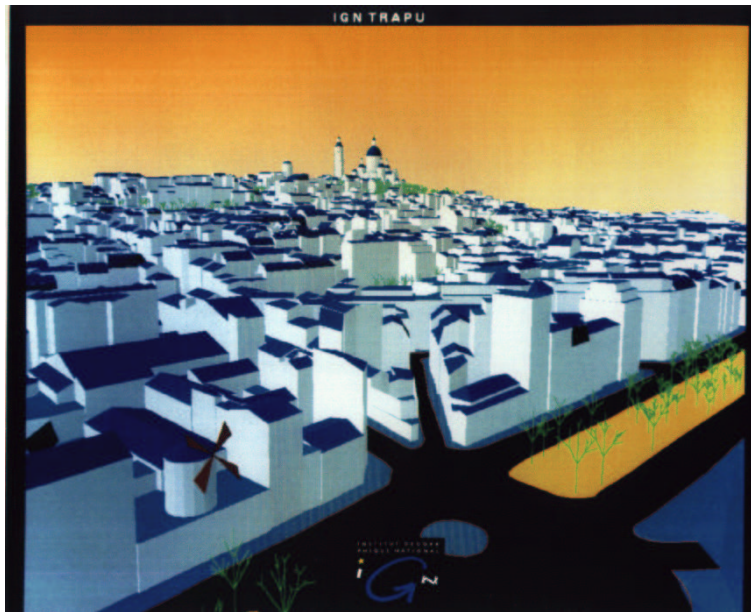


FIG. 1.6 – Vue perspective de données TRAPU.

⁶Des exemples de saisies BD Topo(R) sont donnés page 11.

2.4.2 Objectifs à long terme

Ces reconstructions obtenues sont destinées à être stockées dans des bases de données (BD), puis à servir à un utilisateur final qui les manipulera à l'aide d'un système d'information géographique (SIG). Dans le cas présent, les bases de données visées sont des bases de données urbaines, à forte échelle (l'échelle d'une base de données étant l'échelle typique à laquelle cette base sera vue, étudiée, imprimée ...).

Dans un contexte réel de production, les processus mis en oeuvre seront bien entendu semi-automatiques, car il est illusoire, pour un horizon proche, d'espérer des processus automatiques, à moins d'autoriser des taux de fautes inacceptables pour les utilisateurs des bases de données. Ces travaux constituent donc une étape vers cet objectif à long terme, et le passage de l'automatique au semi-automatique n'est pas traité ici. Toutefois à certaines étapes de la méthode, on entreverra dans quelle mesure les outils élaborés ici pourraient s'insérer dans un cadre semi-automatique. Ce contexte opérationnel ne peut en effet pas être perdu de vue, ce qui mettra par ailleurs en exergue certains points positifs et d'autres, négatifs, des méthodes proposées.

2.4.3 Les toits, rien que les toits

Nous ne nous intéresserons qu'à la reconstruction des toits. Les parties plus basses que les toits (en premier lieu les façades, mais aussi les balcons par exemple) ne seront pas traitées. Ceci se justifie par le contexte proposé : les images utilisées sont des images stéréoscopiques «simples», ce qui ne permet pas d'espérer reconstruire les façades, celles-ci n'étant que peu vues sur les deux images du couple.

Il est en effet aisé de voir que dans le cas d'un bâtiment rectangulaire, on ne peut voir qu'au maximum deux façades sur une image aérienne. Avec deux images (un couple stéréoscopique), on peut donc voir au mieux :

- deux fois deux des façades.
- deux façades une seule fois, et une façade dans les deux images.
- quatre façades, mais systématiquement dans une seule des deux images.

Il est donc illusoire d'espérer représenter les façades des bâtiments à partir de la stéréoscopie, car au mieux, on pourrait espérer en reconstruire la moitié. A titre d'exemple, les figures 1.22 (page 28) et E.1 (page 161) montrent aisément l'impossibilité de voir toutes les façades en stéréoscopie «simple».

La reconstruction des façades est donc exclue du sujet.

Nous signalons, en revanche, qu'une thèse menée partiellement au laboratoire par M. Marchadier ([MME00]), sur la reconstruction et l'habillage des façades à partir de prises de vues terrestres et d'un modèle 3-D approché, est en cours au moment de la rédaction de ce mémoire. Enfin, une thèse visant explicitement l'exploitation «multiscopique» des images de la caméra numérique de l'IGN pour reconstruire le bâti en milieu urbain vient de débiter. Elle est menée par M. Taillandier au laboratoire.

Ce mémoire décrit donc un travail qui prend sa place dans un projet plus global de reconstruction du bâti.

2.4.4 Les toits, tous les toits

Dans le contexte grande échelle, nous recherchons à représenter la forme extérieure (le pourtour) des toits de bâtiments, ainsi que la structure interne des toits, en nous limitant toutefois aux pans de toits.

Malgré ces précautions qui visent à cibler le sujet, il reste et il restera des cas se trouvant à la limite de ce que l'on peut espérer représenter ou pas. Nous tenterons d'aborder alors les questions de complexité du bâti, et la notion de «vérité terrain», qui souvent, posent des problèmes cartographiques pour l'opérateur lui-même, comme nous le verrons en section 2.5.

A titre d'exemple, signalons simplement quelques exemples se trouvant aux limites des superstructures : la figure 1.7 (zone 21) donne un bâtiment à double pan orné de 6 lucarnes, qui correspondent à une surface importante du toit, et la figure 1.8 est un exemple de deux chiens assis occupant plus de 50% de la surface du toit.

Ces exemples montrent partiellement la difficulté du problème cartographique. Dans le cas de la figure 1.8, on peut notamment se poser la question de savoir s'il faut restituer 2 pans de toit s'appuyant sur les gouttières du bâtiment, 2 pans passant par les chiens assis, ou alors 4 pans. Cette question est une question de généralisation, notion déjà évoquée, et qui interviendra dans la définition les objectifs de ces travaux.



FIG. 1.7 – Un bâtiment paré de ses 6 lucarnes

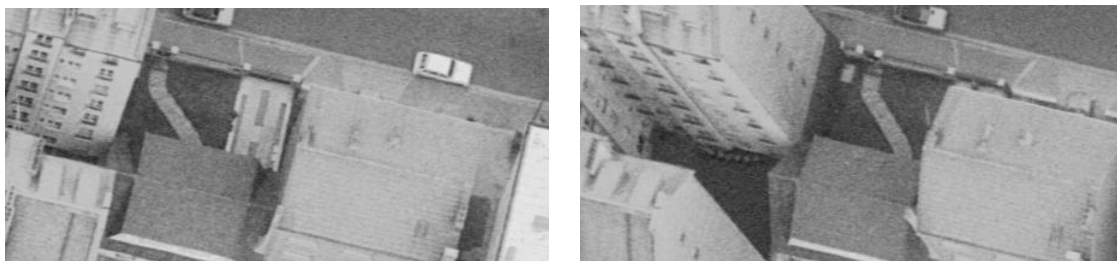


FIG. 1.8 – Un bâtiment (au centre) où une superstructure (chien assis) domine les pans de toit.

2.4.5 Exhaustivité

Il est illusoire d'espérer reconstruire, de façon automatique, l'ensemble d'une ville sans se tromper. L'objectif est, pour un pourcentage important (supérieur à 50 %) de déterminer correctement la forme du bâtiment, et d'effectuer la reconstruction correspondante, avec une certaine précision de généralisation (voir section 2.4.6), et une certaine précision géométrique (voir section 2.4.7).

Par «déterminer la forme du bâtiment», nous entendons : proposer la description vectorielle (définie plus haut) correspondant aux pans de toits du bâtiment.

2.4.6 Généralisation

Nous avons déjà montré que la généralisation est une donnée du problème, puisqu'elle intervient dans les applications souhaitées des résultats obtenus (voir section 2.3) Nous avons déjà précisé que les superstructures des toits sont exclues de la reconstruction cherchée, ce qui est en soi la définition d'une forme de généralisation que nous acceptons. Mais le degré de généralisation que nous tolérerons ne se limite pas à la suppression des superstructures des toits. En effet, le pourtour du bâtiment peut lui aussi être généralisé tout en se montrant satisfaisant pour de nombreuses applications. Il nous faut donc définir quel degré de généralisation nous acceptons.

A titre d'exemple, le bâtiment présenté dans la figure 1.9 présente sur sa façade nord une protubérance dont l'épaisseur est de 1 m. Ainsi, ce bâtiment peut être représenté par un prisme polygonal de 8 points comme sur l'image droite de la figure, ou alors par un rectangle simple, à condition de tolérer une généralisation suffisante de la forme. En l'occurrence, une tolérance de décrochement de 1 m est suffisante pour admettre ces deux reconstructions. Le problème de la reconstruction du bâti devient alors mal posé, puisque dans ce cas simple et pour cette tolérance, deux formes différentes sont admissibles, même si l'une est plus fine et plus précise que l'autre.

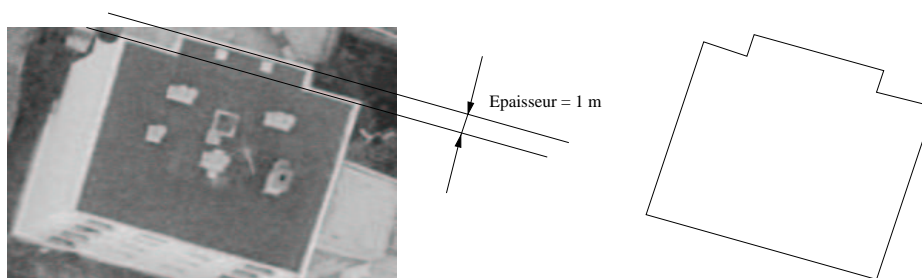


FIG. 1.9 – Protubérance sur un bâtiment rectangulaire et une reconstruction possible de ce bâtiment.

Nous sommes dans l'incapacité de définir plus précisément les tolérances de généralisation acceptées pour le système, puisque bien des cas doivent être définis pour donner des spécifications de généralisation : décrochement toléré, saut d'altitude toléré, surface maximale d'un trou toléré, surface minimale d'un bâtiment à reconstruire.

Nous admettrons donc un principe un peu flou, disant que les phénomènes liés à la forme d'un bâtiment (décrochement, sauts d'altitude, pans de toits ...) doivent être représentés si leur taille caractéristique est de l'ordre du mètre.

2.4.7 Précision géométrique

L'objectif de précision géométrique est décimétrique. Plus précisément, l'objectif de précision géométrique des sommets 3-D sur lesquels s'appuieront les reconstructions, *lorsque le modèle utilisé pour la reconstruction est adapté au bâtiment*, est décimétrique.

Dans le cas où le modèle choisi n'est pas adapté, le problème de la précision géométrique se pose différemment. Dans ce cas, il se peut que même après généralisation de la forme du toit, certains sommets reconstruits correspondent à certains sommets 3D réels, mais pas tous. Nous souhaitons obtenir, pour ces sommets, la même précision géométrique que dans le cas de formes non généralisées.

Ainsi, dans le cas de la figure 1.9, si le bâtiment reconstruit est finalement rectangulaire, comme nous le tolérerons, alors la précision géométrique souhaitée pour les 4 sommets du rectangle est décimétrique.

2.5 Les difficultés de la reconstruction du bâti

Les données initiales utilisées ici sont des images stéréoscopiques et un modèle numérique d'élévation.

La difficulté de définir une représentation unique comme solution à la reconstruction du bâti a déjà été exprimée en section 2.4.6. Nous l'illustrons ici par des exemples réels réalisés par des opérateurs de l'IGN.

Il s'agit de saisies stéréoscopiques d'un échantillon de la ville de Châlons en Champagne, réalisées dans des conditions identiques - à l'aide d'images argentiques prises à l'échelle de 1/25000. Les saisies ont été réalisées sur des stations de restitution photogrammétrique analytiques (utilisant les images argentiques, sans numérisation).

La figure 1.10 présente les résultats de saisie sur 3 parties particulières de la ville. Pour chacune des régions de cette figure, une petite boîte à outil propose une règle graduée donnant l'échelle des longueurs. Le lecteur aura peut-être des difficultés à lire les valeurs de ces règles graduées. Nous les précisons donc ici : il s'agit de 5 mètres pour la première région, 3 mètres pour la seconde, et 8 mètres pour la troisième.

Sur cette figure, le bâtiment de gauche est rectangulaire, et n'a posé aucun problème de généralisation aux opérateurs : la forme est respectée. Seules des imprécisions géométriques subsistent. Pour le second bâtiment (au centre), les opérateurs ont effectué diverses généralisations : l'opérateur «vert» a proposé une restitution plus détaillée que les deux autres. L'opérateur «rouge» a généralisé les murs sud-ouest du bâtiment, et l'opérateur «bleu» a généralisé les murs sud-ouest et sud-est du bâtiment. Dans les trois cas, toutefois, le bâtiment est restitué comme une seule entité. La troisième région, par contre, pose de plus des problèmes de connexité, puisque les opérateurs «vert» et «rouge» ont restitué un grand bâtiment avec les deux trous et un bâtiment plus petit à l'est, alors que l'opérateur «bleu» a restitué un seul bâtiment avec ces deux trous. On notera également que l'opérateur «bleu» a été le seul à saisir une protubérance au sud du bâtiment. Ceci montre toute la problématique des spécifications de saisies d'une base de données, et du caractère variable d'une saisie : un opérateur peu tantôt saisir plus de détails qu'un autre, et tantôt moins.

La notion de vérité terrain, ou de référence terrain, est donc difficile à établir, malgré l'existence de spécifications de saisies pour les bases de données. Nous rencontrerons naturellement les mêmes difficultés à définir la vérité terrain dans notre contexte.

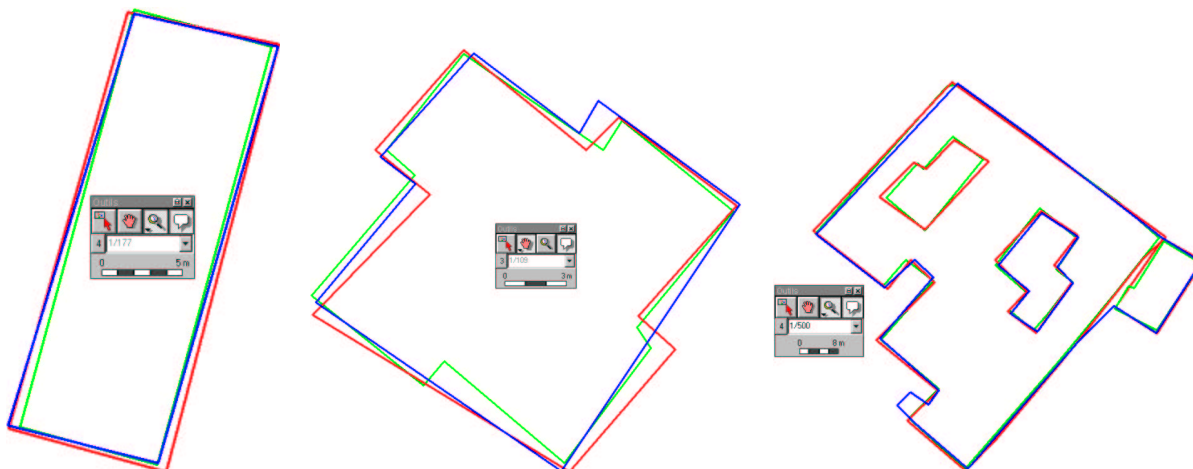


FIG. 1.10 – Variabilité d’une saisie : 3 exemples de saisies stéréoscopiques de mêmes régions, faites par 3 opérateurs différents.

2.6 Le contexte spécifique du bâti en milieu urbain.

Nous relèverons une seule spécificité du milieu urbain. Il s’agit de la densité du milieu. Elle induit des adjacences fortes entre les bâtiments. Ceci provoque le phénomène d’occultation.

Parfois, même, on ne peut pas identifier (en stéréoscopie, en tout cas) si deux bâtiments voisins sont effectivement adjacents ou pas (voir le bâtiment 17 en figure F.1 page 167) : on ne peut pas savoir d’après ces imagettes stéréoscopiques si les deux bâtiments rectangulaires de la zone sont effectivement adjacents ou pas.

2.7 Rappels concernant la stéréoscopie.

2.7.1 Principe

Cette section présente quelques notions fondamentales nécessaires à la compréhension des mécanismes mis en oeuvre en stéréorestitution, qu’elle soit faite par un opérateur, ou par une machine.

La figure 1.11 rappelle quelques éléments intervenant dans la stéréorestitution. Cette figure représente deux centres de prise de vue, C et D, suivant lesquels ont été prises deux photographies, I et J. Les deux emprises au sol des images (en rayures) possèdent une zone de recouvrement.

Cette zone de recouvrement, vue dans les deux images, est celle où le principe stéréoscopique s’applique. En effet, si l’on voit dans les deux images, en P et en Q, un même détail du terrain, comme le point M, alors, il est possible, en calculant l’intersection des rayons perspectifs (CP) et (DQ), de trouver la position 3-D du point M.

Ce calcul nécessite toutefois de connaître préalablement la position dans l’espace des sommets C et D, ainsi que les orientations des caméras (roulis, tangage, lacet)⁷. Ces éléments, qui constituent

⁷Il faut bien évidemment connaître également la calibration interne de la caméra, mais nous n’entrerons pas dans ces considérations qui sont inutiles dans cette partie introductive, ainsi que pour le reste des travaux, car nous supposons tous ces éléments parfaitement connus.

la calibration externe de la caméra, sont calculés par un ensemble de techniques, l'aérotriangulation, que nous n'expliciterons pas.

A titre d'exemple, les figures 1.18 et 1.19 montrent les images gauche et droite d'un couple stéréoscopique (ces images ne constituent pas le couple entier, mais seulement la moitié de la zone de recouvrement). Ces deux images couvrent une scène dont les dimensions sont environ de 380m en ouest-est et 420m en nord-sud, ce qui représente une surface de 16ha.

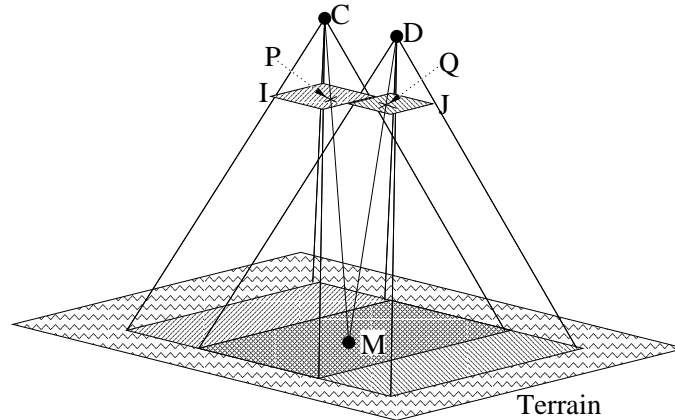


FIG. 1.11 – Rappel des éléments principaux permettant d'exploiter le principe stéréoscopique en imagerie aérienne.

Le problème majeur, et qui occupe la communauté depuis des années, est de déterminer de façon automatique, fiable, et précise, les coordonnées des points P et Q. C'est la question de la mise en correspondance. Cette question ne s'arrête pas à la mise en correspondance de points, mais on peut aussi tenter de mettre en correspondance des segments de droite, des régions, ou tout type d'objet que l'on peut extraire des images.

Notre travail se situe en aval de ces questions de mise en correspondance, comme nous le verrons plus tard. En conséquence, l'état de l'art qui sera proposé dans ce chapitre ne couvrira pas ces questions de mise en correspondance.

2.7.2 Influence base / hauteur

La résolution planimétrique est donnée par l'échelle de prise de vue et la taille du pixel, comme nous l'avons montré précédemment. La résolution altimétrique dépend, elle, d'un autre paramètre : la base, qui est la distance entre les deux sommets de prise de vue. Considérons la figure 1.12, sur laquelle est présentée une coupe d'un couple stéréoscopique. Soit M un détail du terrain, vu en P et Q dans les deux images. Soit Q' situé à une distance d de Q . Le rayon perspectif DQ' intersecte le rayon CP en un point M' , de hauteur dH . Les triangles MCD et $MM'M''$ étant homothétiques, nous avons : $dH/M'M'' = H/CD$. Si d est la taille d'un pixel, alors $M'M''$ est la résolution planimétrique de l'image (en supposant que dH soit petit devant H). Le rapport entre résolution altimétrique et planimétrique est donc : $r_z/r_{xy} = dH/M'M'' = H/d(C, D) = H/B$ où $B = d(C, D)$ est la base.

Dans la pratique, pour les images utilisées, nous avons : $B = 380m$ et $H = 860m$, d'où : $r_z/r_{xy} = 2,3$. La résolution altimétrique est donc 2,3 fois plus forte (moins bonne) que la résolution

planimétrique.

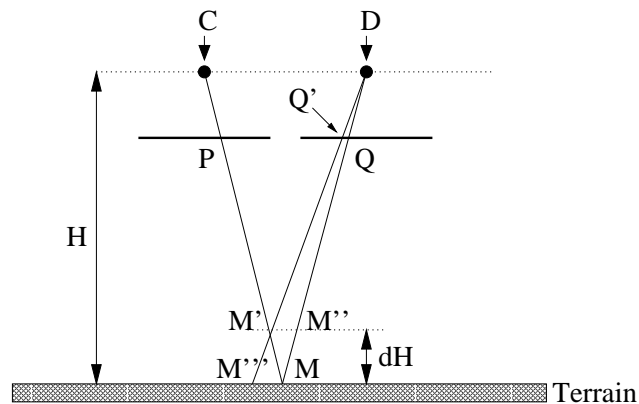


FIG. 1.12 – Lien entre résolutions altimétrique et planimétrique.

2.7.3 Multiscopie

L'extension naturelle de la stéréoscopie, est la vision avec N images, que le lecteur nous autorisera à dénommer multiscopie. L'avantage majeur de la multiscopie est de réduire les occultations, qui sont importantes en milieu urbain. La multiscopie est exclue du cadre de ces travaux.

3 Etat de l'art en reconstruction du bâti.

Nous proposons de faire un tour d'horizon des méthodes de la littérature, concernant la reconstruction du bâti.

Les méthodes sont nombreuses et nous n'en ferons pas une description exhaustive. Par contre, nous tenterons d'analyser l'existant suivant différents points de vue, afin de situer notre propre approche par rapport à cet existant.

3.1 Points communs

3.1.1 Focalisation

Toutes les méthodes (dont la nôtre, comme nous le verrons plus tard), commencent par focaliser leur attention sur les bâtiments. Les techniques de focalisation sont généralement les mêmes. En milieu non arboré, les détections de taches de sursol sont utilisées par tous les auteurs (la figure 1.13 explicite les notions de sol et sursol. Sur cette figure, la surface topographique correspondant au sol et au sursol est en rouge, alors que celle correspondant au sol seul est en bleu). Des techniques de ce type sont données par Baltasvias ([BMS95]) dans le cadre du projet AMOBE de l'ETH Zürich. En milieu arboré, une détection de la végétation est nécessaire afin de ne pas traiter les zones arborées comme du bâti, d'où des travaux de détection de la végétation.

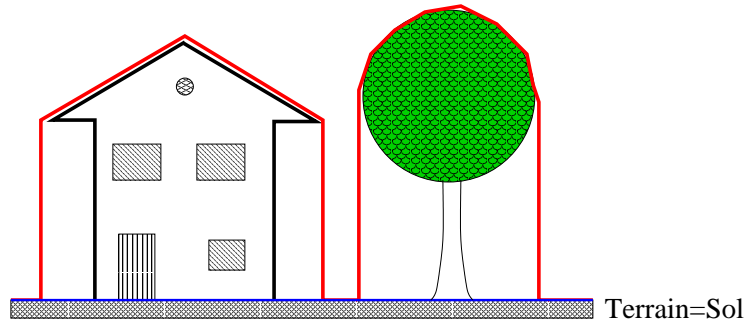


FIG. 1.13 – Distinction entre sol et sursol.

3.1.2 Représentation vectorielle

Cela est naturel, mais encore fallait-il le préciser, toutes les méthodes rencontrées s'attachent à décrire les scènes par un ensemble plus ou moins explicite de sommets, arcs, et faces. L'objectif commun des méthodes est bien entendu le chargement des reconstructions dans des bases de données.

3.2 Monoscopie, stéréoscopie, multiscopie

3.2.1 Analyses monoscopiques

Heuel et Nevatia [HN] s'intéressent à la reconstruction monoscopique du bâti. L'aspect monoscopique rend la tâche bien plus difficile (il n'est pas possible d'obtenir une appréciation sérieuse du relief), si bien que l'auteur se limite à l'extraction de parallélépipèdes. La validation des modèles est possible par l'analyse des ombres. De travaux similaires sont présentés par McGlone et Shufelt ([MS94]). Lang et Förstner ([LF]) présentent également une méthode monoscopique appuyée sur le formalisme de Géométrie Solide Constructive (CSG). Des primitives volumiques (parallélépipèdes, pyramides par exemple) sont placées manuellement par l'opérateur en monoscopie. Le positionnement fin peut être réalisé interactivement ou par recalage automatique. Les volumes primitifs sont assemblés un à un pour décrire une scène.

3.2.2 Analyses stéréoscopiques

Dang ([Dan94]) propose une extraction des bâtiments rectangulaires à partir d'images stéréoscopiques. Il propose l'utilisation de la programmation dynamique le long des lignes épipolaires pour imposer une contrainte globale sur le profil d'altitude le long de la ligne. Après l'extraction du MNE ainsi calculé, l'extraction du bâti rectangulaire est faite grâce à des techniques de groupement perceptuel, inspirées de la Gestalt ([RP90]).

Baillard ([Bai97]) propose une technique de focalisation sur les régions de sursol d'altitude homogène, spécifique au milieu urbain. Après calcul d'un modèle numérique d'élévation respectant les discontinuités de la surface topographique, le MNE est segmenté en régions d'altitude homogène. Par après, une relaxation permet d'attribuer aux régions les labels «sol» ou «sursol». Le modèle numérique du sol seul peut alors être estimé par interpolation des informations d'altitude des régions dont le label est «sol». Cette estimation est réinjectée dans la relaxation pour estimer à nouveau les

labels. La méthode est donc itérative, et s'arrête lorsque les labels sont stabilisés. Cette méthode permet de définir des régions de focalisation. Les régions connexes de sursol peuvent être séparées en sous-régions si leur altitude moyenne diffère d'un seuil critique d'environ 2m.

M. Cord ([Cor98]) propose une stratégie séquentielle après la restitution du Modèle Numérique d'Élévation. Après avoir produit le MNE et identifié les zones de sursol, il procède à une approximation polygonale de celles-ci ([JCCP99]), puis détermine une segmentation en plans de la région polygonalisée. L'approche de segmentation est bayésienne. Le coeur de l'approche nécessite de connaître a priori le nombre de plans. Mais M. Cord effectue des essais avec 2, puis 3, puis N plans et propose une méthode de sélection du bon nombre de plans.

3.2.3 Analyses Multiscopiques

M. Fradkin ([MF99]) propose une détection multiscopique des bâtiments en milieu urbain, appuyée explicitement sur la pré-détection des façades. Cette pré-détection est faite par accumulation sur un plan horizontal des projections de points 3D vus sur plusieurs images.

Henricsson ([Hen96]) propose un système dans lequel les segments de droites détectés sont flanqués d'attributs colorimétriques concernant leur voisinage. Ces attributs contribuent au groupement de ces segments. Par après, une poursuite du groupement permet de produire des hypothèses de plans 3-D.

Vestri ([Ves00]) présente un approche qui, également, passe par une reconstruction de façades. Dans son contexte, le recouvrement longitudinal et latéral est très élevé (80 %), ce qui permet effectivement leur reconstruction. Son objectif n'est pas la reconstruction des toits, mais des plans verticaux, car le contexte applicatif principal est le placement d'antennes pour les réseaux téléphonie mobile, dont le premier besoin est une connaissance approfondie des façades. Les toits des bâtiments sont toutefois extraits et approximés par des plans horizontaux.

3.3 Degré d'interactivité

3.3.1 Systèmes automatiques

Les travaux de Nevatia, Henricsson, Cord, Dang, Baillard cités plus haut sont automatiques.

3.3.2 Approches semi-automatiques

De nombreuses équipes se sont tournées vers de systèmes interactifs en vue d'une production à court terme, notamment à Bonn et Zürich (voir CC-Modeller section suivante).

3.3.3 Approches interactives / systèmes prêts à produire

L'équipe d'ETH Zürich ([GW98]) propose un système, CC-Modeller, qui, à partir de tous les points 3-D du modèle désiré, reconstruit automatiquement la topologie du modèle, en tentant de calculer un ensemble consistant de polygones s'appuyant sur les points saisis. Le système permet de modéliser des objets complexes, et possède un très bon taux de succès.

CC-Modeller est une version améliorée d'un système non générique, utilisant une batterie de modèles, TOBAGO ([Grü98]).

La méthode est très interactive, car tous les points doivent être saisis à la main à l'aide d'une station de stéréorestitution. Par contre, une fois les points saisis, le processus de calcul du modèle 3-D est automatique. Cette approche est très générale et ne fait aucun présupposé sur les formes recherchées (excepté leur forme polyédrique). Le caractère opérationnel de ce projet est important. Les auteurs présentent des taux de réussite très élevés, et des vitesses de productions élevées : 500 toits par jour, sans préciser la complexité de ceux-ci. Des chantiers importants ont déjà été traités.

3.4 Niveau des Informations Externes

3.4.1 Travaux avec un minimum d'informations externes

Les travaux de Vestri, Cord, Baillard, Henricsson cités plus haut n'utilisent que peu d'information externe, hormis le fait que les bâtiments sont polyédriques.

3.4.2 Travaux avec acquisition active du MNE

Les équipes allemandes utilisent souvent les modèles numériques obtenus par balayage laser ([Loh96]).

Brunn ([BW98]) propose une classification entre bâti et végétation appuyée sur un réseau bayésien exploitant les valeurs d'altitude, de normale locale à la surface des données Laser. Pour les bâtiments, le MNE Laser est ensuite segmenté sur un critère d'homogénéité de la normale à la surface. Dans le cas de zones bâties, des plans sont extraits grâce, à nouveau, à un critère d'orientation de normales locales. Cette approche permet de mettre en évidence à la fois les zones de normales homogènes (i.e. les pans de toit), mais aussi celles où les normales sont inhomogènes (les ruptures de pente, c'est-à-dire les arêtes séparant les pans), ainsi que les discontinuités d'altitude, et donc les bords de bâtiments.

3.4.3 Travaux s'appuyant sur des cartes ou plans 2D

Faible résolution

On trouve notamment les travaux de Roux ([RM97, Rou97]), où la carte est utilisée en milieu urbain afin de produire des zones d'intérêt correspondant aux îlots de bâtiments. Pour ceci, le réseau routier urbain est extrait de la carte, et représenté par les segments routiers et les intersections. Les intersections correspondent à des zones relativement dégagées, ce qui permet un calcul d'une altitude de l'intersection par mise en correspondance stéréoscopique robuste. Roux en déduit les altitudes le long des tronçons.

L'interprétation de la carte permet aussi de fournir une information sémantique sur les zones bâties. Ce type d'information peut être utilisé pour guider des reconstructions.

Forte résolution

Jibrini ([Jib00]) mène des travaux de reconstruction du bâti à l'aide d'images aériennes stéréoscopiques et de plans cadastraux vectorisés.

Ces données vectorielles 2D fournissent une quantité d'information qui permet de formuler des hypothèses sur la scène observée, et de plus, elles sont bien souvent déjà disponibles au moment de la saisie des bâtiments à forte résolution.

Toutefois, des problèmes liés à ces données subsistent. Dans le cas des données papier, la vectorisation et l'interprétation sont un problème d'importance, et lorsqu'il est résolu, ou absent (dans le cas des bases de données), il subsiste les incohérences entre l'information fournie et les images, et les objectifs de reconstruction.

Par incohérence on entend notamment les erreurs de saisie dans le cas d'une base de données, mais aussi des incohérences entre les images et la base, même si les objets de la base sont conformes aux spécifications : une base de données voit la scène au travers de ses spécifications (c'est le *terrain nominal*) qui agissent comme autant de filtres. Sans compter le fait que deux opérateurs peuvent produire deux saisies différents, en étant toutefois tous les deux conformes aux spécifications de la base.

A titre d'exemple, la figure 1.14 présente deux imagettes sur lesquelles sont superposées des parcelles cadastrales. Sur l'imagette de gauche, les parcelles cadastrales découpent le bâtiment en 3 parcelles principales, telles que les frontières entre parcelles correspondent effectivement à des différences de hauteur (on voit des portions de mur, en clair, sur cette imagette). Par contre, sur l'imagette de droite, la parcelle cadastrale regroupe plusieurs volumes que l'on pourrait représenter par 3 volumes différents dans le cas d'une représentation 3-D. Ces images sont dues à H. Jibrini, qui étudie l'exploitation du cadastre pour la reconstruction 3D du bâti ([Jib00]).

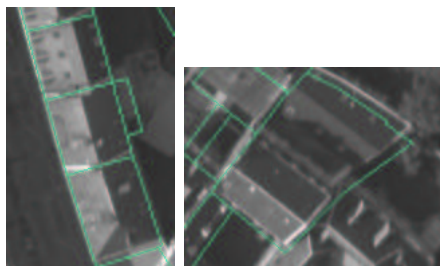


FIG. 1.14 – Exemples d'incohérences entre les objets topographiques et cadastraux.

3.4.4 Sources de données multiples

Dosch ([DM]) propose d'utiliser simultanément Cadastre et MNE afin de produire des descriptions volumétriques du bâti en milieu urbain. Ces travaux se limitent toutefois actuellement à des modélisations prismatiques.

Des travaux utilisant des données 2D vecteur et des MNE obtenus par balayage ont été menés par Haala [HB97, Haa99]. Cet auteur se sert notamment des échos multiples des données laser pour détecter la végétation.

3.5 Généralité des Formes Reconstituées

3.5.1 Approches par modèles

Les travaux de Dang cités plus haut utilisent une forme explicite de modèle rectangulaire. Läbe ([LG98]) utilise l'algorithme général RANSAC pour estimer les formes et positions de modèles paramétriques.

La système TOBAGO ([Grü98]) est lui aussi basé sur une famille de modèles.

Approches paramétriques

Haala propose une méthode de détection des bâtiments par reconnaissance des maxima locaux (excroissances) d'un Modèle Numérique d'Elevation, ce qui lui permet de se focaliser sur des bâtiments isolés. Une extraction de segments 3-D lui permet alors d'effectuer des groupements de ces segments suivant des règles perceptuelles : les droites parallèles permettent d'estimer la (ou les) orientations principales du bâtiment. Enfin, pour un sous-groupe de segments 3-D, les paramètres du modèle sont estimés en considérant des segments 3-D. Les sous-groupes candidats étant très nombreux, cela pose un problème combinatoire justement pallié par le groupement perceptuel des segments parallèles exposé juste avant. Le fait de grouper les segments suivant des règles perceptuelles permet de limiter la combinatoire des hypothèses de formes.

3.5.2 Approches génériques

Ces approches introduisent peu de contraintes sur les formes reconstruites. Elles supposent toutefois généralement que les bâtiments sont polyédriques, ce qui restreint fortement l'espace des formes possibles, mais reste assez général pour représenter des bâtiments. On pourra toutefois trouver des travaux adaptés à des formes non polyédriques ([ML92]).

[Hen96] propose une technique ascendante, par organisation progressive d'objets simples. La résolution des images est forte (8 cm). Il s'attaque toutefois à des régions pavillonnaires (bâti isolé). Toutefois, les reconstructions obtenues sont très fines. A l'aide des mêmes relations géométriques entre objets, [WA96] propose un système de règles générales permettant la reconstruction de polyèdres. Certaines règles ascendantes permettent de produire des hypothèses de structures de moyen ou de haut niveau à partir d'hypothèses de niveau plus bas. Inversement, certaines règles descendantes permettent de pallier les sous-détections des objets de bas-niveau en produisant des hypothèses pour ces objets.

3.6 Etats de l'art

Le lecteur trouvera dans [Ves00] un état de l'art plus vaste, portant notamment sur des équipes travaillant sur des modèles très complexes, afin d'obtenir des représentations 3-D de qualité architecturale, avec habillage des façades. Ces procédés sont toutefois très interactifs, et exploitent des prises de vues terrestres.

4 Présentation de l'approche

4.1 Justification

Nous avons vu que les méthodes les plus performantes (notamment les plus fines) utilisent la multiscopie, ou sont fortement interactives. Nous pensons toutefois qu'il est possible de progresser dans le domaine des reconstructions automatiques. En effet, même si les méthodes automatiques ne permettent pas de résoudre tous les problèmes de reconnaissance et de reconstruction posés par une scène urbaine, dans un contexte de production (qui est notre contexte), le fait de pouvoir reconstruire

automatiquement une proportion importante des bâtiments induit un gain non négligeable dans la constitution des bases de données.

Par ailleurs, nous estimons que toutes les potentialités des modèles numériques d'élévation n'ont pas encore été exploitées dans la littérature.

Nous allons donc proposer une stratégie utilisant abondamment les modèles numériques d'élévation. Par contre, en raison de la difficulté du problème de la reconstruction en milieu dense, et à partir de 2 vues, nous proposerons une approche qui utilisera une redondance importante entre les objets détectés à partir des données initiales. Enfin, un système d'alarme sera proposé, afin de guider l'opérateur humain après l'étape de reconstruction automatique.

Les prochaines sections décrivent les étapes majeures de la stratégie proposée.

4.2 Les données initiales

Les données initiales sont composées d'images stéréoscopiques, et d'un modèle numérique d'élévation (MNE) issu des travaux de [Bai97] [DBMJ97]. Le couple stéréoscopique est issu d'une prise de vue aérienne à grande échelle (1/4000), suivie d'un scannage avec un pas de $20\mu m$. La résolution des images numériques utilisées est donc de 8 cm. Ces images ont la particularité d'avoir été prises sous couvert nuageux, ce qui a pour principal impact l'absence d'ombre. Comme nous le verrons plus tard, cet effet est ressenti par les méthodes d'analyse d'image (voir page 43).

Le modèle numérique d'élévation est exploité dans ces travaux, ainsi que la séparation sol / sursol et l'identification de zones de sursol. L'entité traitée par ces travaux est la zone de sursol. Les zones de sursol sont traitées indépendamment les unes des autres.

Les figures 1.18 et 1.19 montrent les images gauche et droite d'un couple stéréoscopique issu de la prise de vue (ces images ne constituent pas le couple entier, mais seulement la moitié de la zone de recouvrement).

La figure 1.20 présente le modèle numérique d'élévation associé à ce couple d'images. Ce MNE est obtenu par corrélation automatique des images du couples. Cette corrélation utilise une programmation dynamique le long des lignes épipolaires du couple ⁸.

Enfin la figure 6 présente les zones de sursol identifiées par des méthodes mises au points dans ces travaux ([DBMJ97, BM99]).

Ces données permettent d'extraire des zones de focalisation. Une de ces zones est présentée dans la figure 1.22, où se trouvent des extraits du couple stéréoscopique, ainsi qu'un MNE restreint à la zone de sursol.

Ces travaux ont permis de produire des Modèles Numériques d'Elevation fiables et denses à moyenne résolution (résolutions typiques : 40 cm à 1m).

Malgré sa conception dédiée à des résolutions plus grossières que celle utilisée ici, cet outil sera exploité à ces résolutions plus fines, car il continue à donner des résultats très intéressants, et qui, en tout cas, permettent effectivement d'espérer reconnaître la structure interne des toits. Le lecteur n'aura qu'à se reporter aux figures C.1 et E.2 pour se convaincre de ce fait.

Toutefois cet argument n'est pas une preuve, mais nous tenterons de montrer a posteriori que ces modèles numériques sont suffisants.

⁸Pour plus d'efficacité, les images ont donc préalablement été rééchantillonnées de façon à ce que les lignes épipolaires du couple soient des lignes horizontales de l'image.

4.3 Présentation

L'état de l'art proposé ici montre que les méthodes automatiques permettant de se confronter aux formes les plus diverses (voir section méthodes génériques) utilisent des images à forte résolution, mais généralement en multiscopie. Dans notre contexte particulier (images à fortes résolution, milieu urbain, et stéréoscopie simple), nous tenterons de mettre en place un système se plaçant à mi chemin entre les méthodes paramétriques et les méthodes génériques. Par ailleurs, nous tenterons de pallier les lacunes inhérentes aux processus de focalisation, notamment l'impossibilité de séparer des bâtiments voisins et d'altitudes proches. A cet effet, nous proposerons une méthode capable de se «re-focaliser» sur des bâtiments, à l'aide de modèles. L'utilisation de ces modèles nous permettra de qualifier l'approche de structurelle. Cette méthode structurelle utilise des objets 3D simples, détectés dans les données initiales, et appelés primitives. En raison de la faiblesse de la stéréoscopie simple, ces primitives souffriront de sous-détections inévitables. Afin de combler ces sous-détections, nous mettrons en place un système permettant d'exploiter des redondances d'informations entre différents types d'objets simples.

La stratégie générale adoptée est présentée dans la figure 1.15. Dans cette figure, la ligne pointillée sépare les travaux antérieurs (ceux de C. Baillard) de nos propres travaux. Les prochaines sections présentent plus en détail cette approche.

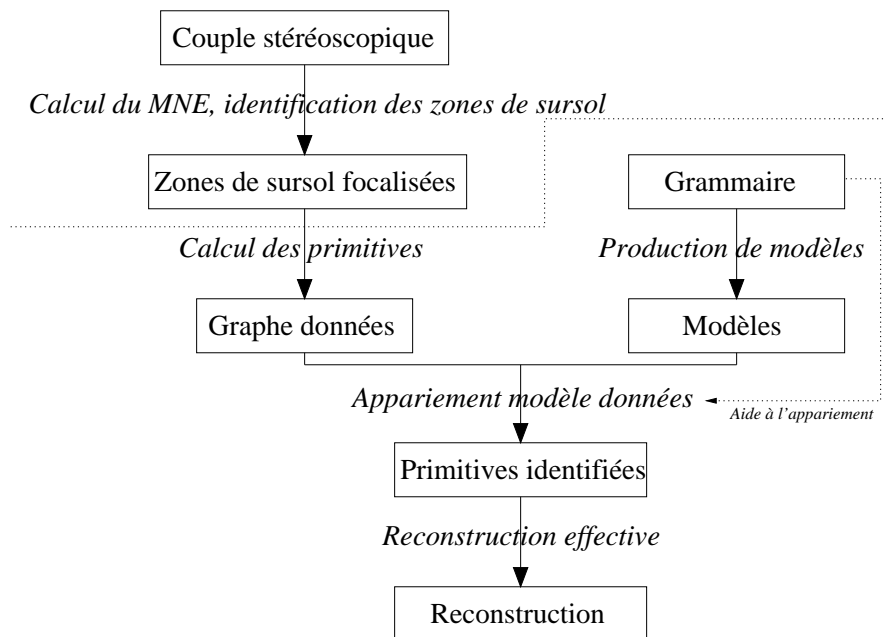


FIG. 1.15 – La stratégie générale adoptée

4.4 Cas de la végétation

En milieu urbain, le sursol est constitué essentiellement du thème bâti et de la végétation. Des études ont été menées afin d'identifier de façon automatique la végétation ([Bai97]), et d'autres

travaux sont en cours (voir les actes [asc01]). Parallèlement, l'utilisation de capteurs dans le domaine du proche infra-rouge, laisse penser que l'identification de la végétation sera un problème «facile» ou aussi bien résolu que l'extraction du thème bâti seul ⁹. Nous estimons donc que l'identification de la végétation est une tâche qui est hors du champ de ces travaux, d'autant plus que la végétation se caractérise essentiellement par sa couleur et ses micro-textures, ce qui nécessite des techniques de reconnaissance assez distantes de ce qui est requis pour le bâti, puisque le bâti se caractérise surtout par sa forme. Nos travaux seront en fait essentiellement des travaux de reconnaissance des formes, assez éloignés de techniques adaptées à la détection de la végétation.

En conséquence, les régions de sursol constituées de végétation seront éliminées par examen visuel. Par contre, les régions de sursol composées de bâti *et* de végétation seront laissées intactes, de manière à appréhender l'impact des arbres sur les techniques de reconstruction proposées ici.

On notera que le cas du bâti *recouvert* de végétation est donc exclu du champ de ces travaux.

Nous considérons donc maintenant qu'une zone de focalisation ne contient qu'un ou plusieurs bâtiments.

4.5 Diviser pour régner

Nous appellerons zones ces zones de focalisation produites par la séparation sol/sursol citée plus haut. Ces zones constituent l'entité sur laquelle nous travaillerons. Les zones seront traitées indépendamment les unes des autres. On distinguera donc la notion de scène (de grande étendue, comme la figure 1.18) de celle de zone (faible étendue, comme la figure 1.22).

4.6 De la zone aux objets simples

Comme de très nombreuses méthodes (sinon toutes), nous aborderons la reconstruction du bâti par une première étape qui consiste à calculer pour une zone donnée, à partir des données initiales, des entités de moyen niveau, que nous appellerons primitives. Ces entités sont des objets 3-D simples. Il s'agit de segments de droites, de portions de plans, et d'objets que nous apparenterons pour l'instant aux façades des bâtiments ¹⁰, et que nous appellerons discontinuités ¹¹.

Ces types d'objets seront détectés à partir des données initiales, mais indépendamment les uns des autres. Ceci nous placera en décalage avec de nombreux auteurs, qui généralement détectent un type d'objet, puis un autre en s'aidant du premier ([Hen96, Cor98]). Cette optique vise à ne pas propager les erreurs d'une primitive sur des primitives d'un autre type. L'inconvénient est alors de ne pas pouvoir combler les lacunes de certaines primitives par les qualités des primitives des types différents, mais on verra au chapitre 3 que nous conserverons tout de même la possibilité d'un renforcement mutuel des primitives lors de la seconde étape de la stratégie. En effet, dans la seconde étape de la stratégie, l'étape d'appariement modèle / données, les primitives seront utilisées simultanément, et mises sur un pied d'égalité.

⁹en supposant que les difficultés de ces deux problèmes très différents soient comparables.

¹⁰Le lecteur prendra bien garde à la précaution verbale prise ici : nous avons dit précédemment que nous ne rechercherons pas les façades car cela était illusoire, et nous prétendons ici rechercher des objets qui s'y apparentent. Le chapitre 2 décrira quels sont réellement les objets présentés ici et lèvera le voile sur ce mystère apparent. Le lecteur curieux de savoir maintenant quels sont ces objets peut se reporter à la section 5 du chapitre 2.

¹¹Encore une fois, le chapitre 2 explicitera l'origine de ce nom.

4.7 Les graphes entrent en scène

La seconde étape de la stratégie consiste à définir des modèles de bâtiments, et à les utiliser afin de produire effectivement des reconstructions. Cette étape donne la qualification de structurelle à la méthode.

Dans la pratique, les modèles sont donnés par des graphes, car les graphes sont des outils de représentation puissants, adaptés aux reconstructions recherchées, et qui permettent de formaliser les problèmes posés ici.

Un graphe est composé de sommets et d'arcs qui les relient. La figure 1.16 présente un graphe simple contenant 4 sommets et 5 arcs.

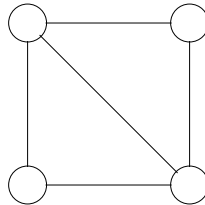


FIG. 1.16 – Notre premier graphe.

Les sommets des graphes modèles représentent les primitives 3-D utilisées. Ils ne jouent donc pas tous le même rôle, puisqu'ils ont des natures différentes. De plus, dans certains cas, une propriété géométrique peut venir s'adjoindre à la nature de l'objet, par exemple l'horizontalité : une gouttière est généralement horizontale et il est intéressant de faire figurer cette propriété dans les modèles afin d'en tirer parti. Les sommets possèdent donc des attributs qui représentent leur nature, et d'éventuelles propriétés géométriques.

Les arcs qui relient les sommets ne sont pas non plus quelconques. Ainsi, la relation qui relie une gouttière à son pan adjacent n'est pas la même que la relation qui relie la faîte à ses pans adjacents. Les arcs possèdent donc aussi des attributs qui représentent des propriétés géométriques que vérifient les deux sommets aux extrémités de l'arc.

Ainsi, formellement, nous utiliserons des graphes relationnels attribués pour représenter les modèles. De la même façon, les primitives calculées peuvent être représentées par un graphe relationnel attribué. La figure 1.17 présente deux graphes relationnels attribués. Les attributs sont représentés par les couleurs. Bien que leurs formes soient les mêmes, ces graphes ne sont pas semblables (isomorphes) à cause d'une différence d'attribut sur un des sommets. De même, il peut y avoir des différences d'attribut sur un arc (symbolisées par une différence de couleur des arcs sur cette figure).

A l'aide d'un graphe modèle, et du graphe représentant les données, nous procéderons à une *identification* qui consiste à reconnaître, dans le graphe des données, le modèle. Techniquement, il faut trouver dans le graphe des données, un sous-ensemble des données (un sous-graphe) qui ressemble le mieux au modèle.

Trois problèmes majeurs se posent :

- Il faut définir la fonction de ressemblance.
- Il faut trouver le sous-graphe le plus ressemblant.
- Si une primitive nécessaire n'a pas été détectée, il faut que le système comprenne qu'elle n'est pas là, et s'adapte.

Nous tenterons de répondre à ces questions par le formalisme des distances d'édition (voir [Mes96]).

Nous appellerons également cette identification l'appariement entre le modèle et les données, puisque cela consiste à établir les correspondances (des paires, d'où le terme d'appariement) entre les sommets du graphe modèle, et certains sommets du graphe appuyé sur les données.

Une fois cette identification faite, la reconstruction proprement dite doit être réalisée. L'identification n'est en effet qu'une recherche, parmi les primitives, des primitives pertinentes au regard du modèle. Or une reconstruction doit être cohérente (les points doivent être aux extrémités des segments, les segments dans les plans ...), une étape de reconstruction simple est donc proposée. Cette reconstruction effective s'appuie uniquement sur le modèle, les primitives identifiées par l'appariement, et le formalisme des distances d'édition déjà cité.

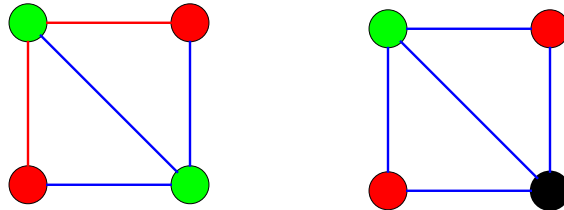


FIG. 1.17 – Nos deux premiers graphes relationnels attribués

4.8 Les modèles grandissent.

L'exploitation d'une grammaire de graphe permettra ensuite de générer de façon automatique et homogène les modèles. En outre, elle permettra de profiter de certaines optimisations de la méthode d'identification.

Cette stratégie n'est toutefois pas suffisante, car deux points n'ont pas été traités : le choix d'un modèle adapté aux données, et, le cas échéant, la reconstruction de plusieurs bâtiments au sein d'une même zone de focalisation. Ces deux points font l'objet de la dernière étape proposée (section 4.9).

La classe des objets reconstruits

Du fait de l'utilisation de cette grammaire, la classe des objets reconstruits est connue a priori. Dans la pratique, la grammaire ne permettra pas d'engendrer toutes les formes de toits possibles. Donc, la question de savoir quelle proportion de bâtiment peut être traitée par les méthodes proposées se posera inéluctablement. A contrario, nous verrons que la méthode est souple, et permet de définir les modèles utilisés de façon *déclarative*, ce qui n'est pas le cas, par exemple des méthodes du type *paramétrique*, pour lesquelles généralement la connaissance est donnée de façon procédurale (le code source des programmes, par sa structure, l'imbrication de ses boucles pour faire varier des paramètres ou minimiser une fonction de coût, contient implicitement le nombre de degrés de liberté présents dans les formes du bâtiment). La méthode se montrera donc extensible à souhait, malgré une certaine lourdeur qu'il faut d'ores et déjà reconnaître. Inversement, comme nous verrons, le fait de se restreindre à une famille bien définie de modèles permet dans une certaine mesure de ne pas se perdre dans des problèmes de bruit et d'artéfacts.

4.9 Et la reconstruction grandit.

La grammaire ne sait pas tout faire. Dans le cas de deux bâtiments voisins, voire d'un îlot de bâtiments, le modèle devient complexe, voire très complexe, ce qui rend sa reconnaissance dans les données illusoire d'un point de vue pratique. En s'appuyant sur des reconstructions élémentaires proposées par les premières étapes de la stratégie, et sur une évaluation de ces reconstructions, nous proposerons une stratégie générale visant à reconstruire petit à petit une zone de focalisation.

5 Plan du mémoire

La suite du document présente des outils s'appuyant successivement les uns sur les autres. Le chapitre 2 présentera les extractions de primitives. Trois types de primitives seront extraits (portions de plans, segments, et discontinuités). Le chapitre 3 présentera les travaux d'appariement de graphe modèle / données (l'identification), ainsi que la méthode de reconstruction qui en découle pour un modèle donné. Le chapitre 4 présentera l'apport des grammaires de graphe pour l'appariement. Le chapitre 5 proposera une stratégie effective de reconstruction d'une zone de focalisation. Le chapitre 6 présentera les résultats et leur évaluation. Le chapitre 7 conclura ce mémoire.

Un certain nombre d'annexes sont portées à ce document, dont notamment un glossaire, un résumé des notations, et l'explicitation de fonctions de coût introduites dans ce mémoire. Plusieurs autres annexes sont des illustrations des méthodes proposées.

6 Conclusion

Ce chapitre a présenté des considérations générales devant permettre au lecteur d'entrer dans la problématique posée ici. Il a également présenté l'approche proposée et situé cette approche par rapport à l'existant. Plus particulièrement, nous avons montré le fait que l'approche proposée n'est ni très générale (elle n'est pas générique), ni très limitée, car elle est déclarative et non procédurale. De plus, nous avons précisé un certain nombre de problèmes auxquels nous tenterons d'apporter des réponses dans ce mémoire : les lacunes des techniques de focalisation, auxquelles nous confronterons une technique d'appariement permettant de voir un modèle dans des données même si le champ des données est plus vaste que l'instance du modèle ; les problèmes de stéréoscopie «simple», auxquels nous tenterons de répondre par une complémentarité des objets de moyen niveau (primitives) utilisés.



FIG. 1.18 – Un extrait du couple stéréoscopique de Colombes (image de gauche).



FIG. 1.19 – Un extrait du couple stéréoscopique de Colombes (image de droite).

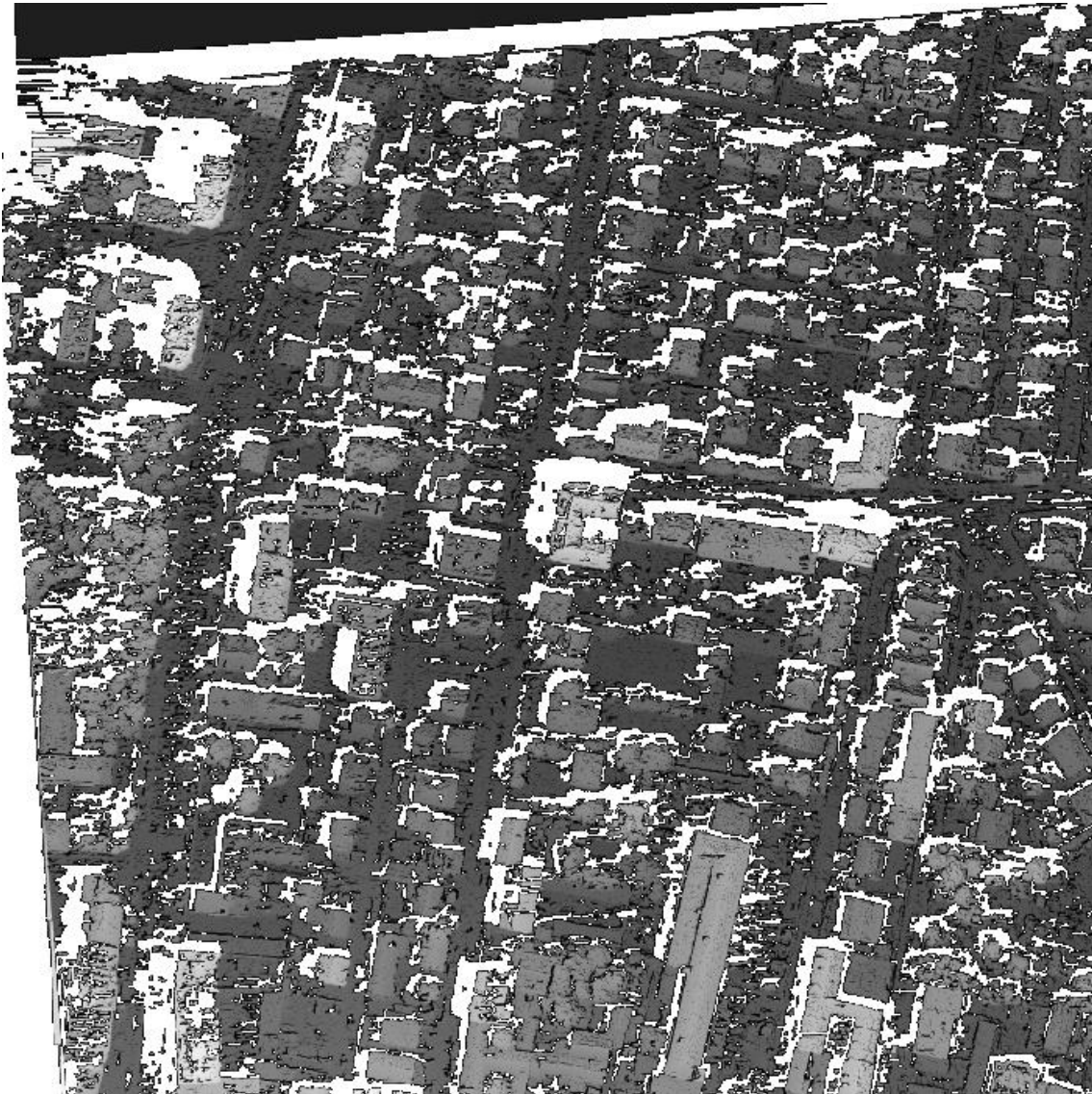


FIG. 1.20 – Le modèle numérique d'élévation (MNE) associé au couple stéréoscopique des figures 1.18 et 1.19.

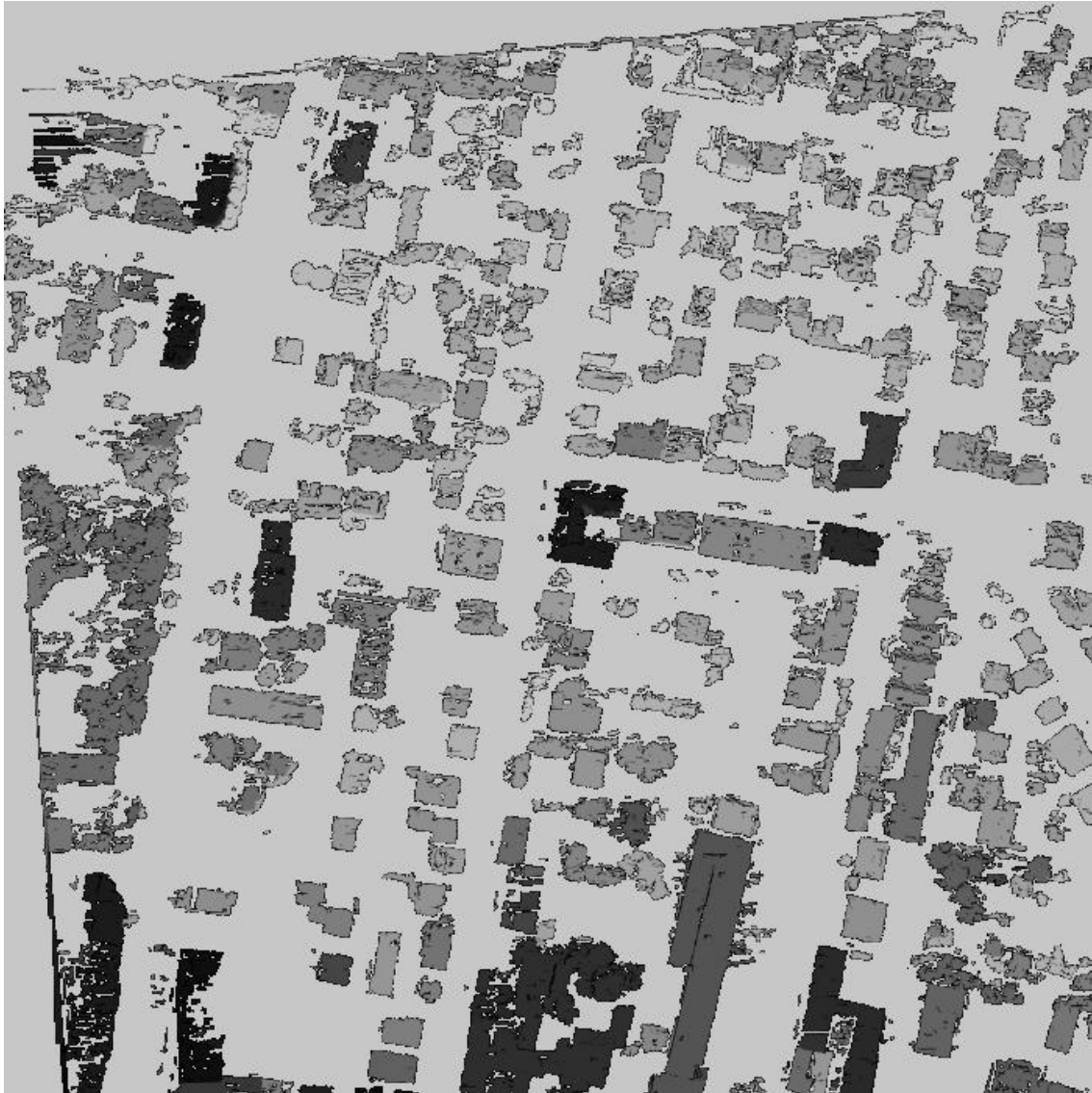


FIG. 1.21 – Les zones de sursol identifiées dans le MNE.



FIG. 1.22 – Vue stéréoscopique du bâtiment 172, et le MNE associé.

Chapitre 2

Détection de Primitives

1 Introduction

Ce chapitre décrit la première étape de la stratégie générale adoptée. Après avoir décrit pourquoi et comment le Modèle Numérique d'Élévation est nettoyé (section 2), il présente successivement les 3 méthodes correspondant aux 3 types de primitives détectés. Ces trois détections sont rigoureusement indépendantes.

Ces objets seront appelés *primitives*, car ce sont les éléments de base sur lesquels s'appuieront les reconstructions des bâtiments.

Un quatrième type de primitive (les sommets 3D) est utilisé dans les modèles, mais n'est pas détecté dans les données. La section 6 précisera pourquoi.

2 Nettoyage du MNE

2.1 Motivation

Le Modèle Numérique d'Élévation utilisé est celui produit par les travaux de Caroline Baillard ([Bai97]).

Nos travaux (notamment la section 3) nécessitent un MNE dense. L'examen du modèle numérique d'élévation montre que celui-ci n'est dense que dans une certaine mesure.

En effet, à moyenne échelle (l'échelle pour laquelle les travaux de C. Baillard ont été menés), le MNE est dense, dans la mesure où l'on reconnaît (visuellement) explicitement la structure de la grande majorité des bâtiments. A grande échelle, par contre, on observe par endroits l'absence d'information. C'est le cas, par exemple de la zone 15 (figure 2.2) où le MNE du bâtiment n'est pas connexe, même si l'on distingue dans ce MNE la structure du toit.

Cette constatation nous pousse à mettre en place une procédure de nettoyage, afin de mettre le MNE en conformité avec le présumé de densité que nous faisons ici. Ce prétraitement est local, afin de préserver l'information structurante du MNE.

La méthode vise à supprimer d'éventuels pixels du MNE qui auraient des altitudes trop basses. Ceci se produit le long des bords de bâtiment dont les façades sont vues sur les images initiales : dans la mesure où le MNE est calculé sur ces images, certains pixels des façades sont mis en correspondance,

et projetés dans le MNE au voisinage des bords de bâtiments, mais à des altitudes allant du sol à la gouttière. On observe donc cet artéfact sur certains bords de bâtiments (voir le zoom sur la figure 2.1).

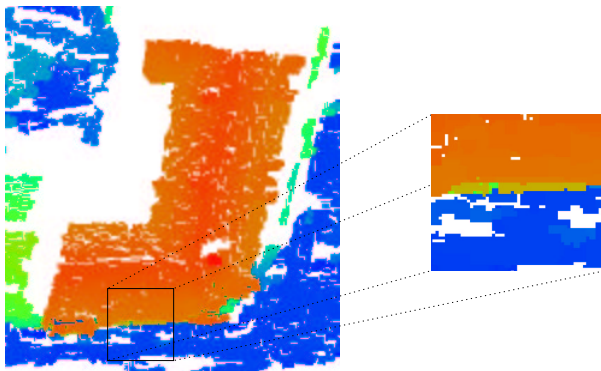


FIG. 2.1 – Modèle Numérique d'Élévation de la zone 19 (bleu = 25m, rouge = 52m). A droite, un zoom sur le bord de toit. On remarque les artéfacts en jaune et vert, liés à l'accumulation de points sur la façade vue de ce bâtiment.

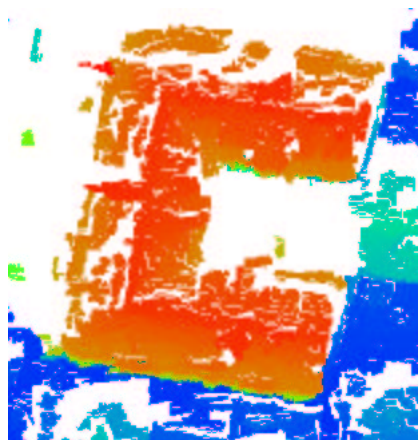


FIG. 2.2 – Modèle Numérique d'Élévation de la zone 15 (bleu = 25m, rouge = 56m)

2.2 Méthode de Nettoyage du MNE.

2.2.1 Etape 1.

Le MNE souffre de quelques petits trous liés au passage de l'espace image à l'espace terrain. Ces trous sont comblés sans difficulté par une fermeture morphologique ¹ de taille minimale : par une

¹Cette fermeture morphologique est une fermeture en niveaux de gris, et non binaire, ainsi que tous les outils morphologiques utilisés dans ces prétraitements

fenêtre carrée 3x3.

2.2.2 Etape 2

Le MNE souffre de trous plus importants liés à des absences d'information. Une fermeture morphologique plus importante (de fenêtre caractéristique 1m) permet de combler ces trous en 2D. Mais en 3D cette fermeture provoque une modification des formes du MNE (les altitudes augmentent). C'est donc une fermeture conditionnelle qui est utilisée pour combler ces lacunes : seuls les pixels ne portant pas de valeur avant ce traitement sont affectés par la fermeture.

Enfin, les pixels ne comportant aucune valeur après cette fermeture conditionnelle sont mis à la valeur du terrain (au sol).

2.2.3 Etape 3

Soit r_f un rayon de filtrage exprimé en mètres.

Soit un pixel p du MNE, d'altitude $h(p)$. Il représente un point 3-D $P(x, y, z)$. Soit $E(p)$ l'ensemble des pixels du MNE, situés à une distance planimétrique de p inférieure à r_f . Soit h_m la valeur médiane des altitudes des points de $E(p)$. Si $h(p) < h_m - r_f$, alors le pixel est jugé trop bas pour appartenir au toit, et est donc mis à la valeur du terrain.

La valeur de r_f est choisie de façon à ne pas mettre au sol des pixels qui appartiendraient à un autre toit de la même zone de focalisation. Or les bâtiments sont regroupés au sein d'une même zone si leurs altitudes caractéristiques sont distantes de moins de 2,5m. Nous prenons donc : $r_f = 3m$.

Cette étape a pour effet majeur de gommer les artéfacts cités plus haut (voir figure 2.1), mais au bout du compte, l'image ainsi produite contient des informations «sol» qu'il faut à nouveau gommer au cours d'une étape de re-focalisation :

2.2.4 Etape 4 : refocalisation

Une refocalisation autour de la zone de focalisation initiale est nécessaire, car au cours des étapes précédentes, le sol et d'autres bâtiments éventuels du voisinage de la zone peuvent réapparaître dans l'image focalisée. L'objectif est donc de ne conserver, parmi les pixels traités au cours des étapes précédentes, que ceux qui appartiennent au bâtiment désiré. Pour chaque pixel p , si aucun pixel du masque initial n'est à une distance 3-D inférieure à r_f , ce pixel est mis au sol.

La figure 2.3 présente le MNE initial pour la zone 19 (A), ainsi que le masque de focalisation (B), puis les résultats des 4 étapes du traitement (C,D,E,F). Sur l'imagette D de la figure, on constatera la présence du MNT. C'est le MNE F qui est utilisé dans la suite des travaux.

On constate principalement une régularisation des bords du MNE, et la suppression de certaines aberrations de faible surface du MNE initial, mises en évidence sur la figure 2.4, qui présente la différence entre le MNE initial et le MNE final. Sur cette figure, les différences fortes sont en bleu (-17m) et les différences faibles en rouge. Cette valeur importante (17 m d'écart localement, en bas à gauche du bâtiment), est liée à un artéfact des données initiales.

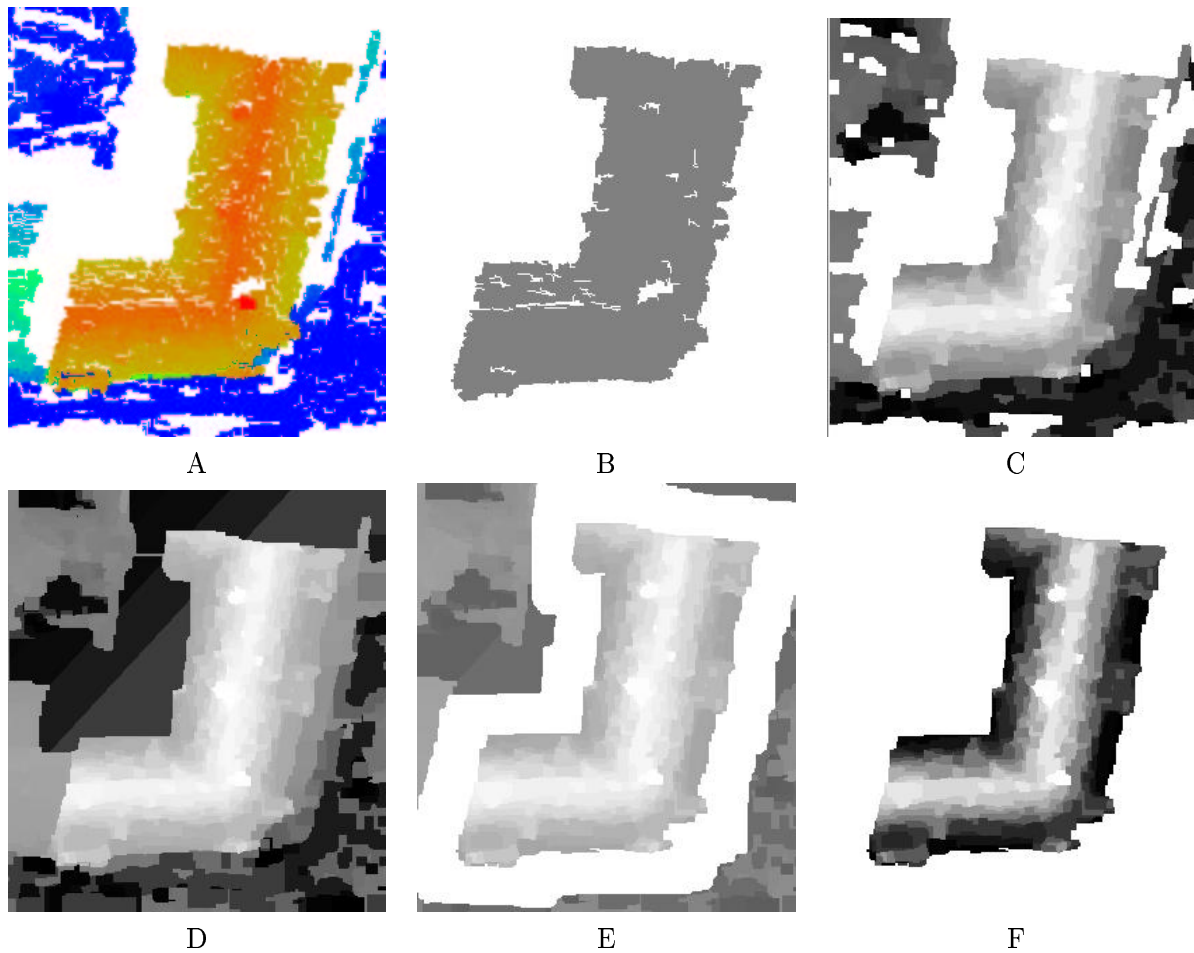


FIG. 2.3 – Modèles numériques d’élévation correspondant au bâtiment 19 après traitement. Chaque image a subi une forte augmentation de dynamique afin de mettre en évidence le fait que la structure du toit est effectivement présente dans le modèle numérique d’élévation. (A) MNE initial. (B) Masque de focalisation (C) Résultat de l’étape 1. (D) Résultat de l’étape 2. (E) Résultat de l’étape 3. (F) Résultat de l’étape 4.

3 Les Primitives Planes

3.1 Définition

Les primitives planes sont les pans de toit, comme nous l’avons déjà exprimé. Nous tenterons des les trouver dans les données en effectuant une segmentation du modèle numérique d’élévation visant à détecter des régions planes, c’est-à-dire obéissant à une loi affine $z = ax + by + c$.



FIG. 2.4 – Différence entre le MNE initial est le MNE final pour la zone 19. Rouge : 0m, Bleu : -17m. Les petites aberrations du MNE initial ont été supprimées.

3.2 Etat de l'Art en Segmentation d'Images de Profondeur.

La segmentation d'images de profondeur fait l'objet d'un certain nombre de travaux. La communauté travaillant sur ce domaine organise même, en 2000, en conjonction avec la conférence ICPR '2000 à Barcelone, un groupe de travail destiné à effectuer l'état de l'art dans le domaine, et à identifier les faiblesses communes de ces méthodes dans le but de définir les futures priorités de recherche en la matière. Ce groupe de travail ('Range Image Segmentation Contest') peut être trouvé sur le site <http://iamwww.unibe.ch/~fkiwww/RIA/RangeSegContest/index.html>

Les résultats de ce groupe de travail auraient dû faire l'objet d'un numéro particulier du journal IJPRAI (International Journal on Pattern Recognition and Artificial Intelligence), mais les résultats de ce workshop n'ont pas été suffisants pour effectuer cette publication (nous citons ici l'organisateur du colloque).

On trouvera par contre dans la littérature des travaux d'ordre général ([JB96, JB94]), ainsi que des applications de ces méthodes à notre domaine ([HB97]).

Roux ([RMG97]) effectue une densification de cartes de disparités, qui, explicitement, s'appuie sur une segmentation des images couleur utilisées. Cette densification passe explicitement par une segmentation en plans.

3.3 Stratégie

Les primitives planes sont issues d'un calcul s'appuyant exclusivement sur le MNE.

Nous allons exposer 3 raisons qui nous poussent à ne pas utiliser de segmentation unique du MNE :

- Raison applicative : on souhaite parfois obtenir plus de détail d'une région particulière d'une image, alors que dans d'autres régions, une représentation plus générale convient. Cette vision nécessite des méthodes très adaptatives, sachant se focaliser sur les zones dans lesquelles des algorithmes spécifiques devront être utilisés. La compression intelligente notamment vise cet objectif. Ce type d'approche est trop vaste pour être traité ici.

- Raison algorithmique : pour un algorithme donné, il est possible d’obtenir un résultat satisfaisant dans une certaine région de l’image pour un certain paramétrage, alors qu’un autre paramétrage est plus adapté à telle autre région de l’image. Ceci est très lié aux idées évoquées juste avant. Dans ce cas, on ne peut pas espérer trouver un paramétrage unique satisfaisant tous les besoins de l’application.
- Raison cartographique : d’un point de vue cartographique il est tout à fait possible de souhaiter plusieurs représentations différentes d’une même scène. Ce fait trouve sa source dans la nature même de la cartographie : il est capital de donner l’information pertinente au regard de l’échelle où l’on se place, et donc de l’application : à petite échelle, la structure générale d’une ville peut suffire, à moyenne échelle, l’identification des bâtiments peut suffire, à grande échelle, on peut souhaiter distinguer les structures internes des toits.

Ces trois points sont probablement sujets à débat. Ils ne visent pas à rejeter en bloc les recherches de segmentations, mais à poser le problème de la validité d’une segmentation *au regard de l’application qui en est faite*. Par contre, dans notre contexte, la troisième raison nous poussera à ne pas utiliser les résultats d’une segmentation unique, contrairement à ce que propose de faire la majorité des auteurs dans la littérature.

Bien au contraire, nous allons proposer une méthode de segmentation a priori sans paramètre. Mais nous ferons varier l’une des constantes de la fonction à minimiser au cours de cette segmentation, afin d’obtenir plusieurs segmentations complémentaires, plus ou moins généralisées. Par après, nous collecterons et exploiterons simultanément les résultats de ces segmentations.

Le double intérêt de cette stratégie sera montré a posteriori : il s’agit de ne provoquer que très peu de sous-détection des primitives planes d’une part (au prix d’une certaine sur-détection, contrée par un filtrage), et de pouvoir produire des reconstructions plus ou moins généralisées d’autre part (en s’appuyant sur des résultats de segmentations plus ou moins généralisées).

3.4 Méthode

La méthode utilisée propose une fonction L , qui pour une partition \mathcal{P} de l’image à segmenter, définit un coût lié à la complexité de la segmentation, et à sa validité au regard des données de l’image. L est un critère s’appuyant sur le principe de longueur minimale de description (Minimum Description Length, MDL, en anglais). Ces outils théoriques ont été posés par Rissanen ([Ris83, Ris86]), puis exploités sur des problèmes de reconnaissance de formes dans des cas de formes géométriques planes ([Li92]), de modélisation des formes ([CC99]), ou dans celui de bâtiments ([WF95, Ves00]).

Nous ferons quelques rappels concernant les idées générales de cet outil, avant de présenter l’utilisation faite pour la segmentation de Modèles Numériques d’Élévation.

Dans notre cadre, nous considérons une simplification du formalisme proposé par Rissanen. Nous nous ramenons à l’expression du coût des données sous forme d’une longueur de codage.

Ainsi, en supposant que les observations obéissent à un modèle M (M est donné par une partition et les approximations planes des valeurs des pixels pour chacune des régions de cette partition), nous notons la longueur de description des observations relativement au modèle M (exprimée en bits) :

$$L(\mathcal{P}) = L_M(\mathcal{P}) + L_e(\mathcal{P}). \quad (2.1)$$

où L_M est la fonction décrivant la longueur de codage du modèle lui-même, alors que L_e décrit les erreurs commises par le modèle.

On voit ici clairement que cette approche réalise le compromis entre la complexité (ou la finesse) du modèle (qui se traduit par une valeur de L_M élevée), et les erreurs commises par ce modèle (le terme L_e).

Dans les 3 prochaines sections, nous présentons quelques notations utilisées pour exprimer L dans notre cas, puis nous détaillons les expressions de L_M et L_e avant d'en déduire le processus de minimisation proprement dit.

3.4.1 Notations

Nous utiliserons les notations suivantes :

- S est l'ensemble des pixels à traiter d'une image (S n'est pas nécessairement égal à l'ensemble de tous les pixels de l'image. En particulier, dans notre cas, nous travaillons sur un sous-ensemble de l'image issu de la focalisation).
- Nous noterons pour chaque pixel p de S , $I(p)$ la valeur du pixel p . I peut représenter la disparité du pixel, ou encore l'altitude elle-même ².
- Nous noterons \mathcal{P} une partition de S , R_i un élément de \mathcal{P} (une région de S). Et pour deux régions R_i et R_j voisines (au sens de la 8 connexité ³), nous noterons $R_i \leftrightarrow R_j$.
- Pour une région R , nous notons P_R le plan de régression linéaire approximant l'ensemble des valeurs des pixels de R . Pour chaque pixel $p \in R$, nous noterons $e_R(p) = I(p) - P_R(p)$ l'erreur de l'approximation plane.
- Nous noterons $O(R)$ la frontière de la région R , et $|O(R)|$ la longueur de cette frontière.
- Pour deux régions R_i, R_j d'une partition \mathcal{P} , nous notons $\mathcal{P}_{R_i, R_j} = \mathcal{P} \setminus \{R_i, R_j\} \cup \{R_i \cup R_j\}$ la partition obtenue après fusion des deux régions R_i et R_j dans \mathcal{P} .

3.4.2 Longueur de description d'une partition

Cette section précise la définition que nous utiliserons pour la Longueur de Description (LD) d'une partition (une segmentation) de S . Nous considérons un système de codage de S pour une partition \mathcal{P} donnée. Ce système de codage doit représenter S et les valeurs de ses pixels, sans perte, de telle sorte que, potentiellement, on puisse décoder l'information ⁴.

Le système de codage choisi suit trois étapes :

- Une représentation de \mathcal{P} elle-même est faite : chaque région est codée par son contour. Les régions étant connexes, ce contour est composé d'une ligne fermée correspondant à son contour, et d'éventuelles lignes décrivant les trous de cette région. Ces trous sont eux-même d'autres régions de la segmentation, qui seront représentées par ailleurs. Il n'est donc pas nécessaire de représenter explicitement les trous d'une région. Le codage d'une région peut donc se contenter d'un codage de son contour. Ce codage est composé du codage d'un des points du contour, qui

²Les travaux présentés dans [FM99] montrent que dans le cas où I est une image de disparités, la méthode est capable dans une certaine mesure, de détecter les façades vues des bâtiments.

³L'étude de l'impact du voisinage utilisé n'est pas faite, mais on peut raisonnablement estimer que l'impact de la connexité sur les résultats finaux est négligeable.

⁴Même si l'objectif ici n'est pas de transmettre réellement l'information puis de la décoder, mais seulement de la représenter.

sert de point initial de codage, puis d'une séquence de codage de Freeman ⁵. La longueur de codage correspondante est donc :

$$L(O(R)) = L_{pt} + 3|O(R)|$$

où L_{pt} est la longueur de codage de la position d'un des points du contour, $|O(R)|$ est la longueur de la séquence, et le terme 3 correspond aux 3 bits nécessaires pour coder l'une des 8 directions possibles pour chacun des termes de la séquence. La longueur de codage d'un point de S est une constante dépendant de la taille de S . Typiquement, il faut $\ln_2 |S|$ bits pour identifier l'un de ses points. L_{pt} vaut donc $\ln_2 |S|$. Plus simplement, L_{pt} peut être vu comme la longueur de deux entiers.

- Une représentation, pour chacune des régions, du plan d'approximation des valeurs des pixels de R . Nous considérerons ici que ce codage est une constante L_P valant typiquement la longueur (en bits) des 3 flottants nécessaires à la représentation des termes a, b, c de l'équation $z = ax + by + z$ d'un plan non vertical. Nous avons donc :

$$L(P_R) = L_p$$

- Une représentation des erreurs commises par l'approximation elle-même. De façon naturelle, cette représentation est la concaténation des représentations individuelles des erreurs $e(p)$ pour $p \in R$. Cette erreur est théoriquement flottante, mais en raison de l'échantillonnage α nécessaire des valeurs de I , on peut en fait la ramener à une valeur entière, moyennant une division de l'erreur commise par le pas d'échantillonnage des valeurs de I (cf. [Ris86]). La longueur de codage de l'erreur s'exprime comme $L(e_R(p)) = 1$ si l'erreur est nulle, ou de valeur inférieure au pas d'échantillonnage, et $L(e_R(p)) = 2 + \ln^*(|e_R(p)|/\alpha)$ si sa valeur absolue lui est supérieure (\ln^* est définie par : $\ln^*(k) = \ln(k) + \ln(\ln(k)) + \dots$ tant que ces termes sont positifs). La constante 2 intervenant dans cette expression correspond à 1 bit spécifiant que l'erreur est supérieure en valeur absolue à α , et à un bit de signe de l'erreur. Dans la pratique, nous prendrons α égal au pas d'échantillonnage en z du MNE (typiquement 10cm), et $\ln^* = \ln$, ce qui nous amène à :

$$L(e_R(p)) = \begin{cases} 1 & \text{si } |e_R(p)| < \alpha \\ 2 + \ln(|e_R(p)|/\alpha) & \text{sinon} \end{cases}$$

La longueur de description totale d'une partition \mathcal{P} est donc :

$$L(\mathcal{P}) = L(|\mathcal{P}|) + \sum_{R \in \mathcal{P}} \left\{ L_P + L_{pt} + 3|O(R)| + \sum_{p \in R} L(e_R(p)) \right\}$$

Le terme $L(|\mathcal{P}|)$ vaut approximativement $\ln_2 |\mathcal{P}|$ et est négligeable devant la somme de termes qui suit. Nous n'en tiendrons donc plus compte dans la suite.

Il faut noter que dans le cadre théorique, les résultats de l'exploitation du principe de MDL dépendent de la famille de distributions choisie. De la même façon, dans notre cas, les résultats

⁵Nous rappelons que le codage de Freeman consiste à coder une chaîne de points voisins par une suite de 8 directions possibles liant deux points successifs de cette chaîne.

dépendent de la façon dont nous représentons les régions. Nous aurions ainsi pu choisir d'autres modélisations, desquelles auraient découlé d'autres longueurs de description.

Au sens de la compression du signal, notre procédé est loin d'être optimal. Mais le but n'était pas de compresser réellement les données, mais uniquement de trouver un cadre qui permette de gérer simultanément les erreurs d'approximation ainsi que la complexité du modèle utilisé.

Variante

En considérant que la première région de la partition \mathcal{P} est codée en partant du premier pixel le plus en haut à gauche de S , et que la seconde région est codée en partant du premier pixel de S privé de la première région, et ainsi de suite, les points de départ des codage des contours n'ont pas besoin d'être explicitement représentés. Ainsi, le terme L_{pt} peut être supprimé des expressions précédentes. D'autre part, le terme $L(|\mathcal{P}|)$ peut être supprimé également car en codant les régions de cette façon il est aisé de voir à quel moment le codage représente une partition de S !

3.4.3 Analyse de cette Fonctionnelle

L'expression de cette fonctionnelle est assez voisine de certains classiques de la littérature. On notera notamment une certaine similarité avec la fonctionnelle introduite par Mumford et Shah ([MS89]) :

$$E(u, B) = \int_{\Omega \setminus B} \|\nabla u\|^2 + \int_{\Omega} |u - g|^2 + \lambda l(B)$$

où u est la fonction continue par morceaux qui approxime la fonction g , Ω est l'espace où g et u sont définies, B est l'ensemble des points où u est discontinue, et λ est un paramètre de pondération entre la régularité de la fonction (le terme en $\|\nabla u\|^2$) et le terme d'attache aux données ($|u - g|^2$) d'une part, et le terme de «complexité de la segmentation» $l(B)$ d'autre part, en supposant que le lieu des discontinuités de u représente effectivement les frontières de la segmentation.

La minimisation de ce type de fonction n'est pas triviale. Toutefois certains articles montrent qu'un processus de fusion de paires de régions qui minimise au mieux la fonction de coût après fusion est capable, pour une certaine famille de fonctionnelles, d'atteindre l'optimum global d'une certaine famille de segmentations (voir [BCG94]).

Par ailleurs, hormis des techniques statistiques lourdes (de type recuit simulé) la méthode de fusion de régions 2 à 2 est la seule méthode à notre connaissance, qui soit développée dans la littérature pour minimiser de telles fonctionnelles (voir [Lee99, ZY98]).

Nous ne sommes pas capables de prouver qu'un processus de fusion par couples de régions atteint le minimum de la fonction que nous avons proposée, mais en revanche les résultats que nous venons de citer nous encouragent à penser que la méthode de fusion proposée dans la prochaine section est adaptée à la fonctionnelle proposée.

3.4.4 Processus de fusion de régions

Il est illusoire de penser trouver la partition \mathcal{P}_0 qui soit un minimum absolu de la fonction L , comme vu précédemment. C'est pourquoi nous utilisons un processus de fusion de régions, tel que à chaque étape de son déroulement, nous cherchons à descendre *le plus vite possible* vers un minimum.

Le processus est composé de trois étapes : l'initialisation, le déroulement et la terminaison.

3.4.4.1 Initialisation. L'initialisation du processus de croissance de régions est arbitraire. Nous divisons S en une maille régulière carrée. Il est possible que la régression linéaire ne puisse pas être faite sur certaines régions (notamment sur les bords de S), dans ce cas, cette région se voit pénalisée d'une coût arbitrairement élevé.

3.4.4.2 Déroulement. On envisage pour tout couple de régions voisines R et T de la partition courante \mathcal{P} , la LD de la partition $\mathcal{P}_{R,T}$ obtenue par fusion de R et T . Parmi tous ces couples, nous fusionnons le couple (ou un des couples s'il y en a plusieurs) tel que la LD de $\mathcal{P}_{R,T}$ soit minimale. Formellement :

$$R, T \in \mathcal{P} \text{ sont fusionnées si et seulement si } R \leftrightarrow T \text{ et } L(\mathcal{P}_{R,T}) = \min_{R', R'' \in \mathcal{P}} L(\mathcal{P}_{R', R''})$$

3.4.4.3 Terminaison. Le processus s'arrête lorsque qu'il n'y a plus qu'une seule région, ou si la fusion des régions n'apporte plus de gain, c'est-à-dire lorsque :

$$\forall R, T \in \mathcal{P}, L(\mathcal{P}_{R,T}) > L(\mathcal{P}) \quad (2.2)$$

Le développement de cette équation donne :

$$3(|O(R \cup T)| - |O(R)| - |O(T)|) - L_P - L_{pt} > \sum_{p \in R} L(e_R(p)) + \sum_{p \in T} L(e_T(p)) - \sum_{p \in R \cup T} L(e_{R \cup T}(p)) \quad (2.3)$$

Le terme gauche de cette équation est le gain (en terme de complexité de la segmentation), et le terme droit est la perte en terme de qualité de la représentation des données. Le critère d'arrêt est atteint lorsque le gain en complexité ne compense plus la perte. Le sens de l'inégalité de l'équation semble aller dans le sens contraire de cette affirmation, mais le terme droit est en fait toujours négatif, car l'erreur commise sur $R \cup T$ est supérieure à l'erreur commise sur R et T indépendamment.

3.4.5 Discussion

On voit dans l'équation 2.3 que la valeur de la constante $L_P + L_{pt}$ influence de façon importante le critère d'arrêt. Plus précisément, plus cette constante est grande, plus le terme de gauche est faible, ce qui implique que le critère d'arrêt est difficile à réaliser : le processus s'arrête donc plus tard, ce qui conduit à une segmentation plus forte, et donc à une plus forte *généralisation* de la scène. Ceci est cohérent avec le fait que cette constante est associée au codage d'une région : plus les régions sont coûteuses, moins il y en reste après le traitement.

Pourtant, d'un point de vue théorique, cette constante devrait effectivement être fixée. Elle équivaut en effet au codage de 3 flottants et 2 entiers, comme le montre la section 3.4.2. Ceci implique donc que pour ce modèle de codage, la méthode ne possède aucun paramètre. Dans la pratique, nous faisons toutefois varier la valeur de $L_P + L_{pt}$ afin de tirer profit de la capacité de généralisation de la méthode. Cette étape est expliquée dans la section qui suit.

On notera par ailleurs, de manière informelle, que L. Guigues, chercheur du MATIS, mène des travaux de segmentation d'images utilisant également le formalisme MDL, et dans lesquels intervient

ce type de constante. Il constate également, avec une exploitation de ce principe un peu différente de la nôtre, que les constantes prévues par la théorie mènent à des sursegmentations. Ceci ne remet pas en cause le principe lui-même, mais probablement les choix effectués dans les codages. En effet, rien ne prouve que justement les choix de codage soient optimaux. Cette notion de codage optimal est à rapprocher de la complexité de Kolmogorov : la complexité d'un signal est représentée par la longueur du plus petit programme pouvant produire ce signal. Il est clair que l'implémentation ad hoc que nous avons faite ici pour segmenter les images d'altitude ne répond pas à ce critère.

Cette expérience menée par L. Guigues nous conforte donc dans l'idée d'une variation de la constante $L_p + L_{pt}$ afin d'obtenir des résultats plus ou moins généralisés.

3.5 Exploitation de la méthode

3.5.1 Résultats bruts

On rappelle ici que d'après les développements précédents, la méthode devrait conduire à une segmentation unique, puisqu'elle ne comporte aucun paramètre. Toutefois, pour des valeurs proches des valeurs théoriques de la constante $L_P + L_{pt}$, les segmentations obtenues ne sont pas pertinentes au regard de notre application. Ce fait est explicité par la figure 2.5. On y voit en effet (image e) que pour les valeurs de $K = L_P + L_{pt}$ valant 32 (32 est la taille de 3 flottants plus 2 entiers), les résultats de la segmentation sont un peu trop fins.

Toutefois, un examen précis permet de constater que pour ce bâtiment particulier, qui possède des prolongements horizontaux de ses pans de toits, les régions désignées dans l'image f de la figure 2.5 correspondent effectivement à ces méplats du bâtiment. Dans l'ensemble, le bâtiment est tout de même sursegmenté. Les faîtes notamment, se traduisent par des régions identifiées comme horizontales. Ce phénomène provient du fait que sur le MNE, les faîtes correspondent effectivement à des bandes d'altitudes constante. Ces bandes, dont la largeur n'est pas nulle, sont détectées.

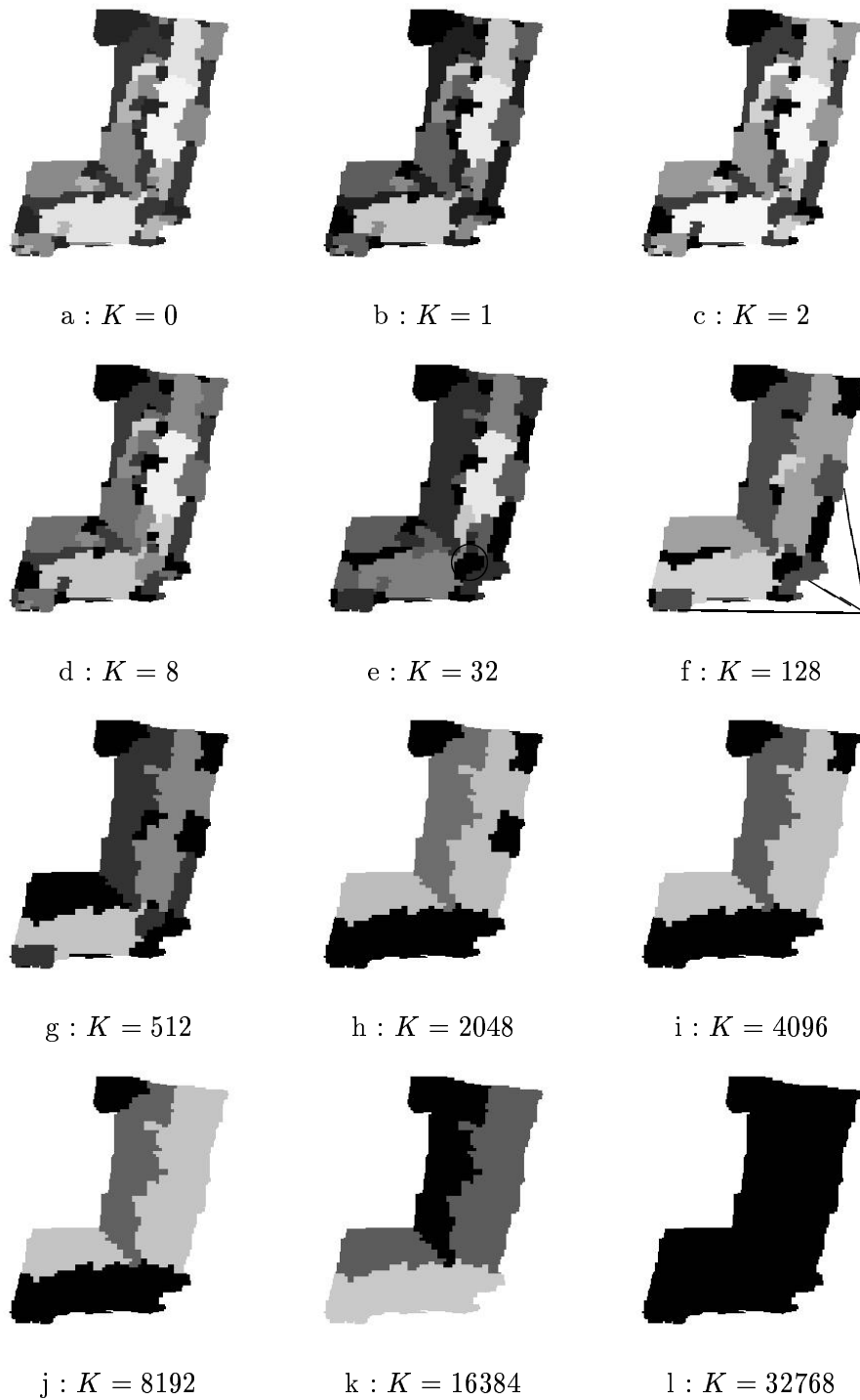
Les images suivantes de la figure montrent en revanche, que pour des valeurs plus fortes de K , le processus franchit des étapes de segmentation plus satisfaisantes, pour aboutir finalement à la fusion complète de S en une unique région, qui est une segmentation très généralisée.

3.5.2 Utilisation simultanée de résultats multiples

Collecte des résultats. Les résultats de segmentation obtenus pour les différentes valeurs de $K = 0, 1, 2, \dots, 2^N$ sont collectés puis filtrés grâce à des critères physiques. Pour éviter de conserver trop de petites régions non significatives, ne sont conservées que les régions possédant une aire supérieure à un seuil S_{aire} . Actuellement la valeur de ce seuil est de $S_{aire} = 5m^2$.

Nous sommes convaincus qu'il faut faire varier les valeurs jusqu'à des valeurs très élevées, mais le mode de variation (toutes les puissances de 2) est tout à fait empirique. Nous n'avons pas pu justifier ces valeurs du point de vue théorique.

Estimation robuste de la position des plans. La question de l'estimation de la position des plans se pose, car l'estimation aux moindres carrés est connue pour sa sensibilité aux aberrations des données. Dans le cas présent, étant donné le grand nombre d'échantillons sur lesquels les plans sont calculés, nous estimons toutefois que les aberrations du MNE peuvent être absorbées.

FIG. 2.5 – Segmentations pour différentes valeurs du paramètre d'arrêt K .

Toutefois, afin de nous en assurer, nous proposons de comparer les écarts quadratiques médians des plans estimés aux moindres carrés et les écarts quadratiques médians des plans estimés par une méthode robuste.

La méthode robuste utilisée est celle présentée par Girard ([GGMR98]). Dans cet article, Girard propose 3 méthodes. Nous utiliserons la première. Cette méthode effectue des tirages aléatoires de triplets de points. Pour chaque triplet non colinéaire, un seul plan passe par ces trois points. La méthode consiste à retenir, pour un grand nombre de tirages de triplets, le plan qui minimise précisément l'écart médian.

Dans la pratique, nous n'utilisons pas l'écart médian pur, mais un écart «médian ajusté» à 0.75 : il s'agit pour un ensemble d'écarts, de ne pas retenir l'écart médian, mais l'écart tel que 75% des valeurs lui sont inférieures, et 25% lui sont supérieures. Ceci permet de considérer dans l'évaluation 75% des données au lieu de 50%.

La figure 2.6 présente une comparaison des écarts. L'axe des x de cette figure correspond à l'écart quadratique médian ajusté pour les plans aux moindres carrés. L'axe des y correspond à l'écart pour les plans estimés de façon robuste. Cette figure montre clairement que les points se répartissent sensiblement suivant une droite d'équation $y = x$, ce qui nous conforte dans l'hypothèse que l'estimation aux moindres carrés est suffisante dans notre cas, en raison du grand nombre de points sur lesquels les plans aux moindres carrés sont estimés, et en raison de la fiabilité du MNE. Toutefois, dans la mesure où ce travail d'estimation robuste est fait, nous conservons l'estimation robuste des plans dans le système. Les figures D.3 (page 159) et F.2 (page 168) présentent d'autres comparaisons entre les deux estimations. Les résultats sont tout à faits similaires à ceux de la figure 2.6.

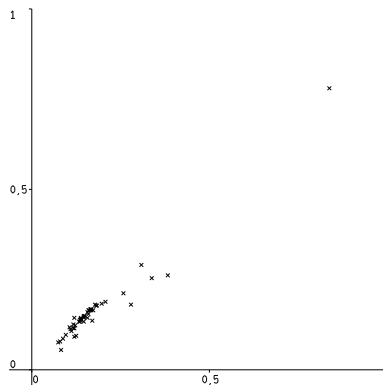


FIG. 2.6 – Comparaison des écarts quadratiques médians pour les plans de la zone 19.

Filtrage des résultats. Les segmentations pour les différentes valeurs de K produisent un grand nombre de régions, dont certaines sont égales, et d'autres proches. Nous proposons donc de filtrer ces régions, afin d'en limiter le nombre. Ce filtrage est toutefois strict, afin de ne pas perdre de région pertinente. Pas ailleurs, le filtrage utilisera un critère très physique, ce qui permet d'appréhender aisément son comportement.

Pour chacune des régions R, Q, \dots collectées au cours du processus de fusion de régions, et

pour toutes les valeurs de K citées plus haut, on associe les plans P_R, P_Q, \dots (l'approximation aux moindres carrés de I). Ce plan permet de plonger le contour de la région dans l'espace objet, ce qui permet alors de former un polygone 3D, plan, noté $pol(R), pol(Q), \dots$.

Le filtrage des régions opère itérativement, en supprimant à chaque itération, la région la plus propice au sens d'une fonction de coût. Il faut expliciter cette fonction. On considère donc qu'une région R peut être supprimée si une autre région Q la contient et lui est proche.

Il faut expliciter ces deux notions :

- R ne peut être supprimée que si une autre région la contient : on notera que par construction, pour deux régions R et Q de l'image I , soit R et Q sont disjointes, soit l'une des deux régions est incluse dans l'autre. Dans le cas où R et Q sont disjointes, rien ne justifie la suppression de R ou Q , car elles ne représentent pas la même information de la scène. Si $Q \subset R$, la suppression de R peut entraîner des 'trous' dans l'information représentée. Seul le cas $R \subset Q$ est donc potentiellement valable pour supprimer R , car dans ce cas Q , qui est maintenue, continue à représenter la région décrite par R .
- La notion de proximité entre R et Q doit être définie. Cette notion est en fait orientée : il s'agit de mesurer à quel point, l'approximation plane de R est proche (au voisinage de R), de l'approximation plane de Q . Pour cela, nous posons la fonction $dh_{RQ}(p) = |P_R(p) - P_Q(p)|$ qui mesure l'écart d'altitude entre les deux approximations P_R et P_Q pour le pixel p . Soit alors $D_{R,Q} = 1/|R| \sum_{p \in R} dh_{RQ}(p)$. Cette moyenne estime le volume contenu dans le cylindre porté par $pol(R)$, et compris entre les deux approximations planes P_R et P_Q , ce qui, ramené à la surface de R , revient à mesurer l'écart géométrique entre ces deux plans, pour la région R , exprimé en mètres. Parallèlement, cette fonction s'exprime en mètres, une grandeur physique adaptée au filtrage. Dans le cas où R n'est pas incluse dans Q , nous poserons $D_{R,Q} = \infty$

L'algorithme de filtrage est donc le suivant :

1. Calculer pour tout couple (R, Q) de l'ensemble $\{R, Q, \dots\}$ des régions, la valeur $D_{R,Q}$.
2. Identifier un couple de régions (R, Q) tel que $D_{R,Q} = \min_{R',Q'} D_{R',Q'}$.
3. Si $D_{R,Q} < S_{fp}$, alors supprimer R de l'ensemble $\{R, Q, \dots\}$ et aller en 1.

La valeur de S_{fp} est fixée à 10 cm, ce qui est un seuil faible, visant à ne supprimer effectivement que les portions de plans très voisines en 3D.

Enfin, nous supprimons des données les plans dont l'écart quadratique médian ajusté est supérieur à un seuil $S_{fpm} = 1m$, car ces plans sont très nettement sur-généralisés pour notre application (pour des écarts quadratiques de 1m, les plans obtenus sont en effet systématiquement des unions de pans de toits que nous souhaitons conserver).

3.6 Résultats

La procédure engendre des primitives qui peuvent se recouvrir. C'est pourquoi leur représentation dans ce document est difficile. Pour plus de lisibilité, nous représenterons donc ces ensembles de primitives en les éclatant en plusieurs sous-ensembles, tels que pour chacun de ces sous-ensembles, les primitives qui s'y trouvent ne se recouvrent pas. Le lecteur notera donc que les résultats de cette procédure sont représentés par *plusieurs* images dans une même figure.

La figure 2.7 présente le résultat de la collecte des régions pour le bâtiment 19. La figure 2.8 présente les plans après filtrage. Tous les pans de toit souhaités sont présents après filtrage. Des

résultats similaires sont présentés dans les figure F.4 et F.5, page 169.

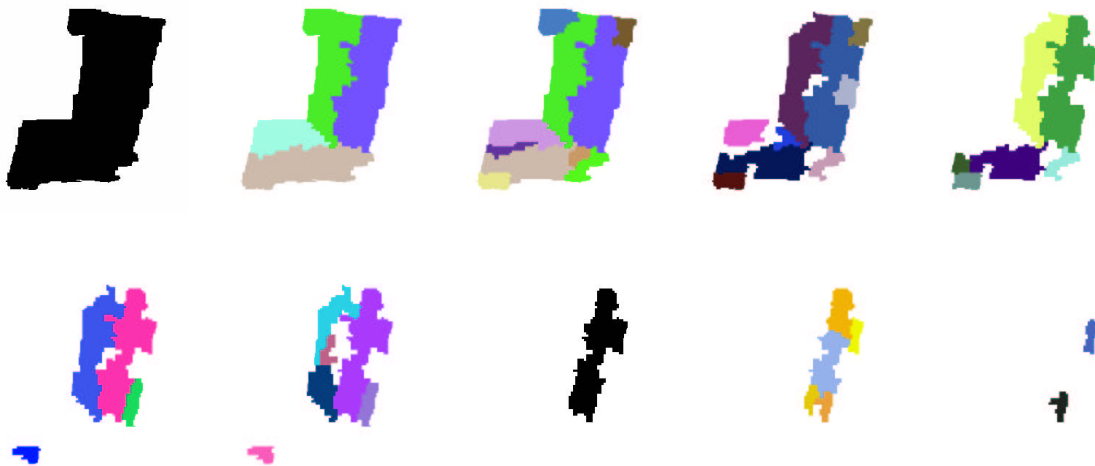


FIG. 2.7 – Les régions collectées pour le bâtiment 19.



FIG. 2.8 – Les régions collectées pour le bâtiment 19, après filtrage.

4 Les primitives linéaires

4.1 Définition

Les primitives linéaires sont des segments de droite de l'espace objet. Nous utiliserons des segments de droites correspondant aux gouttières, aux arêtes faîtières, aux arêtes de rives, etc. Les détecteurs mis en place ne permettront toutefois pas de reconnaître explicitement dans les images les objets qui nous intéressent, car ceci nécessiterait une interprétation d'ordre sémantique des primitives, ce qui est excessivement difficile à mettre en oeuvre. Il est donc déjà possible d'affirmer que ces détections souffriront de nombreuses sur-détections.

Les primitives linéaires en bord de toit se trouvent généralement en contraste avec la façade qui leur correspond, ou bien avec le sol. Dans ce cas, l'hypothèse d'un contraste radiométrique permettant de les détecter est plausible.

Dans le cas des arêtes faîtières, nous sommes en présence de deux portions de plans ne possédant pas la même normale. Dans ce cas, les différences d'apparence liées à l'éclairage valident également l'hypothèse d'un contraste radiométrique.

Toutefois, dans le cas particulier des images de Colombes, il est remarquable de constater qu'en raison des conditions de prise de vue (couvert nuageux, comme nous l'avons dit page 19), l'illumination de la scène est très diffuse, si bien que ce qui vient d'être énoncé n'est pas valable. Par contre, on constatera que de façon générale, les arêtes faîtières sont visibles car elles possèdent une microstructure au sein du toit, laquelle implique des contrastes détectables.

La figure 2.9 présente un exemple où le faîte est très peu visible, en raison de pans de toits de radiométries très voisines, et de l'absence de détails spécifiques au faîte. L'autre exemple de cette figure est un bâtiment pour lequel le faîte peut être détecté grâce à sa radiométrie propre. Enfin, dans le troisième exemple, le faîte est détectable, mais on peut raisonnablement estimer qu'il s'agit ici d'une surface de l'image, bordée de deux lignes de contraste, ce qui peut mener à une surdétection et à une délocalisation par rapport à la position «idéale» du faîte.

Dans la mesure où les primitives linéaires se manifestent donc par des contrastes, la méthode de calcul utilise principalement les images ⁶.

On notera que nous espérons détecter les objets recherchés par ces techniques, mais il est aussi évident que d'autres objets qui ne nous intéressent pas apparaîtront au cours du déroulement de cette méthode. Nous comptons alors sur les appariements modèle / données pour retrouver dans les données celles qui sont réellement pertinentes.

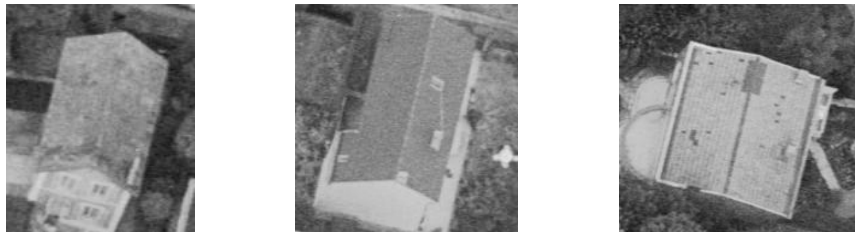


FIG. 2.9 – Exemples de faîtes. À gauche, un faîte non marqué, au centre, un faîte apparent, à droite, un faîte large.

4.2 Calcul

Dans la mesure où le MNE est le plus fiable sur les zones de contraste, nous disposons de l'information nécessaire pour déterminer la position 3-D des segments 2D. Il n'est notamment pas utile de mettre en oeuvre des méthodes générales de recherche de segments 3-D (comme celle présentée dans [SZ97]), mais une exploitation naturelle du MNE sera proposée.

Par ailleurs, nous tentons de tirer profit des deux images stéréoscopiques indépendamment. Ainsi, la méthode utilisée réalise premièrement une détection de segments de droites dans les images du

⁶Dans le cas de faîte non apparent, on peut envisager une voie de calcul par l'intermédiaire du MNE. Mais on peut dire que ces pistes reviennent à estimer d'abord les pans de toits voisins, puis à en déduire la position de l'arête. Cette technique est en fait naturellement mise en oeuvre dans le chapitre suivant, lorsque, en l'absence de primitive linéaire, celle-ci est reconstruite.

couple (segments 2D), puis un passage de ces segments dans l'espace objet.

Ceci présente toutefois un inconvénient : le risque de voir des doublons apparaître, ce qui nous pousse à poursuivre par une étape de fusion des segments 3-D. Par contre, l'avantage est d'exploiter la complémentarité des deux images : si une primitive n'est pas détectée dans une image, elle peut l'être dans l'autre.

La figure 2.10 présente la structure globale de la méthode de détection des primitives linéaires. Chacun des points de cette figure est décrit dans les sections suivantes.

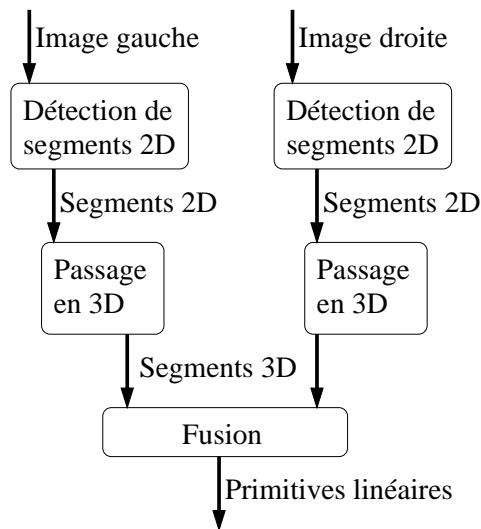


FIG. 2.10 – Méthode d'extraction des primitives linéaires

4.2.1 Calcul des primitives linéaires 2-D dans les images

La procédure de calcul des segments 2D utilise un détecteur de segments de droites. Le détecteur utilisé est celui de Lee et Kweon ([LK97]). Ce détecteur utilise une carte de contours produite par le détecteur de contour de Canny ([Der87]).

Il procède par une détection de contour, puis une suppression des maxima locaux du gradient, ainsi qu'un seuillage par hystérésis. Une fois ce traitement effectué, les directions du gradient sont discrétisées. En raison des variations locales des niveaux de gris, le processus de discrétisation des directions du gradient peut produire des aberrations locales. Un système d'optimisation tend alors à faire disparaître ces différences locales. Les pixels voisins de même direction discrète sont regroupés pour former des contours élémentaires. Ces contours élémentaires sont alors groupés suivant des techniques de groupement perceptuel : deux contours rectilignes élémentaires sont regroupés si ils ne se recouvrent pas, si ils sont suffisamment colinéaires, et si ils sont adjacents ou presque. Ces trois notions font appel à trois paramètres distincts, exprimés en mètres. Pour plus de détails, on se reportera à l'article de l'auteur ([LK97]).

Dans la pratique ce détecteur n'est utilisé qu'au voisinage de la zone de focalisation, comme le montrent les images 2.11 et 2.12 où l'on voit les primitives 2D extraites pour les zones 19 et 172.

4.2.2 Plongement des primitives 2-D en 3-D.

Le passage dans l'espace objet est fait avec l'aide du MNE. Le MNE permet de plonger l'objet en 3-D, mais aussi - et ce fait est important -, de détecter d'éventuels alignements fortuits de contrastes dans les images. Il est en effet probable de trouver des segments de droite alignés sur une image, alors qu'ils ne le sont pas en réalité. Ici on peut tirer profit de cette information 3D. En effet, il est possible de détecter des segments 2D qui correspondent à deux objets 3D différents, mais qui se projettent *fortuitement* sur la même droite d'une image. Un tel exemple est présenté en figure 2.13, où sont désignés deux points de l'image : sur le premier (au centre) on observe la gouttière d'un bâtiment dans l'alignement du faîte d'un autre, pour le second (en bas) on observe l'alignement d'une autre gouttière avec une structure au sol.

Dans ce cas, les deux objets n'ont généralement pas la même altitude. Grâce au MNE, on peut détecter ce fait et découper le segment de droite avant de le plonger en 3D.

En considérant la fonction $a(s)$ qui relie l'abscisse curviligne d'un point le long du segment, et son altitude, ces phénomènes de «décrochement» se traduisent par une discontinuité forte de la fonction a . Bien entendu, dans la pratique, nous discrétisons la fonction $a(s)$ avec un pas égal à la résolution de l'image. Ainsi, nous disposons d'une suite d'altitudes a_1, \dots, a_n . Nous introduisons un seuil s_a exprimé en mètres, correspondant à la différence caractéristique d'altitude entre deux bâtiments : 2m. Si pour $1 \leq i < n$, on a $|a_i - a_{i+1}| > s_a$, alors le segment de droite est coupé à la position i , et

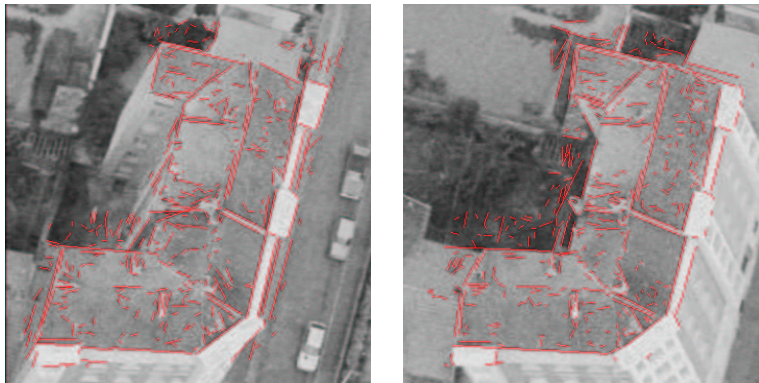


FIG. 2.11 – Primitives linéaires 2D extraites pour la zone 19.

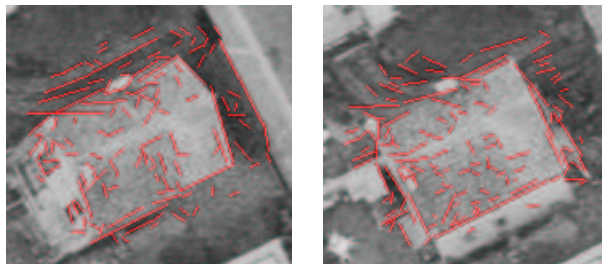


FIG. 2.12 – Primitives linéaires 2D extraites pour la zone 172.

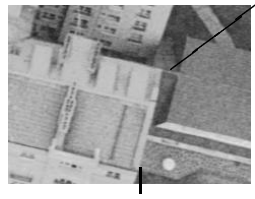


FIG. 2.13 – Exemples d’alignements 2D

le même traitement est répété jusqu’à ce qu’aucune discontinuité supérieure à a_s n’intervienne dans les segments ainsi obtenus.

Dans le même esprit, il aurait été possible de découper les segments de droite autour de leurs points de rupture de pente (ce type d’approche est mis en oeuvre dans des méthodes de segmentation d’images de profondeur par groupement de lignes (scan line grouping en anglais), comme le fait Jiang ([JB99])). Ce cas peut se produire lorsque deux arêtes de rive d’un bâtiment sont alignées fortuitement sur l’image. C’est le cas lorsque le sommet de prise de vue se trouve aux environs du plan portant ces deux arêtes. Un tel exemple se trouve dans l’image droite de la zone 172, page 28. Ce problème n’a toutefois pas été traité dans ce travail. Nous comptons sur la complémentarité des deux images pour produire dans ce cas les arêtes, car si elles sont alignées sur une image, elles ne le sont forcément pas sur l’autre, sauf dans un cas très particulier : lorsque la façade portant ces deux arêtes est parallèle à l’axe de vol (aux lignes épipolaires du couple stéréoscopique plus exactement), et quasiment dans le plan vertical passant par les deux centres de prise de vue du couple. Ce cas doit toutefois être très rare.

4.2.3 Fusion des Segments 3-D

Cette section décrit d’abord la manière dont deux segments 3-D sont fusionnés, puis la technique de fusion générale.

Fusion de deux segments

Soient s_i et s_j deux segments 3-D. Ces deux segments sont échantillonnés avec un pas typique de 10cm, ce qui produit deux ensembles de points 3-D, que nous réunissons en un ensemble E . Soit alors G le centre de gravité de E . Les points de E sont alors centrés autour de G (pour un point p on considère le vecteur $v_p = p - G$). Pour l’ensemble des vecteurs ainsi obtenus, soit v un vecteur normé correspondant à l’axe principal d’inertie de l’ensemble des vecteurs ⁷. Soient alors x_{min} (resp. x_{max}) la valeur minimale (resp. maximale) des produits scalaires $v_p \cdot v$. Cette opération consiste à rechercher les points de E dont la projection le long de la droite portée par v et passant par G sont les plus éloignées. Le segment fusionné est alors donné par $G + x_{min} * v$ et $G + x_{max} * v$. Ce mécanisme est explicité dans la figure 2.14, où l’on voit les deux segments initiaux échantillonnés, le centre de gravité G , et le résultat de la méthode.

Fusion des segments

La méthode de fusion est itérative. Elle consiste à fusionner en premier les couples de segments les plus «alignés». Pour cela, soit D_{ij} la droite porteuse de l’ACP de deux segments s_i et s_j , et

⁷Ce vecteur peut être obtenu par une analyse en composantes principales (ACP)

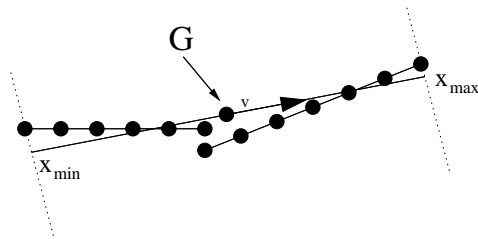


FIG. 2.14 – Fusion de deux segments 3-D

$\sigma_{i,j}$ l'écart quadratique moyen des distances des points d'échantillonnage des deux segments à cette droite.

Les deux segments dont $\sigma_{i,j}$ est minimal sont fusionnés en premier. La fusion s'arrête lorsque pour tout couple de segments, l'écart-type ainsi défini est supérieur à un seuil d'arrêt σ_f . σ_f est fixé à 1m.

Après cette fusion, un recalage des segments 3D est réalisé. L'outil utilisé a été proposé par Jibrini (voir [Jib00]).

4.2.4 Résumé des paramètres

Les processus de détection des segments de droites dans les images comporte donc 8 paramètres :

- Le paramètre du filtre de Canny : α_i
- Les seuils d'hystérésis : H_1, H_2 .
- Le pas de discrétisation des segments : p_s
- Les 3 paramètres de groupement.
- Le seuil de fusion σ_f .

4.3 Résultats

La figure 2.15 présente deux vues 3-D des primitives linéaires détectées pour la zone 172. On constate une bonne détection des gouttières, des deux arêtes de rives au nord-est du bâtiment, mais une absence de faîte et des arêtes de rive sud-ouest. Enfin, de nombreuses primitives présentes ne correspondent à aucun détail réel du bâtiment (comme annoncé en section 4.1), mais sont issues de fusions d'objets plus petits, et alignés par hasard.

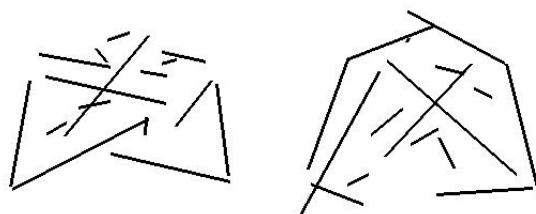


FIG. 2.15 – Vues 3-D des primitives linéaires pour la zone 172. Gauche : vue nord-est, droite : vue sud-ouest.

5 Les discontinuités du MNE

5.1 Définition

On appellera *discontinuité* une portion de plan comprise entre une gouttière ou une arête de rive et les objets situés sous cette gouttière (principalement le sol), comme le décrit la figure 2.16.

Ce type de primitive ne correspond donc pas à des détails tangibles de la zone, car ces portions de plans n'ont généralement pas de réalité physique. Elles ne correspondent notamment pas toujours aux façades des bâtiments. Par contre, en supposant que le MNE décrit la surface topographique sans délocalisation, ces objets se trouvent sur les discontinuités d'altitude détectables dans le MNE (ce qui n'est pas le cas des façades).

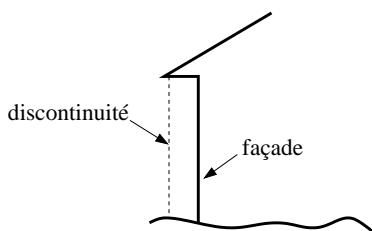


FIG. 2.16 – Définition des discontinuités

5.2 Pourquoi pas une polygonisation du contour ?

On pourrait estimer que ces primitives pourraient être détectées par une polygonisation du contour de la zone de focalisation. Il est vrai que ceci fournirait une grande partie des primitives que nous recherchons mais pas toutes.

La zone 11 est en effet un exemple type de deux bâtiments simples d'altitudes voisines, qui ont été groupés dans une même zone de focalisation. L'intérêt de détecter les discontinuités du MNE est donc de pouvoir déceler ce fait.

Dans ce cas particulier, une polygonisation du contour de la zone aurait nécessairement laissé échapper cette discontinuité particulière.

Nous pensons que cet exemple est suffisant pour montrer l'intérêt de ne pas se limiter à une polygonisation du contour de la zone.

Enfin, il serait tentant d'orienter la notion de discontinuité (la discontinuité sépare son voisinage en deux régions, dont l'une est d'altitude plus élevée que l'autre), ce qui semblerait utile pour la reconstruction du bâti : un bâtiment se trouve du côté le plus élevé de ses discontinuités. Pourtant, dans le cas de deux bâtiments adjacents, la discontinuité qui les sépare est telle que l'un des bâtiments est situé du côté le plus élevé, mais l'autre est situé du côté le plus bas. Un exemple typique de ce cas est présenté dans la zone 17 (page 167), pour lequel une discontinuité sépare les deux bâtiments rectangulaires de la zone. Par rapport à cette discontinuité, le bâtiment nord se trouve plus bas en altitude, alors que le bâtiment sud se trouve plus haut. Nous n'orientons donc pas les discontinuités.

5.3 Calcul

5.3.1 Méthode

Les discontinuités sont extraites du MNE. Leur détection est réalisée avec le même détecteur que la détection des segments de droites dans les images stéréoscopiques. Les paramètres sont toutefois adaptés à ce cas particulier.

5.3.2 Résultats

La figure 2.17 présente les discontinuités de la zone 172. La figure F.3, page 168 présente les discontinuités de la zone 17.

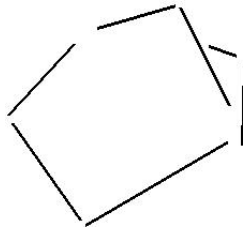


FIG. 2.17 – Les discontinuités détectées pour la zone 172.

6 Les primitives ponctuelles

6.1 Définition

Les primitives ponctuelles utilisées dans ces travaux sont les intersections entre les primitives linéaires décrites plus haut.

6.2 Utilisation

Les primitives ponctuelles sont utilisées dans les modèles. Elles possèdent des relations avec les autres types de primitives. La seule relation utilisée est la relation d'appartenance : un sommet appartient à un segment 3D, une portion de plan, ou une discontinuité.

Il existe dans la littérature en grand nombre de détecteurs de points (dits points d'intérêt : Förstner ([FG87]), Harris [HS88]), ou des détecteurs plus récents ([MS98]) et des études comparatives existent ([Sch96]). Dans notre contexte cartographique, ces points d'intérêt sont généralement utilisés à des fins d'automatisation de l'aérotriangulation ([BFJW01]), ou d'indexation d'images, mais dans le cas de calculs de sommets 3-D, les points utilisés doivent être confortés par des calculs de droites issues de ces points ([Ste97]). Dans la mesure où nous effectuons dans tous les cas les extractions de primitives linéaires, ainsi que des extractions de primitives planes et discontinuités, qui, combinées, permettront (ce point sera justifié a posteriori) le placement des primitives ponctuelles en 3D, nous n'effectuerons pas de détection de ces primitives ponctuelles.

Apportons par ailleurs d'autres justifications (moins fondamentales) à cette absence de détection de primitives ponctuelles :

- L'ajout de ce type de primitive augmenterait la combinatoire du problème d'appariement du chapitre suivant.
- D'un point de vue tout à fait pragmatique, il n'a pas été possible dans le temps imparti, d'implanter dans le système un quatrième détecteur de primitives.

Pour résumer, ces primitives ne sont pas extraites des données. Leur position géométrique est calculée après l'appariement avec des modèles, dans le but de donner une description géométrique complète des objets reconstruits.

7 Conclusion

Ce chapitre a présenté des méthodes de calcul visant à extraire, à partir des données initiales, 3 types d'objets géométriques, qui contribueront à la reconstruction du bâti dans les chapitres suivants.

Ces détections d'objets simples sont indépendantes, afin de ne pas propager d'erreur entre elles. Dans le cas des primitives planes, nous utilisons plusieurs segmentations du modèle numérique d'élévation, afin d'en fusionner les résultats. Ceci a pour objectif de ne pas souffrir de sous-détections pour ces primitives bi-dimensionnelles, très structurantes dans la reconstruction du bâti.

Dans le cas des primitives linéaires, nous avons tenté de tirer profit des deux images stéréoscopiques et du MNE, ce qui peut se résumer par le terme de «coopération 2D - 3D», puisque la détection se fait par l'exploitation des images pour produire des segments 2D, puis par l'exploitation de l'information 3D du MNE pour plonger ces segments 2D en 3D.

Les discontinuités, qui ne sont pas les façades, ne peuvent pas être détectées dans les images. Elles sont donc détectées dans le MNE. Ce MNE est réputé pour sa fiabilité et constitue donc une donnée fiable pour ce type de détection.

La contribution majeure de ce chapitre est probablement la détection des primitives planes, puisque nous sommes dans ce cas en contraste net avec de nombreux auteurs, qui, généralement, se contentent d'une seule segmentation, alors que nous exploitons simultanément une batterie de segmentations du MNE. La seconde contribution est probablement le fait que les détections sont indépendantes, alors que certains auteurs effectuent des détections successives s'appuyant les unes sur les autres.

Chapitre 3

Appariement de graphes

Ce chapitre présente la seconde étape de l’approche structurale proposée. Il s’agit, pour un modèle *donné* et un ensemble de primitives calculé dans la zone, *d’identifier*, dans les données, les primitives principales qui décrivent la forme recherchée, puis d’en déduire une reconstruction.

Ce problème d’identification est posé dans le cadre théorique des graphes, qui est, d’une part très adapté à la représentation des objets que nous manipulons (des objets 3-D accompagnés de propriétés géométriques qui les relient), et qui a, d’autre part, été porteur récemment de solutions nouvelles et efficaces pour résoudre le problème de l’identification lui-même.

Des notations et certaines notions de base sont donc présentées en section 1. La section 2 présentera une vue d’ensemble des problèmes d’appariement de graphes. La section 3 justifiera certains choix faits dans ces travaux. La section 4 présentera la méthode de détection d’isomorphisme de sous-graphe avec tolérance d’erreur, proposée par Messmer ([Mes96, MB98]), qui est au coeur de nos travaux, et qu’il est nécessaire d’explicitier ici. La section 5 présentera notre application des travaux de Messmer au cas de la reconstruction du bâti, ainsi que certaines extensions de sa méthode. Ces extensions seront utilisées largement dans les chapitres suivants.

1 Notations

Ces notations sont proches de celles utilisées par Messmer ([Mes96]), mis à part le fait que les arcs sont des hyper-arcs, ce qui fait de ces graphes des hyper-graphes. Elles rappellent certains éléments concernant les notions de graphes. Ces notions seront largement utilisées par la suite.

1.1 Graphe, ou plutôt, hyper-graphe.

1.1.1 Définition

Nous noterons un graphe $G = (V, E, \mu, \nu, L_V, L_E)$, où :

- V est l’ensemble des sommets v_i du graphe.
- E est l’ensemble des arcs (ou plutôt hyper-arcs) du graphe. Un arc est donné par un ensemble de sommets initiaux (les sources) et de sommets finaux (les cibles). On peut donc le noter $e = (S, T)$ où $S = \{v_1, \dots, v_m\}$ et $T = \{w_1, \dots, w_n\}$. m est le nombre de sources, n est le

- nombre de cibles. La valeur $n + m$ est l'arité de l'arc. Nous n'utiliserons que des arcs binaires ($m = 1, n = 1$), ou ternaires ($m + n = 3$).
- L_V est l'ensemble des labels des sommets.
 - L_E est l'ensemble des labels des arcs.
 - $\mu : V \rightarrow L_V$ est la fonction d'attribution des labels des sommets.
 - $\nu : E \rightarrow L_E$ est la fonction d'attribution des labels des arcs.

1.1.2 Remarques

Cette définition formelle nécessite quelques remarques informelles :

- Certaines relations géométriques ne peuvent pas être modélisées par des relations binaires, mais seulement par des relations ternaires. Par exemple, le faite d'un bâtiment à deux pans se trouve «entre» les deux gouttières qui lui sont parallèles. On pourrait être tenté de dire qu'une des gouttières se trouve d'un côté du faite, et l'autre de l'autre côté, mais ceci nécessite une relation d'ordre sur les droites du plan. Or il n'y a pas d'ordre naturel sur les droites du plan. La relation en question n'est donc pas modélisable par un arc binaire, d'où la nécessité, si le système doit utiliser cette relation, d'introduire les relations ternaires. Or, il s'avère que les relations de type «tel objet est entre tels objets» s'avèreront utiles (un exemple visant à en montrer l'intérêt est présenté en section 3.2.3, page 98).
- Dans ce document, nous n'utiliserons pas le terme de «nœud» pour désigner les sommets d'un graphe, car le terme «nœud» sera utilisé pour désigner un autre élément, en section 4.2 de ce chapitre.
- Ces graphes sont des hyper-graphes. Il n'y a pas de définition unifiée des hyper-graphes. Le lecteur pourra notamment se reporter à [GM95], page 29 pour une définition ensembliste un peu moins générale des hyper-graphes, dans laquelle la distinction entre sources et cibles n'est pas faite, ou alors à [Hab92], pour une définition plus générale, dans laquelle les répétitions de sources et de cibles sont autorisées.
- A l'exception des hyper-arcs mis en place ici, cette définition est très proche de celle proposée par Messmer ([Mes96]). Ainsi, le lecteur n'aura aucune difficulté à faire le lien entre ce document et les travaux de Messmer sur lesquels il s'appuie.
- S'il n'y a pas d'ambiguïté, un graphe sera aussi noté $G = (V, E, \mu, \nu)$, voire $G = (V, E)$ ou même G .
- Le nombre de sommets d'un graphe est appelé son ordre. On le notera $|G| = |V|$. S'il vaut 1, le graphe est un singleton.
- Nous noterons aussi $V(G)$ l'ensemble des sommets d'un graphe G .
- Si deux sommets v et v' sont voisins (s'ils sont extrémités d'un même arc), nous noterons $v \leftrightarrow v'$.

1.1.3 Les vues d'un graphe modèle.

Nous proposerons trois façons de voir les graphes modèles.

- La vue théorique : il s'agit d'une représentation symbolique dans laquelle apparaissent les sommets et les arcs du graphe. Sur les vues théoriques, les sommets sont représentés par des cercles numérotés, en haut à droite desquels les attributs sont précisés. Les attributs des arcs

ne sont pas représentés.

- La vue 3-D : il s'agit d'une représentation d'une forme réelle correspondant au modèle, depuis un point de vue donné. C'est une vue perspective. Dans ces vues, les primitives de type discontinuité ne sont pas représentées, mais la perception de la forme du toit est très bonne.
- La vue 2-D : il s'agit d'une représentation de la projection des primitives au sol. Chaque type de primitive possède une représentation particulière : les segments de droite sont en noir, les discontinuités en hâchures, les sommets ponctuels sont représentés par un disque noir, et les primitives surfaciques sont représentées par une surface grisée.

La figure 3.1 présente les trois vues décrites ci-dessus pour un bâtiment à double pan, appelé modèle D. D'autres vues d'autres modèles sont présentées en figure 4.8, page 97.

La figure 3.2 présente un graphe possédant 6 sommets, 8 arcs simples, et 2 hyper-arcs. Nous appellerons ce graphe le graphe J ¹.

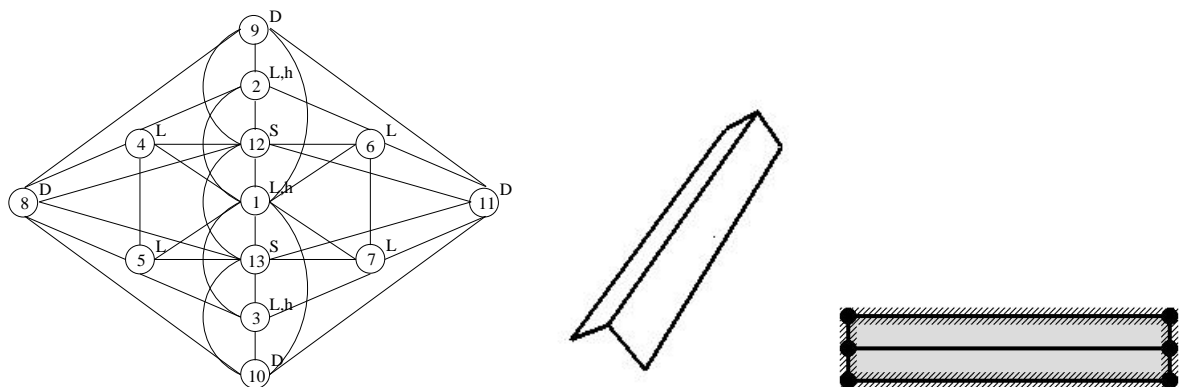


FIG. 3.1 – Les 3 vues d'un modèle. De gauche à droite : vue théorique, vue 3-D, vue 2-D. Pour la vue théorique, les sommets ponctuels ne sont pas représentés.

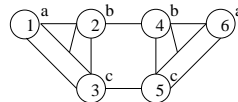


FIG. 3.2 – Le graphe Jouet. Ce graphe porte deux hyper-arcs.

1.2 Sous-graphe

Pour un graphe $G = (V, E)$, et un sous-ensemble de ses sommets $V_S \subset V$, le sous-graphe de G s'appuyant sur ces sommets est défini par : $G_{V_S} = (V_S, E_S, \mu_S, \nu_S, L_V, L_E)$ où :

- $E_S = E \cap (V_S \times V_S)$
- $\mu_S : V_S \rightarrow L_V$ est la restriction de μ à V_S (pour tout v de V_S , $\mu_S(v) = \mu(v)$).
- $\nu_S : E_S \rightarrow L_E$ est la restriction de ν à E_S .

¹pour «jouet» car ce graphe ne correspond à aucun des graphes vraiment utilisés dans ces travaux.

Si G' est un sous-graphe de G alors nous noterons : $G' \subseteq G$. Lorsque G' n'est, de plus, pas égal à G , nous noterons $G' \subset G$.

1.3 Différence

La différence entre un graphe G et l'un de ses sous-graphes G_{V_S} ($V_S \subset V$) est le sous-graphe de G s'appuyant sur le complémentaire de V_S dans V : $G - G_{V_S} = G_{V \setminus V_S}$.

1.4 Union

Pour deux graphes $G_1 = (V_1, E_1, \mu_1, \nu_1, L_V, L_E)$ et $G_2 = (V_2, E_2, \mu_2, \nu_2, L_V, L_E)$, et un ensemble d'arcs $E' \subset (\mathcal{P}(V_1) \times \mathcal{P}(V_2)) \cup (\mathcal{P}(V_2) \times \mathcal{P}(V_1))$, ainsi qu'une fonction d'attribution de labels à ces arcs $\nu' : E' \rightarrow L_E$, l'union de G_1 et G_2 relativement à E' est le graphe $G = (V, E, \mu, \nu, L_V, L_E)$, que l'on notera $G = G_1 \cup_E G_2$ tel que :

- $V = V_1 \cup V_2$: V est l'union *disjointe* des ensembles de sommets de G_1 et G_2 .
- $E = E_1 \cup E_2 \cup E'$
- $\mu(v) = \begin{cases} \mu_1(v) & \text{si } v \in V_1 \\ \mu_2(v) & \text{si } v \in V_2 \end{cases}$
- $\nu(e) = \begin{cases} \nu_1(e) & \text{si } e \in E_1 \\ \nu_2(e) & \text{si } e \in E_2 \\ \nu'(e) & \text{si } e \in E' \end{cases}$

L'union de deux graphes produit un graphe dont l'ordre est égal à la somme des ordres des deux graphes de départ, car l'union des ensembles de sommets est disjointe. Nous avons : $|G \cup G'| = |G| + |G'|$. Cette remarque aura un intérêt dans la section 7.4, où nous autoriserons des unions non disjointes dans certaines circonstances.

1.5 Isomorphisme

Pour deux graphes $G = (V, E, \mu, \nu)$ et $G' = (V', E', \mu', \nu')$ et une fonction $f : V \rightarrow V'$, *bijective*, on définit $\mathcal{F} : \mathcal{P}(V) \rightarrow \mathcal{P}(V')$ l'extension de f aux ensembles de sommets : $\mathcal{F}(\{v_1, \dots, v_k\}) = \{f(v_1), \dots, f(v_k)\}$. f est un isomorphisme entre G et G' si et seulement si :

- Pour tout $v \in V$, $\mu(v) = \mu'(f(v))$: la fonction respecte les labels des sommets.
- Pour tout arc $e = (S, T) \in E$, il existe un arc $e' = (\mathcal{F}(S), \mathcal{F}(T)) \in E'$ tel que $\nu(e) = \nu'(e')$, et inversement, pour tout arc $e' = (S', T') \in E'$, il y a un arc $e = (\mathcal{F}^{-1}(S'), \mathcal{F}^{-1}(T'))$ tel que $\nu'(e') = \nu(e)$: la fonction f respecte la topologie du graphe, et les labels des arcs.

On notera que le graphe J est composé de deux sous-graphes isomorphes (les sous-graphes composés respectivement des triplets de sommets (1,2,3) et (4,5,6)).

Automorphisme

Un automorphisme de graphe est un isomorphisme entre un graphe et lui-même. Il existe toujours un automorphisme trivial (l'identité), mais la présence d'automorphismes non triviaux révèle des symétries internes des graphes. Nous ne ferons pas l'étude des automorphismes d'un graphe, mais cette notion sera toutefois utilisée.

Le graphe J possède deux automorphismes, et le graphe modèle D (voir figure 3.1) en possède 4.

1.6 Isomorphisme de sous-graphe

Un isomorphisme de sous-graphe entre un graphe G et un graphe G' est une fonction $f : V \mapsto V'$ injective, telle que G est isomorphe au sous-graphe de G' s'appuyant sur $\mathcal{F}(V)$. On notera que cette notion est orientée. Nous avons notamment : $|G| \leq |G'|$.

L'isomorphisme de graphe est un cas particulier d'isomorphisme de sous-graphe.

2 Généralités.

L'appariement (ou mise en correspondance) de graphes est un problème théorique général, qui se décline principalement en deux catégories.

2.1 Isomorphisme de graphes et de sous-graphes

La première catégorie est le calcul d'isomorphismes de graphes, ou d'isomorphismes de sous-graphes. L'isomorphisme de graphe est le problème de la détermination d'une fonction f réalisant cet isomorphisme². Dans ce domaine, on citera uniquement l'algorithme d'Ullman ([Ull76]), qui fut le point de départ des méthodes visant à résoudre ce type de problèmes. On notera que l'algorithme d'Ullman est capable de déterminer l'ensemble des automorphismes d'un graphe.

L'isomorphisme de sous-graphe consiste à rechercher pour G et G' tels que $|V| \leq |V'|$, un sous-graphe de G' isomorphe à G . Ceci revient à trouver une fonction $f : V \rightarrow V'$ satisfaisant les propriétés d'isomorphisme de sous-graphe. L'algorithme d'Ullman s'applique également dans ce cas. Il est capable d'énumérer tous les isomorphismes de sous-graphe entre G et G' , et, de façon équivalente, tous les sous-graphes de G' isomorphes à G .

Remarque

Le fait que nous employons ici des hyper-graphes n'est en aucun cas gênant pour les considérations évoquées ici. Le seul obstacle à résoudre est l'implémentation informatique des hyper-arcs, que nous ne détaillerons pas³.

2.2 Isomorphisme de sous-graphe avec tolérance d'erreur

Cette seconde famille de problèmes est plus vaste. L'isomorphisme de sous-graphe est notamment un cas particulier de cette famille. Etant donné une fonction de coût $F(.,.)$, qui, à deux graphes associe une distance, le problème est de trouver, pour deux graphes G_1, G_2 le sous-graphe de G_2 qui est à distance minimale de G_1 (ou les sous-graphes si la solution n'est pas unique). Lorsque le sous-graphe G solution est trouvé, on peut exprimer que G_1 est isomorphe à ce sous-graphe si l'on tolère une erreur valant $F(G_1, G)$.

²On notera au passage que le problème de savoir si deux graphes G et G' sont isomorphes est très différent de celui de savoir si deux graphes ne sont pas isomorphes. En effet, dans le premier cas, il 'suffit' de trouver une fonction f et de vérifier qu'elle réalise l'isomorphisme, alors que dans le second cas, il faut prouver qu'aucune fonction ne convient. Ceci peut par exemple être réalisé en trouvant une mesure M sur les graphes, invariante par isomorphisme, et en montrant que $M(G) \neq M(G')$.

³D'autant plus que nous n'utilisons en réalité que des hyper-arcs possédant 2 sources et une cible, ce qui rend possible une implémentation ad hoc, plus simple qu'une implémentation générale.

Cette opération s'apparente à une projection de G_1 sur G_2 , comme l'exprime la figure 3.3. Considérons que l'attribut d'un arc est égal à la distance entre les nœuds joints par cet arc, alors sur cette figure, on voit que le graphe de gauche se projette sur les 4 sommets les plus en bas du graphe de droite, mais cette projection implique une faible distorsion, qui se traduit par un certain coût, qui est la distance entre ces deux graphes.

Le paradigme de l'appariement proposé dans ce chapitre repose sur cette projection : en considérant que le graphe modèle est petit par rapport au graphe des données, il se projette sur ce dernier, en minimisant une fonction dépendant de l'application. L'isomorphisme de graphe produit par ce calcul permet d'identifier, dans les données, les primitives pertinentes pour la reconstruction effective du modèle en 3-D : le modèle se voit attribuer une géométrie et une réalité physique.

Dans la suite de ce chapitre, le terme « appariement » sera utilisé pour désigner le calcul d'isomorphismes de sous-graphe avec tolérance d'erreur.

Nous avons choisi d'utiliser le formalisme des distances d'édits car il s'accompagne d'un certain nombre de progrès récents liés aux calculs d'isomorphismes de sous-graphe avec tolérance d'erreur. En effet, [MB98] présente des méthodes efficaces pour résoudre ces problèmes ⁴.

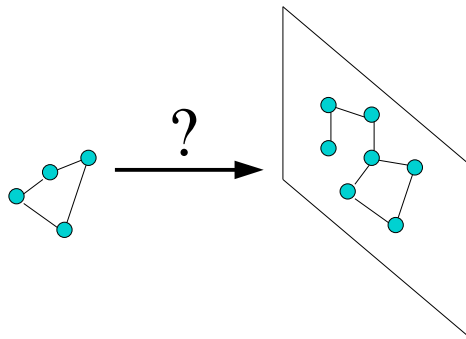


FIG. 3.3 – L'isomorphisme de sous-graphe avec tolérance d'erreur s'apparente à une projection.

Bibliographie

La littérature concernant cette problématique est assez vaste. On trouve notamment les travaux de l'équipe de Hancock ([WH97, WCH98, CH98, FWH98, MWH00]) qui proposent de nombreuses techniques d'appariement souple de graphes. Ses applications principales sont le recalage de graphe routier (vecteur / image) et le recalage image / image. Par ailleurs, une approche d'appariement de

⁴Messmer propose également une méthode de calcul d'isomorphisme de sous-graphe sans tolérance d'erreur très rapide, et qui s'appuie sur une étape de pré-traitement de l'ensemble des modèles considérés. Elle a été adaptée au cas de la tolérance d'erreurs dans le cas de valeurs discrètes des coûts d'édition ([MB96]). Dans notre cas, où les valeurs des coûts d'édition seront flottantes, il serait nécessaire de discrétiser les valeurs, ce qui impose le choix d'un pas d'échantillonnage d . Nous pourrions alors discrétiser les valeurs suivant d , ce qui produirait alors les graphes édités pour les valeurs $0, d, 2d, 3d, 4d$, etc. Il se trouve que le coût de stockage de la méthode proposée est exponentiel en fonction du nombre de valeurs considérées. Dans notre cas, un minimum d'une cinquantaine de valeurs serait nécessaire (car des différences de coût d'édition de $0,1$ peuvent être considérées comme significatives, et, par ailleurs, certaines valeurs peuvent s'élever à plus de 5 , d'où ce minimum de 50 valeurs nécessaires). Il est donc clair que cette approche n'est pas adaptée.

graphes est proposée dans [MM96], mais en milieu urbain. Un graphe modèle est extrait de cartes vectorisées, puis recalé sur des images aériennes ou satellitaires après une détection d'informations sur ces images.

De même, C. Hivernat ([HDRZ98]) a utilisé ces outils dans un contexte routier.

On trouve dans [KK00] une utilisation en reconnaissance de caractères. Enfin, dans le domaine de l'analyse de plans architecturaux, les travaux de Dosch ([Dos00]) proposent également une approche par appariement de graphes visant à recalé entre eux différents plans des étages d'un même bâtiment.

3 Le cas du bâti.

Dans le cas du bâti, et malgré la diversité des formes de bâtiments existantes, il est indéniable que les bâtiments possèdent des sous-parties communes. De même, un bâtiment seul peut fréquemment posséder deux sous-parties de formes identiques. Dans ce cas, les méthodes proposées par Messmer se montrent adaptées car elles permettent de tirer profit de cette situation pour réduire la combinatoire du problème de l'appariement. C'est la raison pour laquelle le graphe J (figure 3.2) possède une forte symétrie interne : il servira à montrer l'impact de cette symétrie sur les méthodes employées.

Par ailleurs, la méthode choisie est déterministe, ce qui permet de découpler l'analyse de la validité de la fonction de coût utilisée dans l'appariement, de l'analyse de la façon dont cette fonction est minimisée. En effet, avec une procédure déterministe, le minimum est sûr d'être atteint, ce qui permet, en cas de doute sur les résultats, de mettre en cause la fonction de coût et non le processus de minimisation.

Par contre, et nous le verrons plus tard, le caractère déterministe provoque des explosions combinatoires importantes, que nous tenterons successivement de contrer.

4 Méthode standard

Il est nécessaire de rappeler au lecteur l'essentiel de la méthode proposée par Messmer car la suite des travaux s'appuie sensiblement sur ceux-ci : nous avons en effet adapté et étendu un certain nombre d'outils proposés par Messmer à notre problème.

4.1 Formalisme

Nous définissons les termes suivants :

Opération d'édition

Une opération d'édition est une opération élémentaire sur un graphe. Elle affecte soit un sommet, un arc, ou un label de sommet ou d'arc. 5 opérations d'édition sont prévues par la théorie :

- Substitution de label de sommet : l'opération $\mu(v) \rightarrow l, v \in V, l \in L_V$ consiste à remplacer le label du sommet v , $\mu(v)$, par la valeur l .
- Substitution de label d'arc : l'opération $\nu(e) \rightarrow l', e \in E, l' \in L_E$ consiste à remplacer le label de l'arc e , $\nu(e)$, par le label l' .
- Suppression de sommet : $v \rightarrow \$$ consiste à supprimer le sommet v et ses arcs incidents du graphe.

- Suppression d'arc : $e \rightarrow \$$ consiste à supprimer un arc e du graphe.
- Ajout de d'arc : $\$ \rightarrow e = (v_1, v_2), v_1, v_2 \in V$: consiste à insérer un arc e entre v_1 et v_2 .

Graphe édité

Pour un graphe $G = (V, E, \mu, \nu)$ et une opération d'édition δ , on définit le graphe édité $\delta(G) = (V_\delta, E_\delta, \mu_\delta, \nu_\delta)$ de la manière suivante :

$$\begin{aligned}
 - V_\delta &= \begin{cases} V - \{v\} & \text{si } \delta = (v \rightarrow \$) \\ V & \text{sinon} \end{cases} \\
 - E_\delta &= \begin{cases} E \cup \{e\} & \text{si } \delta = (\$ \rightarrow e) \\ E - \{e\} & \text{si } \delta = (e \rightarrow \$) \\ E \cap (V_\delta \times V_\delta) & \text{sinon} \end{cases} \\
 - \mu_\delta(v) &= \begin{cases} l & \text{si } \delta = (\mu(v) \rightarrow l) \\ \mu(v) & \text{sinon} \end{cases} \\
 - \nu_\delta(e) &= \begin{cases} l' & \text{si } \delta = (\nu(e) \rightarrow l') \\ \nu(e) & \text{sinon} \end{cases}
 \end{aligned}$$

Séquence d'édits

Une séquence d'édits est un k -uplet $\Delta = (\delta_1, \dots, \delta_k)$ d'opérations d'édits. Pour un graphe G , le graphe édité $\Delta(G)$ correspondant est $\Delta(G) = \delta_k \circ \delta_{k-1} \cdots \circ \delta_1(G)$.

Coût d'édition

On définit pour une opération d'édition δ un coût $C(\delta)$, et pour une séquence d'édits $\Delta = (\delta_1, \dots, \delta_k)$, le coût de la séquence est la somme des coûts individuels :

$$C(\Delta) = \sum_{i=1}^k C(\delta_i) \tag{3.1}$$

Les coûts individuels dépendent de l'application, et seront définis en section 5.

Isomorphisme avec correction d'erreur

Un isomorphisme de sous-graphe avec correction d'erreur entre deux graphes G et G' est un couple (Δ, f_Δ) où Δ est une séquence d'édition telle qu'il y ait un isomorphisme de sous-graphe entre $\Delta(G)$ et G' , et tel que f_Δ est l'un de ces isomorphismes de sous-graphe. Cette notion est orientée.

On notera $ICE(G, G')$ l'ensemble des isomorphismes de graphes avec correction d'erreur liant G et G' .

On notera que pour deux graphes G et G' il existe au moins un isomorphisme de sous-graphe avec correction d'erreur, puisqu'il existe une séquence d'édition triviale qui consiste à supprimer un

à un les sommets de G , ce qui produit un graphe vide, qui est un sous-graphe de G' . $ICE(G, G')$ est donc toujours non vide. Plus exactement, $ICE(G, G')$ est même infiniment grand ⁵.

Distance d'édition

Pour G et G' , la distance d'édition entre deux graphes est le minimum des coûts des isomorphismes de sous-graphe avec correction d'erreur entre G et G' : $d(G, G') = \min_{(\Delta, f_\Delta) \in ICE(G, G')} C(\Delta)$.

La distance d'édition entre deux graphes définit donc un ensemble d'isomorphismes de graphe avec tolérance d'erreur «idéaux», qui permettent d'identifier, dans G' , le ou les sous-graphes les plus proches de G . Dans notre application, ce sous-graphe identifié par l'isomorphisme de coût minimum représente le lot de primitives que nous espérons être le plus adapté pour représenter le modèle G .

Nous cherchons donc, pour un graphe G représentant un modèle (un forme de bâtiment), et un graphe G' construit à partir des primitives (nous verrons comment plus tard), un de ces couples (Δ, f_Δ) de coût minimum :

$$(\Delta, f_\Delta) \in \operatorname{argmin}_{(\Delta', f_{\Delta'}) \in ICE(G, G')} C(\Delta')$$

La figure 3.4 explicite le mécanisme d'isomorphisme de sous-graphe avec correction d'erreur. Considérons comme modèle G le premier des graphes de la figure 1.17 page 23, et comme graphe G' le graphe du bas de la figure 3.4. Considérons enfin que tout changement de label d'arc ou de sommet a un coût unitaire, alors la distance d'édition de G à G' vaut 3, comme l'explique cette figure : 3 éditions successives sont nécessaires pour produire le graphe en bas à droite de cette figure, qui lui-même est isomorphe à un sous-graphe de G' . L'isomorphisme en question est identifié par les flèches pointillées.

4.2 Décomposition d'un graphe

4.2.1 Présentation

La méthode standard propose de décomposer récursivement un graphe en sous-graphes, en partageant éventuellement des sous-graphes isomorphes. Ce partage des sous-graphes isomorphes permet de réduire la combinatoire de l'appariement. Cette partie du document n'est pas nouvelle, car il s'agit des travaux de Messmer. Mais il est nécessaire pour la compréhension de la suite du document, de rappeler les structures de données utilisées dans cette méthode, ainsi que l'algorithme d'appariement associé, car l'un et l'autre feront l'objet de modifications sensibles dans notre cas particulier.

Une *décomposition* est un ensemble de *nœuds* (G, G', G'', E) tels que :

- G, G' et G'' sont des graphes, tels que $G' \subset G$ et $G'' \subset G$. G est appelé le graphe principal du nœud, et G' et G'' sont appelés les ascendants du nœud, ou pères de ce nœud.
- E est un ensemble d'arcs tels que $G = G' \cup_E G''$.
- Pour un nœud (G, G', G'', E) de D il n'y a aucun autre nœud (G_1, G'_1, G''_1, E_1) tel que $G = G_1$.
- Pour un nœud (G, G', G'', E) de D , si G' (resp. G'') est composé d'au moins deux sommets, alors il existe un nœud $(G_1, G'_1, G''_1, E_1) \in D$ tel que $G' = G_1$ (resp. $G'' = G_1$).

⁵Dans la mesure où nous manipulons des attributs à valeurs réelles et non discrètes, comme nous le verrons en section 5, cet ensemble est même non dénombrable.

- Pour un nœud (G, G', G'', E) de D , si G est un singleton, en fait $G' = \$$, $G'' = \$$ et $E = \$$, ce nœud s'écrit donc $(G, \$, \$, \$)$.

Pour un graphe modèle G , on notera $D(G)$ une décomposition dont un des nœuds s'écrit (G_1, G'_1, G''_1, E_1) avec $G_1 = G$. Une telle décomposition n'est pas unique. Par contre, si, pour un graphe G , on a une décomposition D , alors il est possible d'éclater *récurivement* G en sous graphes, jusqu'à obtenir des singletons. Cet éclatement récursif est par ailleurs unique, aux automorphismes de G près. Enfin, il faut remarquer qu'un graphe G peut être l'antécédent de deux ou plusieurs autres graphes de la décomposition, ce qui permet à ces graphes de *partager* des sous-graphes communs. Ce point permet à la méthode d'appariement, qui s'appuie sur cette structure, d'être efficace si la décomposition d'un modèle est bien construite.

La figure 3.5 présente l'exemple d'une décomposition pour le graphe J. La figure 3.6 présente une autre décomposition pour ce même graphe, mais celle-ci tient compte du fait que ce graphe possède une forte symétrie interne : les sous-graphes (1,2,3) et (4,5,6) sont isomorphes, ce qui permet de les partager dans la même décomposition, si bien que les deux antécédents du graphe dans cette décomposition sont le même nœud numéro 5. Dans la pratique, l'expérimentation montre que les décompositions qui exploitent réellement les symétries internes des modèles sont plus efficaces (voir section 5.2.2).

4.2.2 Algorithme de Décomposition

Cette section rappelle la méthode proposée par Messmer pour décomposer récursivement un graphe en sous-graphes (voir [Mes96], page 17), tout en partageant les sous-graphes communs.

De manière informelle, disons que pour un graphe G , et une décomposition D , la méthode consiste à chercher dans D le plus grand nœud S_{max} dont le graphe principal est un sous-graphe de G . Si ce graphe est G lui-même, alors il n'y a rien à faire. Sinon, il suffit d'ajouter à D le graphe $G - S_{max}$, ce

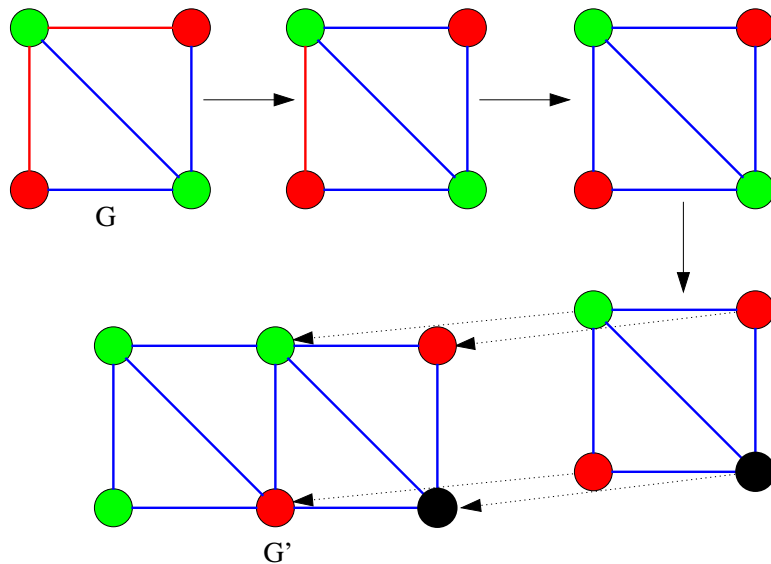


FIG. 3.4 – Explication de la distance d'édition entre deux graphes G et G' .

qui permet alors d'ajouter simplement le nœud $(G, S_{max}, G - S_{max})$. Cette technique est appliquée récursivement aux graphes S_{max} et $G - S_{max}$.

Un cas particulier se produit lorsque G est un singleton : il ne peut pas être décomposé, il est donc ajouté directement à la grammaire. Enfin, si S_{max} est vide, alors G est éclaté aléatoirement.

La méthode de décomposition d'un graphe est explicitée dans la figure 3.7.

Dans la pratique, nous commençons toujours par ajouter tous les singletons possibles à la décomposition, si bien qu'il n'y a jamais de meilleur sous-graphe S_{max} vide pour un modèle donné. L'étape 5 de ce pseudo-code ne se produit donc pas dans la pratique.

4.3 Appariement

Cette section présente l'algorithme d'appariement de Messmer proprement dit ([Mes96, MB98]), qui s'appuie sur la décomposition, et que nous appellerons l'algorithme standard. Cet algorithme

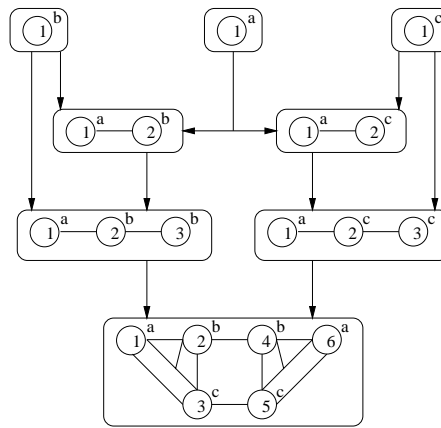


FIG. 3.5 – Exemple d'une décomposition pour le graphe J

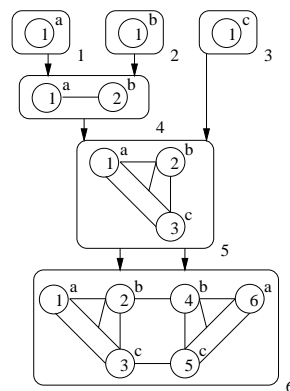


FIG. 3.6 – Une autre décomposition pour le graphe J. Ici, il y a partage de sous-graphes identiques grâce au nœud 5 de la décomposition, qui représente à lui seul les deux antécédents du nœud 6.

garantit l'obtention du minimum de la fonction de coût.

Il propage des hypothèses d'appariements entre les nœuds de la décomposition et les données. Le principe de base de la méthode est de diviser pour régner : on décompose récursivement le problème de l'appariement d'un nœud de la décomposition $S = (G, G', G'')$ en deux sous-problèmes : appairer G' et G'' puis en combiner les résultats.

Toutefois, la méthode ne suit pas rigoureusement un parcours récursif de la décomposition, mais choisit à chaque instant de combiner les hypothèses les plus adaptées à trouver rapidement le minimum de la fonction de coût.

L'algorithme exploite, pour chaque nœud de la décomposition, deux listes d'hypothèses (d'appariements) associant le graphe principal du nœud et les données. La première liste, dite «ouverte», $o(S)$, comprend des hypothèses à combiner, la seconde, dite «fermée», $c(S)$, contient des hypothèses déjà combinées. La méthode assure par construction que la première hypothèse arrivant dans la liste fermée d'un nœud est le meilleur appariement entre le graphe principal de ce nœud et les données. Ceci est en fait garanti par le processus de sélection des hypothèses à combiner. On notera que cette propriété n'est pas vraie pour les hypothèses arrivant dans la liste ouverte d'un nœud. L'algorithme exploite quelques notations supplémentaires, que voici :

- Pour $S \in D$, on note $h(S)$ une fonction dite d'estimation des coûts futurs. Cette fonction produit un minorant de l'expression $min(S) - min(M)$ où $min(S)$ est l'hypothèse de coût minimal produite pour S et M est l'hypothèse de coût minimal pour le modèle M de la décomposition. Nous n'explicitons pas la définition de $h(S)$ ici, mais le lecteur pourra se reporter à la thèse de son auteur ([Mes96]) pour en découvrir les subtilités.
- A un instant donné, $C[o(S)]$ représente le minimum des coûts des hypothèses contenues dans $o(S)$, si $o(S)$ n'est pas vide, ou $+\infty$ sinon.

Algorithme de décomposition d'un graphe.

DECOMPOSE(G, D)

1. Si G est un singleton, ajouter $(G, \$, \$, \$)$ à D .
2. Soit $S_{max} = \$$ le graphe vide.
3. Pour tout nœud (G_i, G'_i, G''_i, E) de D , faire :
 - Si G_i est un sous-graphe de G , et $|G_i| > |S_{max}|$, alors faire $S_{max} = G_i$.
4. Si S_{max} et G sont isomorphes, alors retourner (car G est en fait déjà dans D).
5. Si $S_{max} = \$$, alors choisir S_{max} aléatoirement comme un sous-graphe de G , et décomposer S_{max} .
6. Appeler *DECOMPOSE*($G - S_{max}, D$).
7. Ajouter $(G, S_{max}, G - S_{max}, E)$ à D , où E est l'ensemble des arcs reliant S_{max} et $G - S_{max}$ dans G .

FIG. 3.7 – Pseudo code de la décomposition d'un graphe G .

- On définit un seuil T au-delà duquel les hypothèses sont rejetées. Ce seuil doit être un majorant du coût minimal recherché. Il évolue au cours de l'algorithme. Un majorant trivial du coût recherché est $|G|C(v \rightarrow \$)$ qui est le coût de la séquence d'édition consistant à supprimer un à un les sommets du graphe G .

L'algorithme commence par énumérer les hypothèses pour les singletons. Cela consiste à placer pour chaque singleton S , N hypothèses correspondant aux N sommets du graphe des données. Ceci est explicité dans la figure 3.8.

Puis, l'algorithme sélectionne itérativement un des nœuds de la décomposition, et combine la meilleure de ses hypothèses encore non combinée avec les hypothèses des pères de ses fils, pour produire ainsi des hypothèses pour leurs fils communs. L'algorithme s'arrête lorsque les listes ouvertes sont épuisées.

L'algorithme, se compose de 3 fonctions principales, que nous rappelons dans les figure 3.8, 3.10, et 3.11.

Appariement des singletons

L'algorithme débute par un appariement des singletons, qui consiste simplement à lister, pour les nœuds de la première couche de la décomposition, les primitives des données. Dans notre cas, étant donné les natures différentes des primitives, cet algorithme tient compte de leur natures. Ainsi, pour un singleton de nature n , nous ne proposons que les primitives données de même nature n .

Pour des raisons qui seront explicitées au chapitre 5, nous modifions légèrement l'algorithme standard d'appariement des singletons, de façon à tenir compte d'un besoin qui sera exprimé dans ce chapitre. Cette modification ne viole en rien l'algorithme standard. Pour cela, nous introduisons une fonction $a(\cdot)$ qui, à un sommet du graphe des données associe 1 si ce sommet peut être apparié, 0 sinon.

Algorithme d'appariement pour un singleton.

$VM(G, G_I)$

1. Soit v le seul sommet de G
2. Soit $F = \emptyset$
3. Ajouter l'opération d'édition $v \rightarrow \$$ à F .
4. Pour tout sommet v' de G_I , de même nature que v , si $a(v')$ vaut 1, faire :
 - Ajouter l'opération d'édition $\mu(v) \rightarrow \mu(v')$ à F .
5. retourner F .

FIG. 3.8 – Pseudo code de l'appariement standard pour un singleton

Combinaison d'appariements

Nous rappelons que l'esprit de la décomposition est de diviser les modèles pour régner. Dans l'appariement, à l'inverse, il faut savoir combiner deux appariements de graphes G' et G'' pour en faire un appariement pour le graphe G du nœud (G, G', G'') de la décomposition.

La méthode de combinaison de deux appariements f_1 et f_2 pour deux nœuds S_1 et S_2 de D est présentée en figure 3.10. Le point 2 de ce pseudo-code vérifie que les deux appariements des deux

sous-graphes G_1 et G_2 envoient ces graphes vers des ensembles de primitives disjointes des données. Ceci est nécessaire car formellement un sommet du graphe des données ne peut pas être apparié avec deux sommets distincts d'un graphe modèle.

Techniquement, nous rappelons que les deux graphes G_1 et G_2 de ce pseudo-code forment par union relativement à un ensemble d'arcs E , un graphe $G = G_1 \cup_E G_2$, et que ce code vise à former un appariement pour G à partir d'appariements pour G_1 et G_2 . L'édition correspondant à ce nouvel appariement est la concaténation des éditions pour les deux graphes G_1 et G_2 , et de l'ensemble des éditions nécessaires pour transformer l'ensemble E en l'ensemble d'arcs reliant précisément $f_{\Delta_1}(\Delta_1(G_1))$ et $f_{\Delta_2}(\Delta_2(G_2))$ dans l'ensemble des données. Cet ensemble d'éditions est noté Δ_E .

Son rôle est explicité par la figure 3.9 où l'on voit un graphe G union de deux graphes G_1 et G_2 relativement à un ensemble d'arcs E (voir l'ellipse de gauche), et la transformation de G en un graphe (à droite). Cette transformation agit sur G_1 et G_2 par Δ_1 et Δ_2 , mais on voit que les arcs de E subissent aussi une transformation, notée précisément Δ_E .

Enfin, l'algorithme lui-même est présenté dans la figure 3.11

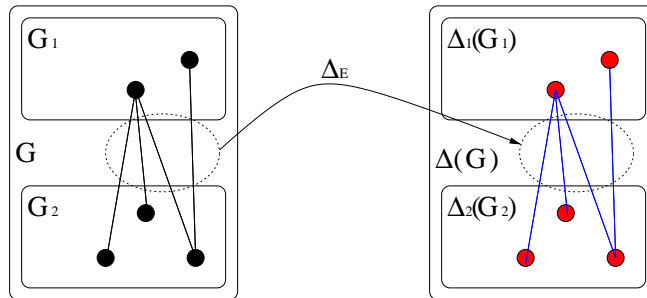


FIG. 3.9 – Edition des arcs reliant deux graphes d'une union.

Algorithme de combinaison standard.

COMBINE($S_1, S_2, E, G_I, f_1, f_2, E$)

1. Notons : $f_1 = (\Delta_1, f_{\Delta_1})$, $f_2 = (\Delta_2, f_{\Delta_2})$, $S_1 = (G_1, G'_1, G''_1)$, $S_2 = (G_2, G'_2, G''_2)$.
2. Si $\mathcal{F}_1(V(\Delta_1(G_1))) \cap \mathcal{F}_2(V(\Delta_2(G_2))) \neq \emptyset$ alors ne rien retourner.
3. Combiner f_1 et f_2 en un isomorphisme avec correction d'erreur défini par :

$$f_{\Delta}(v) = \begin{cases} f_{\Delta_1}(v) & \text{si } v \in V(\Delta_1(G_1)) \\ f_{\Delta_2}(v) & \text{si } v \in V(\Delta_2(G_2)) \end{cases}$$

$$\Delta = \Delta_1 \cup \Delta_2 \cup \Delta_E$$

4. retourner f .

FIG. 3.10 – Pseudo code de la combinaison de deux hypothèses f_1 et f_2

Algorithme d'appariement standard (Messmer).

NSG(D, G_I).

1. Pour tout singleton $(G, \$, \$)$ de la décomposition D , effectuer l'appariement de singleton : $VM(G, G_I)$.
2. Sélectionner S_1 dans D , tel que $C[o(S_1)] + h(S_1)$ soit minimal. Si $C[o(S_1)] + h(S_1) > T$ alors aller en 7.
3. Sélectionner dans $o(S_1)$ une hypothèse de coût minimal, et la placer dans $c(S_1)$.
4. Si S_1 est un modèle, alors affecter : $T = C(f_1)$.
5. Pour tous les nœuds S_2 de la décomposition s'écrivant (G, S_1, S_2, E) ou (G, S_2, S_1, E) , pour tout $f_2 \in c(S_2)$, combiner f_1 et f_2 (*COMBINE*($S_1, S_2, E, G_I, f_1, f_2, E$)) et placer le résultat dans $o(S)$.
6. Aller en 2.
7. Les hypothèses de coût minimal placées dans les listes fermées des modèles sont les meilleures hypothèses, i.e. elles réalisent le minimum de la fonction de coût.

FIG. 3.11 – Pseudo code de l'appariement standard

5 Notre contribution

5.1 Sommets et labels utilisés

Cette partie présente les éléments qui constituent les sommets et les arcs des graphes utilisés dans notre système. Elle fait donc le lien entre les aspects théoriques et la pratique. La définition des distances d'édicions réellement utilisées dans la pratique est également explicitée dans cette section.

5.1.1 Sommets

Les sommets d'un graphe représentent les primitives 3-D énumérées dans le chapitre précédent. Un nœud possède donc un attribut de nature (L pour une primitive linéaire, S pour une primitive surfacique, D pour une discontinuité du MNE, P pour un point), ainsi qu'un attribut désignant des propriétés géométriques de l'objet.

La seule propriété géométrique utilisée est l'horizontalité d'un segment de droite. On peut donc noter $\mu(v) = (n, l)$ où $n \in \{L, S, D, P\}$ et $l \in [0, \infty[$.

Dans le cas des modèles, la valeur de l vaut 0, ou 1, ce qui signifie que la propriété géométrique est requise (nécessaire) ou pas. Dans le cas des données, l'attribut est effectivement une valeur flottante qui représente la compatibilité entre la propriété géométrique représentée par l'attribut, et la géométrie de la primitive elle-même. Dans la pratique, un seul attribut de sommet est utilisé : l'horizontalité des primitives linéaires. La fonction de compatibilité utilisée est alors définie par :

$$h(s) = 2 \sin(z(s)/\|s\|)/\pi$$

où $z(s)$ est la différence (en valeur absolue) des altitudes des extrémités du segment s , et $\|s\|$ est la longueur de s .

On notera qu'une compatibilité est nulle si la géométrie de la primitive respecte rigoureusement la propriété souhaitée, et 1 dans le pire des cas (le cas d'une primitive linéaire verticale). Cette remarque sera valable également pour les compatibilités des arcs du système.

5.1.2 Arcs

Les arcs utilisés représentent les propriétés géométriques reliant des primitives géométriques.

Ces propriétés sont assez nombreuses, elles sont donc réparties en fonction des natures des sommets qu'elles relient. On distinguera ainsi les labels des arcs reliant deux sommets linéaires, un sommet linéaire et un sommet surfacique, etc. Plusieurs propriétés peuvent relier des primitives. Les arcs sont donc réellement multivalués. Dans le cas des modèles, les valeurs des attributs sont, ici aussi, binaires, et valent 0 lorsque la propriété correspondante n'est pas requise, et 1 lorsqu'elle l'est. Dans le cas des données, comme pour les labels des sommets, les labels des arcs mesurent la compatibilité de la propriété décrite par le label, avec la géométrie réelle des primitives reliées par l'arc en question.

Les prochains paragraphes dressent l'inventaire des relations utilisées dans le système. Ces relations ne sont pas détaillées ici, mais une annexe, page 171, détaille leur comportement et les fonctions qui permettent de les calculer, ainsi que les seuils et paramètres qui interviennent dans ces calculs.

5.1.2.1 Arcs PP : ponctuel - ponctuel Il n'y a pas de relation liant deux points dans le système. Mais il pourrait y en avoir (typiquement : p et q sont à la même altitude).

5.1.2.2 Arcs PL : ponctuel - linéaire La seule relation liant un point p et un segment de droite l est : p est une extrémité de s .

5.1.2.3 Arcs PS : ponctuel - surfacique La seule relation liant un point p et une portion de plan s est : p est dans le plan portant s .

5.1.2.4 Arcs PD : ponctuel - discontinuité La seule relation liant un point p est une discontinuité de MNE d est : p se trouve à l'extrémité de d .

5.1.2.5 Arcs LL : linéaire - linéaire Le système utilise plusieurs relations possibles entre deux segments de droite l et l' :

- l_1 est **parallèle** à l_2 .
- l_1 est **plus haut que** l_2 .
- l_1 **intersecte** l_2 .
- l_1 est **en regard** de l_2 .
- l_1 et l_2 sont de **même hauteur**.
- Les projections de l_1 et l_2 sur le plan $z = 0$ sont **collinéaires**.

5.1.2.6 Arcs LS : linéaire - surfacique Le système utilise plusieurs relations possibles entre un segment de droite l et une portion de plan s :

- l est **dans le plan portant** s .
- l et s ont une **même longueur caractéristique**.
- l est **plus haut que** s .
- l est **plus bas que** s .

5.1.2.7 Arcs LD : linéaire - discontinuité Le système utilise plusieurs relations possibles entre un segment de droite l et une discontinuité de MNE d :

- l est **dans** le plan portant d .
- Les projections de l et d dans le plan $z = 0$ sont **parallèles**.
- Les projections de l et d dans le plan $z = 0$ sont **en regard**.

5.1.2.8 Arcs SS : surfacique - surfacique Le système utilise plusieurs relations possibles entre deux portions de plan s_1 et s_2 :

- Les deux vecteurs de **plus grande pente** des deux portions de plan sont **opposés**.
- s_1 et s_2 se **bordent** mutuellement.
- Les projections de s_1 et s_2 dans $z = 0$ ne s'intersectent pas.
- Les vecteurs de plus grande pente de s_1 et s_2 sont **orthogonaux**.

5.1.2.9 Arcs SD : surfacique - discontinuité Le système utilise une relation entre portions de plan s et une discontinuité d :

- d borde s .

5.1.2.10 Arcs DD : discontinuité - discontinuité Le système utilise plusieurs relations entre deux discontinuités du MNE d_1 et d_2 :

- d_1 et d_2 sont **parallèles**.
- d_1 et d_2 sont **orthogonales**.

5.1.3 Distance d'édition utilisée

La théorie prévoit 5 opérations d'édition élémentaires. Dans notre cadre, la suppression d'un arc n'est pas utilisée, car le graphe des données est complet. L'ajout d'arc n'est pas utilisé non plus, car dans notre modélisation, l'existence d'un arc en plus dans les données représente une propriété particulière qui n'est pas nécessairement requise par le modèle, mais qui n'est pas interdite. Typiquement, si l'on cherche deux objets parallèles, et que les deux objets appariés sont parallèles et de même longueur, cela n'a pas d'impact sur le résultat : cette propriété 'de même longueur' arrive par hasard.

Les seules opérations d'édition utilisées sont donc la suppression de sommet, et les substitutions de labels.

Suppression de sommet Dans la mesure où le problème de l'appariement est invariant si l'on multiplie tous les coûts par une constante, on peut fixer arbitrairement le coût d'une suppression de sommet. Ce coût est fixé à 1 (sans unité). Nous avons donc :

$$C(v \rightarrow \$) = 1 \tag{3.2}$$

Substitution de labels de sommets Un label de sommet est composé de la nature du sommet et éventuellement d'une propriété géométrique. Naturellement, les changements de nature de sommet ne sont pas autorisés. Quant à la propriété géométrique, le coût naturel est proportionnel à la compatibilité entre la propriété souhaitée, et la primitive géométrique elle-même.

Ainsi, pour un label $\mu(v) = (n, l)$ du modèle et un label $\mu(v') = (n', l')$ on définit le coût d'édition

$$C((n, l) \rightarrow (n', l')) = \begin{cases} \infty & \text{si } n \neq n' \\ l' & \text{si } n = n' \text{ et } l = 1 \\ 0 & \text{sinon} \end{cases} \tag{3.3}$$

Substitution de labels d'arcs Les labels d'arcs sont clairement multivalués : $\nu(e) = (l_1, \dots, l_p)$. Ces labels représentent implicitement plusieurs arcs pour plusieurs propriétés géométriques différentes. Les mesures de ces propriétés ne s'expriment pas dans les mêmes unités, il faut donc combiner les compatibilités de manière multiplicative. Par ailleurs, cette fonction ne doit pas s'annuler dès qu'une des compatibilités est nulle, car ceci annulerait l'influence des autres propriétés géométriques des objets. Nous proposons donc la fonction de combinaison suivante :

$$C((l_1, \dots, l_p) \rightarrow (l'_1, \dots, l'_p)) = \prod_{i=1 \dots p} (1 + C(l_i, l'_i)) - 1 \quad (3.4)$$

où nous avons :

$$C(l_i, l'_i) = \begin{cases} l'_i & \text{si } l_i = 1 \\ 0 & \text{sinon} \end{cases}$$

Pondération Une pondération des substitutions des labels par rapport aux suppressions de sommets est nécessaire dans la pratique. Le coût d'édition d'une séquence d'édition $\Delta = (\delta_1, \dots, \delta_k)$ s'écrit donc comme une adaptation de l'équation 3.1 :

$$C(\delta_1, \dots, \delta_k) = \sum_{i=1 \dots p} P(\delta_i) C(\delta_i) \quad (3.5)$$

où $C(\delta_i)$ est définie par l'une des équations 3.2, 3.3, ou 3.4 en fonction de la nature de δ_i , et $P(\delta_i)$ est la fonction de pondération :

$$P(\delta_i) = \begin{cases} 1 & \text{si } \delta_i \text{ est une suppression de sommet} \\ K & \text{sinon} \end{cases}$$

où K est la constante de pondération. Cette pondération n'est pas prévue dans la théorie, mais est nécessaire dans la pratique. Elle sera justifiée expérimentalement dans la section suivante.

5.2 Expérimentation

L'expérimentation présentée ici s'articule autour de deux points majeurs : les résultats et la performance. La partie consacrée aux résultats présente des appariements de graphes réalisés. Ces résultats sont indépendants des décompositions choisies ou d'éventuelles heuristiques de calcul. Ils ne dépendent que des primitives, de la modélisation choisie, et de la fonction de coût. La partie consacrée à la performance mettra en évidence certaines propriétés de la méthode, et notamment l'impact de la décomposition sur les temps de calcul.

5.2.1 Premiers résultats

Pour la zone 172, l'appariement du modèle de bâtiment à double pan permet d'obtenir les résultats présentés en figure 3.12. On y voit 3 paires d'images stéréoscopiques, représentant respectivement les discontinuités appariées, les segments de droites, ainsi que les primitives surfaciques. Le calcul a été réalisé avec la valeur $K = 0,2$. Les figures 3.13 et 3.14 présentent les résultats des mêmes appariements, mais pour des valeurs respectives de $K = 0,05$ et $K = 1$.

Le comportement de l'appariement en fonction de la variable K est conforme à la signification de cette variable. En effet, une valeur faible de K induit peu d'influence de la qualité géométrique des primitives, ce qui favorise le nombre de primitives appariées, car le coût relatif de la suppression d'un nœud est élevé. Ainsi, la méthode privilégie des appariements nombreux, même avec des primitives mal positionnées. On voit ainsi sur la figure 3.13 deux primitives linéaires aberrantes, entraînant avec

elles l'appariement d'une discontinuité elle aussi aberrante, alors que pour $K = 0, 2$, les appariements des discontinuités sont corrects. Le faîte du bâtiment, quant à lui, est apparié avec la primitive notée A dans l'image.

A l'inverse, pour $K = 1$, il y a moins de primitives appariées, mais celles-ci ont des compatibilités faibles (donc bonnes) avec les propriétés géométriques qui les caractérisent.

On voit ici l'importance de la valeur de K et la nécessité de l'avoir introduite, car sans pondération (i.e. $K = 1$) la méthode apparie réellement trop peu d'objets. Ce point justifie l'introduction de cette constante de pondération non prévue par la théorie.

On notera qu'aucune figure ne présente les sommets de type ponctuel appariés, puisqu'aucune détection associée à ces sommet n'est proposée dans ce travail.

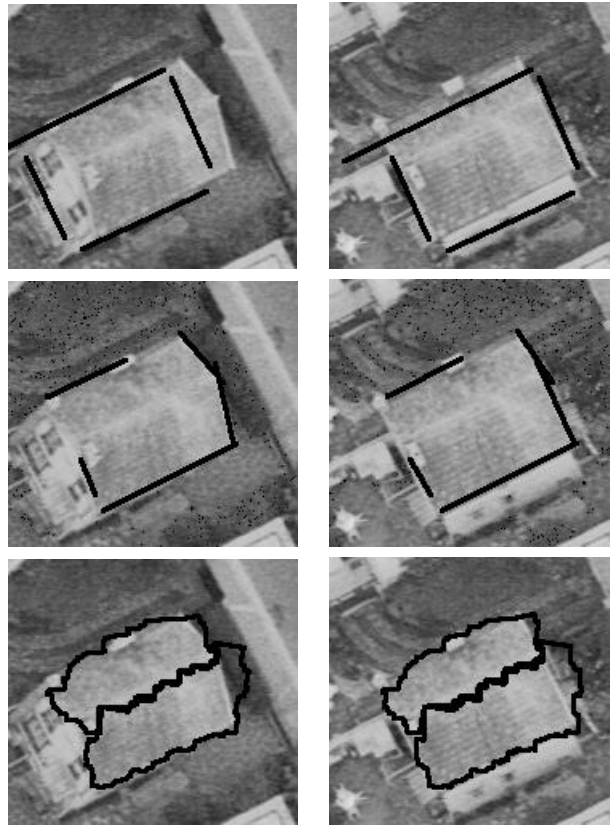


FIG. 3.12 – Discontinuités appariées, primitives linéaires appariées, et surfaces appariées, pour $K = 0, 2$

5.2.2 Impact de la décomposition

La décomposition du modèle joue un rôle primordial dans la performance de la méthode de Messmer. C'est elle qui permet le partage des sous-graphes isomorphes, et la réduction de la combinatoire du problème. Mais elle doit être mise en place de façon à utiliser effectivement cette possibilité.

Cette section montre l'impact d'une décomposition adaptée du modèle.

Tout d'abord, considérons une décomposition n'utilisant aucun a priori sur le modèle utilisé. La décomposition du modèle simple de la figure 3.1 produit le réseau présenté dans la figure 3.15. La figure 3.16 montre une décomposition dans laquelle des graphes simples ont été introduits a priori. Ces graphes sont des éléments fréquents dans les modèles de bâtiments, et leur introduction dans la décomposition est donc une exploitation des connaissances du domaine d'application de la méthode.

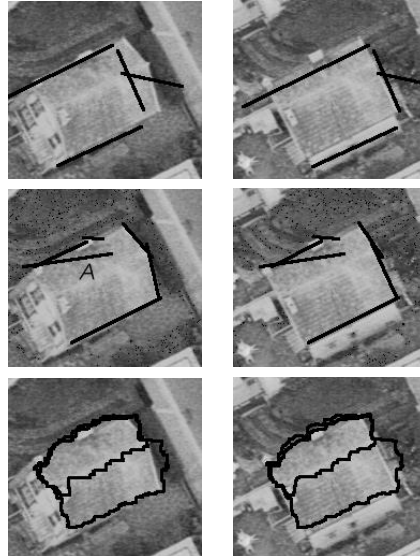


FIG. 3.13 – Discontinuités appariées, primitives linéaires appariées, pour $K = 0,05$

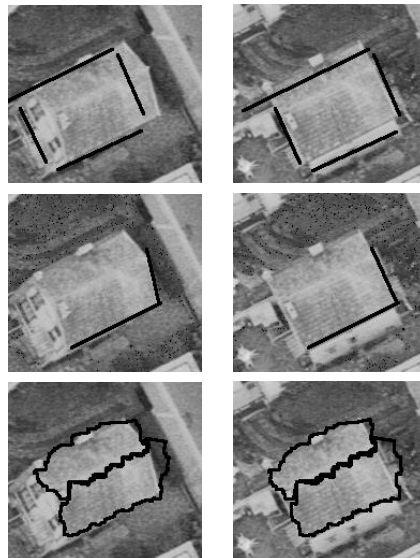


FIG. 3.14 – Discontinuités appariées, primitives linéaires appariées, pour $K = 1$

Pour le bâtiment 172, l'utilisation de la décomposition minimale amène un temps de calcul de 1,7s, alors qu'avec la décomposition améliorée, le temps de calcul est de 0,1s. Le tableau 3.1 présente des temps CPU d'appariement pour certains modèles et certaines données. Les colonnes notées MH correspondent aux temps CPU nécessaires pour que le meilleur appariement apparaisse au cours du déroulement de l'algorithme. Les colonnes «temps pur» correspondent au temps CPU total. La différence entre ces deux colonnes correspond donc au temps nécessaire pour que la méthode soit sûre que le meilleur appariement trouvé est effectivement le minimum de la distance d'édition. Nous pensons que l'étude des chiffres de ce tableau suffit à montrer l'intérêt de mettre en place une décomposition adaptée aux modèles en jeu.

Bâtiment	Modèle	Décomposition simple		Décomposition adaptée	
		Temps pur	MH	Temps	MH
172 (page 177)	P	0,06s	0,04s		
172	D	0,08s	0,08s		
172	VD	0,36s	0,22s		
45 (page 196)	P	0,6s	0,33s		
45	D	1,7s	0,84s	0,6s	0,2s
45	VD	3,83s	5,67	7,3s	7,5s
22 (page 207)	P	1,99s	1,55s		
22	D	8,7s	3,7s	4,8s	1,2s
22	VD	79s	55s	39s	18s
4 (page 216)	P	140s	134s	9,8s	7,2s
4	D	>10min	>10min	167s	108s
4	VD	546	424s	180s	156s
19 (page 209)	P	0,35s	0,30s	0,46s	0,27s
19	D	0,76s	0,43s	0,51s	0,23s
19	VD	5,7s	3,3s	2,24s	0,67s
19	DD	33,5s	20,8s	0,7s	0,48s

TAB. 3.1 – Temps d'exécution des appariements pour certains modèles, certaines données. Les formes des modèles désignés ici sont présentées page 97.

5.3 Reconstruction effective du modèle

L'appariement de graphe permet d'*identifier*, parmi les primitives, celles qui contribuent à une reconstruction finale du bâtiment. Cette identification est réalisée à l'aide de l'injection de l'ensemble des nœuds du modèle vers l'ensemble des nœuds du graphe des données. Les données n'ont aucune raison d'être topologiquement correctes (typiquement, deux segments de droites qui partagent une extrémité dans le modèle ne sont pas jointifs dans les données), car les méthodes présentées au chapitre 2 ne l'assurent pas.

L'obtention réelle d'une description 3-D d'un bâtiment passe donc par une étape de reconstruction après l'appariement.

Cette reconstruction est elle-même composée de 3 étapes.

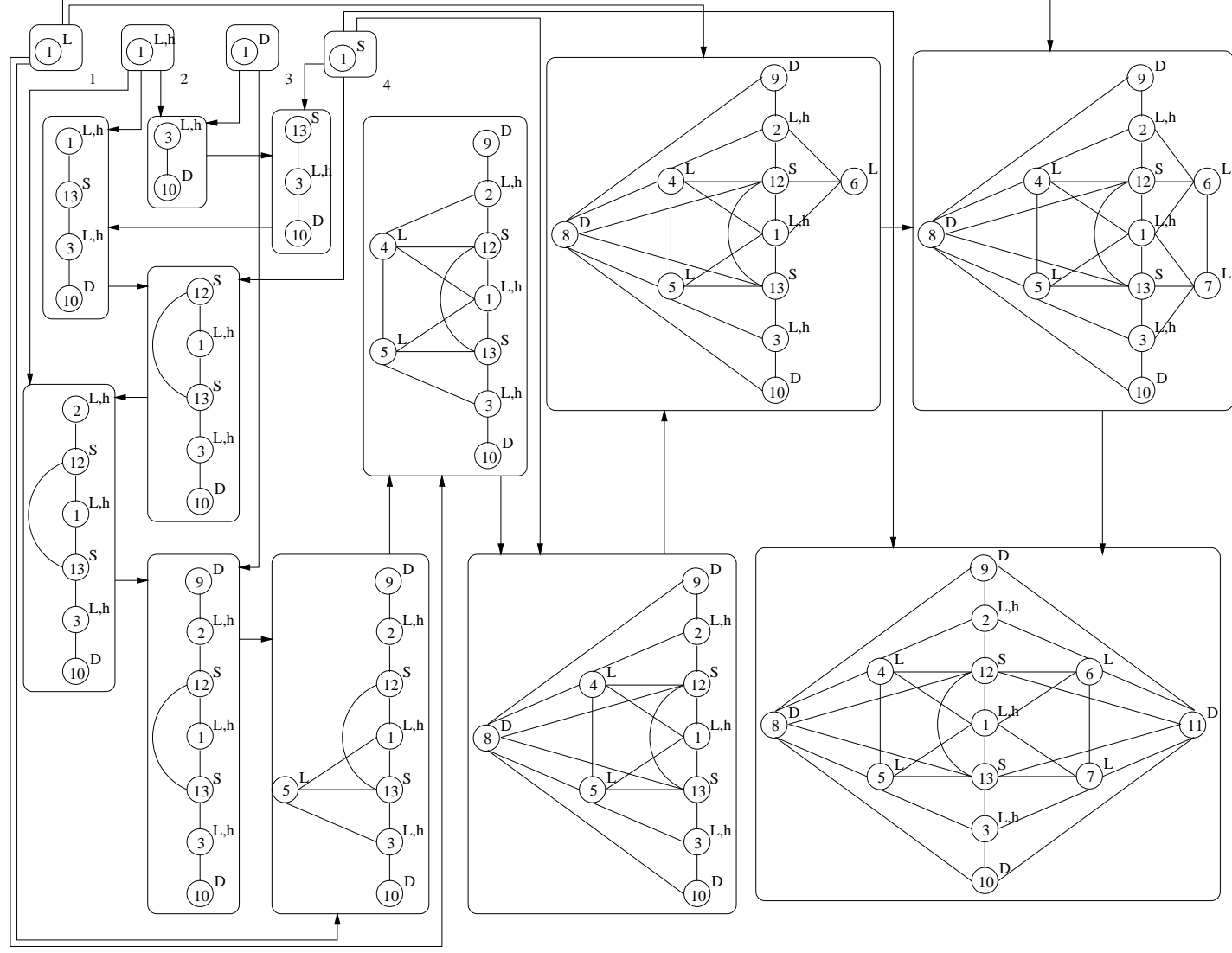


FIG. 3.15 – Décomposition du modèle de la figure 3.1 sans a priori.

5.3.1 Calcul de la faisabilité de la reconstruction

La reconstruction est une représentation vectorielle de type sommets / arcs / faces. Elle a besoin de la position 3-D de tous les sommets présents dans le modèle. Pour cela, il faut donc que les primitives désignées par l'appariement soient suffisantes pour placer chaque sommet en 3-D.

Soit v un sommet ponctuel du modèle. v possède un certain nombre de relations avec les autres sommets du modèle :

- Soit $P_v = \{p_1, \dots, p_n\}$ l'ensemble des sommets de type surfacique avec lesquels s possède une relation de type « est dans ».
- Soit $L_v = \{l_1, \dots, l_m\}$ l'ensemble des sommets de type linéaire avec lesquels s possède une relation de type « est dans ».
- Soit $D_v = \{d_1, \dots, d_l\}$ l'ensemble des sommets de type discontinuité avec lesquels s possède une relation de type « est dans ».

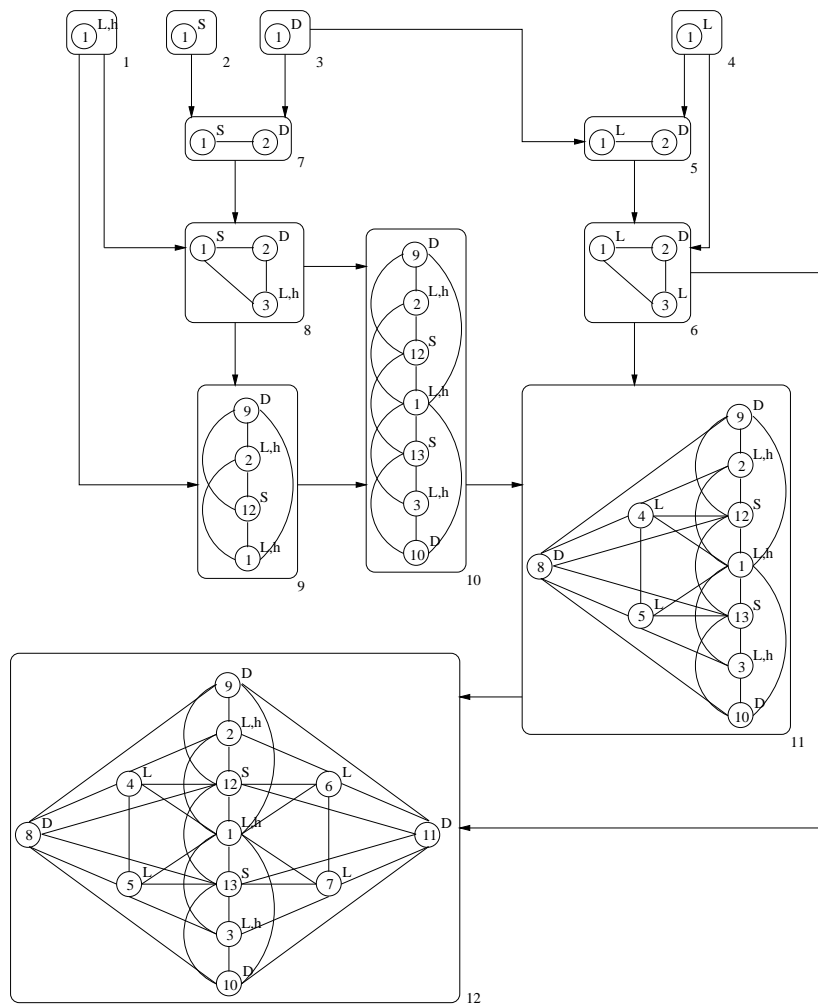


FIG. 3.16 – Décomposition plus efficace du modèle de la figure 3.1.

Soit $app(\cdot)$ la fonction qui à un sommet du graphe modèle associe la valeur 1 si le sommet a été apparié avec un nœud du graphe des données ⁶, c'est-à-dire si $f_{\Delta}(\Delta(v)) \neq \$$, et 0 sinon.

Soit $lp(\cdot, \cdot)$ la fonction qui à un sommet linéaire et à un sommet surfacique associe 1 si ces deux sommets possèdent une relation « est dans », et 0 sinon.

Soit $ld(\cdot, \cdot)$ la fonction qui à un sommet linéaire et à un sommet de type discontinuité associe 1 si ces deux sommets possèdent une relation « est dans », et 0 sinon.

Alors on voit que $\sum_{p \in P_s} app(p)$ (respectivement $\sum_{l \in L_s} app(l)$, $\sum_{d \in D_s} app(l)$) est le nombre de primitives planes (respectivement linéaires, discontinuités) des données auxquelles s doit appartenir.

Un sommet ponctuel a 3 degrés de liberté. Son appartenance à un plan lui supprime un degré de liberté. De même, l'appartenance du sommet à une primitive discontinuité lui supprime 1 degré de liberté⁷. Enfin, l'appartenance du sommet à une droite lui supprime 2 degrés de liberté, sauf si cette droite appartient à un plan auquel ce sommet appartient lui-même, auquel cas cela lui supprime un seul degré de liberté. De même si la droite est dans une discontinuité dans laquelle le sommet se trouve, un seul degré de liberté est supprimé. Enfin, si la droite se trouve simultanément, dans un plan et dans une discontinuité auxquels le sommet appartient, alors cette droite ne supprime aucun degré de liberté, puisqu'elle se trouve précisément à l'intersection de ce plan et de cette discontinuité.

Nous pouvons donc établir une fonction indicatrice DA (pour Degrés de liberté après Appariement), qui indique combien v dispose de degrés de liberté.

$$DA(s) = 3 - \left(\sum_{p \in P_s} app(p) + \sum_{d \in D_s} app(d) + \sum_{l \in L_s} 2app(l) + \sum_{p \in P_s, l \in L_s} lp(l, p) + \sum_{l \in L_s, d \in D_s} ld(l, d) \right)$$

Si cette indicatrice est négative ou nulle, alors le sommet v peut être positionné en 3-D de manière univoque, sans degré de liberté. Si elle est négative, la position du sommet 3-D est même surdéterminée. C'est le cas lorsque le sommet ponctuel se trouve à l'extrémité de 3 segments de droite. Dans ce cas, le sommet ponctuel est placé à une position qui minimise la somme de ses distances aux objets auxquels il doit appartenir (voir section 5.3.2).

La reconstruction pour un modèle G est donc dite « faisable » si et seulement si pour tout sommet s de type ponctuel du modèle G , $DA(s) \leq 0$. Dans ce cas, la reconstruction effective de la zone à l'aide du modèle G est réalisée à l'aide des outils proposés dans les prochaines sections (5.3.2 et 5.3.3).

5.3.2 Positionnement des sommets ponctuels

Nous rappelons qu'une fonction de coût relie le modèle G et le graphe de données G_I . Cette fonction de coût a été minimisée lors de l'appariement, pour identifier les primitives linéaires, planes, et discontinuités. Les primitives ponctuelles, en revanche, n'ont pas contribué à l'appariement, car il n'y en a pas dans les données. En revanche, le modèle en contient, et les sommets correspondant possèdent des relations « est dans » ces sommets aux autres sommets. Pour ces relations, des fonctions de compatibilité existent. Il suffit alors d'injecter dans le graphe des données les sommets 3D correspondant au modèle, à la position pour laquelle la distance d'édition est minimale.

⁶Nous rappelons que cela peut ne pas être le cas, c'est notamment souhaitable en cas de sous-détection d'une primitive.

⁷Nous rappelons qu'une primitive de type discontinuité est en fait un plan vertical.

Ce procédé est explicité par la figure 3.17. Considérons un sommet ponctuel (en rouge) possédant 4 arcs vers d'autres sommets du modèle. Supposons que 3 de ces sommets sont appariés avec des primitives des données : 2 sont appariés avec des discontinuités, et 1 avec un plan. Enfin, le quatrième sommet n'est apparié avec rien. En raison de ces 3 appariements réalisés, la fonction indicatrice DA est nulle, on peut donc le placer en 3D. Dans notre cas, ceci permet de placer le sommet 3D à l'intersection des 3 plans appariés (le plan et les discontinuités).

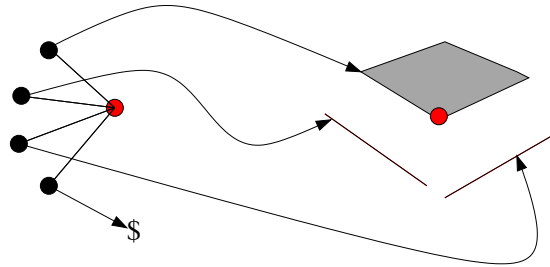


FIG. 3.17 – Explicitation du positionnement des sommets 3D.

La figure 3.18 montre le résultat de ce placement pour le bâtiment 172.

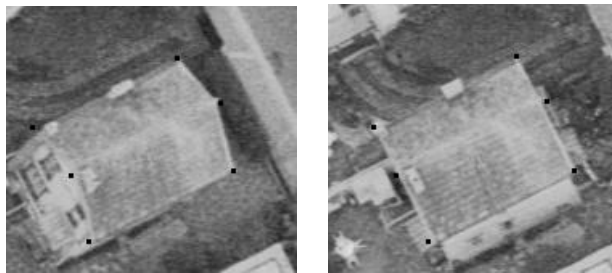


FIG. 3.18 – Sommets ponctuels placés en 3-D après appariement, pour le bâtiment 172.

5.3.3 Complétion du modèle

Une fois les sommets placés, il est possible, toujours à l'aide de l'information contenue dans le modèle, de définir une géométrie propre pour les sommets de type linéaire. On notera que dans le cas où un sommet linéaire a été apparié, cette opération remplace la géométrie de la primitive par une nouvelle géométrie, mais dans le cas où ce sommet n'a pas été apparié, cette opération place un objet 3-D là où l'appariement n'en a pas identifié. Dans le cas d'une sous-détection (i.e. il n'y a effectivement pas de primitive pertinente pour décrire l'objet 3-D correspondant), cette opération comble la sous-détection.

Par exemple, pour la zone 19, la figure E.3, page 162, montre l'appariement du modèle DD avec cette scène. Les primitives linéaires sont sur-élevées de 2 m pour la lisibilité de cette figure. On y voit que certaines primitives linéaires sont manquantes. Mais grâce aux discontinuités appariées (en base de la figure), tous les sommets 3D peuvent être placés, ce qui mène à une reconstruction faisable de ce modèle (figure E.4).

De même, les primitives surfaciques sont recalculées à partir des sommets ponctuels qui les forment. L'ordre de ces sommets est connu dans le modèle grâce au fait que les sommets ponctuels sont aussi les extrémités des segments qui bordent les primitives surfaciques. Il suffit donc de parcourir les sommets 3-D en s'appuyant les sommets linéaires. Le modèle assure par construction la topologie de la forme recherchée.

Dans le cas de la zone 172, cette complétion permet de mettre en place toutes les arêtes du modèle D, puisque la reconstruction est faisable. Le faîte notamment, est mis en place, bien qu'il n'ait pas été détecté dans les images initiales : la méthode comble ainsi cette sous-détection de primitives linéaires. Les deux pans contribuent au positionnement de ces faîtes, car le modèle indique que les deux extrémités de ce faîte font partie de ces pans. Ceci explique a posteriori la remarque du bas de la page 44 : le faîte est recalculé à partir des autres objets détectés dans la zone, et notamment à partir des portions de plans calculées à partir de MNE.

5.4 Le cas des primitives surfaciques

L'expérience montre que les primitives surfaciques ne souffrent que de très peu de sous-détection (au prix d'une sur-détection évidente, que nous avons tenté de pallier par un filtrage). Par ailleurs, les primitives surfaciques sont très structurantes dans les modèles. Comme ce sont elles qui ont la dimension la plus forte (elles sont de dimension 2), elles possèdent un grand nombre de relations avec les autres sommets des modèles. Pour expliciter ce fait, le tableau 3.2 présente pour quelques modèles, le nombre de sommets des différents types, le nombre d'arcs mettant en jeu chaque type de sommet, et la moyenne des nombres d'arcs par type. Ce tableau montre nettement que les primitives surfaciques possèdent plus d'arcs en moyenne que les autres primitives, d'où leur caractère «très structurant» évoqué ici.

Ainsi la perte d'une primitive surfacique au moment de l'appariement engendre la perte d'un grand nombre de contraintes du modèle.

Le coût d'une suppression de sommet surfacique est donc fixé à l'infini, de manière à forcer les appariements des sommets surfaciques des modèles :

$$\text{Si } v \text{ est surfacique, } C(v \rightarrow \$) = +\infty$$

Le lecteur notera, à la lecture du tableau présenté dans cette section, que les sommets ponctuels ont en moyenne 5 à 6 relations avec les autres sommets du modèle. Cette forte redondance est nécessaire pour augmenter les chances d'obtenir des reconstructions faisables.

Modèle	S	L	D	P	aS	aL	aD	aP	maS	maL	maD	maP
P	1	4	4	4	12	24	22	20	12	6	5,5	5
D	2	7	4	6	30	54	28	36	15	7,7	7	6
DD	4	12	6	9	46	65	42	54	11,5	5,4	7	6

TAB. 3.2 – Comparaison des nombres d'arcs en jeu pour les différents types de primitives. S : nombre de sommets Surfaciques du modèle, L : nombre de sommets Linéaires, D : nombre de sommets de type Discontinuité, P : nombre de sommets Ponctuels. aS : nombre d'arcs mettant en jeu un sommet Surfacique, ..., maS : moyenne des arcs pour un sommet surfacique, ...

5.5 Redondance des types de primitives

Cette section discute la redondance existant entre les données. Plus exactement, nous tentons de montrer que les 3 types de primitives mis en jeu dans le système se complètent, et même, sont tous nécessaires. Elle se fera de façon détaillée sur l'exemple 172, mais nous donnerons d'autres exemples.

Primitives surfaciques

Nous avons déjà évoqué l'intérêt que nous portons aux primitives surfaciques (elles sont peu sous-détectées, et très structurantes). En supprimant les sommets surfaciques du modèle D , la zone 172 est reconstruite de manière aberrante.

Primitives linéaires Dans le même cas de figure, la suppression des primitives linéaires suffit pour produire une reconstruction faisable. Il se trouve que les primitives linéaires peuvent être supprimées pour beaucoup de zones de la scène correspondant au modèle D . Toutefois, la zone 101 ne produit pas de reconstruction satisfaisante sans les primitives linéaires du modèle.

Discontinuités Dans le cas de la zone 172, il est possible de se passer des discontinuités du modèle. Par contre, pour la zone 170, qui est correctement reconstruite par ce modèle, l'absence des discontinuités produit une reconstruction non faisable. Les discontinuités sont donc clairement nécessaires à l'application. De même, dans la zone 19, les reconstructions ne sont jamais correctes sans les discontinuités.

5.6 Le cas des automorphismes

Les modèles utilisés sont relativement simples et ne couvrent pas la variété des formes de bâtiments. Les modèles simples possèdent généralement des automorphismes. Typiquement, les modèles D , VDV , et DD possèdent 4 automorphismes, et le modèle VD en possède 2. Dans la pratique, les méthodes d'appariement calculent donc des appariements identiques aux automorphismes près. Le calcul pourrait être N fois plus rapide pour un modèle D , où N est le nombre d'automorphismes du modèle, si nous savions tirer parti de ces automorphismes des modèles.

A l'heure actuelle, nous n'avons toutefois pas pu mettre au point de stratégie permettant de tirer profit de ces automorphismes. Les automorphismes ne sont donc pas utilisés afin de réduire la combinatoire de l'appariement de graphes.

6 Heuristique stochastique

La méthode standard est déterministe. Elle assure l'obtention du minimum de la fonction de coût. Ceci permet de ne pas mélanger le problème de la pertinence de la fonction que l'on minimise et celui du calcul de son minimum proprement dit.

La performance de la méthode montre toutefois ses limites lorsque les modèles sont complexes ou que les données sont nombreuses. Par ailleurs, dans un contexte plus opérationnel, on peut tolérer des solutions sous-optimales si elles sont sensiblement plus rapides. Nous allons donc proposer une heuristique stochastique qui permet d'identifier rapidement des solutions proches du minimum de la fonction de coût, et qui s'intègre dans la méthode déterministe décrite plus haut.

6.1 Heuristique

Cette section décrit l'heuristique pure (par opposition à la méthode proposée dans la section 6.2). Dans la mesure où cette heuristique utilise les structures de données de la méthode de Messmer, elle bénéficie elle aussi directement de la possibilité de partager les sous-graphes communs, ce qui est un avantage par rapport à d'autres méthodes non déterministes de la littérature ([WH97, FWH98]). Nous ne comparerons toutefois pas cette heuristique stochastique aux autres approches non déterministes de la littérature, car leur mise en oeuvre est généralement lourde. Nous avons toutefois conscience que cet argument n'est pas valide d'un point de vue académique.

L'idée principale de cette heuristique consiste à changer le processus de sélection des hypothèses à combiner, afin de privilégier non pas les graphes garantissant l'obtention du minimum, mais afin de privilégier l'obtention rapide d'hypothèses pour le modèle. Ceci est fait simplement en levant ce critère de sélection, et en le remplaçant par une sélection aléatoire. Cette sélection aléatoire permet d'accéder rapidement à des hypothèses pour les graphes les plus gros de la décomposition, et donc pour le modèle.

Techniquement, cela revient à remplacer l'étape 2 du pseudo-code de la figure 3.11 par :

- 2 . Sélectionner aléatoirement S_1 dans D , tel que $o(s_1)$ est non vide.

Par ailleurs, dans le cas d'une méthode stochastique comme celle-ci, la fonction $h(S)$ introduite dans les travaux de Messmer n'est plus valide. Il faut donc poser $h(S) = 0$ pour tous les nœuds de la décomposition.

Comme $h(s)$ n'est plus valide, il est alors intéressant de tenter de fusionner ces deux approches pour que la méthode heuristique profite de l'estimation de $h(S)$, ce qui lui éviterait de stocker des hypothèses inutiles, et pour que la méthode standard profite rapidement des estimations de T faites par l'heuristique stochastique. C'est l'objet de la prochaine section.

6.2 Intégration

L'intégration est possible en faisant travailler l'approche standard et l'heuristique stochastique sur leurs propres listes d'hypothèses, mais en leur faisant échanger leurs valeurs de T lorsque c'est possible. Ceci se traduit par l'algorithme de la figure 3.19 où l'approche standard travaille dans la colonne de gauche, et l'approche stochastique dans celle de droite.

La méthode commence par calculer les appariements pour les singletons, dans chacune des approches (étapes 1 et 2). Puis, chacune des approches effectue une passe de sélection d'hypothèse et de combinaison (étapes 3 à 6). Par après, chacune des méthodes effectue une estimation de T si possible, et ces deux estimations sont fusionnées par $T = \min(T, T')$.

Le tableau 3.3 présente sur ses deux dernières colonnes les temps CPU utilisés pour effectuer les mêmes appariements que ceux du tableau 3.1. On notera, à l'aide de ce tableau, que les performances de la méthode combinée et de la méthode «pure» ne diffèrent que très peu. Par contre, en comparant les temps d'apparition des meilleures hypothèses dans le déroulement de l'algorithme (colonnes MH du tableau), on constate que l'approche stochastique a l'avantage de proposer plus vite des appariements que l'approche standard.

Dans un contexte opérationnel, ceci peut s'avérer important : il est intéressant d'obtenir rapidement des résultats. Nous verrons même dans les chapitres suivants que nous ferons fonctionner les algorithmes d'appariement avec la méthode stochastique, mais en temps limité, afin de maîtriser

l'explosion combinatoire du problème d'appariement.

Algorithme appariement d'un graphe : approche mixte.

$AM(D, G_I)$

- | | |
|---|---|
| <ol style="list-style-type: none"> 1. Pour tout singleton (G, G', G'') de la décomposition D, effectuer l'appariement de singleton : $VM(G, G_I)$ 3. Sélectionner S_1 dans D, tel que $C[o(S_1)] + h(S_1)$ soit minimal 4. Si S_1 est un modèle, affecter $T = C(f_1)$ 7. Faire $T = \min(T, T')$ 8. Combiner S_1 10. Aller en 3 11. Terminer. | <ol style="list-style-type: none"> 2. Pour tout singleton (G, G', G'') de la décomposition D, effectuer l'appariement de singleton : $VM(G, G_I)$ 5. Sélectionner S_1 dans D, aléatoirement, tel que $o(S_1)$ est non vide. Si aucun S_1 ne convient, aller en 8. 6. Si S_1 est un modèle, affecter $T' = C(f_1)$ 9. Combiner S_1 |
|---|---|

FIG. 3.19 – Pseudo code de l'approche mixte d'appariement.

Données	Modèles	Méthode mixte		Aléatoire pur	
		Temps total	MH	Temps total	MH
172	P	0,07s	0,03s	0,05s	0,03s
172	D	0,07s	0,04s	0,08s	0,04s
172	VD	0,4s	0,04s	0,36s	0,05s
45	P	0,38s	0,06s	0,24s	0,09s
45	D	1,61s	0,53s	1,3s	0,35s
45	VD	2,18s	0,06s	1,94 s	0,05s
22	P	2,3s	0,35s	1,27s	0,27s
22	D	13,1s	3,37s	11,7s	2,6s
22	VD	50s	2,1s	60s	1,1s
22	VDV	917s	3,2s	1198s	1,69s
4	P	8,3s	1,49s	7,7s	0,8s
4	D	180s	2,48s	214s	1,7s
4	VD	131s	3,19s	140s	2,1s
19	P	0,26s	0,07s	0,2s	0,04s
19	D	1,24s	0,10s	0,74s	0,06s
19	VD	5,85s	0,25s	6,65s	0,13s
19	DD	32,5s	0,38s	53,4s	0,25s

TAB. 3.3 – Temps CPU pour certains appariements : comparaison des méthodes aléatoire et mixte-aléatoire. Les colonnes MH correspondent aux temps d'apparition des meilleures hypothèses d'appariement pour chacun des méthodes, comme dans le tableau 3.1.

7 Appariement avec connaissance externe

L'idée directrice de cette section est d'exploiter de manière efficace une information externe. Par information externe, on entend ici le fait que dans le processus d'appariement, on force certains sommets du modèle à être appariés avec certains sommets des données.

Ce type d'information peut provenir principalement de deux sources : une saisie d'un opérateur (moyennant un outil de communication performant permettant d'interroger un modèle et des données), ou alors un appariement partiel d'un sous-graphe qui aurait déjà fourni une certaine connaissance sur l'appariement complet du graphe courant.

La méthode de Messmer peut être modifiée de façon à tirer pleinement profit d'une telle information externe. Le formalisme mis en oeuvre dans cette méthode doit toutefois être un peu étendu. Les prochaines sections présenteront de manière progressive la technique permettant d'intégrer cette information externe dans le processus d'appariement. Plus exactement, une méthode «naïve» sera exposée afin de poser certains éléments théoriques, avant d'explicitier la méthode proposée. L'architecture de cette présentation est la même que ce que l'on peut trouver dans la publication [FM00].

7.1 Modélisation de l'information externe.

L'information externe est introduite dans les modèles par une modification des labels des sommets. On appellera un graphe AP un graphe portant une information externe. Un graphe AP est défini comme suit :

- Un graphe AP est un graphe $G^{ap} = (V, E, \mu, \nu)$ où la fonction d'attribution des labels admet un nouvel espace d'arrivée : $\mu : V \rightarrow L_V^{ap}$ où $L_V^{ap} = L_V \times (\{\diamond, \$\} \cup V_I)$, où V_I est l'ensemble des sommets du graphe des données. Pour un sommet $v \in V$ du modèle, son label est $\mu(v) = (l, l_a)$ où :
 - $l_a = \diamond$ signifie qu'il n'y a pas d'information externe associée à ce sommet. L'appariement du graphe et des données devra donc déterminer une correspondance pour ce sommet.
 - $l_a = \$$ signifie qu'il y a une information externe associée à ce sommet, et que cette information stipule que le sommet n'est associé à aucun sommet du graphe des données.
 - $l_a = v_i$ avec $v_i \in V_I$ signifie qu'il y a une information externe pour ce sommet, et que cette information précise que ce sommet est associé au sommet v_i des données.

Un sommet dont le label est du type (l, \diamond) est appelé un sommet *strict*. Pour un ensemble de sommets V , le sous-ensemble de ses sommets stricts sera noté V^s . Son complémentaire est l'ensemble des sommets qui portent une information externe, et il sera noté V^{ap} . On peut donc noter $G^{ap} = (V^s \cup V^{ap}, E)$ où $V^s \subset L_V \times \{\diamond\}$, et $V^{ap} \subset L_V \times (\{\$, \} \cup V_I)$.

L'isomorphisme entre deux graphes AP est défini par extension naturelle de la définition de l'isomorphisme de graphes aux graphes AP. La seule différence réside dans l'égalité des labels des sommets : deux labels de sommets $\mu(v) = (l, l_a)$ et $\mu(v') = (l', l'_a)$ sont égaux si et seulement si $l = l'$ et $l_a = l'_a$. Deux labels de sommets sont donc égaux si ils sont égaux au sens habituel, et s'ils portent la même information externe. Par conséquent, lorsque deux graphes AP, G^{ap} et G'^{ap} sont isomorphes par un isomorphisme f , nous avons : $f(V^s) = V'^s$ et $f(V^{ap}) = V'^{ap}$.

7.2 Exemple

Considérons le graphe J de la figure 3.2. Et supposons qu'il doit être apparié à un graphe de données, en sachant a priori que le sommet 4 de ce graphe correspond au sommet 17 des données, et que le sommet 5 correspond au sommet 23 des données. Cette information externe est modélisée par la modification des labels des sommets 4 et 5 : $\mu^{ap}(4) = (\mu(4), v_{17})$ et $\mu^{ap}(5) = (\mu(5), v_{23})$. La figure 3.20 représente ce graphe.

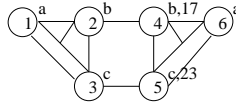


FIG. 3.20 – Exemple de graphe portant une connaissance externe sur ses sommets 4 et 5.

7.3 Première approche.

Cette première approche présente une exploitation directe de la décomposition de Messmer en utilisant les graphes AP. Il est en effet possible de le faire, en remplaçant les graphes par les graphes AP, ainsi que leurs isomorphismes. La décomposition du graphe de la figure 3.20 produit ainsi le réseau proposé dans la figure 3.21. Dans cette décomposition, on suppose que le noeud 5 était présent avant d'introduire le graphe 6. Sur cette décomposition, on voit très clairement que le graphe 5, qui est le sous-graphe (1,2,3) du graphe 6, ne profite pas de l'information apportée par les sommets 4 et 5. Plus exactement, comme on sait que le sommet 4 du graphe 6 doit être mis en correspondance avec le sommet 17 des données, le sommet 2 ne peut être mis en correspondance qu'avec un des voisins du sommet 17 des données. De même le sommet 3 ne peut être mis en correspondance qu'avec un des voisins du sommet 23 des données. Cette information n'est pourtant pas répercutée dans le noeud 5 de la décomposition, si bien que l'élagage de l'arbre de recherche de la méthode d'appariement n'est fait que partiellement. La prochaine section montre qu'il est possible, en changeant le mode de construction de la décomposition, de propager l'information externe sur l'ensemble de la décomposition, rendant ainsi efficace l'usage d'information externe.

7.4 Approche efficace.

Cette section présente une technique permettant de propager l'information externe dans l'ensemble de la décomposition d'un modèle, ce qui permet d'élaguer l'arbre de recherche du problème à toutes les étapes.

Pour cela, il faut se munir d'une nouvelle définition de l'union de deux graphes. Au sens habituel, l'union de deux graphes produit un graphe dont le nombre de sommets est égal à la somme des nombres de sommets des deux graphes de départ. L'union des ces deux ensembles est généralement écrite comme une union *ensembliste*, ce qui implique que l'on suppose que les deux ensembles de sommets des deux graphes sont disjoints. C'était l'objet de la remarque de la section 1.4.

Dans le cas des graphes AP, nous levons cette considération, en autorisant la fusion des noeuds AP portant les mêmes labels.

Union de deux graphes AP

L'union de deux graphes AP $G_1^{ap} = (V_1^s \cup V_1^{ap}, E_1)$ et $G_2^{ap} = (V_2^s \cup V_2^{ap}, E_2)$, relativement à l'ensemble d'arcs $E \subset V_1 \times V_2 \cup V_2 \times V_1$ est le graphe AP $G = (V, E)$ où $V = V_1^s \cup V_2^s \cup V_1^{ap} \cup V_2^{ap}$ où nous précisons que les deux ensembles V_1^{ap} et V_2^{ap} ne sont pas nécessairement disjoints : les noeuds AP de même labels sont fusionnés.

La figure 3.22 présente ainsi deux unions de graphes. L'union des deux graphes A et B, en ajoutant les arcs (2-4) et (3-5) produit le graphe D, et, de même, l'union des deux graphes C et B (sans arc supplémentaire) produit aussi de graphe D.

Avec cette définition, on voit que dans une décomposition, il est possible d'avoir deux antécédents d'un noeud qui partagent des sommets. Dans la mesure où la construction d'une décomposition se fait récursivement des grands graphes vers les petits, il devient donc possible, pour un graphe G , de transmettre à ses antécédents plus petits toute l'information externe nécessaire, ce qui n'était pas le cas dans le formalisme habituel.

Mais outre la possibilité de pouvoir partager l'information entre deux antécédents, il faut propager cette information vers les antécédents, ceci est possible grâce à une dernière notion. Ceci est précisé dans la suite.

Extension d'un graphe AP

Pour un graphe AP $G^{ap} = (V = V^s \cup V^{ap}, E)$, et l'un de ses sous-graphes $G_S^{ap} = (V_S = V_S^s \cup V_S^{ap}, E_S)$, on définit l'extension du graphe G_S^{ap} dans G^{ap} comme le sous-graphe de G^{ap} s'appuyant sur V_S et l'ensemble des voisins AP des sommets de V_S .

Formellement, on notera $E(G_S^{ap}) = G_{V_S \cup E(V_S)}^{ap}$ où $E(V_S) = \{v \in V^{ap} | \exists v' \in V^s, v \leftrightarrow v'\}$.

Si G_S^{ap} est un antécédent de G^{ap} , alors son extension lui ajoute l'ensemble des informations externes qui interviennent dans son appariement avec les données. Ainsi, en considérant que le graphe A de la figure 3.22 est un sous-graphe du graphe D de cette même figure, on voit que l'extension de A dans D est le graphe C. En effet, les sommets 4 et 5 du graphe D sont des sommets AP voisins du sous-graphe A, l'extension de A dans D consiste donc à rajouter ces deux sommets à A. Par ailleurs, si l'on se pose le problème de l'appariement du graphe D avec des données, et que l'on éclate ce

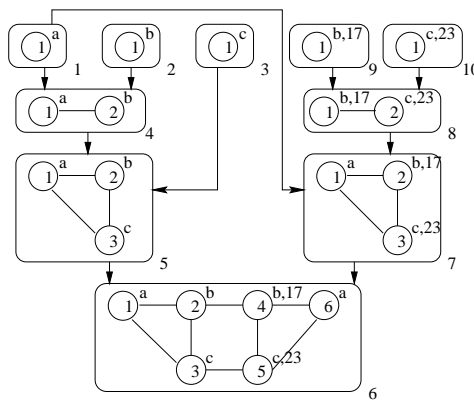


FIG. 3.21 – Décomposition simple du graphe portant une information externe.

graphe en deux sous-graphes A et B, que l'on tente d'apparier individuellement avant de fusionner les résultats d'appariement, il est clair que l'appariement de A produira une des solutions qui ne sont pas pertinentes au regard des informations externes. Par contre, si l'on tente d'apparier C (qui est l'extension de A dans D) et B, alors l'appariement de C se restreindra aux solutions compatibles avec les informations externes portées par les sommets 4 et 5 du graphe D.

Pour faire profiter la méthode de Messmer de cette technique, il est nécessaire de revoir le processus de construction de la décomposition d'un modèle, et d'adapter l'appariement proprement dit.

La figure 3.23 donne le pseudo-code de la décomposition d'un graphe AP dans une décomposition D.

Ainsi, en supposant que l'on utilise une décomposition contenant le graphe 5 de la figure 3.24, et qu'il faille lui ajouter le modèle avec connaissance a priori de la figure 3.20, alors l'algorithme se déroule comme suit :

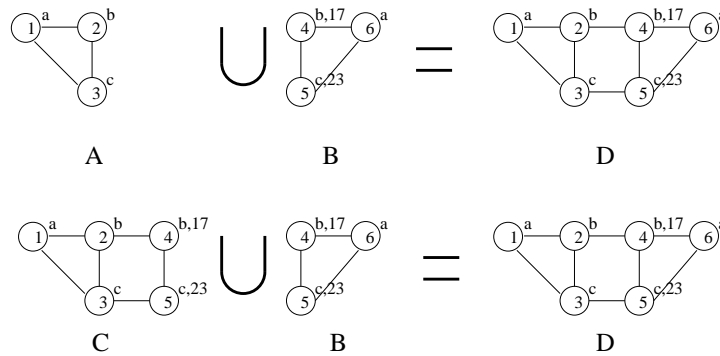


FIG. 3.22 – Exemples de graphes AP dont l'union produit le même graphe AP : l'union des graphes A et B, ainsi que l'union des graphes C et B produisent toutes deux le graphe D.

Algorithme de décomposition d'un graphe AP.

DECOMPOSE_AP(G, D)

1. Si G^{ap} n'a qu'un seul sommet strict, alors ajouter $(G^{ap}, \$, \$, \$)$ à D .
2. Dans D calculer le plus grand graphe S_{max}^{ap} qui soit sous-graphe de G^{ap} , et tel qu'un noeud de D s'écrive $(S_{max}^{ap}, G^{tap}, G''^{ap}, E)$.
3. Soit S_{max}^{tap} l'extension de S_{max}^{ap} dans G^{ap} .
4. Si S_{max}^{tap} est isomorphe à G^{ap} , alors poser $S_{max}^{tap} = H^{tap}$ où H^{tap} est un antécédent de S_{max}^{ap} . Sinon, ajouter S_{max}^{tap} à D .
5. Soit $S_{max}''^{ap} = E(G^{ap} - S_{max}^{tap})$.
6. Ajouter $S_{max}''^{ap}$ à D .
7. Ajouter noeud $(G^{ap}, S_{max}^{tap}, S_{max}''^{ap})$ à D .

FIG. 3.23 – Pseudo code de la décomposition d'un graphe AP.

- Recherche le plus grand sous-graphe du graphe : c'est le noeud 5 de la décomposition.
- Son extension est le graphe 9 de la décomposition.
- Il faut donc ajouter ce graphe 9, ce qui est fait comme suit :
 - Le plus grand sous-graphe du graphe 9 est le graphe 5, à nouveau.
 - Son extension est le graphe 9 lui-même, on considère un fils du graphe 5, comme le graphe 4.
 - L'extension du graphe 4 dans le graphe 9 est le graphe 10, qu'il faut ajouter :
 - Le plus grand sous-graphe du graphe 10 est le graphe 4 dont l'extension est le graphe 10 lui-même, donc on passe à un fils du graphe 4, par exemple le graphe 1.
 - L'extension du graphe 1 est le graphe 1 lui-même, rien n'est donc à faire.
 - L'extension du complémentaire du graphe 1 est le graphe 11 qu'il faut rajouter.
 - Le graphe 11 ne possède qu'un sommet à appairer. On l'ajoute.
 - L'extension du complémentaire du graphe 10 dans le graphe 9 est le graphe 12 qu'il faut ajouter. Comme il ne possède qu'un sommet à appairer, il est simplement ajouté.
- L'extension du complémentaire du graphe 9 dans le graphe 6 est le graphe 7 qui n'a qu'un noeud à appairer. Il est donc ajouté à la décomposition.

On notera que les graphes AP à un seul sommet *strict* sont l'équivalent des singletons dans les graphes standards.

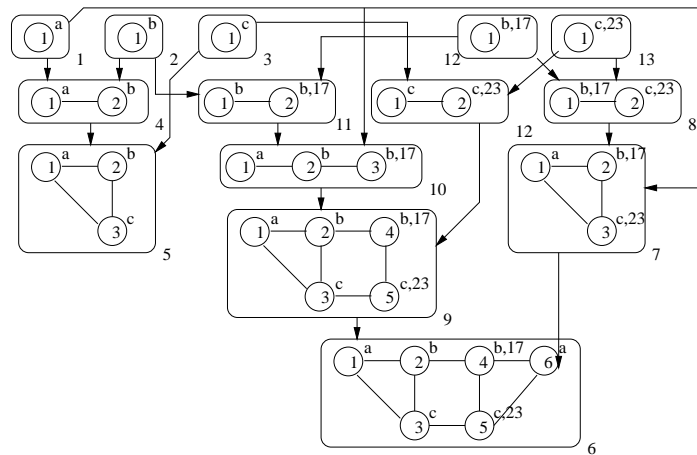


FIG. 3.24 – Décomposition efficace d'un graphe AP : l'information a priori est partagée au maximum entre les sous-graphes.

7.5 Appariement de graphes AP

Toutes les méthodes d'appariement proposées dans ce chapitre (méthode standard, heuristique stochastique, méthode combinée) s'adaptent naturellement aux graphes AP à l'aide de la procédure de décomposition proposée dans la section précédente.

Si d'un point de vue technique, l'adaptation doit être faite avec minutie, la procédure ne pose aucun problème d'ordre théorique. L'implémentation doit simplement veiller à ne pas appairer des

sommets AP, et à tenir compte du fait que l'union de deux graphes n'est pas disjointe, comme nous l'avons déjà dit.

La prochaine section proposera une analyse de la complexité de la méthode proposée.

7.6 Analyse de la complexité

Lorsque l'on considère une décomposition D ne contenant qu'un seul modèle, la méthode standard possède une complexité - dans le pire des cas - exprimée par $O(m^n n^3)$, où m est le nombre de sommets dans le graphe des données, et n est le nombre de sommets dans le graphe modèle. Dans le cas d'une connaissance a priori, certains sommets du modèle sont déjà appariés : ce sont les sommets sur lesquels porte cette connaissance. Le nombre de sommets à appairer est donc $q = n - |V^{ap}|$. La complexité dans le pire des cas devient donc : $O(m^q q^3)$.

Cette complexité est la complexité dans le pire des cas pour les deux approches proposées ici.

Pourtant, il est possible de montrer que l'approche dite «efficace» est justement plus efficace. Cette justification se fait en deux temps.

Premièrement, considérons le cas où le graphe des données n'est pas complet. Soit v_i un sommet du modèle G . Si v_i est voisin d'un sommet possédant une information d'appariement a priori, alors soit V'_i l'ensemble des sommets des données avec lesquels les voisins de v_i possédant une connaissance a priori sont appariés. Les appariements possibles pour v_i sont : le non appariement (§) ou un appariement avec un sommet des données voisin d'un des sommets de V'_i . Soit $m_i = |V'_i|$. Nous avons alors la complexité dans le pire des cas : $O(m_1 \dots m_q q^3)$. En fonction du degré de complétude du graphe des données, les valeurs des m_i peuvent être bien plus faibles que m , ce qui conduit à une baisse importante de la combinatoire du problème.

Deuxièmement, considérons le graphe des données réellement utilisé dans le système. Nous avons dit qu'il est complet. Le raisonnement tenu juste avant semble donc ne pas s'appliquer. Pourtant, nous allons montrer qu'il s'applique dans une certaine mesure. Les attributs des arcs sont évalués par des fonctions de compatibilités. Ces fonctions possèdent un certain nombre de seuils au-delà desquels les compatibilités sont fixées à $+\infty$. Ceci permet d'élaguer l'espace de recherche lorsque deux primitives ne respectent absolument pas une propriété géométrique recherchée. Typiquement, lorsque deux droites doivent s'intersecter et que leur distance réelle est supérieure à un seuil fixé à 5m, alors l'attribut correspondant à la propriété géométrique «intersection» vaut $+\infty$, et ces deux droites ne seront jamais appariées simultanément à deux sommets linéaires nécessitant cette propriété géométrique.

Ce procédé introduit donc dans le graphe complet une pseudo connexité : bien que tous les sommets soient voisins, il arrive que deux sommets voisins ne soient jamais appariés ensemble en raison de leur incompatibilité géométrique au regard de certaines propriétés requises par le modèle.

Ainsi, pour un sommet v_i du modèle et un de ses voisins AP, v'_i apparié au sommet w'_i des données, l'arc (v_i, v'_i) définit dans les données un sous-ensemble de sommets auquel v_i est potentiellement appariable, et ce sous-ensemble n'est pas le graphe des données dans sa totalité, mais seulement les sommets w_j voisins de w'_i tels que l'édition $\nu(v_i, v'_i) \rightarrow \nu(w_j, w'_j)$ possède un coût fini.

En résumé, le raisonnement sur les graphe non complets reste valide.

La complexité dans le pire des cas de la méthode efficace est donc celle annoncée.

7.7 Expérimentation

Dans cette expérimentation, nous utiliserons des données externes sans nous soucier de leur provenance (le prochain chapitre du document proposerons des moyens de produire de telles informations externes).

Le tableau 3.4 présente, dans un cas particulier, l'apport de l'utilisation d'informations externes sur les temps de calcul. Cet apport est indéniable.

Bâtiment	Modèle	Information externe	Temps d'appariement
4	VD	aucune	180s
4	VD	1 pan	2,0s
4	VD	1 autre pan	3,5s
4	VD	pan en V	3,6s
4	VD	2 pans	0,53s
4	VD	3 pans	0,09s

TAB. 3.4 – Temps d'exécution des appariements pour certains modèles, certaines données.

8 Conclusion

Dans ce chapitre, nous avons présenté le principe d'appariement de graphe qui permet d'identifier, dans les données, les primitives formant un sous-graphe le plus ressemblant à un modèle.

Dans la pratique, ce principe peut se formaliser par les distances d'édition, pour le calcul desquels une méthode déterministe de la littérature est utilisée.

Ce formalisme permet d'intégrer des fonctions décrivant la compatibilité entre les géométries des primitives et les propriétés géométriques souhaitées par les modèles. Ceci rend la procédure d'appariement «floue», et en tout cas tolérante aux erreurs des détections de primitives, et à leurs sous-détections.

Ce caractère «flou» est une adaptation naturelle de la méthode de calcul d'isomorphisme avec tolérance d'erreur au cas continu.

Nous avons également tenté de montrer qu'une approche de type stochastique peut se montrer plus efficace que la méthode déterministe. Ceci est principalement dû au fait que la méthode stochastique proposée permet d'obtenir rapidement des estimations du minimum de la fonction de coût.

Enfin, un formalisme dédié à l'intégration, dans le processus d'appariement, d'informations externes a été proposé. Ce formalisme propose notamment de tolérer des unions de graphes non disjointes (avec partage des sommets), ce qui est le point clé de l'efficacité de la méthode proposée, car cette singularité permet de propager l'information externe dans l'ensemble de la structure de données utilisée par les méthodes d'appariement.

Chapitre 4

Exploitation d'une grammaire de graphes.

La production des modèles de graphes est longue, fastidieuse, et semée d'embûches, car les fichiers décrivant ces modèles sont excessivement pénibles à écrire. Ceci nous donne un premier argument pour utiliser une grammaire de graphes : elle permet d'engendrer des graphes modèles de façon automatique, homogène, et sans erreur. Leur écriture est donc plus rapide. Par homogénéité, on entend le fait qu'il est très simple de semer de petites différences entre deux modèles ou deux parties de modèles lorsqu'on les écrit à la main. La grammaire assure que deux portions de modèles censées être identiques (isomorphes) le seront effectivement.

Par ailleurs, un tel outil est un outil de plus haut niveau que les modèles eux-mêmes. En effet, un modèle décrit une forme que nous recherchons, alors que la grammaire décrit comment on obtient cette forme. La grammaire synthétise donc mieux la connaissance du domaine d'application (les formes de bâtiments) que les modèles eux-mêmes.

Enfin, et nous le verrons, la grammaire mise en oeuvre permettra d'augmenter la performance de la méthode exposée dans le chapitre 3, car elle permet notamment d'introduire dans les décompositions des graphes l'information qui rend ces décompositions efficaces, et de piloter la reconstruction d'un modèle au moyen d'appariements partiels.

La section 2 présente le formalisme des grammaires de graphes fonctionnant par remplacement d'hyper-graphes. La section 3 présente la grammaire réellement utilisée dans le système aujourd'hui. La section 4 présente un outil, le réseau grammatical, permettant d'obtenir une décomposition (au sens de Messmer) des modèles (voir section 5), et de proposer une technique d'appariement exploitant la grammaire de graphe (cette technique est proposée en section 6).

1 Tour d'horizon des grammaires de graphes

Cette famille de formalisme n'est pas unifiée. Elle a fait l'objet d'un intérêt croissant depuis 30 ans, comme le montre la figure 4.1, qui représente l'histogramme des publications présenté sur un site bibliographique important (plus de 1000 références) :

<http://liinwww.ira.uka.de/bibliography/Theory/gragra.html>

2 Formalisme

Nous nous plaçons donc dans le cadre des grammaires de graphes agissant par remplacement d'hyper-graphes.

Cette section reprend le formalisme proposé par Annegret Habel ([Hab92]) pour mettre en place les grammaires de graphes par remplacement d'hyper-graphes. Les notions d'hyper-graphe, de sous-hypergraphe, et de morphisme ont été introduites au chapitre précédent. Nous introduisons ici les notions spécifiques aux grammaires.

2.1 Règle

On définit une production, ou une règle (pour une grammaire par remplacement d'hyper-graphes), par un triplet $p = (L, K, R)$ tel que :

- $K \subseteq L$: K est un sous-graphe de L . L est appelé le terme gauche de la règle.
- $K \subseteq R$: K est aussi un sous-graphe de R . R est appelé le terme droit de la règle.
- K est appelé le graphe de 'fixation' ou d'identification, ou encore noyau (Kernel en anglais).

Trois exemples de règles sont fournis en figure 4.2. Les règles proposées dans cette figure ne sont pas les règles du système. Ces dernières seront présentés dans les sections suivantes. Les règles de cette figure sont destinées à présenter les mécanismes de production de modèles (Cette production passe par l'application d'une règle sur un graphe. C'est l'objet de la prochaine section). On voit que les règles sont composées de 3 graphes L , K et R , séparés par des barres verticales. Des flèches relient les sommets de K à certains sommets de L et R . Elles représentent l'inclusion de K dans ces deux graphes, en identifiant la correspondance entre les sommets.

Dans ces règles, le label N indique les non-terminaux. Les graphes terminaux sont donc ceux ne comportant aucun sommet portant ce label.

La première règle (à gauche) produit un graphe à 4 sommets (dont deux non terminaux) à partir d'un axiome. La seconde étend un graphe, et la troisième effectue une 'terminaison' du graphe en remplaçant le non-terminal par des terminaux.

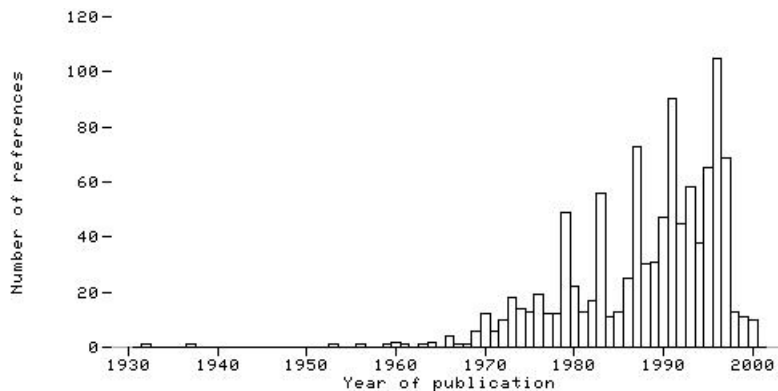


FIG. 4.1 – Histogramme des publications dans le domaine des grammaires de graphes

2.2 Application d'une règle

On applique une règle $p = (L, K, R)$ sur un graphe H de la façon suivante :

- Il faut sélectionner un morphisme h de L vers H . On notera que ce morphisme envoie aussi K sur H , ce qui permet d'identifier les nœuds d'attachement dans H .
- Il faut vérifier les conditions suivantes :
 - Si l'image $h(v)$ d'un sommet $v \in V_L$ est voisine d'un sommet de $E_H - E_{h(L)}$, alors $v \in v_K$. Ceci signifie que si deux sommets de H sont voisins, et que l'un d'eux est l'image d'un sommet de L et l'autre pas, alors le premier des deux sommets est en fait l'image d'un sommet de K . Ceci est fondamental, car par la suite, le graphe H sera privé de l'image de K par h . Il est donc important que les arcs ne soient pas coupés. Cette condition, quoique difficile à exprimer formellement, est facile à entrevoir par la figure 4.3. Sur cette figure, on voit un graphe H (l'ellipse de gauche), et trois de ses sommets. Deux sous-parties sont également identifiées dans : l'image de K ($h(K)$) et l'image de $L - K$ ($h(L - K)$). Deux arcs relient les 3 sommets présentés dans H . L'arc supérieur relie un sommet de $E - h(L)$ et un sommet de $h(L - K)$. L'autre arc relie le sommet de $E - h(L)$ et le sommet de $E - h(K)$. L'opération de suppression de $h(L - K)$ produit le graphe de droite de l'image. On voit qu'il est possible de conserver l'arc inférieur reliant le sommet de $E - h(K)$ au sommet de $h(K)$, mais l'arc supérieur reliant le sommet de $E - h(L)$ et le sommet de $h(L - K)$ est coupé. Cette coupure provoque une indétermination qu'il est préférable d'éviter. Ceci justifie cette condition.
 - Si deux objets (sommet ou arc) de L sont envoyés sur le même objet de H , alors ces deux objets sont dans K . Ceci est nécessaire pour les mêmes raisons que précédemment : lors de la suppression de $h(L - K)$, il ne faut pas supprimer d'objet de $h(K)$, ce qui pourrait se produire justement si un objet de K et un objet de $L - K$ se projettent tous deux sur le même objet de H .
- Définir le graphe $D = H - (h(L) - h(K))$. D est le graphe H privé de la partie à remplacer du terme gauche de la règle.
- Ajouter à D la partie du terme droit à rajouter : $X = D + (R - K)$ tel que :
 - $V_X = V_D + (V_R - V_K)$
 - $E_X = E_D + (E_R - E_K)$
 - Les labels des sommets sont conservés.
 - Les labels des arcs de E_D sont conservés.
 - Les arcs possédant leurs sources et cibles dans $V_R - V_K$ sont conservés.
 - Les arcs possédant une partie de leurs sources et cibles dans V_R et l'autre dans $V_R - V_K$ sont

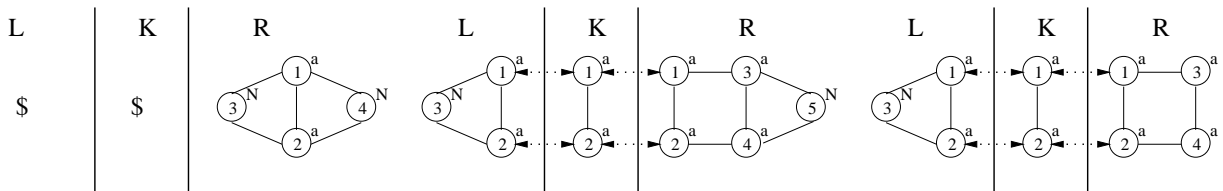


FIG. 4.2 – Exemples de règles par remplacement d'hyper-graphe.

transposés au nouveau graphe :

Pour $e \in E_R - E_K$, si une de ses sources (ou cible) est dans $V_R - V_K$, alors la nouvelle source est celle qui a été transposée dans V_X , si elle est dans V_K , alors cette nouvelle source est celle qui existait déjà dans D .

On notera une telle application : $H \Rightarrow_p H'$.

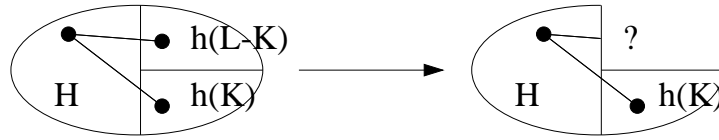


FIG. 4.3 – Détails de la justification de la première condition d'application d'une règle grammaticale.

L'identification proposée par $h : L \rightarrow H$ n'est pas nécessairement unique. En effet il y a au moins autant d'identifications possibles que de sous-graphe de H isomorphes à L .

Le processus mis en place pour produire H' à partir de H fait appel à 6 graphes au total, qui respectent 7 morphismes synthétisés dans la figure 4.4. Une explicitation des morphismes entre ces graphes est explicitée en figure 4.5. Cette figure illustre plus précisément l'application de la seconde règle de la figure 4.2, sur le graphe R de la première règle de cette même figure :

- Le morphisme h est présenté par les 3 flèches courbes liant L à H .
- La suppression de $h(L - R)$ produit le graphe D . On notera qu'aucun arc de h ne relie de sommet de $H - h(L)$ à $h(L - K)$.

Les arcs pointillés de couleur forment des quadrilatères qui permettent d'identifier les sommets qui se correspondent dans les graphes L, K, H, D , ou L, R, D, H' . Ces quadruplets de sommets sont la réalisation au niveau des sommets des quadruplets de morphismes exprimés dans la figure 4.4. Ceci met en évidence la vie des sommets lors d'une production à l'aide d'une règle.

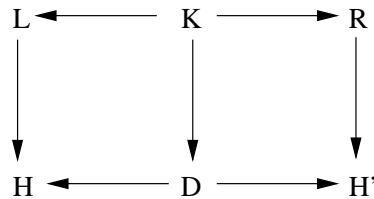


FIG. 4.4 – Résumé des morphismes liant les graphes intervenant dans l'application d'une règle.

2.3 Grammaire d'hyper-graphes

Une grammaire d'hyper-graphes est un quadruplet $HG = (N, T, P, \$)$ où : N est le sous-ensemble des sommets non-terminaux, et T est l'ensemble des sommets terminaux, P est un ensemble fini de règles de production, et $\$$ est l'axiome de la grammaire.

Le langage engendré par la grammaire est noté $L(HG)$ et est constitué de l'ensemble des hyper-graphes terminaux qui peuvent être dérivés de $\$$ par P .

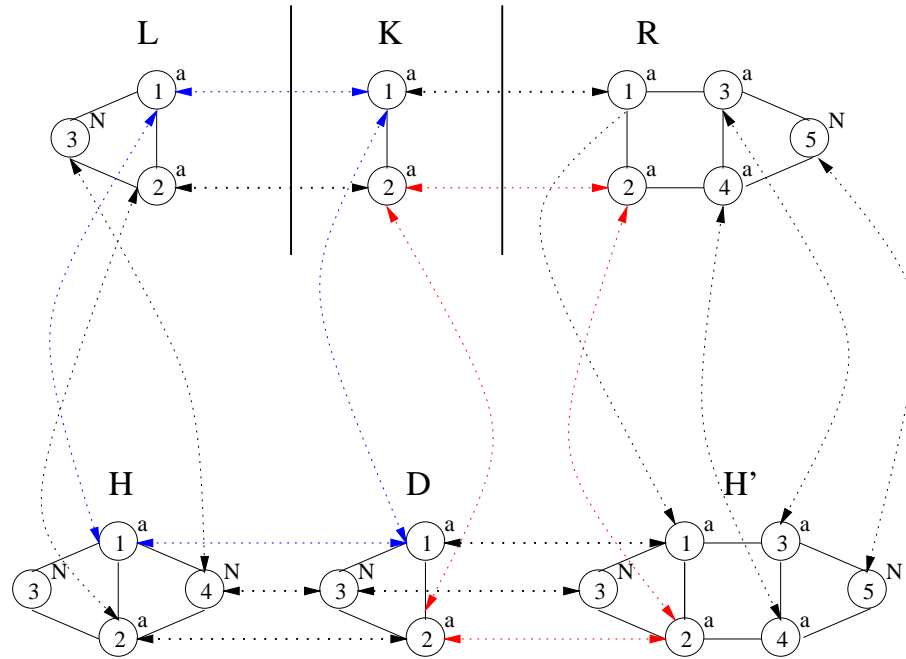


FIG. 4.5 – Explication de l’application d’une règle pour la grammaire de la figure 4.2.

A titre d’exemple, la figure 4.6 présente les 3 plus petits graphes terminaux du langage de la grammaire de la figure 4.2.

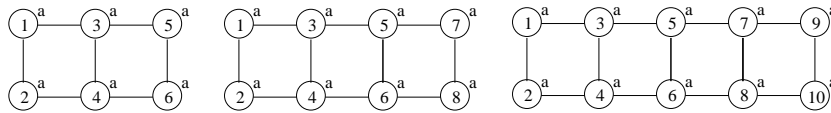


FIG. 4.6 – Les 3 plus petits graphes terminaux produits par les règles de la figure 4.2.

3 La grammaire du système

3.1 Engendrer les modèles de bâtiments

Il est nécessaire d’expliciter quelque peu la grammaire réellement utilisée dans le système. Pour ceci, nous n’utiliserons pas les vues théoriques des graphes, comme dans la section précédente, car les graphes sont trop complexes pour être aisément représentés. Nous utiliserons les vues 2-D, plus adaptées à la compréhension du mécanisme de production des modèles. La figure 4.7 présente 4 règles permettant de produire des bâtiments à 2 pans, 3, ou 4 pans, en L à 4, 5 ou 6 pans, un U à 6, 7, ou 8 pans, ainsi de suite. Cette figure ne présente pas toutes les règles de la grammaire. D’autres règles engendrant des toits à 1 pan existent en effet, ainsi qu’une règle permettant de biseauter un angle de toit.

On remarquera que ces règles correspondent assez précisément aux «composants de bâtiments» introduits par Steinhage ([Ste97]). Cet auteur distingue en effet des éléments terminaux et «connectants» des bâtiments. La correspondance directe avec les règles est claire. Les éléments terminaux correspondent aux règles dont le terme droit ne contient pas de sommet terminal, et les «connectants» correspondent aux autres règles.

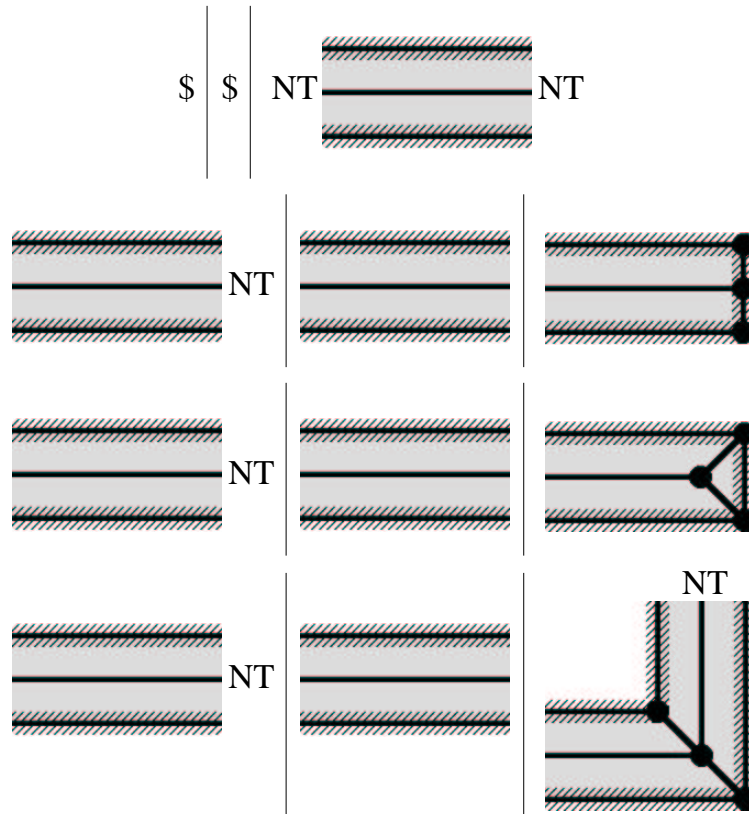


FIG. 4.7 – Des règles de la grammaire réelle

3.2 Un cas particulier : les relations ternaires

3.2.1 Rappel

Nous avons déjà explicité le fait que certaines relations ternaires étaient intéressantes pour représenter des relations géométriques de type «est entre». Dans le cas de la première règle de la figure 4.7, il est possible d'introduire d'emblée des relations de ce type entre les sommets du terme droit de la grammaire, notamment la relation entre le faîte et ses gouttières.

Toutefois, il peut se montrer intéressant d'introduire également dans les modèles finaux des relations triples stipulant que certains pans de toits sont entre les extrémités du bâtiment. Typiquement dans le modèle D (voir figure 4.8), il est souhaitable de savoir exprimer ce fait, comme dans la figure 4.9. Ce type de relation peut s'établir entre un pan de toit et les primitives linéaires ou discontinuités

de ses extrémités.

3.2.2 Mécanisme de mise en place

Il est possible d'introduire ces relations par l'intermédiaire de la grammaire, en modifiant les règles de base, et en ajoutant trois règles dédiées à la mise en place de ces arcs ternaires.

L'idée de base est de faire persister la notion d'extrémité du bâtiment, comme le font les sommets non terminaux dans la règle 1 de la figure 4.7. Les règles réelles perdent cette notion lorsque les non terminaux s'effacent. En ajoutant au terme droit de la règle 1 un graphe dédié comme celui présenté en figure 4.10, la notion d'extrémité des pans de bâtiment persiste grâce à la présence de deux nouveaux sommets carroyés matérialisant ces extrémités. La règle d'ajout d'une extrémité doit alors être elle aussi adaptée, comme en figure 4.11.

Enfin, une règle comme celle présentée en figure 4.12 met réellement en place la relation ternaire

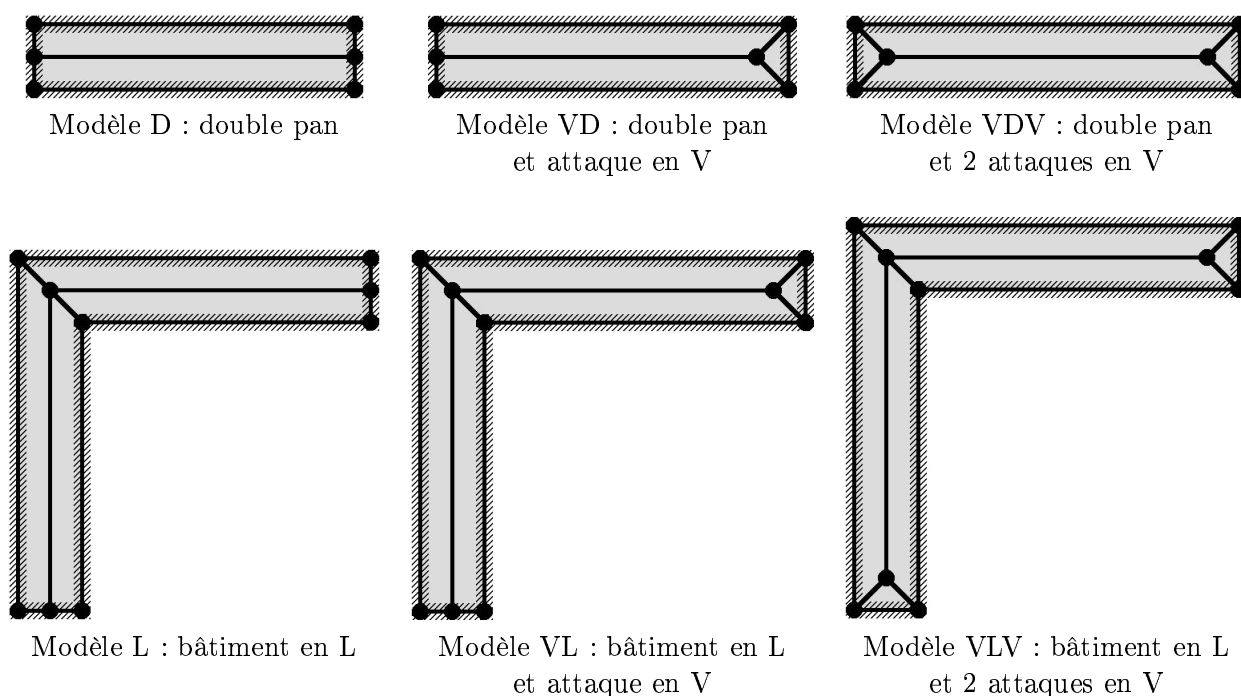


FIG. 4.8 – Les modèles principaux engendrés par la grammaire 4.7.

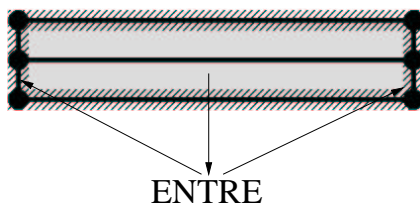


FIG. 4.9 – Une relation souhaitable pour le modèle D

souhaitée. Il y a bien entendu 3 règles de ce type : une associant un plan et 2 sommets linéaires, une associant un plan et deux sommets de type discontinuité, et une associant un plan, un sommet de type linéaire et un de type discontinuité.

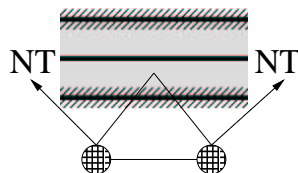


FIG. 4.10 – Une relation souhaitable pour le modèle D

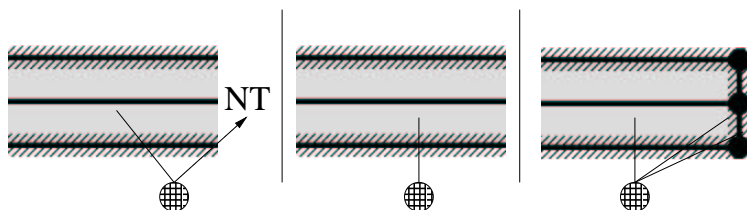


FIG. 4.11 – Une règle adaptée à la mise en place de règles ternaires impliquant les extrémités d'un bâtiment.

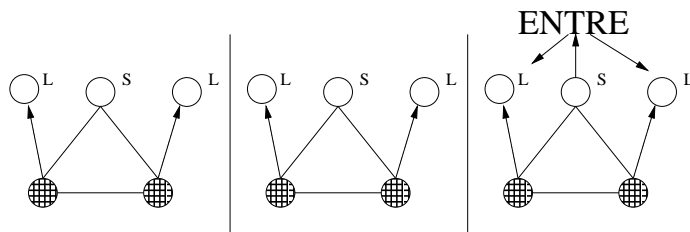


FIG. 4.12 – Une règle de mise en place de relation ternaire impliquant les extrémités d'un bâtiment.

3.2.3 Intérêt

Nous allons montrer par un exemple concret l'intérêt de l'introduction de ces relations triples reliant les extrémités des bâtiments. Considérons la zone 45 de la scène. Elle est composée de deux bâtiments représentables par le modèle D. Un premier appariement produit une reconstruction du bâtiment le plus au sud des deux. Un second appariement peut alors se focaliser (nous verrons comment un appariement peut se focaliser sur une autre région de la zone *après* un premier appariement), sur le second bâtiment de la zone.

Si le modèle D n'est pas muni de relations triples reliant les pans et les extrémités du bâtiment, alors l'appariement produit est montré en figure 4.13, ce qui mène à la reconstruction de cette figure, à droite. Cette reconstruction est dégénérée car en fait elle est repliée sur elle-même et localisée à

l'est du bâtiment. Ceci provient du fait que l'appariement a identifié en réalité deux arêtes de rive à l'est du bâtiment, et une discontinuité correspondant à l'autre extrémité du bâtiment, mais elle aussi située à l'est. Ainsi, les deux extrémités se trouvent positionnées à l'est du bâtiment, ce qui conduit naturellement à cette dégénérescence.

Par contre, si le modèle est muni de ces relations triples, alors l'appariement produit est celui présenté en figure 4.14, où l'on voit que trois discontinuités ont été appariées, dont une à l'est et une autre à l'ouest. De ce fait les extrémités du bâtiment sont bien reconstruites *de part et d'autre* du bâtiment, comme le montre l'image de droite de cette figure.

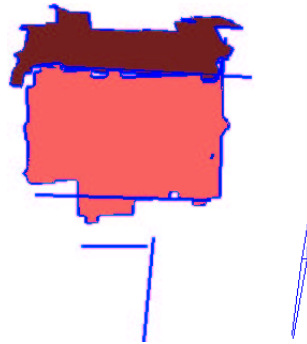


FIG. 4.13 – Appariement du modèle D muni, la zone 45 (à gauche), et sa reconstruction à droite.

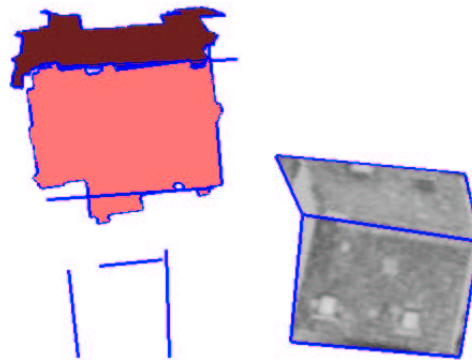


FIG. 4.14 – Appariement du modèle D muni des relations triples entre pans et extrémités, pour la zone 45.

4 Réseau grammatical

4.1 Généralités

La grammaire proposée est ambiguë : certains modèles peuvent être engendrés par la grammaire de plusieurs façons.

Les différentes façons de produire ces modèles peuvent être représentées dans une structure de réseau, dans lequel les nœuds sont des graphes modèles, et les arcs des applications de règles grammaticales.

4.2 Exemple

Les différentes façons possibles de construire le modèle LV (en bas du réseau) correspondent aux différents chemins permettant de passer de l'axiome (noté \$, en haut du réseau) au modèle LV lui-même, en suivant les flèches. Un rapide examen du réseau permet de voir qu'il y a 4 chemins possibles¹, et donc 4 façons de produire ce graphe.

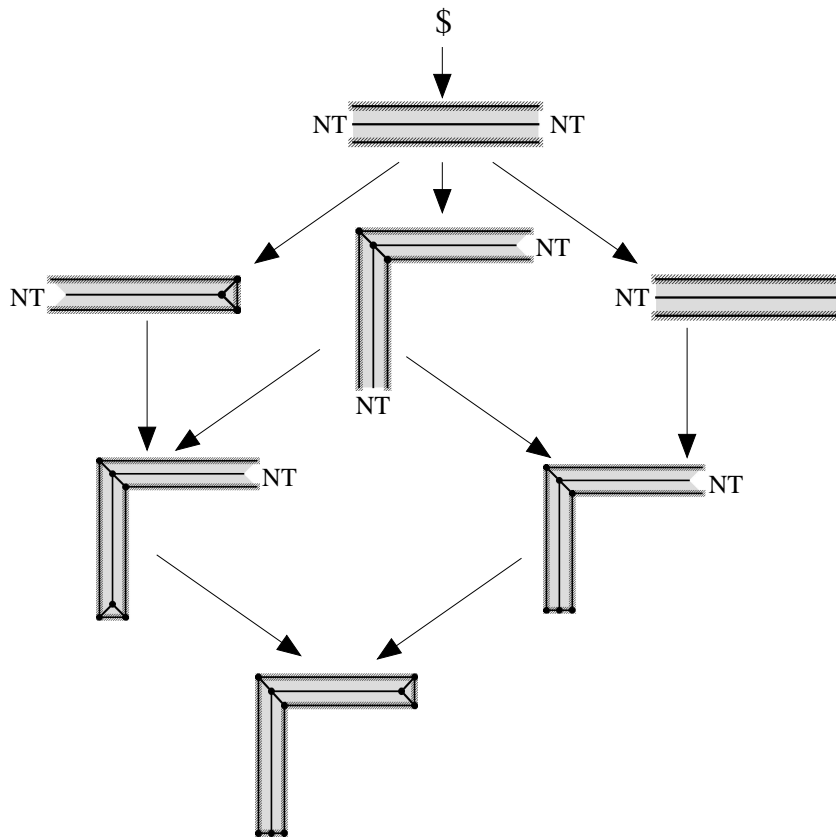


FIG. 4.15 – Le réseau grammatical du modèle LV.

4.3 Applications équivalentes

Il faut noter que pour passer du modèle de la seconde ligne du réseau, qui comporte 2 sommets non terminaux, il y a en fait 4 façons de passer à chacun des modèles de la 3^{ème} ligne. En effet, la

¹pour chacun des trois modèles de la 3^{ème} ligne du réseau, il n'y a qu'un chemin, et les deux modèles de la 4^{ème} ligne, il y en a 2, et donc 4 pour le modèle LV

seconde règle de la figure 4.7 peut s'appliquer en fait aux deux extrémités de ce modèle, et pour chacune des extrémités, il y a deux façons d'appliquer la règle, car en raison de la symétrie interne du modèle, l'identification au moment de l'application de la règle produit deux solutions.

Il se trouve que ces deux solutions produisent deux graphes isomorphes, ce qui nous conduit à proposer la notion d'application équivalentes (de règle) :

Deux applications d'une règle sont dites équivalentes si les deux graphes H' et H'' produits par ces deux applications sont tels qu'il existe un isomorphisme de H' vers H'' , qui laisse leur sous-graphe $H - h(L - K)$ invariant.

Dans le réseau grammatical, nous devrions donc en fait représenter autant d'arcs qu'il y a d'applications non 2 à 2 équivalentes, mais dans ce cas précis le nombre d'applications non deux à deux équivalentes est toujours égal à 1.

L'objectif de cette notion est d'éviter d'exploiter plusieurs fois les mêmes informations a priori au moment des appariements qui seront proposés plus tard.

5 Décomposition

La grammaire permet de proposer une décomposition (au sens de Messmer) du modèle. Plus exactement, elle permet de proposer plusieurs décompositions du modèle.

En effet, pour un modèle donné G , il y a N chemins permettant de produire le modèle. Chacun de ces chemins est une séquence (G_1, \dots, G_M) , où $G_1 = \$$ et $G_M = G$.

En introduisant successivement les $G_i, i = 2, \dots, M$, privés de leurs sommets non terminaux, dans une décomposition, nous obtenons une décomposition qui est adaptée au modèle G .

Cette décomposition est adaptée dans la mesure où, par construction, les modèles G_i cités ici sont en fait des sous-graphes du modèle G , de taille croissante. Leur introduction dans la décomposition permet donc d'introduire dans cette décomposition la manière dont G est construit étape par étape : la connaissance du domaine est donc introduite dans celle-ci. La différence avec les décompositions adaptées vues en section 5.2.2 est que dans cette section nous avons introduit à la main les sous-graphes fréquemment présents dans les modèles. Ici, par contre, ces sous-graphes sont tirés indirectement des règles grammaticales, qui deviennent alors la seule source contenant la connaissance des formes.

Il y a N décompositions possibles. Il aurait été judicieux de savoir dans quelles mesures elles sont équivalentes ou non, voire laquelle est la plus efficace. Cette étude n'a pas été menée. Dans la pratique, nous nous cantonnons donc à choisir arbitrairement *une* des décompositions possibles proposées par le réseau grammatical, donc par la grammaire.

L'intérêt majeur de la grammaire n'est toutefois pas de proposer une décomposition pour un modèle. La section suivante montrera qu'il est possible d'exploiter la grammaire pour effectuer des appariements progressifs en exploitation l'outil développé au chapitre précédent, concernant l'appariement avec connaissance a priori.

6 Appariement progressif

6.1 Réseau et modèles

Le lecteur aura perçu que le réseau grammatical est très similaire à une décomposition au sens de Messmer. D'une part, il a la forme d'un réseau, et d'autre part, il peut être vu en couches de sous-modèles et modèles de plus en plus complexes.

Par contre il y a une différence fondamentale entre la décomposition de Messmer et le réseau grammatical utilisé ici : dans la décomposition, chaque nœud du réseau (chaque graphe) possède deux parents exactement, est l'union (au sens des graphes) de ces deux parents, alors que dans ce réseau grammatical, chaque nœud (chaque graphe) possède un nombre quelconque de parents, et est le résultat d'applications de règles grammaticales sur ces parents.

Messmer propose d'introduire plusieurs modèles dans sa décomposition, et d'arrêter l'appariement dès qu'un meilleur appariement a été trouvé pour l'un des modèles, ce modèle étant alors désigné comme meilleur modèle pour la reconstruction. Or nous avons déjà montré que ce principe n'était pas valable dans notre cas particulier, car il favorisait les modèles les plus petits dans notre contexte, et c'est la raison pour laquelle nous travaillons toujours avec un seul modèle par décomposition.

En revanche, l'introduction de plusieurs modèles est aussi possible, et pertinente, dans le cas du réseau grammatical, car elle permet, comme nous allons le montrer, de tirer à nouveau profit des sous-graphes communs aux modèles. Ainsi, la figure 4.16 présente un réseau grammatical dans lequel les modèles L, LV, et VLV sont représentés simultanément dans ce réseau, sans en augmenter sensiblement la complexité.

Nous avons déjà vu en section 6 que malgré l'efficacité de la méthode proposée par Messmer, et malgré une heuristique stochastique, la complexité de l'appariement devenait assez rapidement trop importante pour être raisonnable.

Nous allons donc proposer une méthode qui combine le réseau grammatical et les travaux d'appariement avec connaissance a priori vus en section 7 du chapitre 3 pour appairer efficacement plusieurs modèles voisins, tels que les modèles V, VL et VLV.

6.2 Appariement multi-modèles

L'idée de cette méthode est à nouveau similaire à celle proposée par Messmer, dans laquelle les hypothèses d'appariements de deux parents sont combinées pour proposer une hypothèse d'appariement pour le fils.

Dans notre cas le point clé est de fournir à l'aide du meilleur appariement d'un père, une information pertinente a priori concernant l'appariement du fils.

Cet appariement est réalisé en utilisant la fonction $NET_AP(n)$ décrite dans la figure 4.1 pour chaque nœud n du réseau grammatical.

L'extraction de l'information a priori utilisée dans ce pseudo code est réalisée de la façon suivante : considérons un nœud n et un de ses pères n_i . Si le graphe principal G_i de n_i a été apparié au cours d'un précédent appel à $NET_AP(n_i, G_I)$, produisant ainsi un appariement $f_i = (\Delta_i, f_{\Delta,i})$, alors considérons un plus grand sous-graphe commun à G et G_I : $MCS(G, G_I)$. Soit alors f'_i la restriction de f_i à $MCS(G, G_I)$. Cette restriction f'_i identifie précisément l'information a priori souhaitée pour l'appariement de G et G_I . Nous notons cette information a priori $priori(n, n_i, f_i)$.

La fonction *NET_AP* parcourt en apparence plusieurs fois les nœuds du réseau puisqu'elle contient un appel récursif à elle-même, portant sur les pères de chaque nœud. Mais en réalité, l'appariement n'est réalisé qu'une seule fois sur chaque nœud du réseau, car lorsque la fonction a déjà été appelée sur un nœud, ce fait est mémorisé, et évite à la fonction de s'appliquer une seconde fois (point 2 du pseudo code).

Malheureusement, dans la pratique, cette méthode ne fonctionne guère, car pour les sous-modèles très simples, les meilleurs appariements au sens de la distance d'édition ne sont pas toujours les meilleurs au sens d'une reconstruction 3D, ce qui brouille les cartes d'entrée de jeu. Il convient donc, au moins pour les modèles simples, d'effectuer tout de même un appariement standard, c'est-à-dire sans information a priori.

Malheureusement, la limite entre sous-modèles simples et sous-modèles non simple ne peut pas être fixée. L'appariement standard doit donc avoir lieu pour tous les modèles du réseau. Ceci produit donc une méthode plus complexe que les appariements sans connaissance a priori, car elle fait plus de calculs que les appariements sans connaissance a priori, tout en prétendant exploiter la grammaire de graphe!

Cette perte dramatique est en fait contrée par la façon dont les appariements sont réalisés. Plus

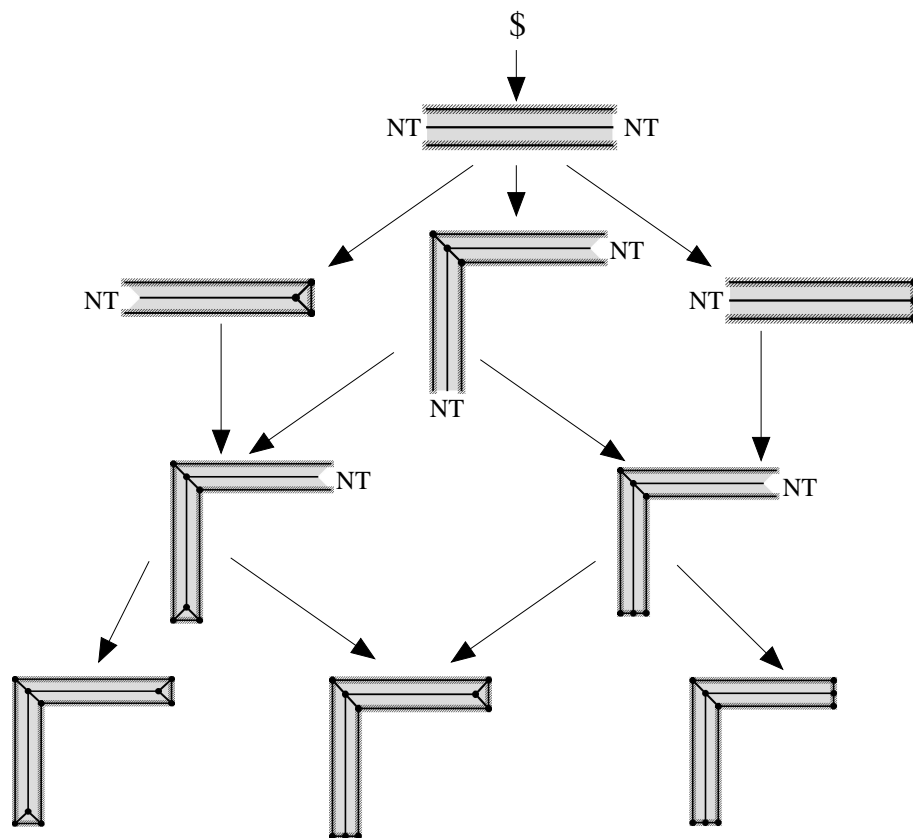


FIG. 4.16 – Le réseau grammatical du modèle LV.

Fonction $NET_AP(n, G_I)$.

1. Si n est l'axiome, ne rien faire, donc sortir de la fonction.
2. Si NET_AP a déjà été appliquée à n , retourner le résultat déjà calculé.
3. Pour chaque père n_1, \dots, n_p de n , faire :
 - faire $f_i = NET_AP(n_i)$.
 - extraire l'information d'appariement de n_i utile comme information a priori pour n : $priori(n, n_i, f_i)$
 - Attribuer à G cette information a priori. Ceci produit un graphe G^{ap}
 - appairer G^{ap} et les données G_I en utilisant la décomposition produite par $DECOMPOSE_AP(G^{ap}, D)$. Ceci produit un appariement f_i .
4. Parmi les appariements f_i a priori réalisés, retourner le meilleur.

TAB. 4.1 – Pseudo code d'appariement d'un nœud de réseau grammatical

exactement, dans la mesure où les appariements des modèles simples sont rapides, et, où, dans le cas des modèles complexes, l'information majeure est elle-même récursivement tirée d'appariements préalables de sous-modèles, il s'avère qu'à tous les étages du réseau les problèmes sont en fait de complexité modérée, ce qui permet, d'un point de vue opérationnel, d'effectuer les appariements en temps limité. Nous proposons donc d'effectuer tous les appariements cités dans ce pseudo code EN TEMPS LIMITÉ. L'appariement en temps limité consiste à stopper les passes d'appariement de la fonction d'appariement AM au bout d'un temps de calcul t_m , et à extraire des listes d'appariements (ouvertes et fermées) les meilleurs appariements stockés.

La durée d'appariement autorisée pour les appariements est la même pour tous, car nous ne saurions pas exprimer une durée variable en fonction de la complexité des modèles. Dans le cas d'une durée limitée pour tous les modèles, il est raisonnable de penser que si l'appariement brutal d'un nœud du réseau n'a pas fonctionné (i.e. n'a proposé aucune hypothèse d'appariement), alors dans la même durée, l'appariement brutal de ses fils, qui sont - rappelons le - plus complexes, ne fonctionnera pas non plus. Il est donc inutile d'appairer brutalement un nœud du réseau si aucun appariement brutal de ses fils n'a proposé d'hypothèse d'appariement.

Ces considérations conduisent à une version modifiée du pseudo-code présentée par la figure 4.2

Les temps de calcul pour un modèle G sont bornés par $(V_r + E_r)t_m$ où V_r est le nombre de nœuds du réseau grammatical contenant G , et E_r est le nombre de ses arcs. On voit que pour un réseau grammatical contenant plusieurs modèles, comme le réseau de la figure 4.16, le fait d'appairer plusieurs modèles n'est pas nécessairement beaucoup plus coûteux que le fait d'appairer un seul modèle, car l'ajout de quelques modèles ne fait qu'ajouter quelques nœuds et arc dans le réseau grammatical.

Fonction $NET_AP2(n)$.

1. Si n est l'axiome, ne rien faire, donc sortir de la fonction.
2. Si NET_AP2 a déjà été appliquée à n , retourner le résultat déjà calculé.
3. Pour chaque père n_1, \dots, n_p de n , faire :
 - faire $f_i = APP_AP2(n_i)$.
 - extraire l'information d'appariement de n_i utile comme information a priori pour n : $priori(n, n_i, f_i)$
 - Attribuer à G cette information a priori. Ceci produit un graphe G^{ap}
 - appairer en temps limité t_m , G^{ap} et les données G_I , en utilisant la décomposition produite par $DECOMPOSE_AP(G^{ap}, D)$. Ceci produit un appariement f_i .
4. Parmi les appariements f_i a priori réalisés, soit f le meilleur appariement trouvé.
5. Si au moins un des appariements brutaux $AM(n_i, G_I)$ des pères de n a fonctionné, alors :
 - appairer brutalement n , (i.e. produire une décomposition de n et l'appairer avec la méthode mixte $AM(n, G_I)$) en temps limité t_m .
 - si cet appariement a produit un résultat, f' , et que f' est meilleur que f , retourner f' .
6. retourner f

TAB. 4.2 – Pseudo code d'appariement d'un nœud de réseau grammatical

7 Conclusion

Ce chapitre propose d'engendrer les modèles utilisés à l'aide d'une grammaire de graphes. Cette grammaire devient alors *la* connaissance déclarative introduite dans le système, et les modèles en découlent.

De l'exploitation de cette grammaire de graphes, il découle une procédure d'appariement NET_AP2 qui exploite la grammaire pour appairer un modèle et les données G_I en effectuant des appariements partiels des modèles, et en exploitant ces appariements pour produire une information a priori pour l'appariement du modèle lui-même.

Enfin, d'un point de vue opérationnel, un fait important a été introduit : tous les appariements mis en oeuvre dans cette exploitation d'une grammaire se déroulent en temps limité, ce qui borne a priori le temps de calcul.

Nous n'avons pas proposé de partie dédiée à la mise en oeuvre de cette algorithmie dans ce chapitre, car toutes les reconstructions faites dans la suite du mémoire, et notamment toutes les reconstructions proposées dans le chapitre consacré à l'évaluation, seront réalisées avec la fonction

d'appariement proposée ici.

Chapitre 5

Vers une reconstruction

1 Introduction

Les chapitres précédents décrivent la détection de primitives, la façon dont nous formalisons les modèles à l'aide de graphes, comment il est possible d'apparier un modèle et des données afin de produire une reconstruction de la zone s'appuyant sur un modèle. Il subsiste toutefois deux questions majeures auxquelles nous n'avons pas répondu, et qui sont intimement liées : définir la *meilleure* reconstruction pour une zone donnée, et, le cas échéant poursuivre la reconstruction si aucun modèle ne *suffit* à représenter la zone complète. A cet effet, nous introduirons des *combinaisons* de reconstructions.

Par ailleurs, et dans l'absolu, le coût d'édition n'est pas nécessairement une bonne mesure de la qualité de reconstruction, car il mesure l'écart entre le modèle et des données (les primitives) qui sont issues de traitements, et donc entachées de bien plus d'erreurs que les données initiales que sont le MNE et les images elles-mêmes. La qualification des reconstructions fera donc un retour aux données initiales, et plus précisément, elle exploitera le MNE.

Nous proposons donc dans ce chapitre :

- Une analyse des qualités que doivent posséder les reconstructions (section 3).
- Deux procédés permettant de produire plusieurs hypothèses de reconstruction.
- Une mesure de qualité de reconstruction.
- Une procédé d'optimisation visant à atteindre l'optimum de la mesure de qualité définie.

2 Le cas des repliements

Nous rappelons que les toits recherchés ne peuvent pas se «replier» sur eux-mêmes. Plus précisément, ils représentent des surfaces du type $z = f(x, y)$. Ainsi, les projections au sol des portions de plan reconstruites ne doivent pas s'intersecter. Mais l'appariement ne garantit pas que la reconstruction obtenue après appariement vérifie ces propriétés.

Il faut donc vérifier cette propriété a posteriori, afin de rejeter les reconstructions «repliées».

A titre d'exemple, la figure E.8 montre un modèle à 3 pans, ainsi que l'appariement résultant sur les données du bâtiment 19. On constate que les primitives surfaciques appariées ne s'intersectent pas. Pourtant, la reconstruction déduite de cet appariement (les deux dernières images de la figure) montre

clairement qu'en projection au sol, le plus petit des 3 pans intersecte clairement les pans principaux du toit, violant ainsi la modélisation. Cette reconstruction n'est clairement pas admissible.

De même, nous n'autoriserons par les pans de toits trop verticaux car ils correspondent aussi à des dégénérescences de reconstruction. La figure C.9 montre un tel exemple pour le bâtiment 172 : lorsqu'une reconstruction à partir d'une modèle à 4 pans est tentée, deux dégénérescences sont produites. Elles sont facilement filtrées à l'aide du critère de pentes.

Nous supposerons dorénavant que les reconstructions utilisées ne comportent ni repliement, ni pans verticaux. Dans la pratique il suffit d'éliminer les reconstructions repliées ou possédant des pans verticaux pour satisfaire cette hypothèse.

3 Analyse

3.1 Rappels de la stratégie

Nous rappelons ici que nous ne disposons pas d'un prédicteur de modèles, qui permettrait, à partir des données initiales, de sélectionner, parmi les modèles possibles, des candidats particuliers. Cette stratégie est envisageable mais n'a pas été abordée.

Dans l'état actuel des travaux, il est donc possible de tester «tous»¹ les modèles disponibles, de les reconstruire, puis tenter une sélection pertinente de la meilleure reconstruction. Par ailleurs, dans la mesure où des zones sont composées de plusieurs bâtiments, nous combinerons des reconstructions utilisant certains modèles afin de reconstruire une zone. Nous appellerons donc combinaison un groupe de reconstructions individuelles, chaque reconstruction individuelle s'appuyant sur une seule instance d'un modèle.

3.2 Contexte

Nous rappelons que le MNE utilisé dans les données initiales est une donnée assez fiable et relativement précise. Il sera utilisé comme indicateur fondamental de la pertinence des combinaisons au regard des données. Par ailleurs, et c'est ce que tend à montrer ce chapitre, l'analyse des MNEs permet à elle seule d'extraire des informations pertinentes pour évaluer les reconstructions et leur combinaison.

Nous avons toutefois conscience que les images initiales devraient être intégrées dans ce processus d'évaluation, car elles constituent les données les plus en amont du processus, le MNE étant lui-même calculé grâce à ces images². Dans cet ordre d'idée, des travaux permettant la comparaison de données 3-D et d'images 2-D existent ([BW]), mais n'ont pas été étudiés.

Afin de comparer les données produites par les reconstructions, et le MNE initial d'une zone, il est naturel de produire un MNE dit «Reconstruit» à partir d'une reconstruction. La prochaine section précise les deux MNEs réellement comparés.

¹En fait ce n'est pas vrai. Dans la pratique, nous limiterons la complexité des modèles testés. La famille des modèles qui peuvent être engendrés par la grammaire est naturellement infinie!

²Dans un contexte utilisant des MNEs obtenus de manière active cette remarque serait à discuter.

3.3 Les 2 MNEs considérés

Le MNE en entrée

Il faut distinguer le MNE brut (la donnée brute en entrée, sur laquelle on s'est focalisé), du MNE réellement utilisé dans le processus. En effet, le MNE brut est travaillé afin d'être plus propre. Cette étape a été décrite au chapitre 2.

A titre d'exemple, la figure C.1 présente un des bâtiments souvent vu dans nos travaux, numéroté 172, ainsi que le MNE associé en niveau de gris et sur une palette de couleurs. La figure C.2 présente le MNE retravaillé (en couleur).

On dénommera donc MNE le MNE retravaillé. Pour un pixel p , nous noterons $h(p)$ son altitude.

Le MNE Reconstruit

Une reconstruction d'une partie de la zone par un modèle donné permet de produire un autre MNE, que l'on notera MNER (R pour Reconstruction). La figure C.3 représente une projection d'une reconstruction pour le bâtiment 172 (à l'aide d'un modèle de double pan), ainsi que le MNER déduit de cette reconstruction.

Les trois figures D.1, D.2 et D.4 présentent les mêmes éléments, pour une zone plus complexe (zone 22).

Pour un MNER appuyée sur une reconstruction r , nous noterons $h_r(p)$ l'altitude d'un pixel de ce MNE.

3.4 Evaluation d'une combinaison : analyse

Il y a deux problèmes sous-jacents à cette étude :

- choisir une meilleure combinaison parmi les combinaisons possibles.
- savoir si une combinaison est admissible (si ce qu'elle représente correspond aux objets présents dans la scène), et, le cas échéant produire une alarme.

La première question implique de pouvoir comparer des reconstructions. Une reconstruction, ou combinaison de reconstructions individuelles, doivent posséder différentes qualités. Elle doit notamment :

1. représenter toute la zone à reconstruire, c'est-à-dire ne pas commettre d'omission. Ceci se justifie par le fait que le MNE est fiable et assez précis. Une partie de la zone non reconstruite correspond donc à une partie de bâtiment non traité.
2. ne pas représenter plus que la zone à reconstruire. En effet, un zone de sursol est issue d'une technique de focalisation s'appuyant sur un MNE réputé fiable. L'ensemble des bâtiments voisins de la zone courante, et d'altitude proche, est donc réputé présent dans la zone de focalisation à traiter. La reconstruction n'a donc pas de raison de s'étendre sensiblement au-delà de la zone.
3. représenter correctement la surface topographique de la zone de focalisation. Ici, des critères typiques d'attache aux données initiales interviennent. Dans la mesure où le MNE est réputé fiable et relativement précis, nous l'utiliserons comme référence pour mesurer la qualité d'une reconstruction.

4. être consistante. Ici, nous entendons le fait que deux reconstructions individuelles d'une combinaison ne doivent pas s'intersecter.
5. être simple : entre deux reconstructions représentant de manière sensiblement équivalente la même surface topographique, la plus simple synthétise le mieux l'information, et est donc a priori meilleure.

Certains de ces critères peuvent être explicités à l'aide de cartes servant à analyser les reconstructions.

L'omission (critère 1) est en effet liée à l'existence de pixels qui sont au sol d'après le MNE reconstruit, et sur un toit d'après le MNE initial. De même, les extrapolations (critère 2) sont liées à des pixels au sol du le MNE initial et sur un toit d'après le MNE reconstruit.

Enfin, les erreurs d'approximation sont les différences entre les deux MNEs. Ici il convient de distinguer deux types de différences, que nous appellerons erreurs et aberrations. En effet, les MNE initial ne respecte pas rigoureusement la surface topographique qu'il représente, en raison de problèmes connus (par exemple, l'outil de mise en correspondance stéréoscopique utilisé pour produire le MNE initial a tendance à produire des marches d'escalier au voisinage des zones de contraste). Par ailleurs, les superstructures des toits ne sont pas reconnues par les formes utilisées dans ce système de reconnaissance. Elles sont donc le lieu de différences entre les deux MNEs considérés ici. Enfin, comme le montrera le chapitre suivant, les reconstructions elles-mêmes souffrent de délocalisations. De faibles différences d'altitude entre le MNE initial et le MNE reconstruit sont donc tolérables. En revanche, des erreurs plus fortes, et réparties sur des zones de grande étendue sont selon toute vraisemblance des signes d'une mauvaise reconnaissance de la forme à reconstruire. Nous appellerons donc aberrations les différences entre les deux MNEs supérieurs à un seuil d'altitude donné. Ce seuil d'altitude doit être inférieur à la différence caractéristique choisie par Baillard pour distinguer deux bâtiments voisins d'altitudes différentes, soit 2,5m. Par ailleurs, il doit être assez grand pour ne pas être trop alarmiste. Dans le système proposé, sa valeur est donc de 1m.

Ces 4 éléments (omission, extrapolation, erreurs, et aberrations) peuvent être représentés par des images comme celles présentées dans la figure 5.1 (il s'agit des cartes liées à la reconstruction de la zone 172 par le modèle D).

On constate sur cette figure que, bien que la reconstruction soit la meilleure reconstruction que le système proposé est capable de produire, aucune de ces cartes n'est vierge : la reconstruction présente une faible omission, un faible extrapolation, et quelques aberrations.

Dans la mesure où le système doit produire des alertes en cas de problème, ces cartes ne sont donc pas adaptées. Nous proposons d'opérer des filtrages morphologiques sur ces cartes avant de produire d'éventuelles alertes. C'est l'objet de la prochaine section.

3.5 Production d'alertes

En raison des imperfections des données comparées (délocalisations du MNE initial, délocalisations des reconstructions, et donc du MNE reconstruit), les cartes d'omission et d'extrapolation ne sont jamais vierges. Pour cette raison, dans un contexte opérationnel, le déclenchement d'alertes visant à retenir l'attention d'un opérateur ne peut pas être fait en utilisant directement ces cartes trop sensibles.

Nous appliquons donc une érosion morphologique sur les cartes d'extrapolation, d'omission, et d'aberration avant de produire des alertes. Le but de cette érosion est d'absorber précisément les



FIG. 5.1 – Les 4 cartes d’analyse de la reconstruction du bâtiment 172 : de gauche à droite : omission, extrapolation, erreurs, aberrations

petites aberrations d’altitude (dus aux cheminées et petites superstructures non reconstruites), ainsi que les délocalisations des contours des reconstruction ou du MNE.

L’érosion est réalisée avec un élément structurant carré, de rayon σ_a . La valeur de σ_a sera choisie dans l’évaluation (chapitre suivant). L’érosion n’est évidemment pas réalisée sur la carte des erreurs. Mais celle-ci sera tout de même utilisée par la suite pour qualifier les combinaisons.

A l’issue des érosions des 3 cartes citées ici, une alerte est produite si il subsiste des pixels sur au moins l’une des cartes. Le système d’alerte ainsi proposé sera réglé et évalué dans le chapitre suivant.

3.6 Evaluation d’une combinaison

Dans cette section, nous proposons 5 fonctions de qualification des combinaisons, liées aux 5 critères explicités dans la section 3.4. Ces fonctions sont ensuite combinées afin d’obtenir une qualification globale d’une combinaison.

On notera que contrairement au mécanisme de production d’alerte qui est binaire (le système produit ou ne produit pas d’alerte à l’issue des érosions des cartes d’analyse), ces fonctions renvoient des valeurs réelles. Ceci permet d’obtenir des appréciations «floues» des combinaisons au regard de ces critères.³

Nous proposons donc les 4 fonctions suivantes, où r_{i_1}, \dots, r_{i_k} représentent des reconstructions individuelles :

1. $S_w(r_{i_1}, \dots, r_{i_n})$ est le volume de la carte d’omission $S \setminus S_{i_1, \dots, i_p}$, i.e. $\rho^2 \sum_{p \in S \setminus S_{i_1, \dots, i_p}} h(p)$
2. $E(r_{i_1}, \dots, r_{i_k}) = \rho^2 \sum_{p \in S \cap S_{i_1, \dots, i_k}} |h(p) - h_{i_1, \dots, i_k}(p)|$ Il s’agit du «volume» des erreurs.
3. $S_b(r_{i_1}, \dots, r_{i_k})$ est le volume de la carte d’extrapolation $S_{i_1, \dots, i_k} \setminus S$, i.e. $\rho^2 \sum_{p \in S_{i_1, \dots, i_k} \setminus S} h_{i_1, \dots, i_k}(p)$
4. $S_o(r_1, \dots, r_n) = 1/2 \rho^2 \sum_{i, j \in \{1, \dots, n\}, i \neq j} h_{r_i}(p) + h_{r_j}(p)$ et le terme de recouvrement entre les reconstructions individuelles.

³L’objectif étant de ici de trouver la meilleure reconstruction au regard des 5 critères utilisés, cette appréciation floue des critères est nécessaire car leur utilisation binaire (comme dans le cas du mécanisme de production d’alerte) entraîne l’impossibilité de comparer deux combinaisons.

En effet, considérons une combinaison C_1 , et une autre combinaison C_2 . Considérons que C_1 a été jugée suffisante (sans omission) et que C_2 n’a pas été jugée suffisante. C_1 est donc a priori meilleure que C_2 . En revanche, C_2 peut mieux représenter la surface topographique, en terme de différences d’altitude, que C_1 . C_1 et C_2 ne peuvent donc pas être comparées. Ceci justifie l’introduction de cette qualification globale «floue».

Ces quatre fonctions sont homogènes (elles s'expriment en m^3), et peuvent être combinées de manière additive. Nous posons donc :

$$F_D(r_{i_1}, \dots, r_{i_k}) = (S_w + E + S_b + S_o)(r_{i_1}, \dots, r_{i_k})$$

La complexité d'une combinaison (critère 5) est évaluée par la fonction suivante :

$$F_C(r) = 3P(r) + 2L(r)$$

où $P(r)$ est le nombre de sommets de type ponctuel de la combinaison $r = \{r_{i_1}, \dots, r_{i_k}\}$, et $L(r)$ est le nombre de ses sommets linéaires. Le terme 3 de F_C correspond au fait qu'un sommet ponctuel a 3 coordonnées, codées par exemple par des flottants, et le terme 2 correspond au fait qu'il faut deux entiers pour représenter un segment, ce deux entiers «pointant» vers les extrémités du segment.

Enfin, les deux fonctions F_D et F_C sont combinées au moyen d'un facteur de pondération K_e , explicité dans le paragraphe suivant :

$$F(r_1, \dots, r_k) = (K_e F_C + F_D)(r_1, \dots, r_k)$$

Interprétation de la constante de pondération

Considérons un bâtiment parallélépipédique reconstruit de façon exacte par une reconstruction r appuyée sur un modèle parallélépipédique. Nous avons alors $S_w = E = S_b = S_o = 0$, d'où $F(r) = 20K_e$. Considérons maintenant que ce bâtiment n'est pas reconstruit du tout ($r = \$$), alors $S_w(r)$ est égal au volume \mathcal{V} du bâtiment.

Il est donc préférable de reconstruire ce bâtiment si $20K_e < \mathcal{V}$.

K_e est donc lié à la précision de généralisation adoptée pour le système.

Dans la pratique, nous avons $K_e = 0.1$ ce qui induit une précision de généralisation de $2m^3$.

Dans la suite de ce chapitre, nous proposons deux méthodes permettant d'engendrer, pour une même zone, un ensemble de reconstructions individuelles (la première de ces méthodes et une méthode «arborescente» naturelle mais très combinatoire et la seconde méthode est plus efficace), et une technique de recherche de l'optimum de la fonction d'évaluation F , c'est-à-dire la recherche de la meilleure combinaison de reconstructions.

4 Vers une stratégie de reconstruction : approche arborescente

Nous avons vu qu'il n'était pas simple de comparer deux reconstructions prises individuellement.

Pourtant une approche naturelle pour reconstruire une zone complète consiste à essayer un modèle, puis 2, puis 3. Ceci mène à un processus de reconnaissance fonctionnant comme un parcours d'arbre. Chaque sommet de cet arbre est une reconstruction r basée sur p modèles où p est la profondeur du nœud dans l'arbre. Chaque nœud possède des fils qui sont des reconstructions basées sur r , et complétée par un modèle de la grammaire. Il peut donc y avoir un nombre important de fils, ce qui induit une grande largeur de l'arbre.

Dans la pratique, nous avons expérimenté ce type de méthode, mais nous n'en expliciterons pas les détails, car ces types d'approches posent deux problèmes :

- C'est excessivement combinatoire. Or les appariements, pris individuellement sont déjà bien coûteux.
- Il est difficile de limiter la profondeur de l'arbre de recherche.

La figure 4 explicite cette recherche arborescente où nous supposons que 3 reconstructions sont possibles individuellement. Une recherche arborescente consiste à détecter r_1, r_2, r_3 indépendamment, puis à propager ces reconstructions. Ainsi, à partir de r_1 , il est possible de retrouver r_2 ailleurs dans la zone, et ainsi de suite. On voit nettement que des calculs doivent être faits plusieurs fois. Ceci nous pousse alors à mémoriser les calculs déjà réalisés pour en tirer profit.

Une heuristique a effectivement été mise en place, pour tirer profit des calculs déjà faits, mais malgré ceci, cette technique reste très inefficace. La section suivante propose donc une stratégie moins combinatoire.

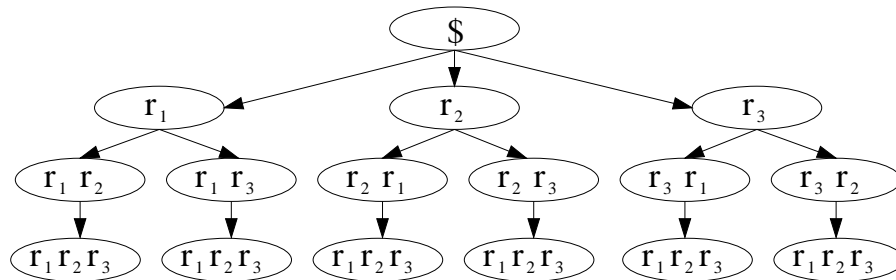


FIG. 5.2 – Explicitation de la recherche arborescente pour la reconstruction d'une zone

5 Vers une reconstruction : approche linéaire

5.1 Génération d'hypothèses

Cette section propose un procédé un peu moins général, mais plus efficace pour produire un ensemble de reconstructions. Par ailleurs, comme précédemment, la question du choix de la meilleure reconstruction se posera, et nous tenterons d'y apporter une solution tenant compte de la complexité de la reconstruction et de l'adéquation aux données.

5.1.1 Génération d'hypothèses à 1 modèle

Tout d'abord, il est évident qu'une instance d'un modèle peut apparaître deux fois dans une zone. Donc, nous devons être capable de détecter plusieurs instances d'un même modèle dans une zone.

Pour cela, nous pouvons envisager deux options.

Augmentation des modèles par juxtaposition Pour un modèle G , cette technique propose d'identifier le modèle G dans les données, puis d'identifier le modèle $G \cup G$, puis $G \cup G \cup G$, et ainsi de suite, jusqu'à ce qu'aucun appariement ne se produise. L'explosion combinatoire qui en dérive est évidente. La parade consiste à figer le modèle G une fois apparié, avant d'apparier le modèle $G \cup G$. Ceci peut être fait en flanquant le modèle G d'une information a priori définie par son premier appariement, ce qui produit un modèle G^{ap} , puis à apparier $G^{ap} \cup G$. Dans le cas où certains

sommets de G ne sont toutefois pas appariés, ceci engendre tout de même des calculs parasites pour produire finalement toujours les mêmes résultats.

En fait, nous exposerons dans la suite une technique voisine, plus simple et plus efficace.

Suppression des données Etant donné un premier appariement d'un modèle G , noté $f = (\Delta, f_\Delta)$, nous savons que f_Δ identifie un ensemble de sommets du graphe de données G_I . Il est possible à ce moment d'envisager de supprimer ces données identifiées du graphe G_I , ce qui produit un graphe G'_I , puis de tenter à nouveau un appariement entre G et G'_I . Nous allons toutefois montrer par des exemples que 2 types de sommets ne peuvent pas être supprimés car ils sont susceptibles d'être partagés par deux instances d'un modèle dans une zone.

Les discontinuités, tout d'abord sont susceptibles de séparer deux instances d'un modèle dans une scène. Nous rappelons que la détection des discontinuités a servi précisément à ceci (voir section 5.2, page 49). Les discontinuités ne peuvent donc pas être supprimées des données avant de tenter un nouvel appariement.

On pourrait alors imaginer ne supprimer que les discontinuités se trouvant au bord de la zone (sur son contour), mais là aussi, dans le cas d'alignements fortuits, même les discontinuités en bord de zone peuvent être partagées. A nouveau la zone 17 en est un exemple de ce type de cas, car le bord ouest de la zone est partagé entre les deux bâtiments.

Les primitives linéaires ne peuvent, elles non plus, être supprimées de la sorte, et pour les mêmes raisons.

Les primitives surfaciques, quant à elles sont les candidates idéales à la suppression. Nous avons déjà dit qu'elles étaient très structurantes et très peu sous-détectées. Il est donc très peu probable qu'elles soient partagés entre deux instances d'un même modèle, d'autant que par leur nature bi-dimensionnelle (ce qui n'est pas le cas des discontinuités et des segments de droite), elles ne se trouvent pas à *la frontière* entre deux bâtiments.

Conclusion. Ainsi l'option de suppression des données consiste à supprimer des données les primitives *surfaciques* identifiées par f_Δ . Toutefois, le paragraphe suivant expliquera d'un point de vue pratique, nous n'effacerons ces primitives que virtuellement.

Suppression virtuelle des données surfaciques Il s'agit de la seconde option proposée dans cette section. Pour supprimer une primitive des données, il n'est pas nécessaire de reconstruire un graphe G'_I , mais il est simplement suffisant d'interdire son appariement. C'est à cet effet que la fonction a a été introduite dans la fonction VM , page 65. La suppression d'un sommet v' surfacique des données revient à mettre la valeur de $a(v')$ à 0.

La méthode de suppression virtuelle des données surfaciques d'un appariement f est donc présentée en figure 5.3.

Suppression virtuelle des primitives surfaciques d'un appariement

$SVPS(G, f = (\Delta, f_\Delta))$.

1. Pour tout sommet surfacique v de G , $v' = f_\Delta(\Delta(G))$ est sont correspondant dans les données. Mettre la valeur de $a(v')$ à 0.

FIG. 5.3 – Pseudo code de la suppression virtuelle des données surfaciques après un appariement.

Génération d'hypothèses

Nous sommes maintenant prêts pour générer les hypothèses d'un modèle. De manière informelle, il s'agit d'apparier un premier modèle, puis d'effacer virtuellement les primitives surfaciques des données grâce à l'appariement, puis d'apparier le même modèle, et ainsi de suite.

On arrête le processus dès qu'aucun appariement n'est trouvé. Il faut noter ici que nous ne tenons pas compte du fait qu'un appariement peut produire une reconstruction faisable ou pas. Il serait en effet dommage de se priver, par exemple, de la seconde reconstruction faisable parce que la première ne l'était pas, ce qui pourrait arriver.

La fonction $PH(G, G_I)$ qui réalise ce travail est précisée en figure 5.4. Elle produit au maximum $E(N_s(G_I)/N_s(G))$ où $N_s(G)$ désigne le nombre de sommets surfaciques d'un graphe, et E est la partie entière d'un nombre.

Production d'hypothèses pour un modèle G

$PH(G)$

1. Soit $R = \emptyset$.
2. Si $NETAPP(G)$ ne produit pas de résultat, alors retourner R .
3. Sinon, soit $f = (\Delta, f_\Delta)$ ce meilleur résultat.
4. Appeler $SVPS(G, f)$.
5. Si f produit une reconstruction faisable r , ajouter r à R .
6. aller en 2.

FIG. 5.4 – Pseudo code de la propagation d'une reconstruction à partir d'une reconstruction (M, f)

5.1.2 Génération d'hypothèses à N modèles

En suivant le principe du plus facile d'abord, et surtout parce que il est peu probable de trouver des modèles de profondeur $P + 1$ de la grammaire si aucun appariement n'a été trouvé pour les modèles de profondeur P , la production de reconstructions individuelles pour plusieurs modèles de la grammaire travaille en utilisant les modèles de profondeur croissante de la grammaire.

La méthode $PH(G_I)$ qui réalise ce processus est explicitée dans la figure 5.5.

A ce stade, la production d'hypothèse est terminée. Elle a été baptisée approche linéaire, car en quelque sorte elle remplit de façon linéaire une matrice de reconstructions dont les lignes correspondent aux modèles de profondeur croissante, et les colonnes aux reconstructions produites pour ces modèles, comme en figure 5.6.

Cette figure présente les graphes G_1 pour la profondeur P_0 , G_2 et G_3 pour la profondeur P_1 , puis G_4 , G_5 et G_6 pour la profondeur P_2 . Chaque graphe G_i produit un ensemble de reconstructions individuelles $r_{i,j}$.

Production d'hypothèses.

$$PH(G_I)$$

1. Soit $R = \emptyset$.
2. Soit $C = C_0$
3. Soit $R_C = \emptyset$
4. Pour tout modèle $G \in \mathcal{G}_C$, faire :
 - (a) Soit l'ensemble des reconstructions $R_G = PH(G, G_I)$
 - (b) Faire : $R_C = R_C \cup R_G$
5. Faire $R = R \cup R_C$
6. Si $R_C = \emptyset$, renvoyer R
7. Faire $C=C+1$
8. Aller en (1)

FIG. 5.5 – Pseudo code de génération d'appariement pour un ensemble de données.

P_0	G_1	r_{11}	r_{12}	r_{13}	r_{14}	r_{15}	
P_1	G_2	r_{21}	r_{22}	r_{23}	r_{24}	r_{25}	
	G_3	r_{31}	r_{32}	r_{33}	r_{34}		
P_2	G_4	r_{41}	r_{42}	r_{43}	r_{44}	r_{45}	
	G_5	r_{51}	r_{52}	r_{53}			
	G_6	r_{61}	r_{62}	r_{63}			

FIG. 5.6 – La matrice des reconstructions dans le cas de l'approche linéaire.

5.2 Combinaison d'hypothèses

Nous proposons de rechercher une combinaison des composantes, satisfaisant un critère global tenant compte de l'adéquation aux données et de la complexité du modèle.

Nous disposons maintenant d'un ensemble $R = \{r_1, \dots, r_n\}$ de reconstructions individuelles, basées chacune sur un modèle de la grammaire.

La section 3.6 a proposé une fonction d'évaluation globale qui qualifie une combinaison (un sous-ensemble de R), en tenant compte de la complexité de la combinaison, et de son adéquation aux données. Nous cherchons maintenant à trouver la combinaison minimisant cette fonction.

5.3 Minimisation de la fonction d'évaluation

Pour N reconstructions, il y a a priori 2^N combinaisons possibles. Il est impossible d'explorer cet espace de recherche explicitement, car N peut être assez grand (jusqu'à 50). Dans la pratique, nous interdirons donc de manière stricte de combiner deux reconstructions individuelles qui se recouvrent sur plus de $10m^2$. Ceci réduit un peu la combinatoire du problème. Mais surtout, il est important de noter que même si N est grand, le nombre de reconstructions individuelles réellement combinées pour former la reconstruction optimale au sens de la fonction $F(r_{i_1}, \dots, r_{i_k})$ est beaucoup plus faible que N lui-même.

Dans la pratique, nous proposons donc de résoudre la minimisation de N en considérant les combinaisons contenant au plus 1 reconstruction, puis les combinaisons contenant au plus 2 combinaisons, ainsi de suite.

Dès que pour un certain N_0 , la meilleure reconstruction contenant au plus N_0 reconstructions individuelles n'en contient en fait que $N_0 - 1$, alors la combinaison optimale ne contient que $N_0 - 1$ reconstructions individuelles, et celles-ci ont déjà été déterminées.

Cette procédure, résumée dans la figure 5.7 reste très combinatoire, mais accessible aux machines actuelles. Les temps de calcul restent toutefois très importants. Ceci provient de la taille de l'espace à explorer, et du fait que chaque évaluation nécessite le calcul et la comparaison de modèles numériques d'évaluation.

L'approche proposée n'est pas certaine d'obtenir le minimum de la fonction F . Nous ne proposons pas explicitement d'expérimentation de la méthode de combinaison, car il n'est pas possible de présenter les reconstructions individuelles proposées par le système, et la combinaison finale qui en est déduite. Les cas intéressants sont en effet ceux pour lesquels il y a de nombreuses reconstructions individuelles, ce qui serait lourd à présenter. Par contre, nous avons fait l'étude de la pertinence de la méthode de combinaisons, dans un grand nombre de cas, et il s'avère que cette méthode est tout à fait pertinente : elle produit de façon systématique la meilleure combinaison possible pour un ensemble de reconstructions proposé.

Combinaison d'hypothèses.

$CH(r_1, \dots, r_n)$

1. Soit $k = 1$.
2. Calculer la meilleure combinaison de reconstructions R_k (au sens de la fonction F), contenant au plus k reconstructions individuelles prises parmi r_1, \dots, r_n .
3. Si R_k contient k reconstructions individuelles, faire $k = k + 1$ et aller en 2.
4. Renvoyer la meilleure combinaison trouvée à l'étape $k - 1$.

FIG. 5.7 – Pseudo code de la combinaison de reconstructions.

6 Conclusion

Ce chapitre a tenté de mettre en évidence des propriétés que doivent respecter les meilleures reconstructions pour une zone de bâti.

Nous nous sommes orientés vers une méthode de reconstruction visant à produire un ensemble d'appariements pour un ensemble de modèles de taille croissante. Ces reconstructions individuelles sont alors combinées pour proposer une reconstruction finale.

La méthode de combinaison reste au moment de la rédaction du mémoire, très combinatoire, mais elle produit des résultats tout à fait satisfaisants.

Les limites de la méthode résident en fait dans la recherche des reconstructions individuelles, et non dans la méthode de combinaison.

Chapitre 6

Résultats et évaluation

1 Évaluation a priori

Ces travaux utilisent une grammaire de graphes pour représenter les modèles de bâtiments gérés par son système de reconstruction. Cette grammaire n'est naturellement pas en mesure d'engendrer toutes les formes possibles, mais seulement un sous-ensemble. Pour évaluer le domaine d'emploi a priori de la méthode en milieu urbain nous proposons d'utiliser comme référence des données urbaines précises (données TRAPU). A partir de ces données de référence nous pouvons générer des primitives « parfaites » et tester la capacité du système à reconstruire des formes réelles de bâtiments à partir de données idéales. Nous ne cherchons donc pas ici à évaluer les performances globales du système mais uniquement sa capacité à décrire des scènes urbaines complexes.

Ce travail d'évaluation a été confié à Grégoire Maillet. Nous remercions très vivement Grégoire pour cette contribution.

Les données TRAPU proposent des descriptions complètes et volumiques des bâtiments. Mais les primitives nécessaires à la reconstruction (linéaires, surfaciques et les discontinuités altimétriques) sont basées sur la géométrie et la radiométrie de la scène observée et n'ont pas toujours de lien simple avec les données beaucoup plus sémantiques de la base de données. La saisie de la BD fait apparaître des découpages arbitraires dans les bâtiments complexes avec l'ajout de faces verticales invisibles, des fonds de bâtiments et d'arêtes fictives nécessaires pour assurer la convexité des faces. Pour limiter l'influence de cette « modélisation TRAPU » nous avons préféré générer les primitives planes et les discontinuités altimétriques à partir d'un MNE calculé à l'aide de la BD. Seules les primitives linéaires proviennent donc directement de la BD.

L'évaluation s'est faite sur six groupes de bâtiments choisis sur les données disponibles sur le centre ville d'Amiens. Ces six extraits sont de complexités variées mais nous avons fait en sorte qu'aucun ne corresponde à un modèle élémentaire de la grammaire. L'objectif est de voir jusqu'à quel point le système de reconstruction est capable de combiner des modèles et de généraliser les structures qu'il ne sait pas modéliser.

Le premier extrait (figure 6.1) est constitué d'un bâtiment à toit plat, sur deux niveaux avec un contour planimétrique assez complexe. Le système a bien détecté deux modèles de bâtiment à toit plat. La reconstruction du plus grand des modèles donne une version généralisée mais correcte du bâtiment, par contre le petit modèle ne se reconstruit pas convenablement et prend une forme

totalelement dégénérée. Cette dégénérescence est probablement due à l'absence de deux des discontinuités altimétriques pour ce modèle : en effet la différence d'altitude entre les deux niveaux n'est pas suffisante pour donner lieu à une discontinuité. Dans tous les cas la stratégie de combinaison de modèles recherche à limiter les recouvrements donc il est probable que si ce deuxième petit modèle avait été convenablement reconstruit il aurait été éliminé à cause d'un trop fort recouvrement avec le premier.

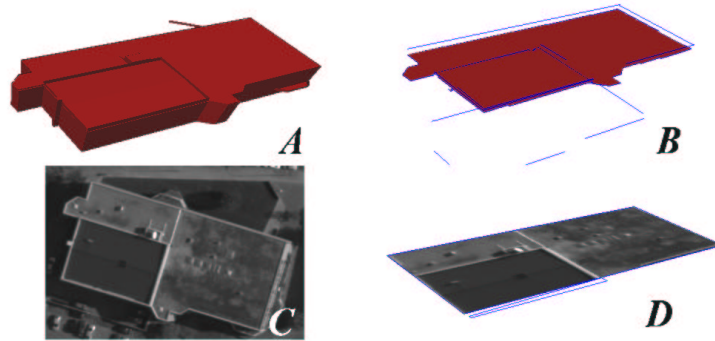


FIG. 6.1 – A. extrait n°1. B. primitives utilisées lors de l'appariement (deux modèles à toit plat). C. extrait de l'image correspondante. D. résultat de la reconstruction du toit (sans les façades).

Le deuxième extrait (figure 6.2) est constitué d'un assemblage de quatre toits à deux pans. La grammaire utilisée ne prévoit pas la connexion « en V » présente entre les deux toits du centre mais le système a proposé une généralisation intéressante avec la connexion « en L » dont il dispose.

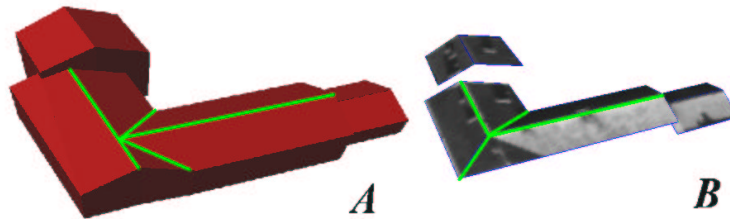


FIG. 6.2 – A. extrait n°2. B. résultat de la reconstruction du toit (sans les façades).

Le troisième extrait (figure 6.3) montre une limite de notre méthodologie d'évaluation. La reconstruction obtenue est une généralisation sous forme d'un toit plat (figure 6.3.B). Mais cette généralisation n'est pas due à une incompatibilité avec la grammaire de modèles, elle a été introduite par le système d'extraction des primitives planes à partir du MNE TRAPU©. En effet cette recherche de primitives se fait par regroupement successifs de zones planes et donne non seulement les pans inclinés mais aussi le plan horizontal global (figure 6.3.C). On constate que si l'on interdit à la grammaire le modèle à toit plat, le système choisit une juxtaposition de modèles à deux pans (figure 6.3.D). Le trou dans ce dernier modèle est dû à un échec dans la reconstruction de l'un des

6 modèles à deux pans détectés. Cet extrait montre donc la difficulté de dissocier l'évaluation du domaine d'application de la grammaire de modèles et celle des performances globales du système de reconstruction.

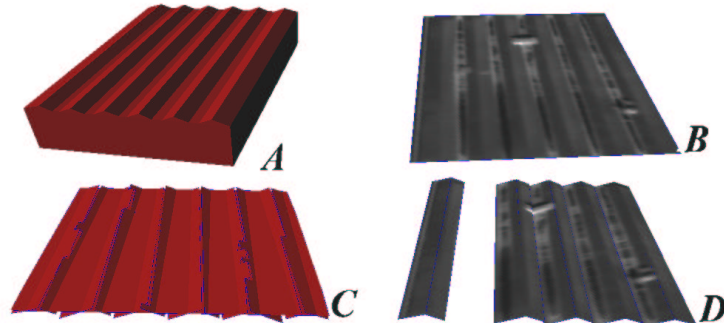


FIG. 6.3 – A. extrait n°3. B. résultat de la reconstruction du toit (généralisation en toit plat). C. résultat de la recherches des primitives planes. D. résultat de la reconstruction en interdisant le modèle à un pan.

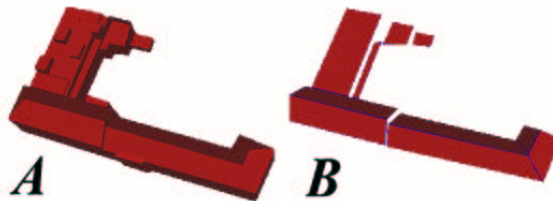


FIG. 6.4 – A. extrait n°4. B. résultat de la reconstruction du toit.



FIG. 6.5 – A. extrait n°5. B. résultat de la reconstruction du toit.

Les figures 6.4 et 6.5 montrent des reconstructions réussies même si certaines super-structures sont généralisées. Sur le sixième extrait choisi (figure 6.6) une série de bâtiments à deux pans encercle quelques bâtiments à toit plat. On observe que la présence de nombreuses mansardes n'a pas perturbé la reconstruction. On note également que la série de trois bâtiments en bas à gauche a été appariée avec un seul modèle en forme de « U ». Par contre l'angle en haut à gauche n'a pu être reconstruit à

cause de sa forme atypique (figure 6.6.C). De plus, deux des petits bâtiments plats de la cour n'ont pas été reconstruits. On peut aussi observer que la grammaire utilisée ici n'autorise ni les coudes sur les bâtiments à toit plat, ni les attaques en V sur les bâtiments à double pan. Enfin, on peut souligner que la stratégie de combinaison de modèles fonctionne correctement même sur des groupes de bâtiments complexes et de taille étendue.

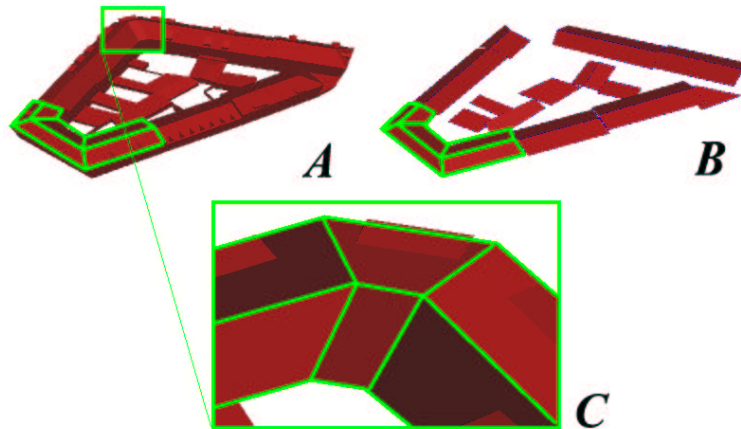


FIG. 6.6 – A. extrait n°6. B. résultat de la reconstruction. C. Coude non modélisable.

Même si les extraits traités sont peu nombreux et pas nécessairement représentatifs ils regroupent un grand nombre de difficultés et permettent de faire un bilan sommaire du domaine d'application a priori de la grammaire de graphes utilisée en milieu urbain dense.

N° d'extrait	NB de Composantes TRAPU	NB de Composantes détectées	NB de reconstructions "parfaites"	NB de généralisations	NB de dégénérescences	NB d'omissions
1	2	2	0	1	1	0
2	4	4	2	2	0	0
3	1	1	0	1	0	0
4	8	7	1	6	0	1
5	3	3	3	0	0	0
6	22	16	11	5	0	6
TOTAL	40	33	17	15	1	7
%		82.5%	42.5%	37.5%	2.5%	17.5%

TAB. 6.1 – Bilan des reconstructions

Conclusion

Cette évaluation du domaine d'emploi de la méthode en milieu urbain dense montre que les modèles utilisés, quoique peu nombreux, sont en mesure de représenter une proportion non négligeable des objets topographiques qui nous intéressent. Plus précisément, le taux d'omission est relativement faible (20%), et en acceptant un certain degré de généralisation (que nous n'avons pas évalué), nous pouvons affirmer que les modèles sont suffisants pour représenter une forte proportion des bâtiments d'une scène complexe.

Mais une zone complexe n'est pas représentée par un modèle, mais une juxtaposition de modèles.

Enfin, nous souhaitons insister sur le fait que cette évaluation du domaine d'emploi n'a pas été réalisée en évaluant visuellement les objets susceptibles d'être représentés par les modèles, mais en effectuant réellement leur reconstruction.

2 Évaluation réelle

L'évaluation réelle a porté sur un extrait de couple stéréoscopique. Cet extrait n'a pas été choisi par nous. Le couple stéréoscopique complet a en effet été traité par blocs, et l'évaluation a porté sur le premier de ces blocs. Nous n'avons notamment pas choisi le bloc «où les méthodes marchent le mieux».

On remarquera notamment que ce bloc se situe au nord du couple stéréoscopique, ce qui a pour conséquence d'augmenter les occultations et les déformations perspectives par rapport à un bloc situé à la verticale de l'axe de vol de l'avion. Le bloc étudié n'est donc pas a priori le plus favorable pour ces traitements.

Nous rappelons que les zones traitées sont présentées exhaustivement dans l'annexe I.

2.1 Quelques résultats clés.

L'objectif de cette section est de proposer quelques exemples de reconstruction visant à montrer les avantages et les inconvénients de la méthode proposée.

La zone 172, si couramment citée dans ce mémoire, est en fait le cas typique où «tout s'est bien passé». La reconstruction est parfaitement fidèle à la réalité, il n'y a aucun problème d'évaluation quant à la forme du bâtiment : il a deux pans. La forme adaptée a été reconnue par la méthode de combinaison d'hypothèses, et la précision géométrique est par ailleurs bonne.

La zone 169 pose un problème de géométrie : une gouttière est très sensiblement déviée par rapport à la réalité. La cause est une discontinuité qui a déporté un sommet 3-D. La forme est globalement reconnue correctement, mais la géométrie est très mauvaise, car le point nord-ouest est à plus d'un mètre de sa position vraie.

La zone 162 est un bâtiment à 8 pans dont 4 sont quasiment horizontaux, et 4 autres quasiment verticaux, accolée à une structure parallélépipédique. Cette zone est reconstruite comme un seul bâtiment parallélépipédique. Il y a donc une généralisation d'une partie de la zone, et une lacune sur la partie nord de cette zone.

Nous retrouvons le même type de bâtiment sur la zone 152, mais cette fois-ci la méthode retrouve deux structures parallélépipédiques accolées.

La zone 152, précisément, est le cas d'une zone de focalisation présentant une discontinuité d'altitude, ce qui engendre la reconnaissance de deux modèles juxtaposés : le système se focalise sur un bâtiment, puis sur l'autre. Des cas similaires sont présentés dans les zones 45 ou 17 par exemple.

La zone 130 est un cas typique où le bâtiment, quoique simple, est trop généralisé par la méthode.

2.2 Protocole d'évaluation

La première partie de l'évaluation concernera le respect des formes. Pour cela, nous précisons le degré de généralisation que nous tolérons, exemples à l'appui. La zone généralisée avec les tolérances fixées constitue l'objectif de reconstruction en terme de forme.

Ensuite, nous évaluerons la précision géométrique de la méthode. Tous les sommets 3D ne sont pas évaluable, comme nous le verrons. Enfin, nous évaluerons le système d'alarme mis en place.

L'évaluation proposée ici porte sur 172 zones de sursol extraites de manière automatique. Certaines zones de sursol ne sont pas traitées car elles correspondent à des zones non bâties (typiquement :

végétation) où à des zones trop près du bord de la scène. Dans ce cas, nous savons que le MNE peut ne pas être correct à cause, précisément, d'effets de bord. Les travaux ne peuvent donc être évalués que sur les bâtiments situés suffisamment loin des bords de la scène considérée pour le calcul du MNE.

Pour une zone donnée, nous considérons d'abord le nombre de bâtiments de la zone. Plus exactement nous considérons les portions de toit pour lesquelles il est possible d'aller d'un point à un autre sans saut d'altitude. Ces surfaces peuvent présenter des sauts d'altitudes. Par exemple, la zone 101 est composée de 3 pans de toits, dont les deux plus au sud sont séparés par un saut d'altitude. Pourtant, il est possible de parcourir tout ce toit sans franchir ce saut. Ce bâtiment est donc bien considéré comme un seul bâtiment, ce qui est naturel, mais il justifie la précision proposée dans ce paragraphe. La zone 97 donne un autre exemple de ce type. Nous noterons BA le nombre de bâtiments ainsi défini pour chaque zone.

Pour ces bâtiments nous considérons une «reconstruction idéale» qui exclut toutes les superstructures (chiens assis, lucarnes, cheminée, cages d'ascenseur). Nous tolérons alors une généralisation de la forme des pans de l'ordre du mètre. Les décrochements (comme sur la zone 101), protubérances (comme sur la zone 43), sont virtuellement généralisés (supprimés) pour fournir le bâtiment à reconstruire de la zone. Nous dirons qu'un bâtiment est modélisable par la grammaire si il est modélisable rigoureusement par la grammaire après cette généralisation de la forme des pans. Nous noterons BM le nombre de bâtiments ainsi modélisables, et PA le nombre de pans de tous les bâtiments de la zone.

Par après, nous tolérons une généralisation de la forme globale du bâtiment si les dénivelées des pans sont de 1m ou légèrement supérieures (jusqu'à 1,5m), comme sur la zone 152. Sur cette zone, composée de 2 bâtiments, le bâtiment nord est composé de 8 pans. 4 d'entre eux sont quasi verticaux, et les autres forment une pyramide dont la dénivelée est de 80 cm. Nous admettrons donc de représenter cette pyramide par un pan simple, rectangulaire. Enfin, les pans quasi verticaux dont la plus petite dimension horizontale est inférieure à 1m (comme c'est le cas des 4 pans quasi-verticaux de cette zone), alors nous admettrons une suppression de ces pans. Ainsi, nous tolérons pour la figure 152 un représentation par 2 pans horizontaux, rectangulaires. De même, la zone 162 pourra être vue comme un seul pan horizontal, ainsi que la zone 147 (dénivelée des pans : 1m). Nous noterons PAG le nombre de pans restant après ces généralisations.

Tous ces éléments fixent les objectifs de reconstruction de la méthode.

L'évaluation de la reconstruction se fait alors d'après les éléments suivants. En premier lieu, nous dénombrerons le nombre de bâtiments de chaque zone tels que le modèle de représentation choisi corresponde au modèle du bâtiment généralisé toléré, ou à un modèle plus fin (ce qui peut arriver). Ce nombre est noté BR. Nous dénombrerons le nombre de bâtiments reconstruits, mais abusivement généralisés. Ce nombre est noté BRG. Il arrive que certains bâtiments modélisables ne soient pas reconnus, mais éclatés en plusieurs pans dont la juxtaposition décrit correctement le volume du bâtiment. Nous noterons BRE le nombre de ces bâtiments. Enfin, pour certains bâtiments, il peut manquer des pans après reconstruction. Nous les dénombrerons. Leur nombre sera noté BRP (P pour Partiellement reconstruit).

Par ailleurs, le nombre PR sera celui des pans reconstruits (à comparer avec PAG). Le nombre AV correspond au nombre de pans reconstruits sur de la végétation : ce sont des Artéfacts liés à la Végétation. De même, des pans sont reconstruits sur des bâtiments voisins de la zone de focalisation. Le nombre AB (pour Artéfact lié à un Bâtiment) les dénombrera.

Les tableaux 6.2, 6.3 et 6.2 présentent les dénombrements développés ici.

Les colonnes SAG de ces tableaux dénombrent les sommets 3-D des modèles généralisés (corres-

pendant à PAG). Après reconstruction, la position 3-D de certains sommets sera évaluée (voir section suivante). Pour les sommets, nous noterons STR le nombre de sommets reconstruits (à l'exception des artéfacts). Plusieurs cas se produisent : certains sommets seront jugés aberrants, et ne seront pas évalués (leur nombre sera noté SA), d'autres seront masqués (notés SM), d'autres, enfin, ne pourront pas être évalués, car ils ne correspondent pas à un détail physique du terrain. Ceci peut se produire lorsque la forme reconstruite est généralisée par rapport à la réalité. Nous noterons SGE ces cas là. Enfin, le reste des points (noté SC) sera comparé à la réalité avec un protocole explicité dans la prochaine section.

Nous aurons pour chaque ligne des tableaux : $STR = STE + SGE + SM + SA + SC$.

2.3 Evaluation générale des formes

Sur les 103 zones évaluées, comportant 141 bâtiments, 111 sont modélisables, 86 bâtiments voient leur forme correctement reconnue par le système, soit un taux de 61%. Si l'on considère les bâtiments reconnus complètement, mais éclatés par sous-modèles, (colonne BRE), pour lesquels la description du volume est correcte, mais pas la topologie, on peut estimer le nombre de bâtiments correctement reconstruits à 101, soit un taux de 71%.

20 bâtiments sont surgénéralisés, soit 14%.

Ces résultats nous semblent conformes aux objectifs initiaux, même si nous verrons dans la section suivante que la précision géométrique atteinte n'est pas tout à fait satisfaisante.

2.4 Précision géométrique des sommets 3-D

2.4.1 Protocole d'évaluation

Pour chaque zone reconstruite, le système propose une vue 3-D de la reconstruction obtenue par combinaison de reconstructions individuelles (voir le chapitre précédent). Les sommets 3-D sont par ailleurs projetés dans les images stéréoscopiques de la zone.

L'opérateur peut alors choisir sur les images du couple les points à évaluer et effectuer une saisie bi-monoscopique du point 3-D réel.

Le choix se fait par click au voisinage d'un des points projetés de la reconstruction. La saisie de la position 3D servant de référence pour le point se fait ensuite par click sur les deux images du couple. Une fenêtre de zoom variable est disponible pour aider le pointé. Un fois les points ainsi saisis, les données de référence (les saisies) et les points 3-D de la reconstruction correspondant sont stockés pour analyse.

Afin de ne pas être tenté d'améliorer les résultats de l'évaluation, les points projetés vus par l'opérateur ne sont pas rigoureusement les projections des points de la reconstruction. Le système ajoute en effet un bruit aux points projetés avant de les afficher. Ce bruit est important (jusqu'à 10 pixels), rendant impossible toute tentative de saisir la référence précisément sur la reconstruction.

Le dispositif de pointé est présenté dans la figure 6.7.

2.4.2 Résultats

Pour les bâtiments pris individuellement, les écarts quadratiques moyens des positions 3D des sommets sont présentés dans les 3 tableaux de la section précédentes. Ces résultats sont très fluctuants

(de 0,37 m à plus de 3m).

En considérant l'ensemble des points 3D mesurés sur tous les bâtiments, l'étude montre un biais de 10,7 cm en X (Ouest-Est), de 1cm en Y (Nord-Sud), et de 0,9cm en Z. Le biais en X reste inexpliqué au moment de la rédaction de ce mémoire. Il correspond à 1,2 fois la résolution planimétrique des images.

Les histogrammes des écarts en X, Y, et Z, ainsi que les histogrammes des distances planimétriques et des distances 3D sont présentés dans les figure 6.8 à 6.12.

La précision géométrique est globalement faible par rapport à la résolution planimétrique et altimétrique des images. Ceci s'explique par le mécanisme de reconstruction qui place les sommets 3D à une position «moyenne» autour des primitives auxquelles ils sont censés appartenir. Par ailleurs, les primitives elles-mêmes, notamment les discontinuités, souffrent de délocalisations.

2.5 Evaluation du mécanisme d'alerte

Nous avons défini, pour les zones faisant partie de l'évaluation une «vérité alarme» précisant quelles zones sont estimées correctement reconstruites, et quelles zones sont estimées incorrectement reconstruites. Nous rappelons que le système d'alarme proposé dépend d'un paramètre qui est le rayon de la fenêtre d'érosion des cartes d'analyse présentées au chapitre précédent. Nous proposons d'étudier le comportement du protocole d'alarme en fonction de ce paramètre.

A cette fin, la figure 6.13 présente 4 courbes, tracées à partir de l'étude de 98 zones de focalisation. Sur ces zones, 54 ont été jugées correctement reconstruites, et 44 non. A ce titre, l'idéal serait que le mécanisme d'alarme produise des alertes en même temps que la «vérité alarme». Sur cette figure, l'axe des X correspond au paramètre σ_a qui permet d'éroder les cartes d'analyse et produire des alertes. L'axe des Y correspond au nombre de zones correspondant aux 4 cas possibles suivant que la «vérité alarme» et le mécanisme d'alarme produisent une alarme ou pas. Dans ces 4 cas, 2 cas montrent des cohérences entre le mécanisme d'alerte et la «vérité alerte» : la courbe bleu foncé correspond au cas des zones correctement reconstruites (pas d'alarme), et pour lesquelles le mécanisme d'alerte ne produit pas d'alarme. La courbe rouge correspond ou cas les deux produisent des alarmes. La courbe jaune correspond aux sous-alarmes : les cas où le mécanisme n'a pas produit d'alarme là où il aurait dû. Enfin, la courbe bleue correspond aux sur-alarmes.

On constatera que pour $\sigma_a = 0$, le mécanisme ne fait que produire des alarmes. Ceci provient du fait qu'après érosion de rayon nul, il subsiste toujours au moins 1 pixel sur les cartes d'analyse, ce qui engendre des alertes systématiques.

Lorsque σ_a croît, le nombre d'alarmes baisse, en particulier celui des sur-alarmes. Les sous-alarmes augmentent assez peu jusqu'à une valeur de 1m.

Pour cette valeur de 1m, précisément, le nombre de sous-alarmes est de 3 (soit bien moins de 10%), alors que les deux courbes où la vérité et le mécanisme automatique sont cohérents sont globalement les plus élevées. Le nombre total d'alarmes est de 67 dont 26 suralarmes. Le rapport sur-alarmes / alarmes est donc de 26 / 67, soit 40%. Ceci signifie que sur 5 alarmes, deux sont fausses, mais 3 sont vraies pour $\sigma_a = 1m$.

Par ailleurs, pour $\sigma_a = 1,4m$, il y a environ 50% d'alarmes. Le rapport entre sur-alarmes et alarmes vraies est alors également d'environ 40%, mais le nombre de sous-alarmes monte à 15, soit $15 / 54 = 30\%$.

Le mécanisme d'alarme est donc un peu «alarmiste» dans la mesure où d'une part pour un taux de sous-alarme acceptable, il a tendance à produire trop de sur-alarmes, et d'autre part si l'on souhaite limiter le nombre d'alarmes à 50%, alors il y a toujours trop de sur-alarmes, et donc de sous-alarmes.

On notera toutefois que ces évaluations sont biaisées par la présence de végétation dans bon nombre de zones, ce qui induit précisément des sur-alarmes.

3 Extensions

La méthode souffre d'un manque de précision géométrique. A cette fin, des travaux abordables sont envisageables. Typiquement, la zone 169 a été reconnue comme le modèle D, mais sa géométrie réelle est très mauvaise. Il serait possible, grâce au modèle, de modifier la position de la gouttière du bâtiment, qui est censée être parallèle au faite de ce bâtiment. Cet exemple montre que nous n'avons pas encore tiré profit de toute l'information du modèle.

Par ailleurs, une caractérisation des reconstruction tenant compte des images et non seulement du MNE est envisageable. Des travaux existent dans ce domaine, par exemple ceux de Basri ([BW]) ou Shufelt ([Shu99]).

4 Conclusion

Ce chapitre a montré que les modèles utilisés sont susceptibles de représenter une grande proportion des bâtiments du milieu urbain.

Sur un grand nombre de zones de test réelles, nous avons ensuite évalué la performance de la méthode en terme de reconnaissance des formes, de précision géométrique, ainsi que le mécanisme d'alarme.

En terme de reconnaissance des formes, les résultats peuvent être jugés satisfaisants, puisque 60% des zones sont reconstruites avec les modèles corrects. En ajoutant à cet effectif les volumes correctement reconstruits, mais éclatés en sous-modèles (colonne BRE des tableaux), nous parvenons à 70%.

En terme de précision géométrique, les résultats ne sont pas à la hauteur de la précision planimétrique des images initiales, mais ce système n'est pas complet, il lui manque cruellement un système de recalage a posteriori, qui est réellement envisageable.

Enfin, le mécanisme d'alerte a tendance à alarmer un peu trop, mais nous tenons à signaler que certaines zones recouvrent de la végétation, qui biaise de système de reconstruction dans son ensemble.

Globalement, le système possède donc des caractéristiques intéressantes, même s'il est encore perfectible.

L'inspection visuelle des zones reconstruites permet également d'observer des propriétés du système non quantifiables par les procédés mis en oeuvre ici, et sur lesquelles nous souhaitons insister.

Premièrement, on constatera des échecs patents sur certaines zones de difficultés variables (zones 165, 121, 119, 93, 76, 57). En revanche, dans certains cas, la complémentarité des primitives a permis de contrer les occultations liées soit à des bâtiments, soit à la présente ce végétation (zones 108 (gouttière partiellement occultée sur l'image de droite), 98, 79, 60, 51, 46, 21, 17).

Enfin, pour certaines zones, composées d'une surface topographique non continue, les reconstructions effectives tiennent compte des discontinuités de la surface pour produire une juxtaposition de reconstructions individuelles (zones 152, 109, 108, 98, 87, 82, 73, 48, 46, 45, 37, 17, 11), ce qui est une qualité importante du système mis en oeuvre ici.

N	ba	bm	pa	pag	br	brg	bre	brp	pr	av	ab	sag	sc	sa	sm	sg	str	e
172	1	1	2	2	1				2			6	6				6	0,62
170	1	1	2	2	1				2			6	6				6	0,75
169	1	1	2	2	1				2			6	5	1			6	1,8
165	1	0	8-10			1			1			15		4			4	
164	1	0	3	3		1			1			10	1	1	2		4	0,75
163	1	1	2	2		1			1			6	4				4	0,65
162	1	0	8	1	1				1			10	4				4	1,3
155	1	1	1	2	1				2	1	1	4	4				4	0,70
152	2	2	9	2	2				2			8	8				8	0,76
151	1	1	2	2	1				2			6	6				6	1,6
150	1	1	2	2	1				2			6	3			3	6	0,6
147	1	1	2	1	1				1			4	4				4	0,98
143	1	1	2	2	1				2	1		6	6				6	1,1
142	1	1	1	1	1				1	2		4	2	2			4	0,67
141	1	1	2	2	1				2			6	6				6	0,70
132	1	0	4	4				1	4			10	7			2	9	0,65
131	1	1	2	2	1				2			6	6				6	0,88
130	1	1	2	2		1			1		1	6	2		2		4	2,7
128	1	1	2	2	1				2			6	6				6	0,65
127	1	1	2	2	1				2			6	6				6	0,49
123	1	0	8	1	1				1			4	3		1		4	0,84
122	3	2	12		1				1				2	2			4	2,4
121	1	1	2	2					1	1		6	2	2			4	0,75
120	1	0	5	5				1	2			12		12			12	1,3
119	1	1	1	1					1	1		4		4			4	
116	1	0	2	1	1				1			4	3	1			4	1,1
113	1	1	2	2		1			1			6	4				4	2,3
112	1	1	2	2	1				2			6	6				6	0,56
110	1	1	2	2	1				2			6	6				6	1,1
109	1	1	2	2	1				2		1	6	6				6	0,53
108	2	2	4	4			1	1	3			12	9		3		12	0,81
107	1	1	2	2	1				2	1		6	6				6	0,98
106	1	1	2	2	1				2			6	6				6	0,77
104	1	1	2	2	1				2	2		6	6				6	0,82
102	1	1	2	2	1				2			6	6				6	0,79
101	1	0	3	2	1				2			12	6				6	1,1
98	2	1	3	2	2				2			8	6		2		8	1,1
97	2	2	3	3			1	1	2			10	8				8	1,4
96	1	0	7	7				1	3			12	5			5	10	0,75
95	1	0	15				1		2	2		20	1			5	6	0,70

TAB. 6.2 – Evaluation pour les zones de la scène (première partie).

N	ba	bm	pa	pag	br	brg	bre	brp	pr	av	ab	sag	sc	sa	sm	sg	str	e
94	2	0	14	2	2				1			8	4				4	1,0
93	1	0	1	1					0			4	0					
92	1	0	1	1	1				1			4	4				4	0,93
88	1	1	2	2	1				2			6	6				6	0,37
87	2	2	4	4	1		1		4			12	9	5			14	0,94
84	1	1	2	2	1				2	1		6	5	1			6	1,4
82	2	2	2	2	2				2		1	8	8		4		12	1,5
81	1	1	2	2			1		2			6	8				8	1,29
79	1	1	4	2	1				2			6	3		1	2	6	0,64
76	1	1	2	2	1				2	1	1	6	4		2		6	1,2
75	1	1	2	2	1				2			6	5		1		6	0,43
74	1	1	4	4	1				4		2	6	6				6	0,81
73	3	3	4	4	2			1	3	1		14	10	2			12	1,5
71	2	0	12	4	1				1			12	4				4	0,90
68	2	2	3	2	1				2		2	6	6				6	1,3
67	1	1	2	2			1		2	1		6	4	4			8	0,91
66	1	1	4	4	1				4	1		6	6				6	1,0
64	2	2	4	4	2				2		1	10	6				8	1,0
63	1	1	2	2	1				2			6	6				6	0,73
62	2	2	5	5		2			3	3		10	4		1	1	6	1,1
61	1	0	8	1	1				1	2		4	4				4	0,93
60	1	1	3	3	1				3			6	2	1	3		6	0,66
59	1	1	2	2	1				2			6	6				6	1,3
58	2	2	3	3	2				3	1		10	8	2			10	0,60
57	1	1	2	2				1	2	2		6	0	8			8	
55	1	1	4	4		1			3			6	2		2	2	6	0,54
54	1	0	7	4			1		4			9	7	2		3	12	1,5
53	1	0	8	1	1				2	1		4	4			2	6	0,94
51	1	1	3	3		1			1	1		5	1		1	2	4	1,8
50	1	1	2	2	1				2	2		6	1	5			6	0,33
49	2	2	4	4	1			1	3			12	2	7	1		10	1,0
48	1	1	2	2	2				4	1		6	10	2			12	0,55
47	1	0	3	3		1			2			8	5			1	6	1,0
46	1	2	4	4	2				4			12	11		1		12	0,74
45	1	2	4	4	2				4			12	12				12	0,69
42	1	1	3	3	1				3			6	3	1	2		6	1,8
41	1	0	1	1	1				1			4	4				4	1,8
40	1	1	2	2			1		2			6	3	5			8	1,3
39	1	0	4	1	1				1			4	4				4	0,48
38	1	1	4	4	1				2			5	4			2	6	0,63

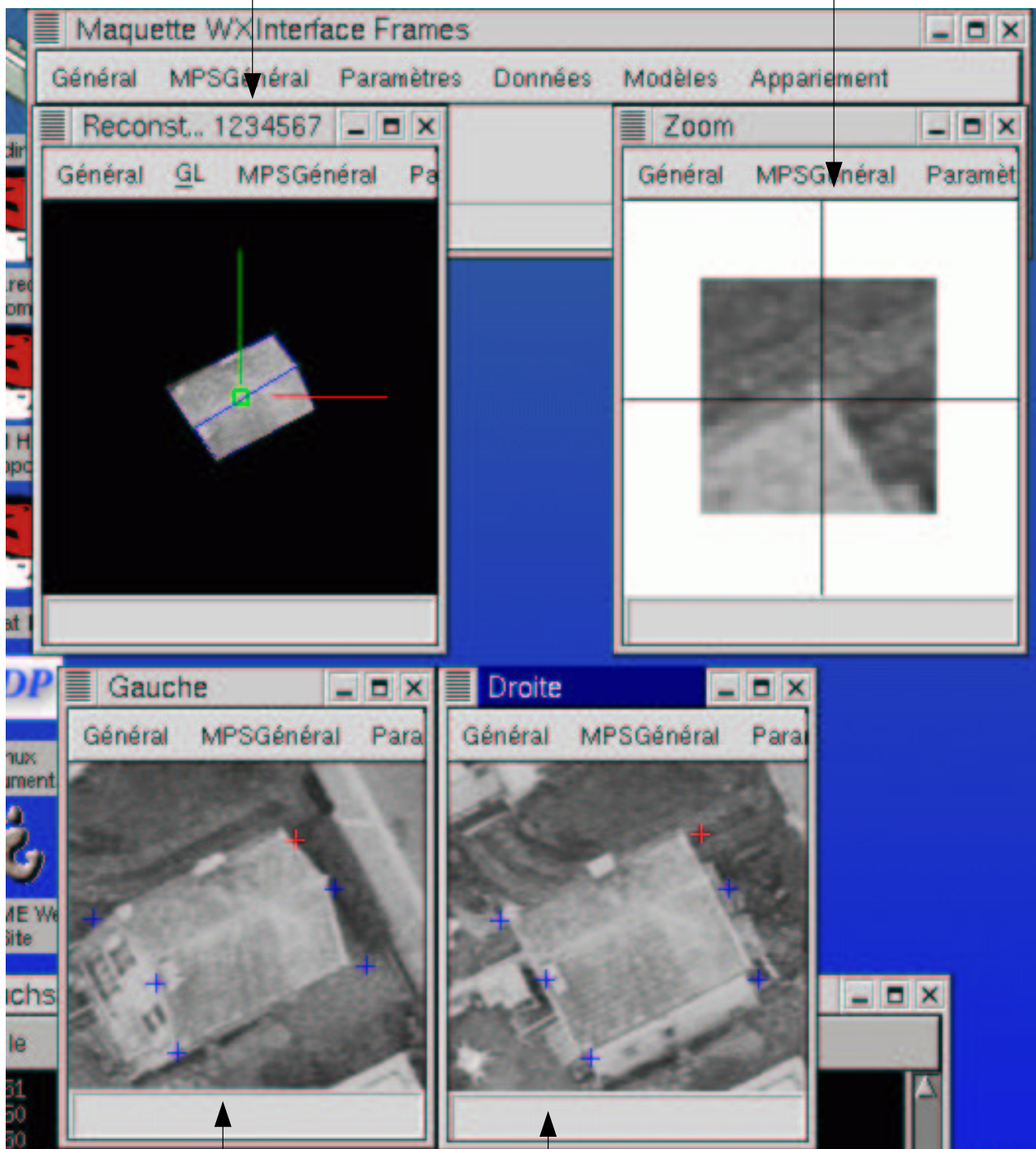
TAB. 6.3 – Evaluation pour les zones de la scène (seconde partie).

N	ba	bm	pa	pag	br	brg	bre	brp	pr	av	ab	sag	sc	sa	sm	sg	str	e
37	4	4	4	4	3		1		6			19	22				22	0,98
36	1	1	2	2	1				2	4		6	6				6	1,1
35	2	2	4	4		1	1		3			12	6	2	4		12	0,83
34	1	0	6					1	4				4	8			12	1,5
33	3	3	5	5		3			1			16	3	1			4	1,8
31	1	0	7	7		1			3			12	3	3			6	0,78
30	1	1	1	1	1				1			4	4				4	0,57
29	4	3	8	8	1	3			4			24	10		2		12	1,4
28	1	1	1	1	1				1		2	4	3	1			4	1,2
27	1	1	2	2				1	2	1		6	2	8			8	0,69
24	6	6	10	10	1	1	1	1	5			27	18			2	20	1,8
23	1	1	1	1	1				1			4	4				4	0,69
22	4	4	7	7	2			1	5		1	22	13	3			16	1,1
21	1	1	2	2	1				2	2		6	4		2		6	1,3
20	1	1	2	2			1		2	1		6	1	7			8	3,1
19	1	0	6	6		1			5			13	10			3	13	1,2
18	1	0	1	1			1		3			6	8			4	12	0,83
17	2	2	2	2	2				2			8	6		2		8	1,3
15	1	0	7	6			1					13	11					
13	1	1	6	4	1		1		4			12	12			2	14	
11	2	1	3	3	1	1			3			14	10				10	
10	2	2	1	1	1				1			4	2	2			4	
4	2	1	4	4	1				3			6	4	2			6	

TAB. 6.4 – Evaluation pour les zones de la scène (troisième partie).

Vue 3D de la reconstruction

Dispositif de zoom



Dispositif de pointé bi-monoscopique

FIG. 6.7 – Dispositif d'évaluation la précision des sommets 3-D des reconstructions.

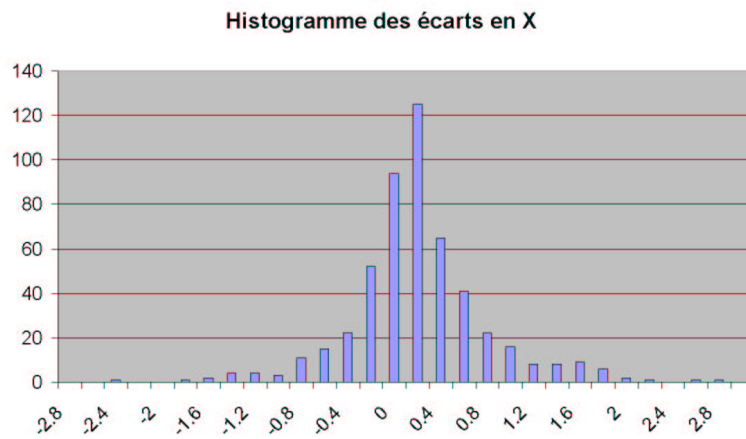


FIG. 6.8 – Histogramme des écarts mesurés en X.

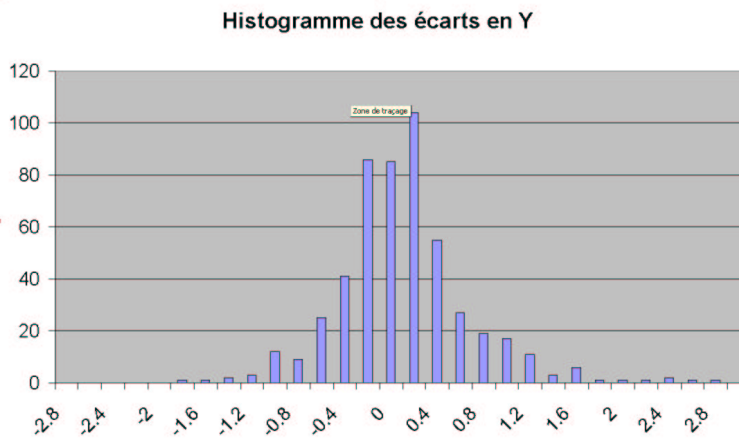


FIG. 6.9 – Histogramme des écarts mesurés en Y.

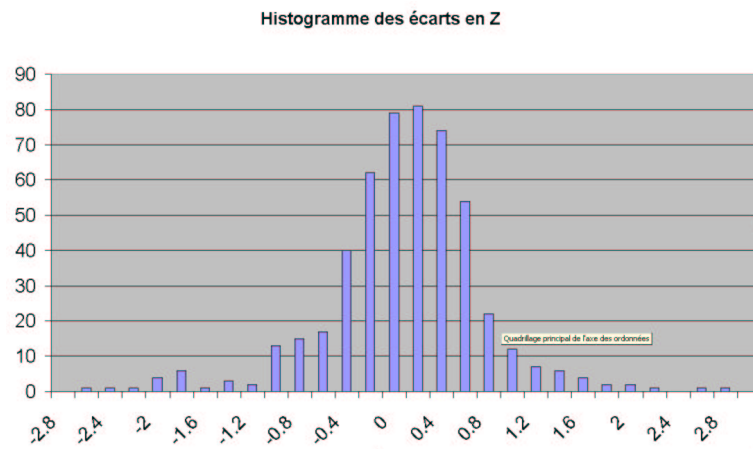


FIG. 6.10 – Histogramme des écarts mesurés en Z.

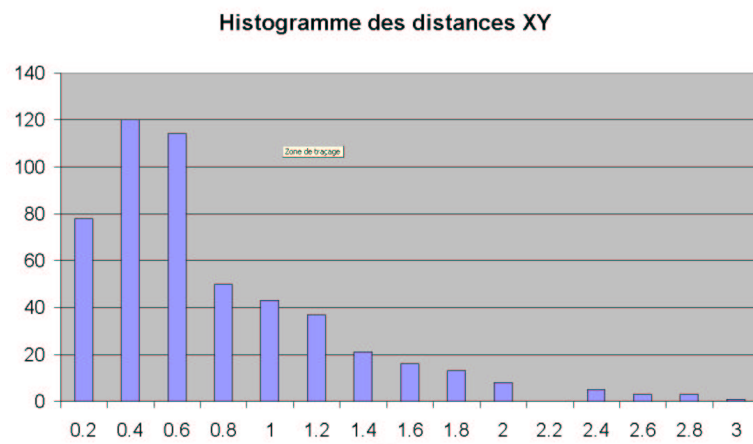


FIG. 6.11 – Histogramme des écarts planimétriques mesurés.

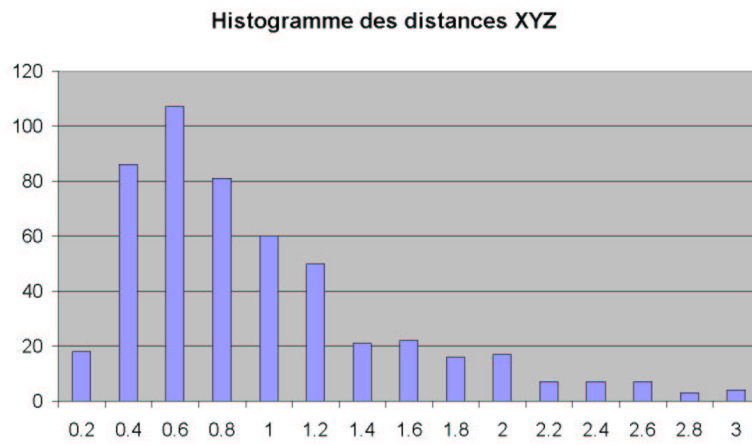


FIG. 6.12 – Histogramme des écarts 3D mesurés.

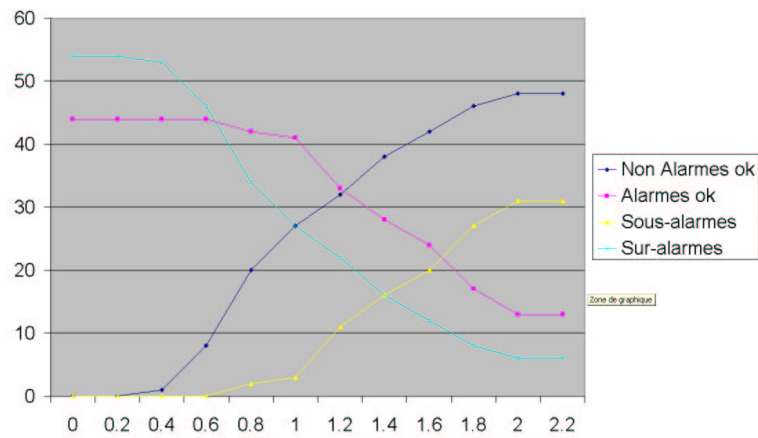


FIG. 6.13 – Courbes d'évaluation du système d'alarme.

Chapitre 7

Conclusion

1 Analyse de la contribution

Comme l'évoque le titre de ce mémoire, ces travaux sont une contribution à la reconstruction du bâti en milieu urbain. La méthode proposée ne permet effectivement pas de résoudre tous les problèmes liés à ce milieu.

Par contre, nous avons tenté de répondre à un certain nombre des questions liées à notre contexte spécifique : le milieu urbain et la stéréoscopie avec deux vues.

Après avoir analysé les inconvénients de la méthode, nous tenterons de mettre en valeur les points précis qui ont trouvé des réponses dans ces travaux, et qui représentent à notre avis la contribution.

1.1 Inconvénients

Lourdeur

Ces travaux font appel à un certain nombre d'outils théoriques que nous avons dû adapter à nos besoins propres. Les outils mis en oeuvre peuvent sembler lourds, et nous le concédons. D'un point de vue opérationnel, la lourdeur de la méthode est certaine, les appariements sont assez gourmands en temps machine. Toutefois, nous avons proposé des heuristiques permettant de limiter les temps de calcul, notamment dans la phase la plus difficile : l'appariement modèle / données.

Faible généralité

La méthode proposée n'est pas générique. Elle ne permet pas de représenter toutes les formes de bâtiments, mais seulement un sous-ensemble. Toutefois, les évaluations ont montré que même avec un sous-ensemble de formes, il est possible de reconstruire une proportion non négligeable de bâtiments.

Faible précision géométrique

Nous avons vu que la précision géométrique finale obtenue n'est pas à la hauteur des images initiales. En effet, les écarts quadratiques moyens et distances moyennes des sommets 3D mesurés sont bien au-delà de la résolution des images. Toutefois nous pouvons préciser que ces délocalisations proviennent du mode de reconstruction : les primitives utilisées souffrent de délocalisations propres, auxquelles s'ajoute le fait que les sommets 3D sont placés à des positions «moyennes» au voisinage de ces primitives, ce qui ne va pas nécessairement dans le bon sens. Il existe dans la littérature des techniques de calcul d'objets de moyen niveau (notamment linéaires) avec des précisions

sub-pixellaires. Il était inutile de mettre en oeuvre ce type d'approche car le positionnement des sommets 3D engendre par après des délocalisations, ce qui dégraderait de manière systématique les précisions obtenues pour les primitives. Par ailleurs, le MNE est connu pour souffrir d'un certain nombre d'erreurs planimétriques au voisinage des zones de contraste, ce qui induit naturellement des délocalisations de primitives de type discontinuité. Ces raisons expliquent certainement la précision finale obtenue.

Par contre, des techniques de recalage sont tout à fait envisageables, a posteriori, même si elles n'ont pas été traitées dans ce mémoire. Ces techniques devraient pouvoir profiter de l'information contenue dans les modèles (notamment le parallélisme de certains objets, et l'horizontalité des gouttières). La précision géométrique peut donc, à notre avis, être sensiblement améliorée dans un nombre important de cas.

Qualification approximative des reconstructions

Le système de qualification des reconstructions, qui permet également leur combinaison, n'utilise que le MNE. Celui-ci est certes réputé fiable et assez précis, mais il n'y avait pas de raison d'exclure a priori les images initiales, qui sont une source d'information plus «pure» que le MNE, dans le sens où le MNE est déduit de ces images, et donc tout de même entaché de lacunes.

1.2 Les contributions

Caractère déclaratif

Nous avons proposé une méthode structurelle et déclarative. Dans cette méthode, les connaissances sont représentées par un ensemble d'informations compact : les règles grammaticales, qui précisent de quoi un bâtiment est formé.

Extensibilité

Ce système déclaratif se montre extensible. Il est possible d'ajouter des règles, moyennant l'apprentissage de leur écriture. Les contraintes géométriques introduites dans les arcs sont elles aussi extensibles, car informatiquement, nous avons implanté une «base de contraintes» qu'il est aisé de modifier, voire d'augmenter.

Il est également possible d'ajouter d'autres types primitives. Ceci aura toutefois pour impact d'augmenter la combinatoire des problèmes d'appariement. Mais c'est envisageable.

Caractère sémantique

Les modèles peuvent être désignés de manière sémantique : «C'est un bâtiment en L», «C'est un bâtiment à double pan», «il y a 1 attaque en V», etc.

Le système passe donc par des étapes de niveaux très différents, des images au MNE, puis à des objets géométriques simples, et enfin à des reconstructions vectorielles topologiquement cohérentes, que l'on peut désigner par des mots de notre langage. Le niveau d'interprétation des zones s'arrête donc au-delà de la description vectorielle pour arriver au niveau sémantique.

Complémentarité des primitives

Nous avons proposé des types de primitives qui sont complémentaires (ce point a été justifié au chapitre 3). Les 3 types de primitives calculés sont détectés *indépendamment* les uns des autres, et se confortent mutuellement lors des appariements et des reconstructions effectives.

Re-focalisation

Nous avons montré que l'appariement peut focaliser son attention sur une sous-partie de la zone de focalisation issue du MNE, et reconstruire un bâtiment qui ne correspond pas à l'ensemble de

la zone de focalisation. Ceci est à notre sens un point fondamental, car en milieu urbain, en raison de la densité du milieu, cette propriété est très importante. Il n'est en effet pas rare que les critères habituellement utilisés pour déterminer les zones de focalisation (un critère d'altitude) ne soient pas suffisants pour délimiter des zones ne correspondant effectivement qu'à un seul bâtiment.

Ceci justifie l'intérêt de pouvoir identifier dans la zone, les bâtiments présents : nous ne supposons pas qu'il n'y a qu'un seul bâtiment dans une zone de focalisation, et, en particulier, la surface topographique de la zone peut être discontinue.

Stéréoscopie «simple»

Nous avons vu que la stéréoscopie simple souffre d'occultations nombreuses, qu'il est possible de contrer par la multiscopie. Dans notre cas, nous avons contré ces occultations par la redondance des primitives extraites, notamment par l'utilisation des discontinuités du MNE. Ceci a permis de reconstruire des bâtiments, malgré certains obstacles majeurs. En effet, en cas d'occultation, une discontinuité d'altitude apparaît nécessairement. La discontinuité ainsi détectée permet de placer des gouttières cachées.

Les deux contributions citées ici (refocalisation et lutte contre les occultations) nous paraissent être les deux problèmes fondamentaux auxquels l'approche structurelle proposée a apporté des réponses.

Système d'alarme

La comparaison du MNE initial et du MNE engendré par la reconstruction de la zone permet d'émettre des alarmes.. Nous avons réglé ce système de manière à conserver un minimum de sous-alarmes (moins de 10%), et à focaliser l'attention de l'opérateur humain sur les bâtiments mal reconstruits.

Evaluation

Nous avons proposé, grâce à une contribution importante de G. Maillet (MATIS / IGN), une évaluation a priori du domaine d'emploi de la méthode, à partir de données de référence complexes (milieu urbain dense), ainsi qu'une évaluation a posteriori (à partir des données initiales réelles : images et MNE), faite sur un grand nombre de zones, traitées avec des paramétrages identiques. La scène traitée n'a pas été spécialement choisie, puisque nous avons utilisé la première scène proposée dans le couple stéréoscopique, et ceci afin de rendre les performances de la méthode plus crédibles. En particulier, la zone du couple stéréoscopique traitée se trouve très à l'écart du nadir des prises de vues, ce qui est a priori défavorable, en raison des angles de vue plus forts que sur le nadir.

Pour cette évaluation, nous avons proposé (en annexe I) l'ensemble des reconstructions effectuées sur cette scène, ce qui permet d'illustrer abondamment le propos.

En cas de présence de végétation au voisinage du bâti, l'appariement proposé est ainsi en mesure de faire abstraction de la végétation pour reconnaître le ou les bâtiments. La végétation est bien entendu source d'artéfacts mais elle ne perturbe que peu la reconnaissance du bâti proprement dit.

2 Extensions possibles

Nous avons déjà cité le recalage a posteriori des reconstructions, qui est probablement la première chose à faire à l'issue de ces travaux, afin d'améliorer la précision géométrique finale des reconstructions.

Par ailleurs, il serait intéressant d'étudier l'importance relative des contraintes géométriques

introduites dans le système, voire, de définir un ensemble de contraintes minimal et suffisant pour résoudre tous les problèmes solubles par ce système. En quelque sorte, il s'agit ici de rechercher une base (au sens mathématique du terme : un sous-ensemble orthogonal et qui engendre l'espace entier) de contraintes. Cette tâche intéressante nous paraît toutefois ardue aujourd'hui.

L'apport de données externes n'est bien entendu pas à exclure. Dans le contexte opérationnel de l'IGN, notamment, où l'on envisage des coopérations avec les services du cadastre, une exploitation des données cadastrales est tout à fait envisageable. Les études menées par H. Jibrini (MATIS / IGN) montrent en effet que les bords de bâtiments du cadastre possèdent de très bonnes orientations 2D, ce qui en fait de bons candidats pour remplacer ou compléter les discontinuités détectées dans ces travaux.

D'un point de vue tout à fait général, ces travaux nécessitent encore des études, notamment l'étude du paramétrage, que nous n'avons que peu abordé ici, ainsi que des évaluations complémentaires afin de vérifier la stabilité des résultats obtenus.

Des études visant à exploiter des images de la caméra numérique de l'IGN devront également être menées. Ces images disposent en effet d'un rapport signal / bruit bien supérieur aux images utilisées ici, mais les résolutions obtenues actuellement sont un peu plus faibles (20 cm).

Perspectives

Un certain nombre de perspectives peuvent être vues dans les paragraphes précédents, sans oublier une extension majeure qui est compatible avec la démarche de la méthode et les techniques utilisées. Il serait en effet intéressant de travailler à la reconnaissance des modèles à partir des données initiales, à l'aide de la grammaire, en utilisant des méthodes de type syntaxique, ce qui éviterait de «tester» les modèles avant de choisir les meilleurs. Ceci renforcerait le caractère structurel de l'approche.

3 Conclusion

Cette approche structurelle pour la reconstruction du bâti en milieu urbain, à partir de deux vues, est une contribution qui a visé à avancer sur une piste encore peu explorée actuellement. Il se place dans un projet plus général, en exploitant notamment abondamment les travaux de C. Baillard visant à détecter les zones de sursol en milieu urbain.

La combinatoire de l'appariement a été un problème assez difficile, et récurrent. Il a fallu la contrer à plusieurs reprises. En conséquence, nous estimons que les méthodes de calcul utilisées ne sont pas nécessairement adaptées à des problèmes mettant en jeu des graphes plus grands (plus de 100 sommets pour les modèles, et plus de 1000 sommets pour les données). Mais nous espérons avoir montré que dans ce contexte, les méthodes utilisées sont concluantes, malgré les craintes qu'il est possible d'avoir a priori concernant ce type d'approche, notamment en raison du caractère NP-complet du problème de l'appariement.

Sur un certain nombre de points, nous nous sommes placés en décalage par rapport à l'existant dans le domaine. La représentation des modèles par des graphes n'a en effet pas été appliquée à des fins de reconstruction du bâti à notre connaissance, même si ce formalisme a été utilisé en analyse d'images, mais d'en d'autres domaines, en se limitant à des objets bi-dimensionnels (analyse de documents, reconnaissance de réseaux routiers). De même, nous avons exploité simultanément trois types de primitives, ce qui a été peu fait dans la littérature, notamment par ce qu'il est difficile de gérer

les interactions entre plusieurs types de primitives. Cette gestion des interactions est précisément réalisée par le formalisme unifié des graphes, qui intègre naturellement les relations entre objets de types variés.

Nous pensons avoir répondu à un certain nombre de problèmes d'ordre académique, en formalisant l'approche de reconstruction du bâti tel que cela a été fait, et d'un point de vue opérationnel, en adaptant les méthodes de calcul de manière à faire rentrer ces méthodes a priori excessivement combinatoires dans des ordinateurs standards, et en évaluant l'approche (les reconstructions et le système d'alarme) sur un grand nombre de zones test.

Nous remercions le lecteur parvenu au terme de ce mémoire pour l'intérêt porté à ces travaux.

Liste des publications

F. Fuchs, H. Le Men. *Detecting planar patches in urban scenes*. Proceedings of SPIE, Visual Information Processing VIII, vol. 3716, pages 167-176, avril 1999, Orlando.

F. Fuchs, H. Le Men. *Building reconstruction on aerial images through multi-primitive graph matching*. Proceedings of the TC-15 IAPR Workshop on Graph-based Representations, mai 1999, Vienne.

F. Fuchs, H. Le Men. *Efficient subgraph isomorphism with 'a priori' knowledge. Application to building reconstruction for cartography*. Proceeding of the IAPR Workshop on Syntactic and Structural Pattern Recognition 2000 (SSPR'2000), Alicante, août-septembre 2000. Lecture Notes in Computer Science n° 1876.

Annexe A

Glossaire

Cette annexe ne prétend pas définir tous les termes proposés, mais, pour certains, il ne fait que préciser le sens dans lequel ils sont utilisés dans ce mémoire.

- appariement : (*Petit Robert*) « Action d'apparier ». Apparier : « Assortir par paires ou par couples ». C'est donc un acte ou un objet définissant une relation bijective entre deux ensembles. Dans notre cas, la relation bijective cherchée se place entre un sous-graphe d'un modèle, et un sous-graphe du graphe des données.
- arc : (*Graphes et algorithmes, p. 2*) Couple de sommets d'un graphe. Un arc définit une relation particulière entre deux sommets du graphe. Il peut être orienté. Dans ce cas, on distingue le sommet *initial* et le sommet *final* de l'arc. Par exemple la relation *être plus haut que* est une relation orientée, et doit être modélisée par des arcs orientés.
- arc attribué : arc d'un graphe auquel est associé une valeur ou un ensemble de valeurs (les attributs) décrivant de manière symbolique un ensemble de propriétés.
- axiome : (*Petit Robert*) Vérité indémontrable mais évidente pour quiconque et comprend le sens (principe premier), et considérée comme universelle.
- base de données :
- bâti : thème cartographique dans lequel sont représentés les bâtiments.
- bit : (*Petit Robert*) : « acronyme anglais de **binary digit unit** ». Il s'agit de la plus petite entité qui puisse être représentée par un ordinateur. C'est en fait une case électronique pouvant représenter les valeur 0 ou 1. Cette unité est utilisée pour représenter des quantités d'information : 1 octet = 8 bits, 1 ko = 1024 octets, 1 Mo = 1024 ko, ...
- caméra : classiquement en cartographie, on appelle caméra un appareil photographique (analogique ou numérique). Il ne s'agit pas de caméras vidéo, car la durée entre deux photographies est généralement supérieure à la seconde.
- carte : représentation à échelle réduite d'une portion de la surface du globe.
- cartographie : (*Petit Robert*) Etablissement du dessin et de l'édition des cartes et plans. Dans [Rua99], page 16 : « Le cartographe représente l'espace géographique sous différentes formes ».
- chien assis : superstructure d'un toit composée d'un pan principale et de deux pans latéraux obliques.
- complémentaire (d'un ensemble) : Pour un ensemble E et un sous-ensemble $F \subset E$, le com-

plémentaire de F dans E est l'ensemble des éléments de E qui ne sont pas dans F . On peut le noter $F \setminus E$.

- composante : nous utiliserons ce terme pour désigner, dans une reconstruction, une sous-partie correspondant à un seul modèle engendré par une grammaire. Ceci s'entend lorsque la reconstruction est issue de l'utilisation simultanée de plusieurs graphes engendrés par la grammaire.
- composante connexe : (*Graphes et algorithmes*, p.)
- correction d'erreur : plus exactement : tolérance d'erreur. Propriété d'un système de reconnaissance des formes, qui tolère des erreurs ou fautes dans les données sur lesquelles il opère.
- couple stéréoscopique : paire d'images (aériennes ou non) prises de deux points de vue différents, et telles que des détails identiques du terrain.
- décomposition : structure de données exploitée dans l'appariement de graphes. Une décomposition éclate récursivement un graphe en sous-graphes. Voir le chapitre 3.
- discontinuité : sans entrer dans le formalisme de la continuité d'une fonction ¹, disons qu'une discontinuité est un point où une fonction subit une variation abrupte de sa valeur. Dans notre cas, nous considérons les discontinuités du modèle numérique d'élévation d'une scène, et nous remarquons simplement que le passage des gouttières des bâtiments au sol se traduit par une discontinuité de la fonction altitude.
- distance d'édition : distance entre deux graphes, suivant un formalisme particulier, consistant à modifier petit à petit l'un des graphes, par des éditions.
- échelle : (*Petit Robert*) « Rapport entre la taille d'un objet représenté sur une carte et la taille de l'objet dans le monde réel ». Ainsi, plus l'échelle est grande, plus la représentation des objets du terrain est grande, et peut donc être détaillée.
- édition : transformation élémentaire d'un graphe. Voir le chapitre 3.
- généralisation : Dans [Rua99], page 34 : (Définition de l'Association de Cartographie Internationale) : « la sélection et la représentation simplifiée de détails en fonction de l'échelle et des objectifs de la carte »
 Dans [Rua99], page 38 : « Généraliser est le moyen pour fournir une information dont le contenu et le niveau de détail correspondent à l'information nécessaire et suffisante pour faire des raisonnements géographiques spécifiques »
- géographie : (*Petit Robert*) « Science qui étudie et décrit la Terre à sa surface, en tant qu'habitat de l'homme et de tous les organismes vivants. La géographie est une science naturelle et humaine ».
 Dans [Rua99], page 16 : « Le géographe analyse l'espace géographique à partir d'informations géographiques pour en déduire de nouvelles connaissances géographiques ».
- grammaire de graphe : ou plutôt : grammaire par remplacement d'hyper-graphes. Une grammaire est un ensemble composé d'un axiome (le point de départ), et de règles, permettant d'engendrer un langage (dans notre cas : le langage est composé des graphes modèles).
- graphe : ensemble de sommets et d'arcs les reliant. Une définition plus précise des graphes utilisés dans ce mémoire est donnée page 53.
- graphe complet : (*Graphes et algorithmes*, p. 4) Graphe tel que pour tout couple de sommets, il existe un arc les reliant.
- graphe données : graphe relationnel attribué complet, comportant N sommets, où N est le

¹ $\forall \epsilon > 0, \exists h > 0, \forall x \in [x_0 - h, x_0 + h], |f(x) - f(x_0)| < \epsilon$

- nombre de primitives calculées à partir des données initiales, et où chaque arcs portent des labels à valeur continue, décrivant à quel point la propriété géométrique correspondant au label est vérifiée par les primitives.
- graphe modèle : graphe relationnel attribué décrivant de façon symbolique la forme d'un bâtiment. Les sommets définissent des primitives utilisées pour représenter le bâtiment, et les arcs décrivent des propriétés géométriques vérifiées par ces primitives.
 - graphe relationnel attribué : graphe dans lequel les sommets et les arcs portent des labels, ou attributs, représentant un information symbolique. Ces attributs peuvent être valués (i.e. porter une valeur numérique) qui permet de quantifier les informations symboliques.
 - graphe terminal : graphe ne portant pas de sommet non terminal.
 - image argentique : photographie traditionnelle, sur support argentique. Une image argentique peut être scannée pour produire une image numérique.
 - image numérique : matrice (ou tableau) de pixels. Une image numérique peut être obtenue par numérisation d'image argentique (image scannée), numérisation d'une image sur papier, par acquisition numérique directe (à l'aide d'une caméra numérique), ex nihilo (par création à l'aide d'un logiciel), par synthèse (images de synthèse), ...
 - image scannée : image numérique obtenue par scannage (numérisation) d'une image argentique.
 - isomorphisme : fonction associant de manière bijective deux ensembles d'objets (typiquement les sommets de 2 graphes), et qui respecte les formes des deux ensembles.
 - lucarne : superstructure d'un toit
 - maillé : représenté par une grille régulière. Cette notion s'applique dans le cas d'images, qui sont des grilles régulières de pixels, ou dans le cas de modèles numériques d'élévation, qui sont des grilles régulières d'altitudes dans notre mémoire.
 - mal posé (problème mal posé) : problème n'admettant pas une solution unique
 - modèle numérique d'élévation : représentation maillée de la surface topographique, c'est-à-dire de l'ensemble des objets vus depuis les airs.
 - modèle numérique de terrain : représentation de la forme du sol, à l'exclusion des objets topographiques (arbres, bâtiments, ponts, etc).
 - nœud : habituellement, les nœuds désignent les premiers objets représentés dans les graphes, que nous appelons les sommets dans ce mémoire. Le terme nœud sera réservé à des entités un peu plus complexes. Ces entités sont les nœuds de réseaux. Deux types de réseaux sont utilisés dans ce mémoire : les décompositions au sens de Messmer, et les réseaux grammaticaux.
 - ortho-image : image superposable à une carte. Les pixels d'une ortho images sont géoréférencés : il y a une correspondance non ambiguë entre les pixels d'une ortho-image et un grille posée au sol.
 - pan de toit : (*Petit Robert*) « Face d'un objet, d'une construction polyédrique »
 - pixel : «Picture element». Il s'agit de la plus petite entité constitutive d'une image numérique. Un pixel est représenté par une valeur numérique qui peut être binaire, entière à 256 valeurs ou plus (images en niveaux de gris), multi-dimensionnelle dans le cas des couleurs, ...
 - postulat : (*Petit Robert*) «Principe indémontrable qui paraît légitime, incontestable».
 - primitive : ici, objet tri-dimensionnel simple, soit calculé à partir des images (dans le cas des données), soit défini ex nihilo, dans le cas des modèles.
 - primitive discontinuité : portion de plan
 - primitive linéaire : segment de droite 3-D.

- primitive ponctuelle : point 3-D
- primitive surfacique : portion de plan 3-D
- raster (anglicisme) : Voir maillé.
- rayon perspectif : droite passant par le centre de prise de vue d'un dispositif à sténopé, et un point particulier de l'image.
- reconstruction : représentation vectorielle d'une zone de focalisation.
- relation binarire : relation entre deux objets (par exemple : le parallélisme de deux droite).
- relation ternaire : relation entre trois objets (par exemple : un objet est *entre* deux autres).
- réseau : graphe orienté sans cycle. Il s'agit d'un graphe orienté dans lequel aucun chemin ne peut passer deux fois par le même sommet.
- réseau grammatical : réseau dont les nœuds sont des graphes engendrés par une grammaire, et dont les arcs représentent les passages possibles entre graphes, par application d'une des règles de la grammaire.
- résolution : (*Petit Robert*) «Aptitude à mesurer une petite valeur d'une grandeur physique». Il s'agit, dans le cas d'images numériques, de la taille d'un objet au sol, dont la projection dans l'image a la dimension d'un pixel.
- surface topographique : surface $z = f(x, y)$ passant par les points les plus hauts des objets topographiques : les toits, les arbres, notamment, et le sol, lorsqu'aucun objet topographique ne se trouve au-dessus du sol.
- sursol : ensemble des objets situés au-dessus du sol.
- terme : l'un des 3 graphes intervenant dans un règle grammaticale (voir le chapitre 4).
- scène : région de grande étendue, vue par un couple stéréoscopique (voir le chapitre 1).
- singleton : ensemble contenant un et un seul élément. Par extension, graphe ne contenant qu'un seul sommet.
- sommet : élément constitutif d'un graphe. Les sommets sont reliés par des arcs.
- sténopé : situation de prise de vue dans laquelle tous les rayons perspectifs convergent vers un point. Historiquement, ce cas correspond à des boîtes parallélépipédiques percées d'un trou au centre d'un des faces. Ainsi, les rayons lumineux traversent ce trou et arrivent sur la face opposée à celle percée. A notre connaissance, c'est ainsi que les premières observations du soleil on été faites. Dans le cas des caméras argentiques utilisées à l'IGN, le matériel utilisé est de qualité telle que l'on se trouve dans une situation très proche de la sténopé, malgré des optiques lourdes et épaisses.
- sous-graphe : graphe appuyé sur un sous-ensemble de l'ensemble des sommets d'un graphe donné.
- stéréorestitution : technique visant à restituer la forme tri-dimensionnelle d'un objet à l'aide de deux vues de cet objet.
- stéréoscopie : « vue dans l'espace » situation dans laquelle deux images permettent de voir le même objet depuis deux points de vue différents, ce qui permet d'en percevoir la forme tri-dimensionnelle
- structurel : se dit des techniques utilisant des modèles donnés de manière déclarative, soit explicitement, ou à l'aide d'outils de plus haut niveau, tels que les grammaires.
- superstructure :(*Petit Robert*) Partie (d'une construction) située au-dessus du sol, d'un niveau. Nous détournerons ce terme pour désigner les volumes se situant au-dessus des pans de toit, tels que les lucarnes, cheminées, chiens assis.

- vecteur, vectoriel : En simplifiant, un vecteur est un couple (ou un triplet, le cas échéant), de coordonnées. Une représentation vectorielle est une représentation (d'une forme 2D ou 3D), appuyée sur les coordonnées des points constitutifs de cette forme, moyennant la mise en place des arcs reliant ces points, et éventuellement les faces.
- zone : portion de l'espace sur laquelle nous travaillons. Les zones sont obtenus à partir des scènes après une étape de focalisation (voir le chapitre 1). Les zones sont constituées de 1 ou quelques bâtiments.
- zone de recouvrement : région du terrain couverte par les deux images d'un couple stéréoscopique, et, par extension, on dénommera ainsi les parties des deux images correspondantes.

Annexe B

Résumé des notations

Certaines notations sont utilisées plusieurs pour désigner des entités distinctes. En fonction du contexte, le lecteur saura toutefois reconnaître sans ambiguïté quel objet est désigné par la notation.

- B : base stéréoscopique
- C, D : centres de prise de vue
- C() : la fonction de coût d'une opération d'édition, ou d'une édition
- D : une décomposition
- E : l'ensemble des arcs d'un graphe
- e : échelle, ou un arc d'un graphe
- f : un morphisme ou isomorphisme de graphe, ou de sous-graphe. f est en fait une fonction liant les ensembles de sommets de deux graphes : $f : V \rightarrow V'$.
- \mathcal{F} : l'extension d'une fonction $f : V \rightarrow V'$ à $\mathcal{P}(V)$. C'est-à-dire : $\mathcal{F}(v_1, \dots, v_k) = \{f(v_1), \dots, f(v_k)\}$.
- G, H, G', G'', ... : un (hyper-)graphe
- (G, G', G'', E) : un nœud de décomposition
- $|G|$: l'ordre d'un graphe (son nombre de sommets)
- H : hauteur de vol
- h(S) :
- I, J : images
- K : la constante de pondération entre le coût de suppression d'un sommet et les autres coûts d'édition.
- L(.) : la fonction de longueur de description
- $lp(.,.)$
-
- L_V : l'ensemble des labels de sommet d'un graphe
- L_E : l'ensemble des labels d'arc d'un graphe
- \mathcal{P} : une partition
- (n, l) : l'attribut d'un sommet : n est la nature, l l'attribut numérique.
- O(R) : la frontière d'une région
- P, Q : points d'une image
- $\mathcal{P}(E)$: l'ensemble des sous-ensembles d'un ensemble E.
- P_s : l'ensemble des sommets po

- R, R_i, R_j : une région
- $R_i \leftrightarrow R_j$: l'adjacence de deux régions
- S : un ensemble de pixels
- $S=(G, G', G'', E)$: un nœud de décomposition
- V, V' : l'ensemble des sommets d'un graphe
- v : un sommet d'un graphe
- $\$$: le néant
- δ : une opération d'édition
- Δ : une séquence d'édition
- (Δ, f_Δ) : un isomorphisme de graphe avec correction d'erreur
- μ : la fonction d'attribution des labels pour les sommets d'un graphe
- ν : la fonction d'attribution des labels pour les arcs d'un graphe

Annexe C

Figures pour la zone 172.

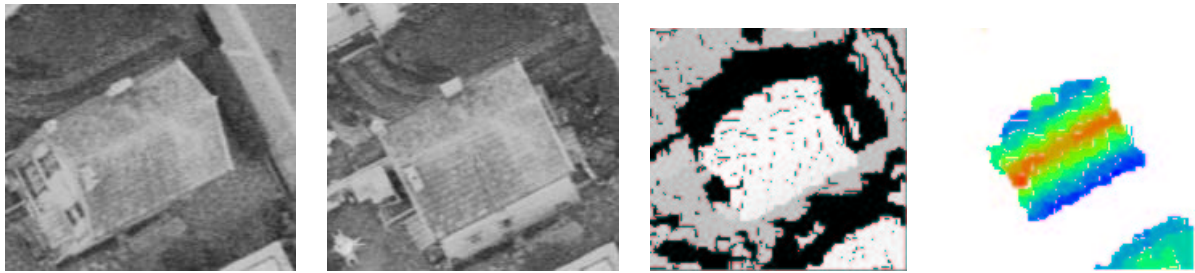


FIG. C.1 – Vue stéréoscopique du bâtiment 172, avec son MNE brut et son MNE de travail (en couleur : bleu pur = altitude 35 m, rouge pur = altitude 37,5 m)

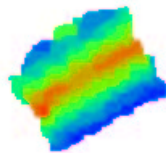


FIG. C.2 – MNE retravaillé pour le bâtiment 172.

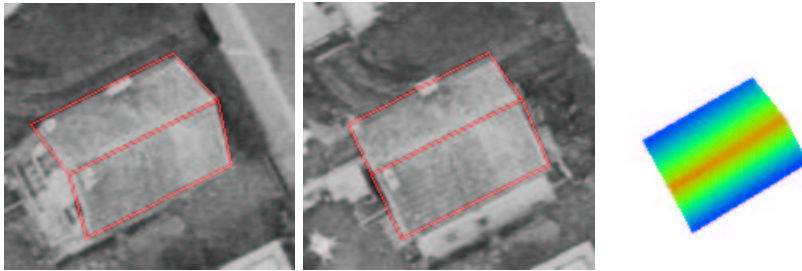


FIG. C.3 – Reconstruction du bâtiment 172, avec le MNE déduit (MNER)

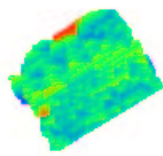


FIG. C.4 – Carte des erreurs pour le bâtiment 172 (bleu : -0,6 m, rouge : + 0,9 m)



FIG. C.5 – Cartes pour le bâtiment 172, de gauche à droite : extrapolation, omission, aberrations

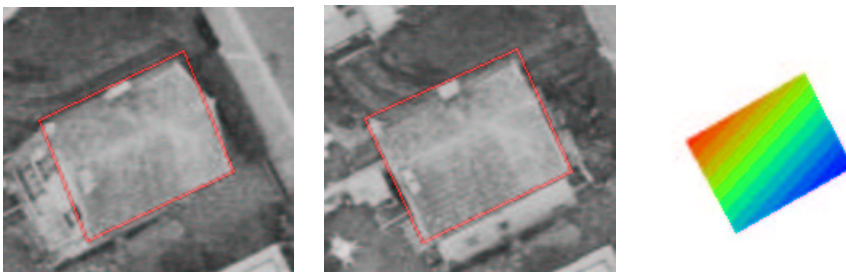


FIG. C.6 – Reconstruction du bâtiment 172 avec un modèle de parallélépipède. A droite, le MNER (bleu : 35,3 m, rouge : 36,5 m)

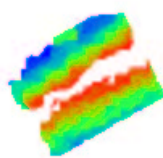


FIG. C.7 – Carte des erreurs pour le bâtiment 172, avec pour un modèle parallélépipédique (bleu : -0,9 m, rouge : + 0,9 m)

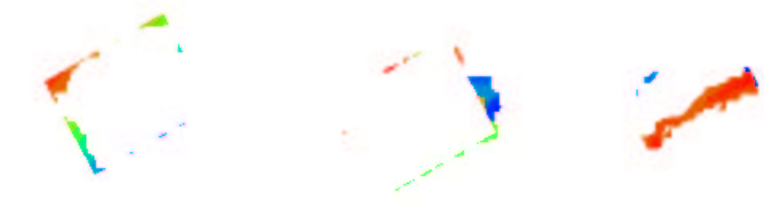


FIG. C.8 – Cartes pour le bâtiment 172 avec reconstruction pour un modèle parallélépipédique, de gauche à droite : extrapolation, omission, aberrations

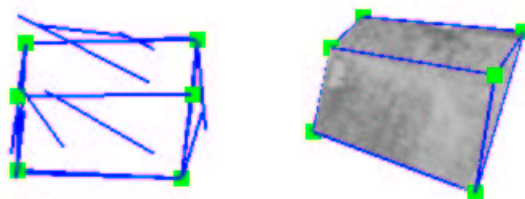


FIG. C.9 – Reconstruction vecteur du bâtiment, superposée aux données linéaires (gauche), et vue perspective (à droite) de la zone 172 reconstruite avec un modèle à 4 pans : cas de dégénérescence de certains pans de toit.

Annexe D

Figures pour la zone 22.

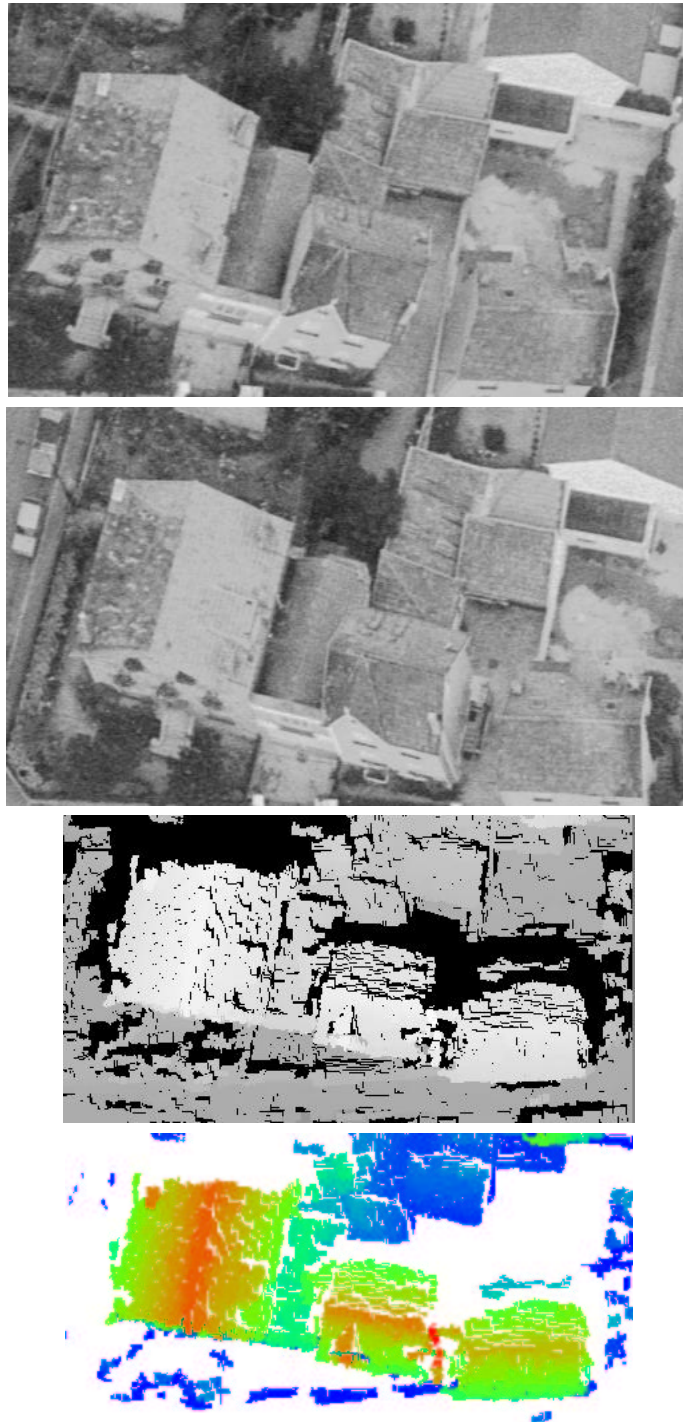


FIG. D.1 – Vue stéréoscopique du bâtiment 22, avec le MNE brut et le MNE de travail (en couleur : bleu = 30 m, rouge = 41 m)

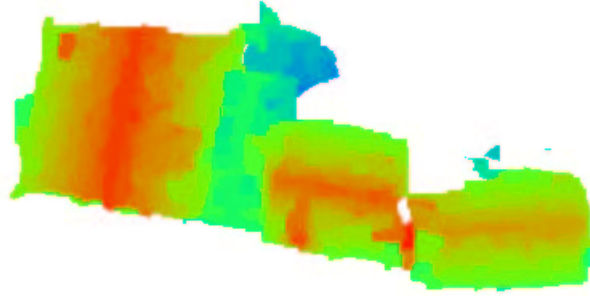


FIG. D.2 – MNE retravaillé pour le bâtiment 22.

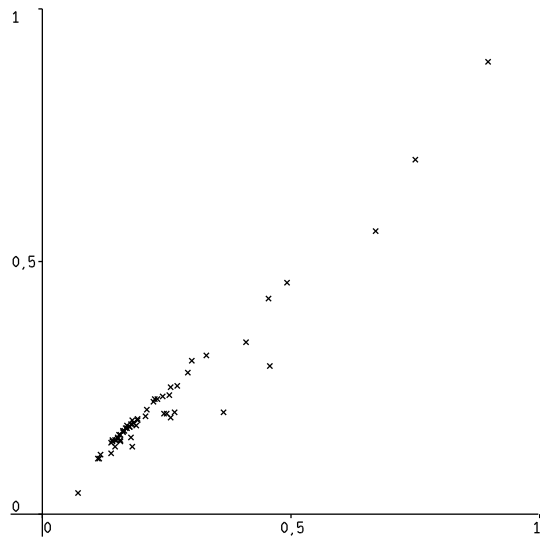


FIG. D.3 – Ecarts quadratiques médians pour la zone 22.



FIG. D.4 – Reconstruction du bâtiment 22, avec le MNE déduit (MNER)

Annexe E

Figures pour la zone 19.



FIG. E.1 – Vue stéréoscopique du bâtiment 19.

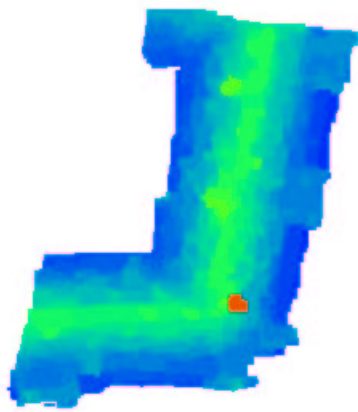


FIG. E.2 – MNE de travail pour le bâtiment 19 (en couleur : bleu = altitude 46 m, rouge = altitude 52 m). On distingue en rouge une petite aberration du MNE

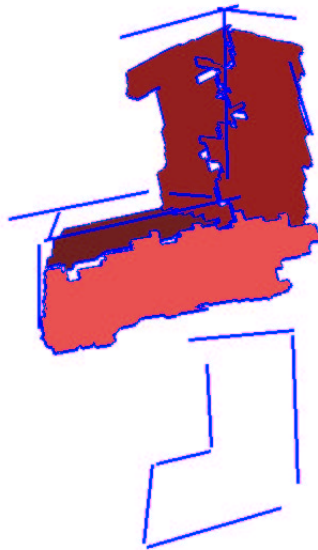


FIG. E.3 – Appariement des primitives pour le modèle DD.



FIG. E.4 – Vue stéréoscopique d’une reconstruction du bâtiment 19.

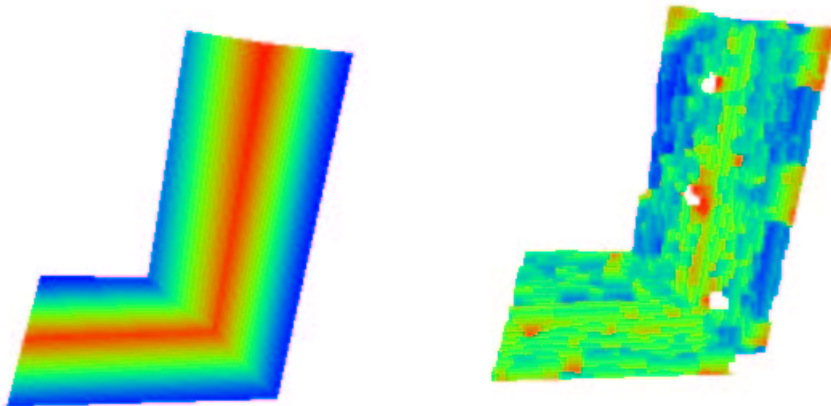


FIG. E.5 – MNER du bâtiment 19, et sa carte des erreurs (bleu : -0,7m, rouge : +0,9m)



FIG. E.6 – Vue stéréoscopique d'une reconstruction du bâtiment 19, avec un angle biseauté du bâtiment.

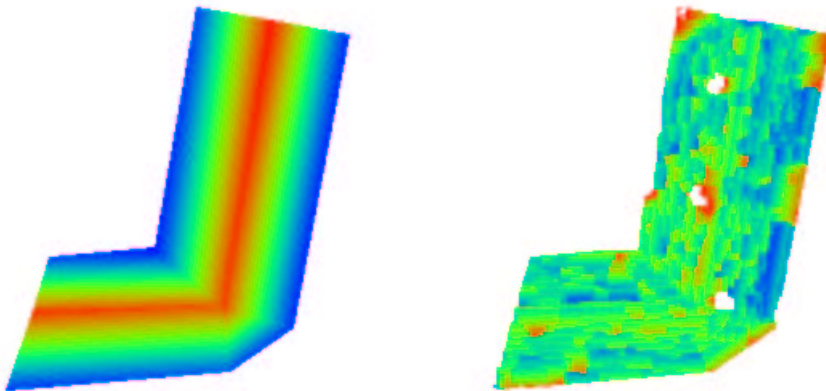


FIG. E.7 – MNER du bâtiment 19, et sa carte des erreurs (bleu : -0,6m, rouge : +0,9m)

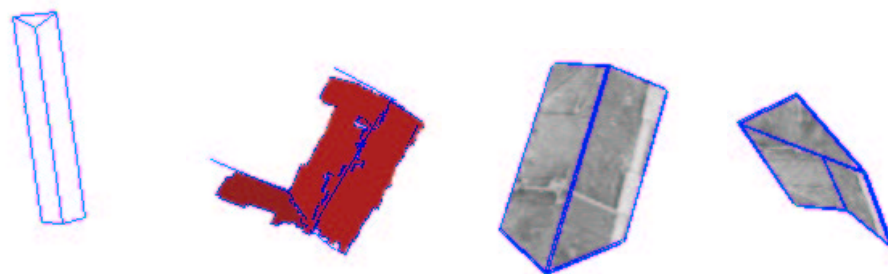


FIG. E.8 – Un modèle à trois pans (gauche), son appariement avec les primitives du bâtiment 19, la reconstruction déduite, une autre vue de la reconstruction déduite. On constate que le troisième pan, après reconstruction, passe sous les 2 pans principaux.

Annexe F

Figures de la zone 17

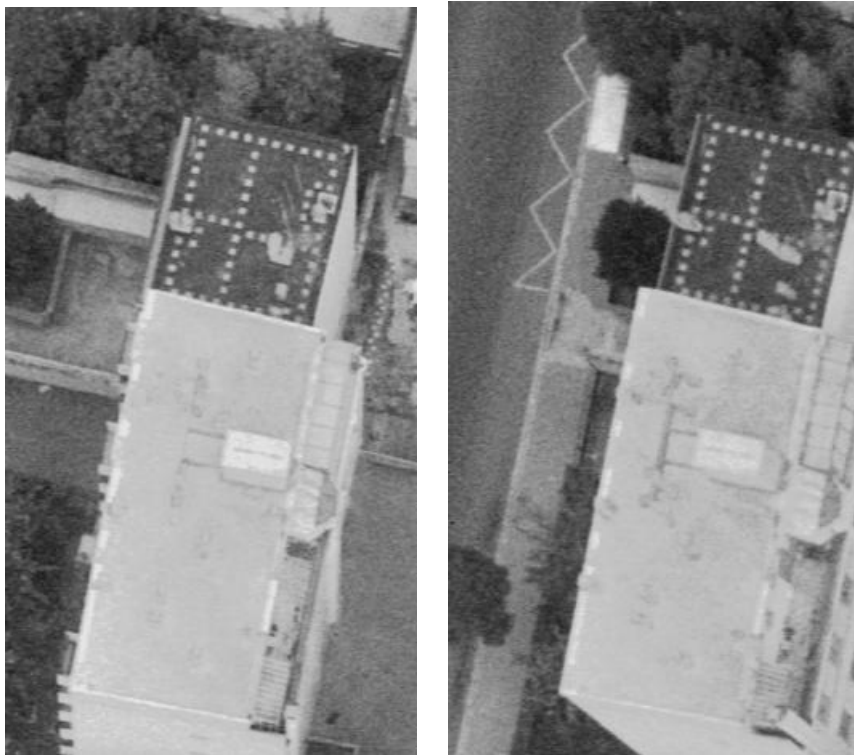


FIG. F.1 – Zone de travail 17.

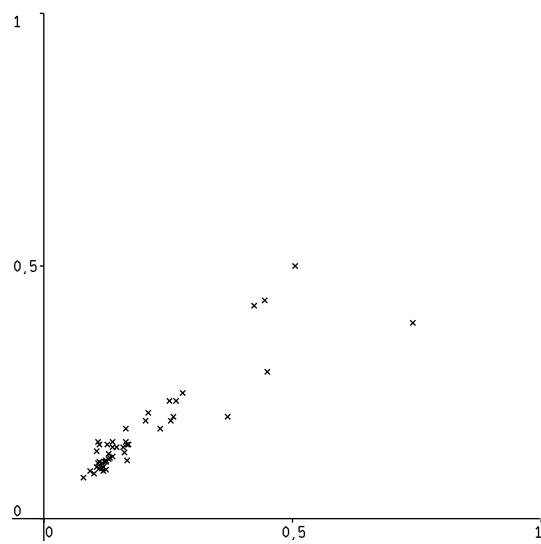


FIG. F.2 – Ecarts quadratiques médians pour la zone 17.



FIG. F.3 – Discontinuités de la zone de travail 17

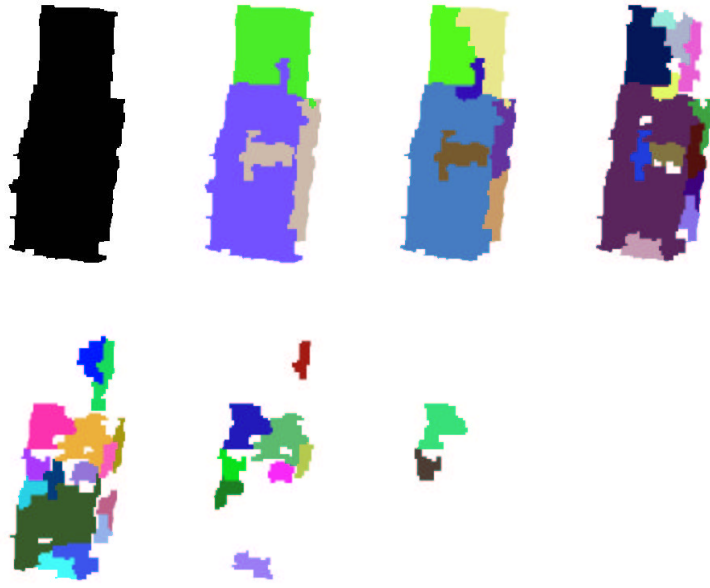


FIG. F.4 – Primitives surfaciques de la zone

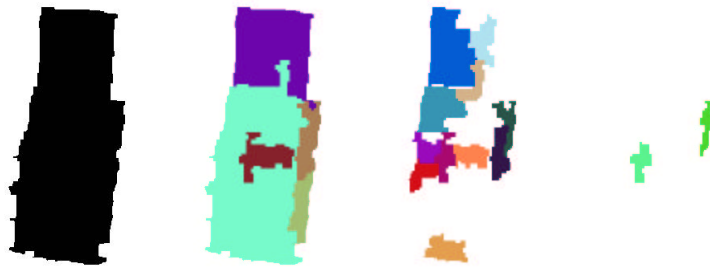


FIG. F.5 – Primitives surfaciques filtrées de la zone

Annexe G

Annexe : fonctions géométriques utilisées

Cette annexe décrit les fonctions géométriques (fonctions de compatibilité) utilisées dans le système.

Leur explicitation dans le mémoire aurait rallongé inutilement le débat, car elles ne contribuent pas de façon essentielles à l'articulation des outils mis en place.

Par contre elles jouent un rôle primordial dans l'appariement modèle / données, car ce sont elles qui, en définitive, définissent la fonction de coût (distance d'édition) qui est minimisée lors de l'appariement. Leur description était donc nécessaire. La voici.

Les relations sont réparties - comme dans le chapitre 3 - en fonction des types des sommets incidents.

1 Relations Ponctuel / Ponctuel

Aucune relation de ce type n'est utilisée.

2 Relations Ponctuel / Linéaire

Relation «est dans»

Cette relation traduit que le sommet ponctuel p se trouve à l'extrémité d'un sommet linéaire l . Les positions des extrémités des segments le long de leur droite porteuse n'étant pas très fiable, nous n'utilisons pas dans cette relation la position du point le long du segment, mais uniquement sa distance à la droite porteuse.

Le coût de cette relation est donc :

$$C_{pl_est_dans}(p, l) = d(p, l)$$

où $d(p, l)$ est la distance de p à la droite porteuse de l .

3 Relations Ponctuel / Surfactive

Relation «est dans»

Pour un sommet ponctuel p et un sommet surfactive s , le coût de cette relation est :

$$C_{ps_est_dans}(p, s) = d(p, s)$$

où $d(p,s)$ est la distance de p au plan portant s .

4 Relations Ponctuel / Discontinuité

Relation «est dans»

Pour un sommet ponctuel p et un sommet de type discontinuité d , le coût de cette relation est :

$$C_{pd_est_dans}(p, d) = d(p, d)$$

où $d(p,d)$ est la distance de p au plan vertical portant d .

A compter de maintenant, nous noterons c la compatibilité calculée.

5 Relations Linéaire / Linéaire

Relation «est parallèle»

En général, la gouttière d'un pan et le faîte qui lui est opposé sont parallèles.

Soient v_1 et v_2 des vecteurs directeurs normés des deux segments. Alors, posons $c = 1 - |v_1 \cdot v_2|$. Si c est supérieur à un seuil s_{pl} alors posons $c = +\infty$. c est alors la compatibilité pour cette propriété géométrique. s_{pl} vaut 0,2, ce qui signifie que la relation de parallélisme est jugée incompatible si l'angle entre les deux segments est supérieur à $\arccos(0,8) = 36^\circ$, ce qui est une valeur élevée. Ce seuillage est donc lâche.

Relation «est plus haut que»

Soit h_1 l'altitude du milieu du premier segment, et h_2 celle du milieu du second segment. Soit $c = 1 - (h_1 - h_2)$. Si $c < 0$, alors posons $c = 0$. Si $c > 1$, alors posons $c = +\infty$. Ceci signifie que la relation est tout à fait compatible si le sommet h_1 a une altitude moyenne supérieure de plus d'un mètre à celle de h_2 , et qu'elle est totalement incompatible si l'altitude moyenne de h_1 est inférieure à celle de h_2 .

c mesure la compatibilité de la propriété en question.

Relation «intersecte»

Cette relation est utilisée pour les segments possédant une extrémité commune.

Soit d la distance des deux droites porteuses des segments. Soit $c = d/s_{li}$. Si $c > 1$, alors posons $c = +\infty$.

c est la compatibilité souhaitée. Dans la pratique nous avons $s_{li} = 5\text{m}$, ce qui rend la relation incompatible pour les segments donc les droites porteuses sont très distantes.

Relation «est en regard de»

Cette relation traduit la notion souvent utilisée en groupement perceptuel, et dénommée «overlap» en anglais. Il s'agit de définir que l'un des segments, une fois projeté orthogonalement sur l'autre, possède un recouvrement avec lui, et inversement.

Soit $IR(s, s')$ la fonction qui pour deux segments, associe la longueur du recouvrement entre s et le segments s' projeté sur la droite porteuse de s , comme sur la figure G.1

Soit $c_1 = IR(s, s')/||s||$, et $c_2 = IR(s', s)/||s'||$. Posons alors $c = \min(c_1, c_2)$. c mesure la compatibilité de la propriété.

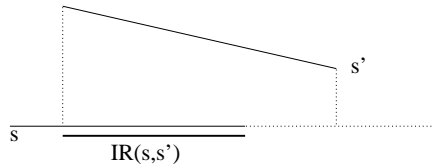


FIG. G.1 – Illustration de la fonction $IR(s, s')$

Relation «est colinéaire en 2D»

Cette relation traduit le fait que des primitives linéaires appartiennent à un même plan vertical, ce qui est généralement le cas des arêtes de rives d'un bâtiment avec un attaque simple.

Soient d_1 et d_2 les distances des 2 extrémités du premier segment à la droite porteuse du second segment. De même soient d_3 et d_4 les distances des extrémités de l'autre segment à la droite portant le premier. Soit $c = (d_1 + d_2 + d_3 + d_4)/s_{uc}$. Si $c > 1$, posons $c = +\infty$. s_{uc} vaut 20m ce qui signifie que la propriété est jugée incompatible si en moyenne chaque segment est distant de plus de 5m de la droite porteuse de l'autre, ce qui est très lâche.

6 Relations Linéaire / Surfaccique**Relation «est dans»**

Cette relation est liée à la distance des extrémités d'un segment l au plan s .

Soient d_1 et d_2 les distances des extrémités du segment au plan portant la primitive surfaccique.

Soit $c = (d_1/d_2)/s_{lse}$. Si $c > 1$, posons $c = +\infty$. s_{lse} vaut 4m, ce qui signifie que la propriété est jugée incompatible si les deux extrémités du segment sont en moyenne à plus de 2 mètres du plan porteur de la primitives sufaccique, ce qui est très lâche.

Relation «borde»

Cette relation est un des relations les plus fondamentales du système. Elle traduit le fait qu'un sommet linéaire l de trouve à la frontière d'un sommet surfaccique d .

Soit d la droite porteuse du segment projeté sur le plan portant la primitives surfaccique.

Soient d_i les distances des points d'échantillonnage du polygone bordant la primitives surfaccique (échantillonné avec un pas égal à la résolution des images).

Soit d' l'ensemble des d_i inférieures à un seuil s_{lsd} . Soit c la moyenne des valeurs de d' . Si le nombre de valeurs de d' est insuffisant (c'est-à-dire si le nombre de points du polygone à distance de d inférieur à s_{lsd} est insuffisant), alors posons $c = +\infty$.

Dans la pratique $s_{lsd} = 3m$, est le seuil de suffisance est fixé à $1/r$ où r est la résolution des images, ce qui signifie que le polygone et la droite porteuse du segment doivent posséder des points distantes de moins de 3m sur au moins un mètre de longueur, ce qui n'est pas très contraignant pour accepter la propriété géométrique.

Relation «est plus haut que»

Cette relation permet de différencier les faîtes des gouttières, les faîtes étant situés plus haut que les gouttières, les relations qui les lient aux pans de toits ne sont pas identiques. Cette relation se calcule par différence d'altitude entre le centre de gravité de la surface et le milieu du segment.

Soit d la différence d'altitude entre le centre de gravité du segment et le centre de gravité de la primitive plane. Soit $c = 1 - d$. Si $c > 1$, nous posons $c = +\infty$.

Relation «est moins haut que»

Cette relation est rigoureusement le contraire de la relation précédente.

Relation «même direction»

Cette relation signifie que la direction de plus grande pente de la portion de plan et celle du segment sont sensiblement les mêmes. C'est ce qui se passe généralement entre un pan de toit et ses arêtes de rive.

Soit v_1 le vecteur normé de la direction de plus grande pente de la primitive plane, et v_2 le vecteur directeur normé du segment.

Soit $c = 1 - |v_1 \cdot v_2|$. Si $c > s_{lsmd}$, posons $c = +\infty$.

s_{lsmd} vaut 0.5, ce qui signifie que la propriété est jugée incompatible si l'angle entre les deux vecteurs est supérieur à 60° , ce qui est très lâche.

7 Relations Linéaire / Discontinuité

Relation «est parallèle à»

Cette relation est la même que la relation du même nom reliant deux primitives linéaires, mais en considérant les vecteurs bidimensionnels v_1 et v_2 où v_1 est composé des deux premières coordonnées du vecteur porteur du segment, et où v_2 est le vecteur directeur de la droite portant la projection de la discontinuité au sol (nous rappelons qu'une discontinuité est en fait un plan vertical).

Relation «est en regard de»

Cette relation est la même que la relation du même nom reliant deux primitives linéaires.

Relation «est dans»

En considérant que les primitives discontinuité sont en fait des plans verticaux, cette relation est en fait la même que la relation du même nom reliant une primitive linéaire et une primitive surfacique.

8 Relations Surfacique / Surfacique**Relation «a une pente perpendiculaire à»**

Cette relation traduit que les directions de plus grande pente de deux plans sont sensiblement non parallèles. C'est ce qui se passe entre un pan de toit et une attaque en V adjacente.

Soient v_1 et v_2 les vecteurs directeurs normés des directions de plus grande pente des primitives planes.

Soit $c = |v_1.v_2|$. Si $c > s_{sspp}$, posons $c = +\infty$.

s_{sspp} vaut 0,7, ce qui signifie que la propriété est jugée non compatible si l'angle entre les deux directions de plus grande pente est inférieur à 45° , ce qui est lâche.

Relation «à une pente opposée à»

Cette relation traduit que les directions de plus grande pente de deux plans sont sensiblement parallèles, et, de plus, opposées. C'est ce qui relie deux pans de toit opposés dans le modèle D.

Relation «borde»

Si les deux normales des surfaces sont quasi parallèles, alors la compatibilité associée est $c = +\infty$. Sinon, soit d la droite d'intersection des deux plans porteurs des primitives surfaciques. Soient c_1 et c_2 les compatibilités du même nom calculées entre d et les deux primitives planes.

Si l'une des deux compatibilités ainsi calculées vaut $+\infty$, alors $c = +infy$.

Sinon, $c = (c_1 + c_2)/2$.

9 Relations Surfacique / Discontinuité**Relation «borde»**

Cette relation traduit le fait que deux portions de plan se bordent mutuellement (sont adjacentes).

L'utilité principale de cette relation est de pallier l'absence éventuelle des relations «borde» entre un faîte et ses deux pans adjacents, lorsque ce faîte est absent des données. Nous exploitons ici la redondance des relations pour contrer les lacunes des détections.

Soit d l'intersection entre les deux primitives (la primitive plane n'est pas verticale, et la discontinuité est un plan vertical, donc cette intersection existe). La compatibilité est alors la même que la compatibilité de même nom entre une primitive linéaire et une primitive surfacique.

10 Relations Discontinuité / Discontinuité

Relation «est parallèle à»

Cette compatibilité est la même que la compatibilité du même nom entre deux primitives linéaires.

Relation «est perpendiculaire à»

Soit s le produit scalaire des normales aux plans verticaux. Alors posons $c = |s|$. Si $c > s_{ddp}$, posons $c = +\infty$.

Ceci signifie que la propriété est incompatible si les normales ont un angle inférieure à 60° , ce qui est lâche.

11 Commentaires

Comme on l'aura remarqué, ces relations sont truffées de paramètres. Ces paramètres interviennent dans ce que nous jugeons être des circonstances aberrantes. Ils sont donc généralement réglés de façon à ne pas être critiques, ce que nous avons souvent dénommé par le terme «lâche».

Annexe H

Les reconstructions

Cette annexe présente *toutes* les reconstructions ayant contribué à l'évaluation de la méthode. Pour chacune des zones, nous présentons le couple stéréoscopique focalisé sur la zone, le MNE brut correspondant, et une vue perspective de la reconstruction effectuée par le système.

Ces vues perspectives sont généralement verticales, sauf dans quelques cas où il a été intéressant de changer de point de vue.



FIG. H.1 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 172

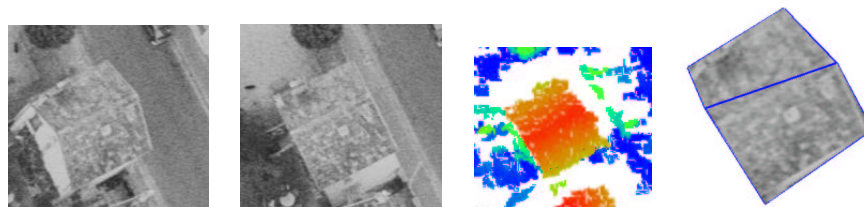


FIG. H.2 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 170

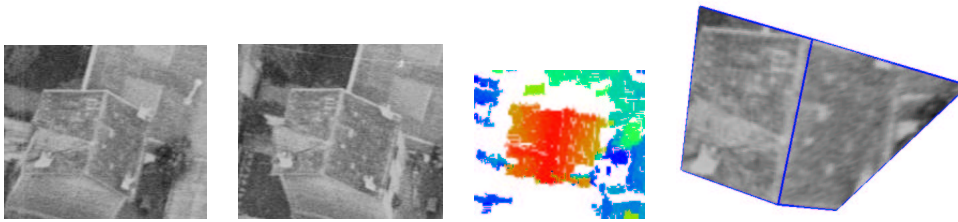


FIG. H.3 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 169

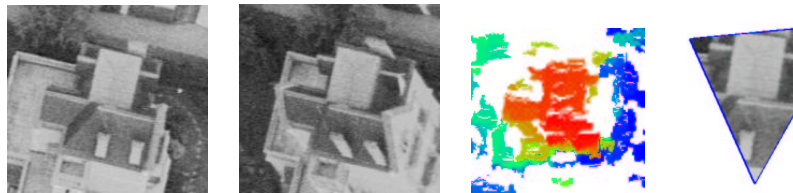


FIG. H.4 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 165

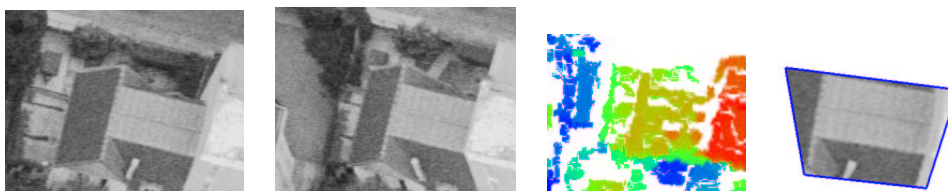


FIG. H.5 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 164

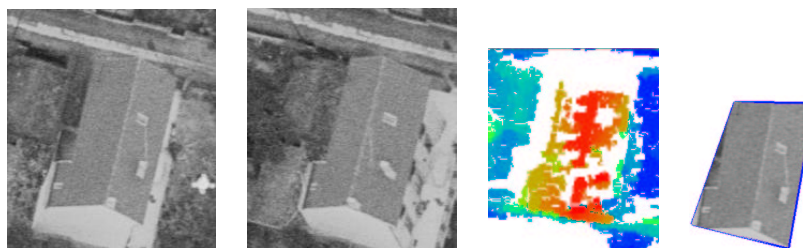


FIG. H.6 – Couple stéréoscopique MNE, et reconstruction finale pour la zone 163

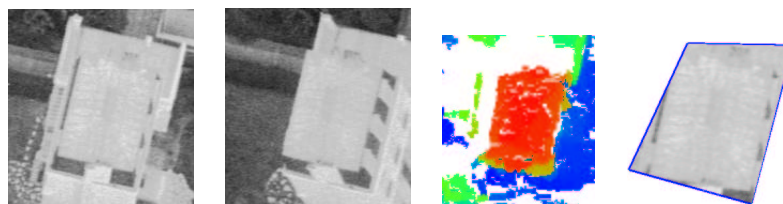


FIG. H.7 – Couple stéréoscopique MNE, et reconstruction finale pour la zone 162

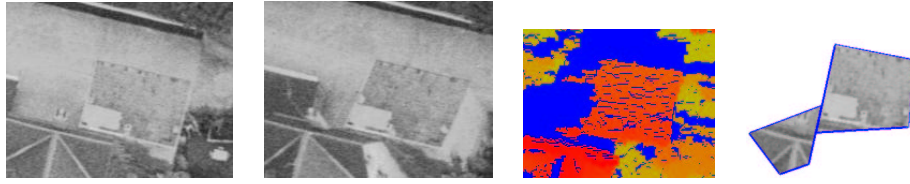


FIG. H.8 – Couple stéréoscopique MNE, et reconstruction finale pour la zone 155

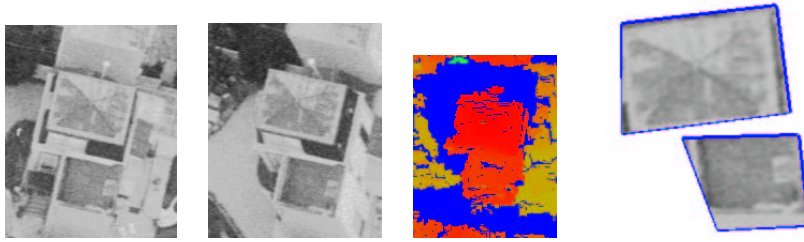


FIG. H.9 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 152

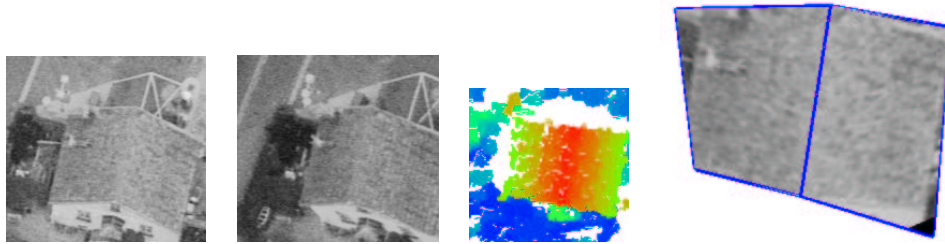


FIG. H.10 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 151

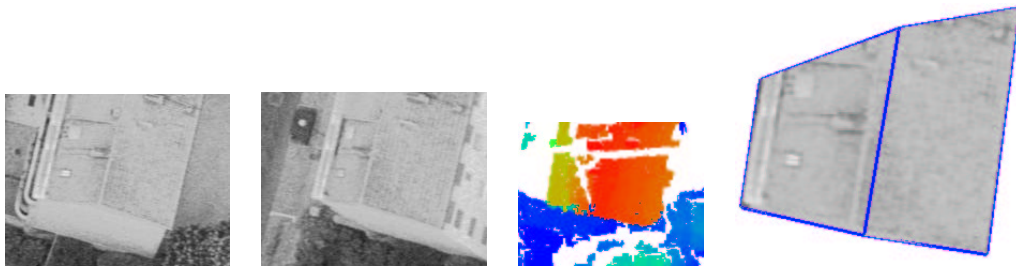


FIG. H.11 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 150

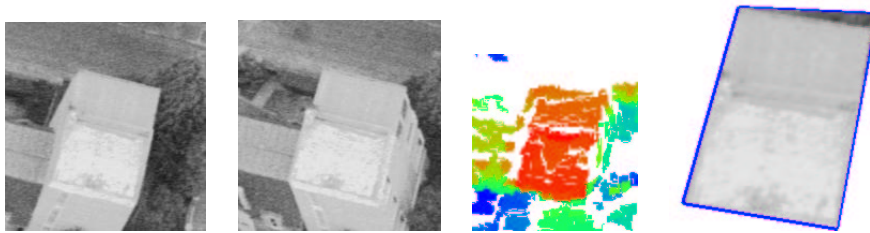


FIG. H.12 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 147

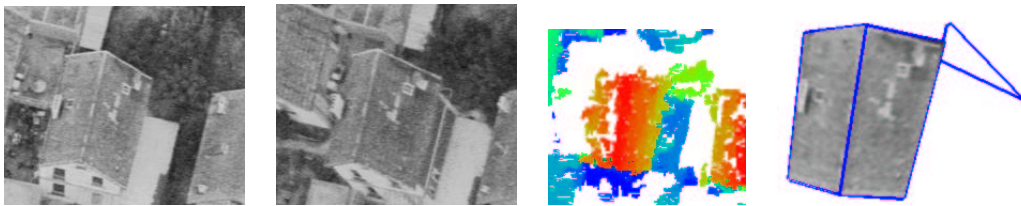


FIG. H.13 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 143

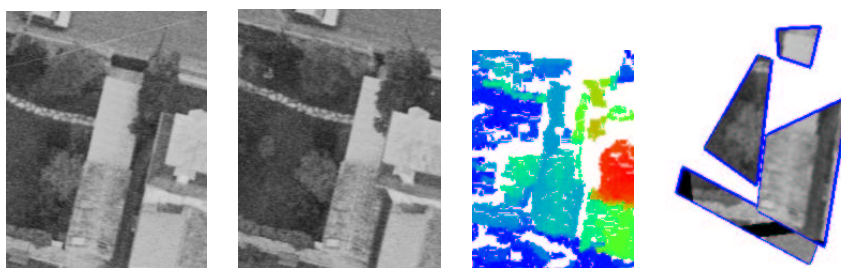


FIG. H.14 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 142

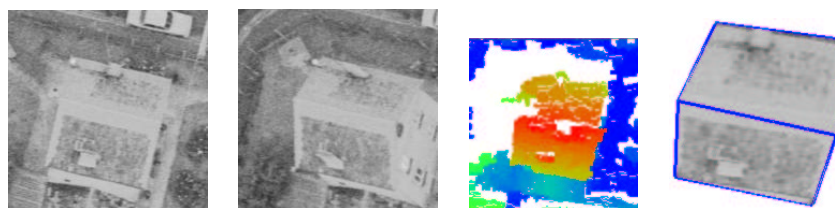


FIG. H.15 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 141

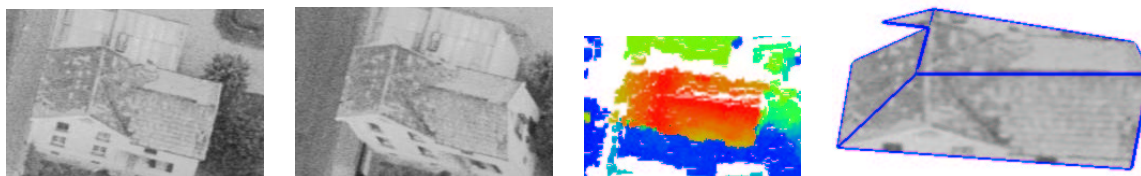


FIG. H.16 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 132

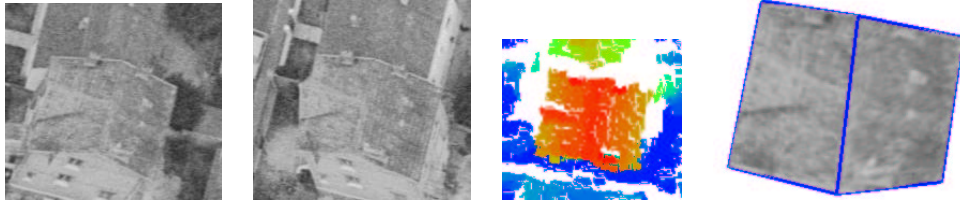


FIG. H.17 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 131

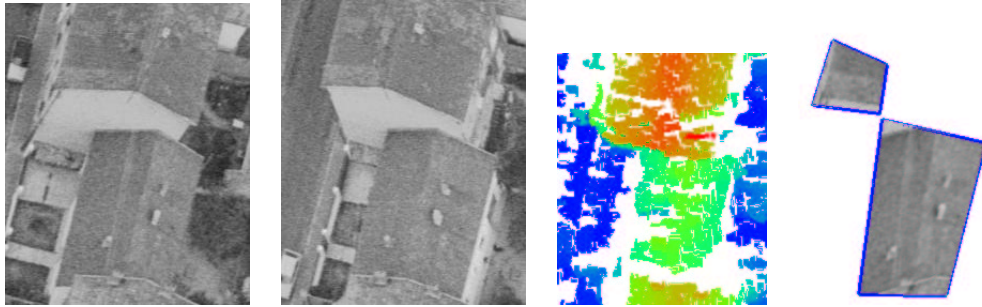


FIG. H.18 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 130

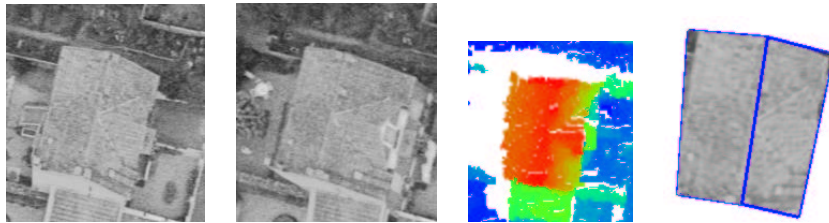


FIG. H.19 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 128

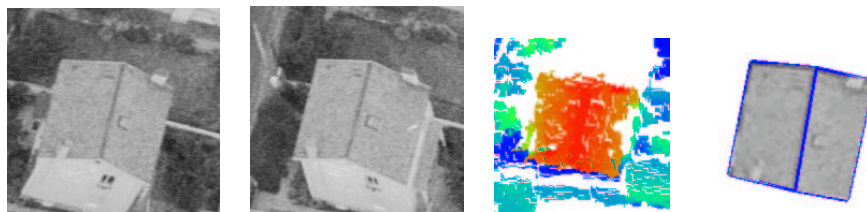


FIG. H.20 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 127

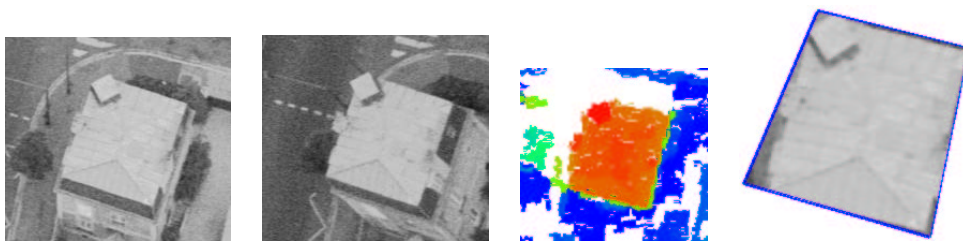


FIG. H.21 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 123

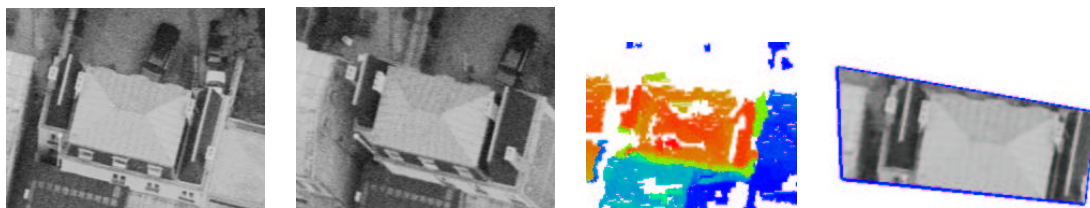


FIG. H.22 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 122

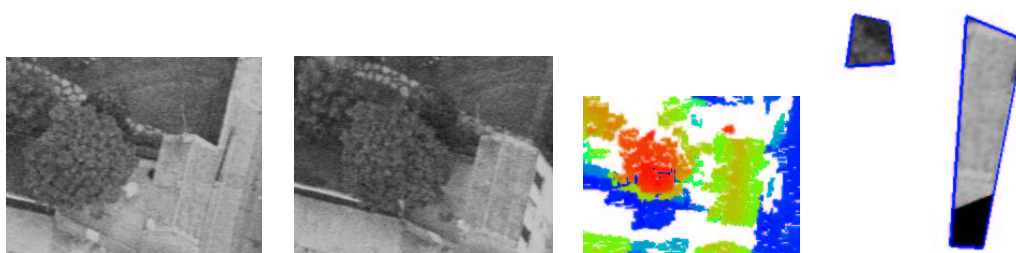


FIG. H.23 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 121

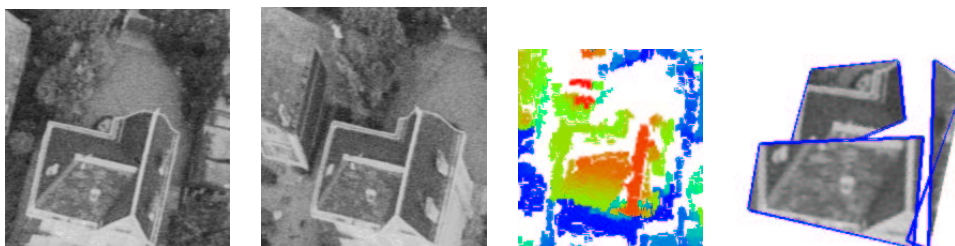


FIG. H.24 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 120

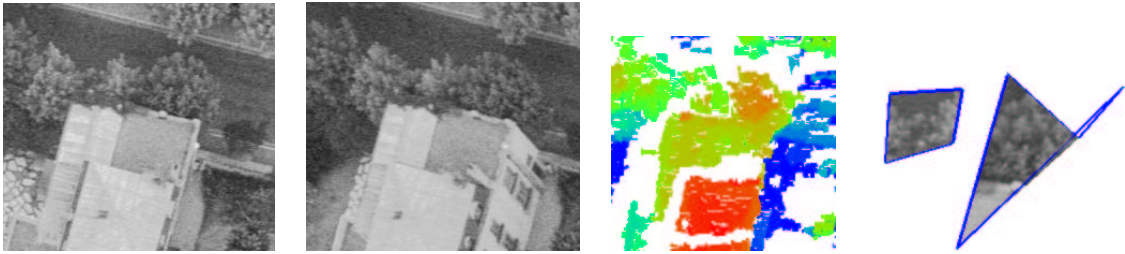


FIG. H.25 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 119

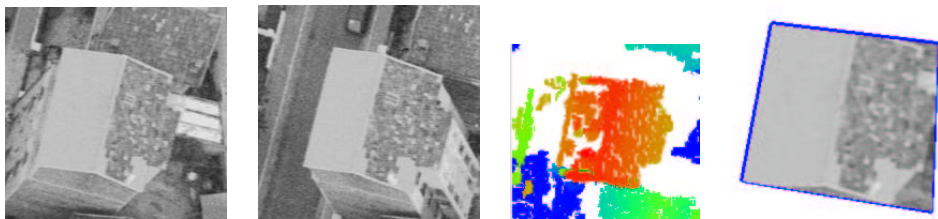


FIG. H.26 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 116

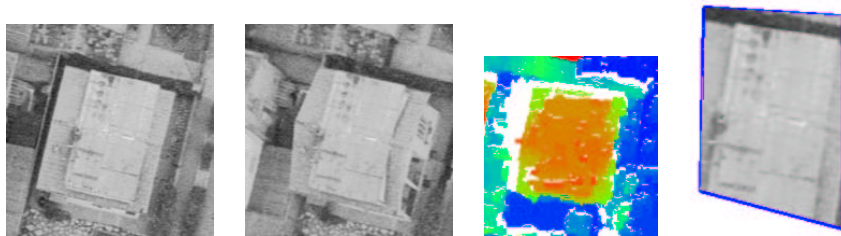


FIG. H.27 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 113

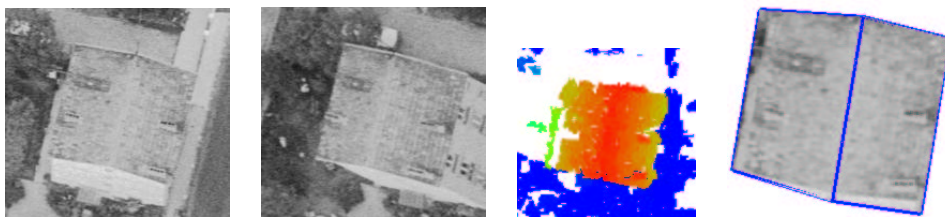


FIG. H.28 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 112

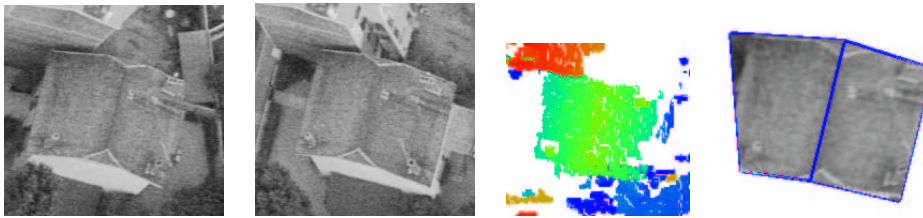


FIG. H.29 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 110

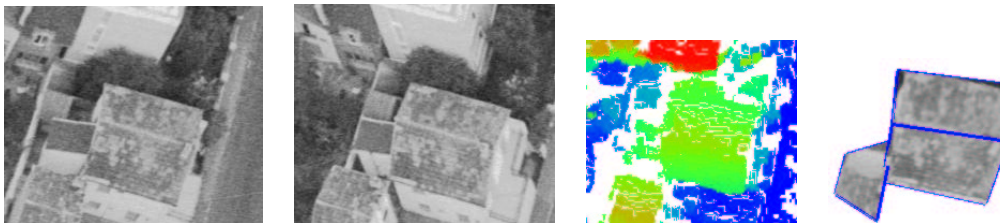


FIG. H.30 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 109

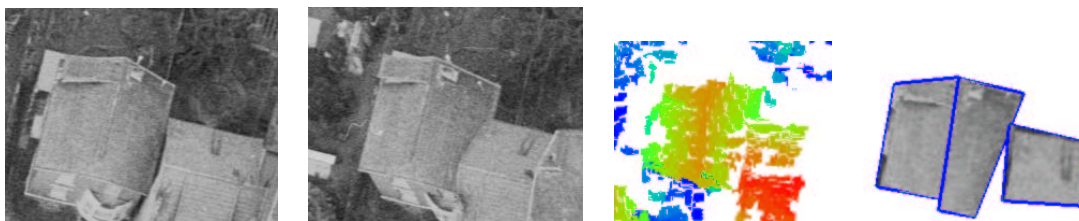


FIG. H.31 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 108

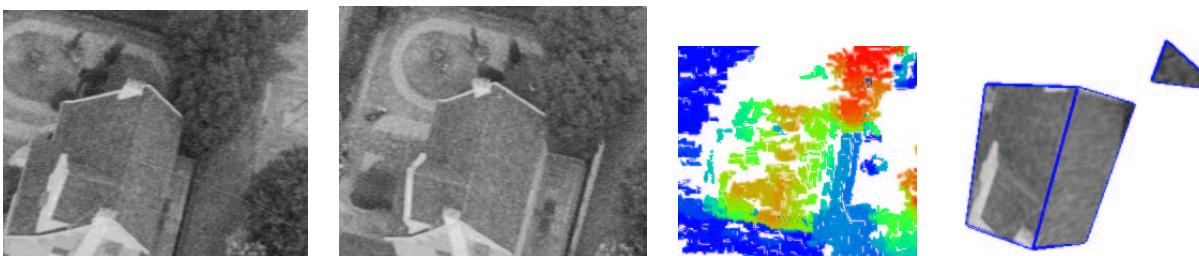


FIG. H.32 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 107

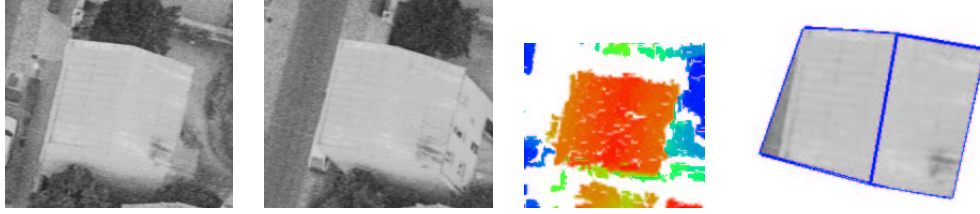


FIG. H.33 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 106

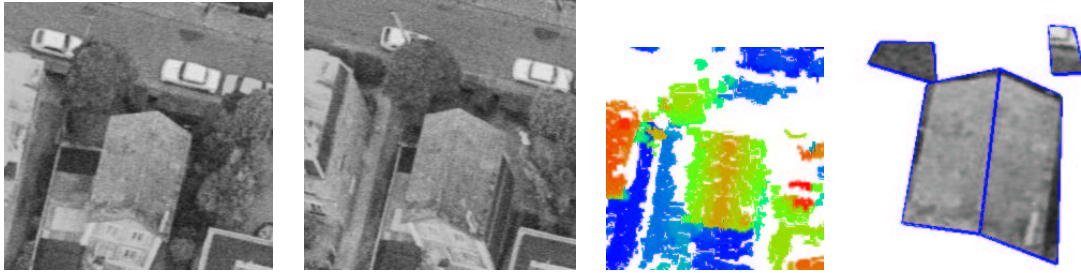


FIG. H.34 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 104

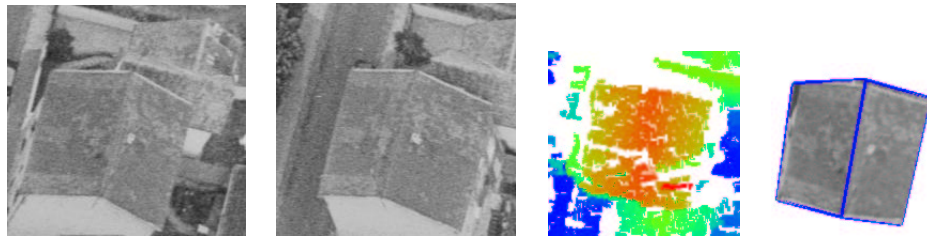


FIG. H.35 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 102

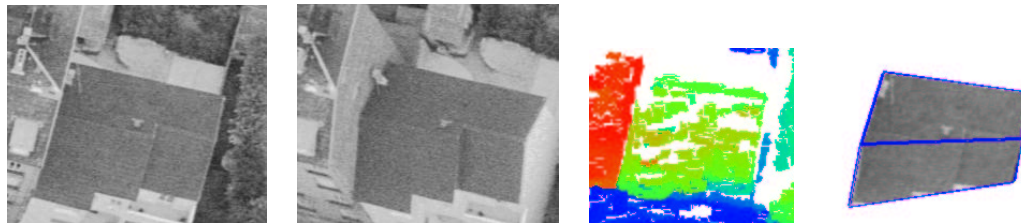


FIG. H.36 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 101

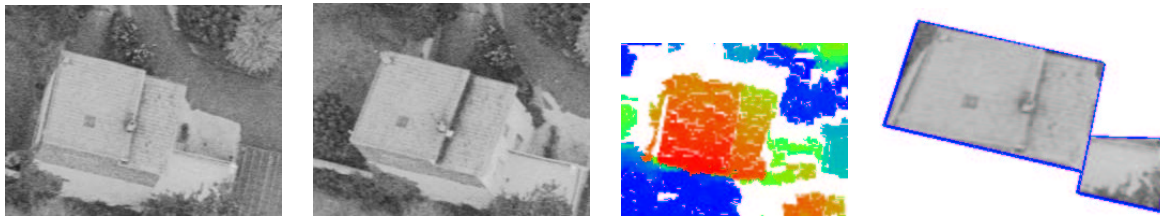


FIG. H.37 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 98

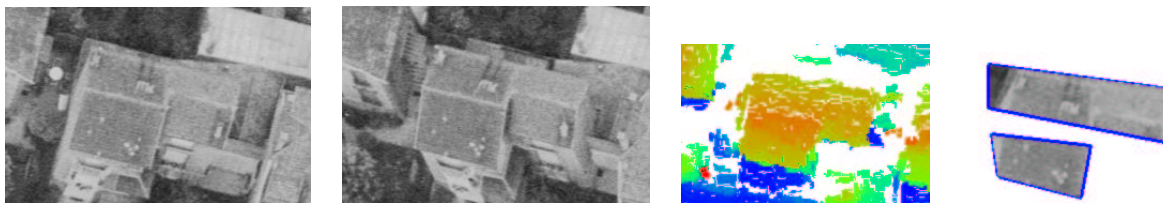


FIG. H.38 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 97

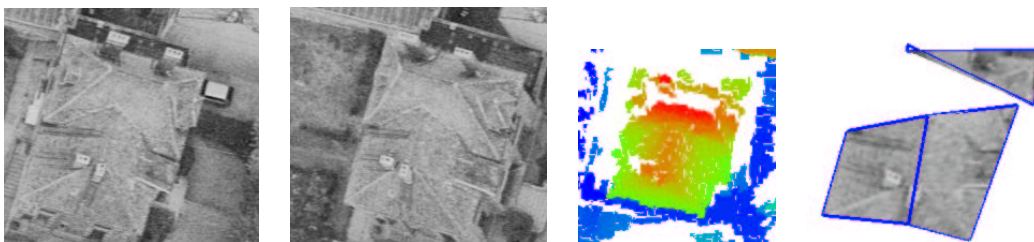


FIG. H.39 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 96

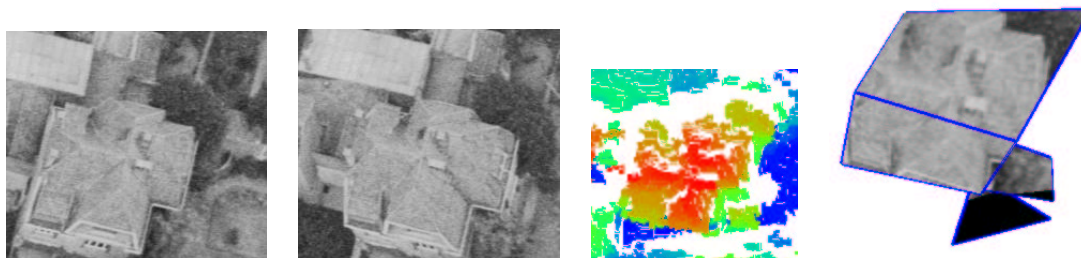


FIG. H.40 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 95

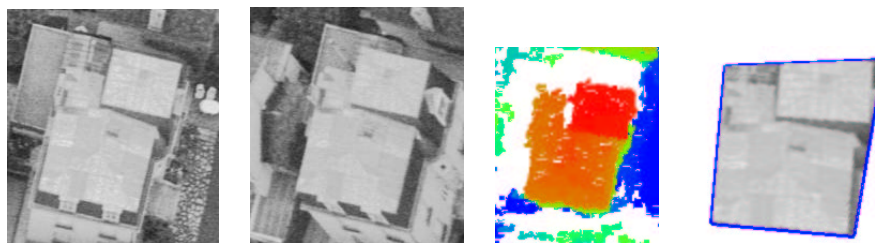


FIG. H.41 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 94



FIG. H.42 – Couple stéréoscopique, MNE, pour la zone 93. Pas de reconstruction pour cette zone.

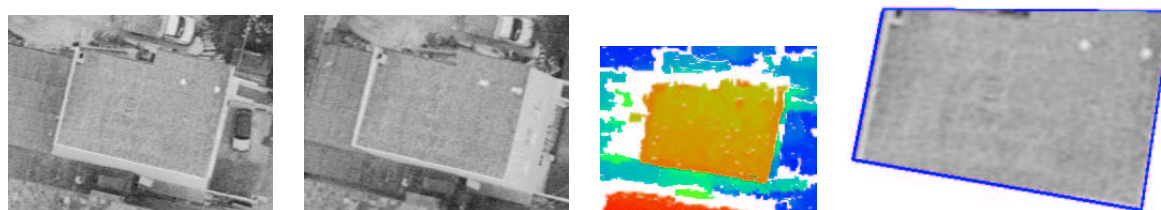


FIG. H.43 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 92

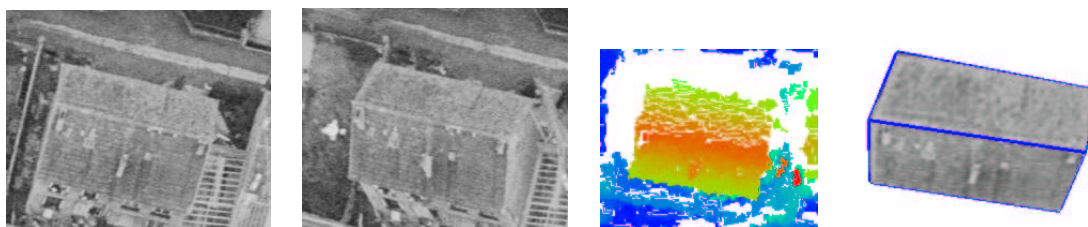


FIG. H.44 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 88

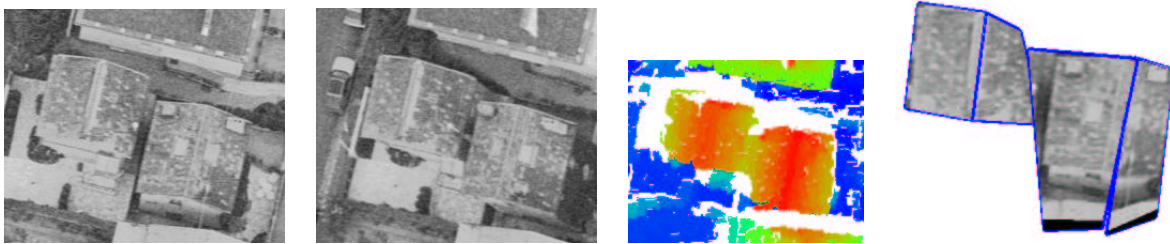


FIG. H.45 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 87

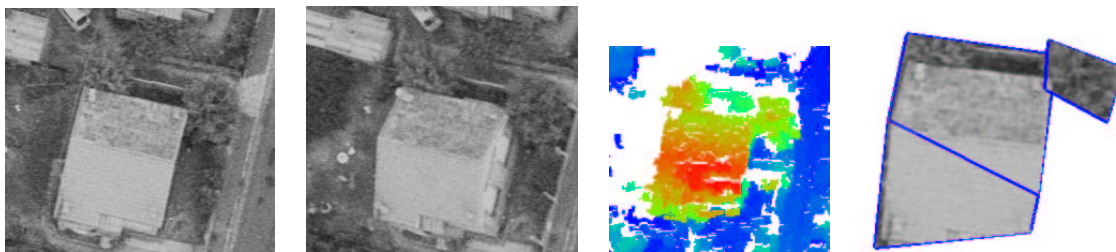


FIG. H.46 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 84



FIG. H.47 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 82

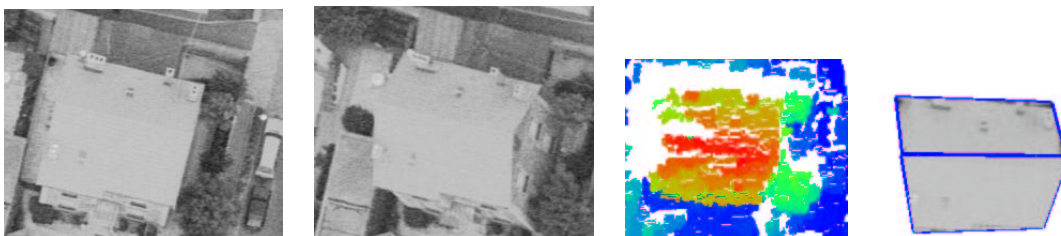


FIG. H.48 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 81

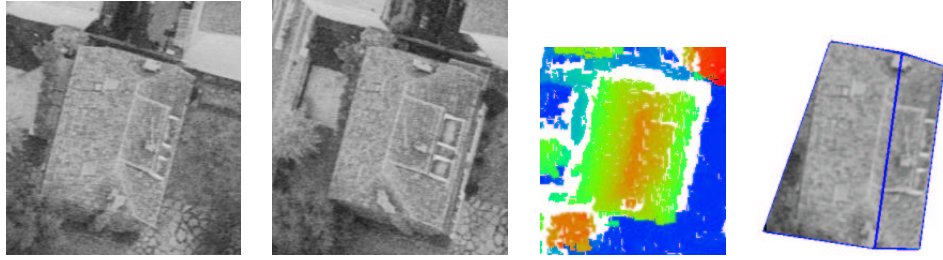


FIG. H.49 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 79

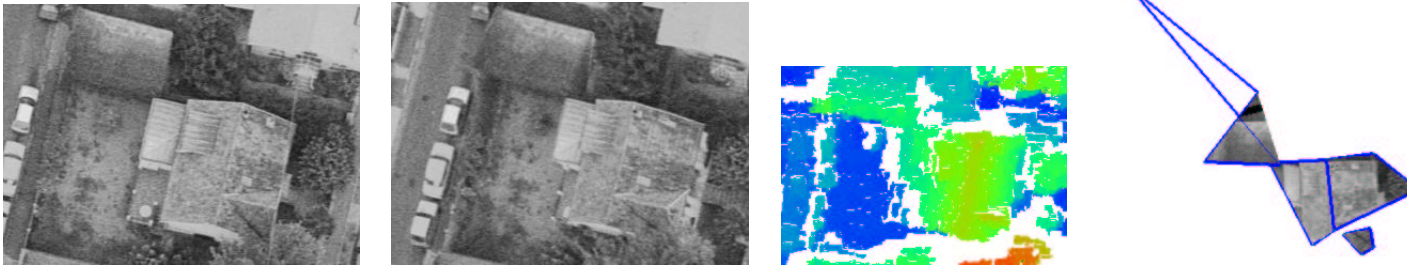


FIG. H.50 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 76

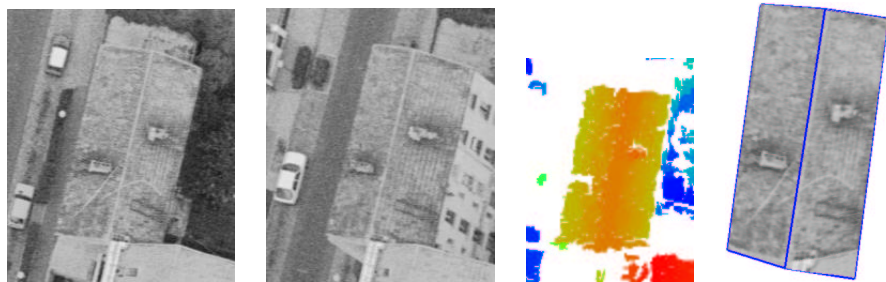


FIG. H.51 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 75

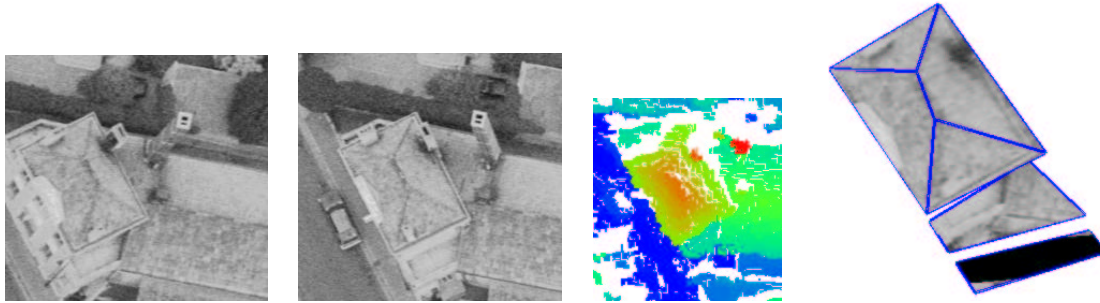


FIG. H.52 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 74

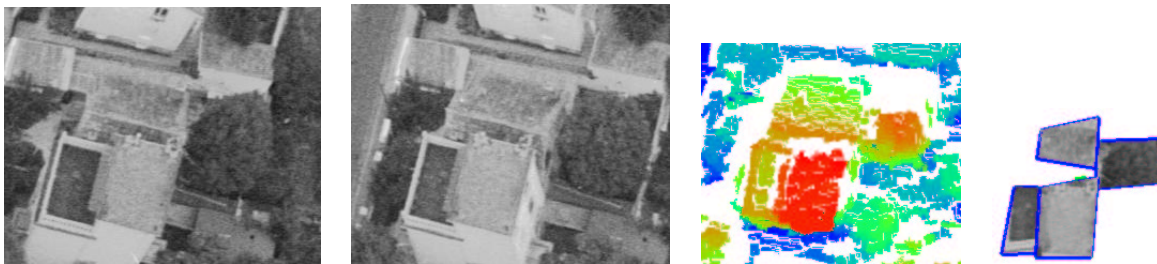


FIG. H.53 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 73



FIG. H.54 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 71

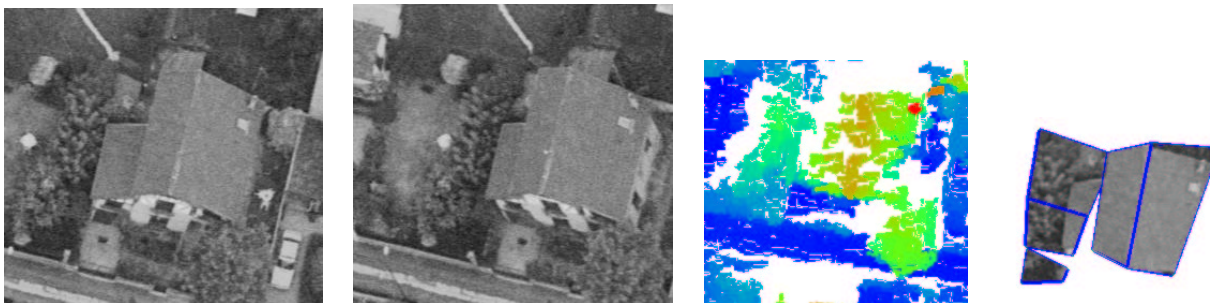


FIG. H.55 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 68

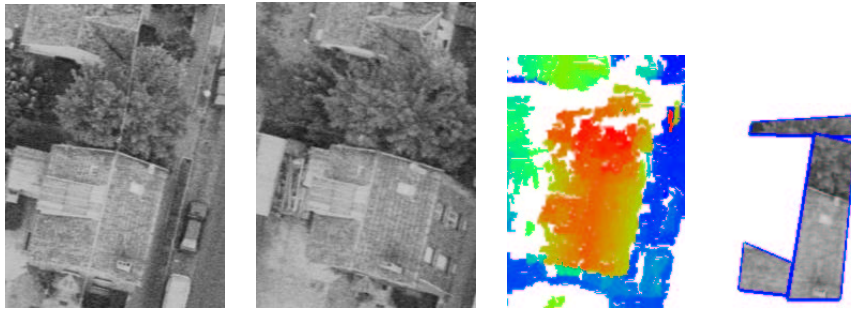


FIG. H.56 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 67

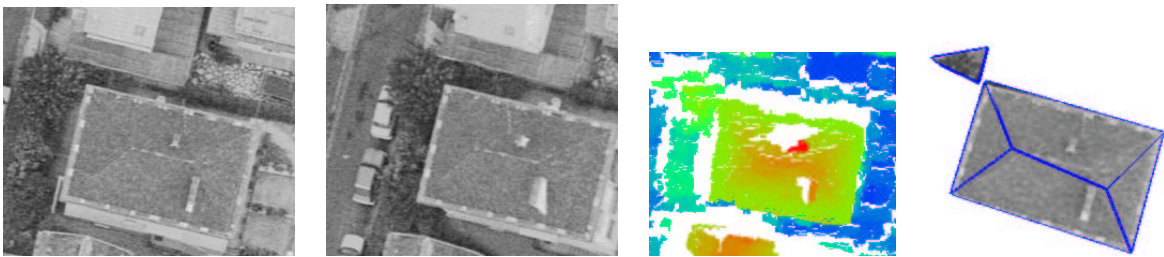


FIG. H.57 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 66

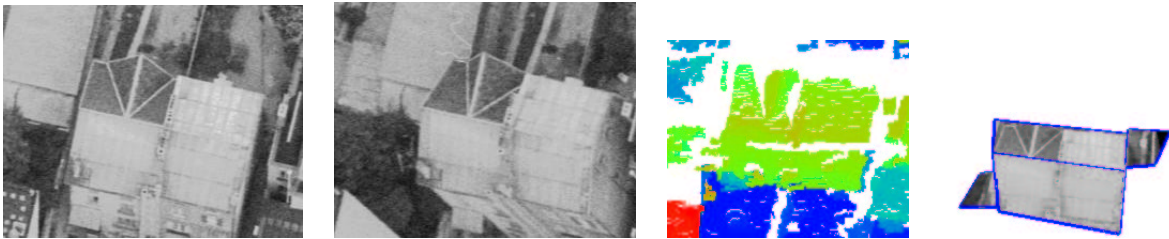


FIG. H.58 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 64

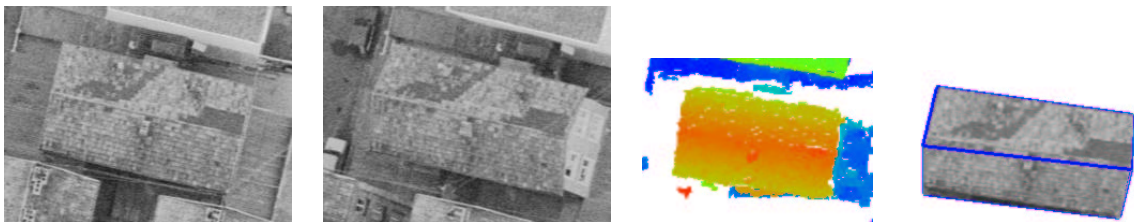


FIG. H.59 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 63

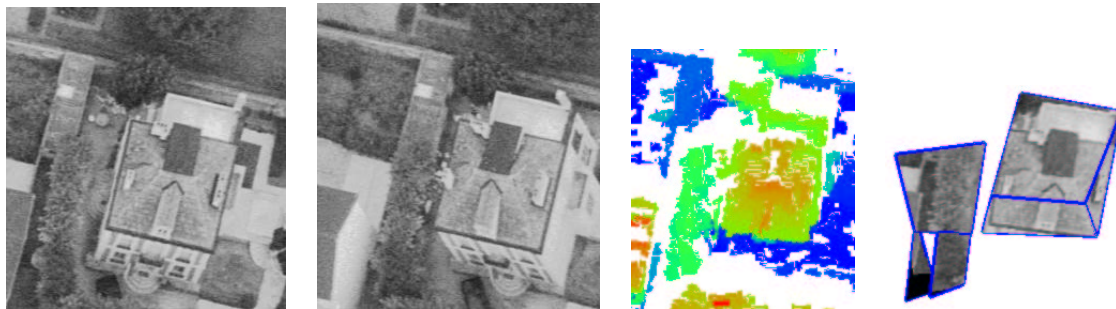


FIG. H.60 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 62

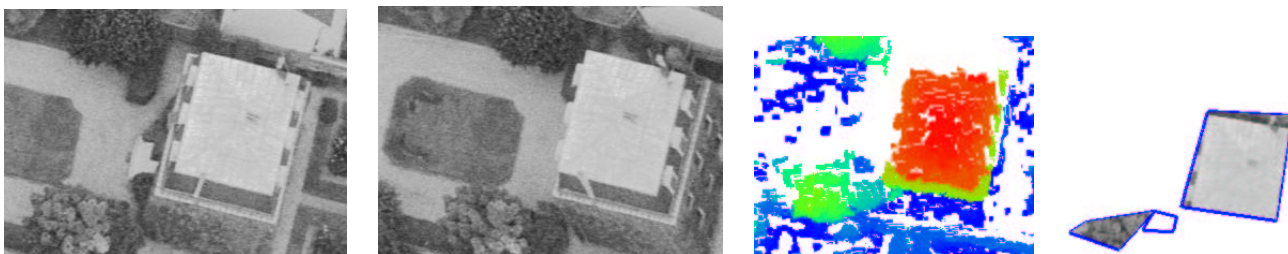


FIG. H.61 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 61

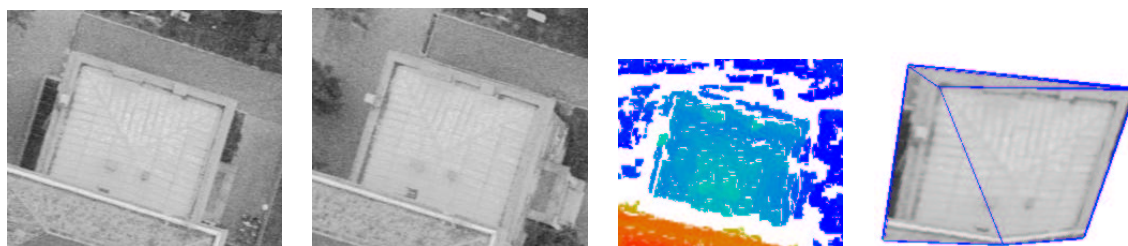


FIG. H.62 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 60



FIG. H.63 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 59

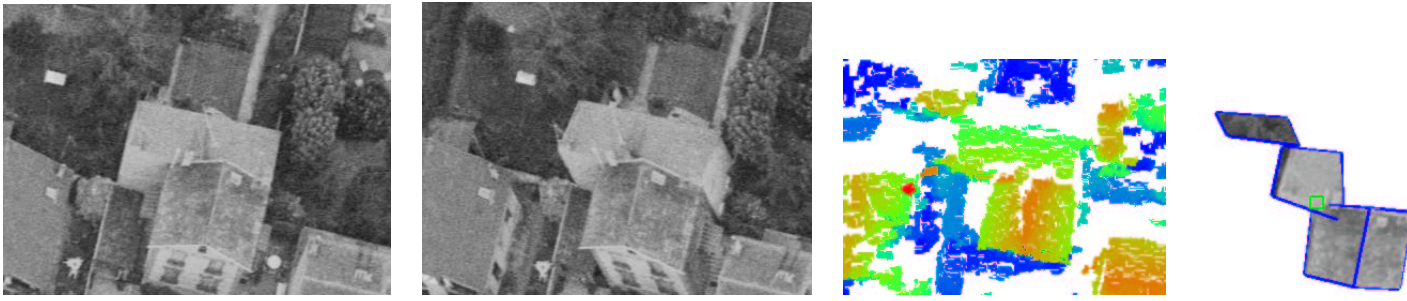


FIG. H.64 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 58

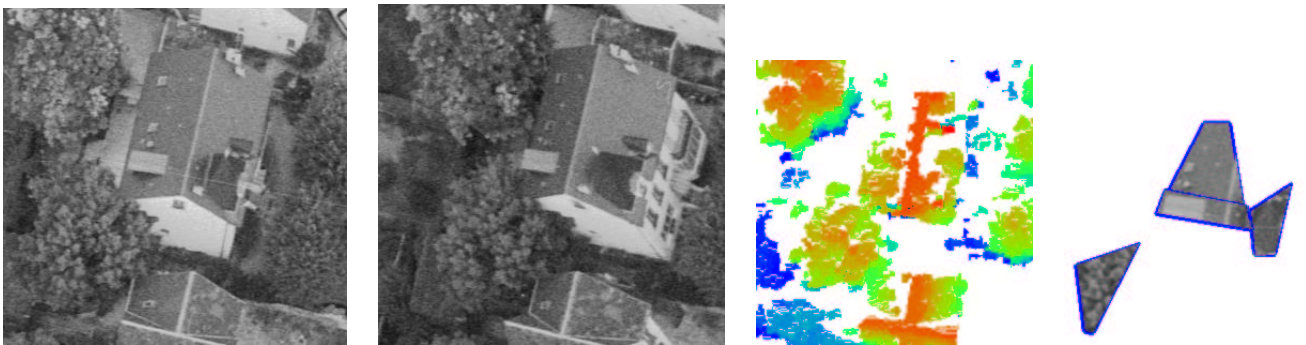


FIG. H.65 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 57

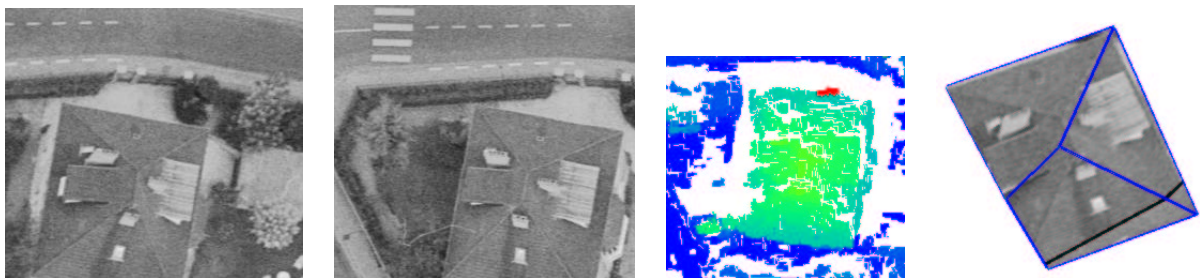


FIG. H.66 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 55

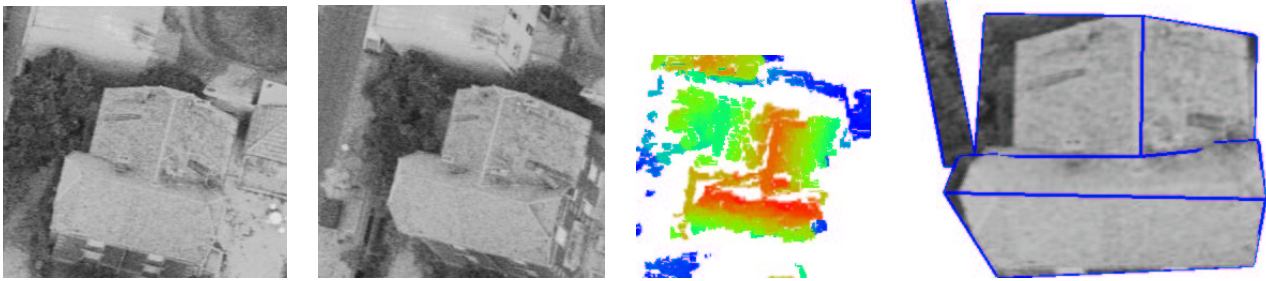


FIG. H.67 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 54

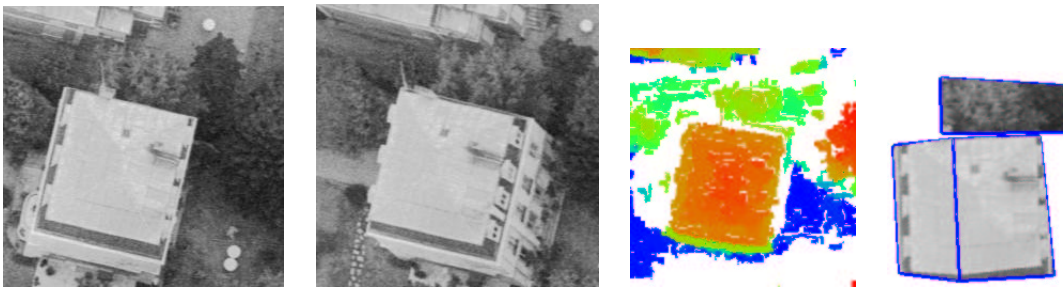


FIG. H.68 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 53



FIG. H.69 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 51

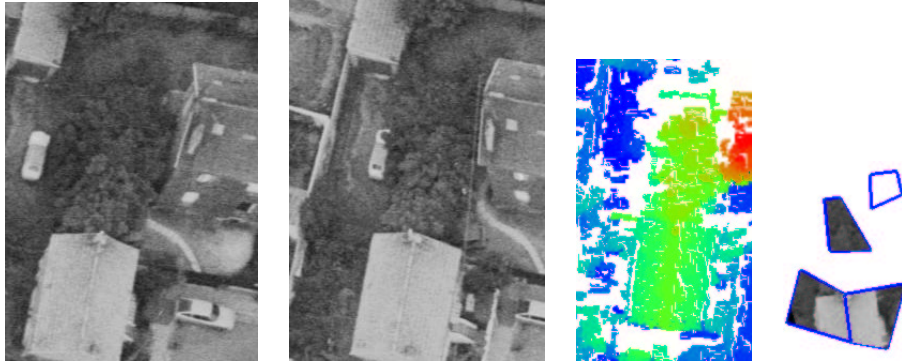


FIG. H.70 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 50

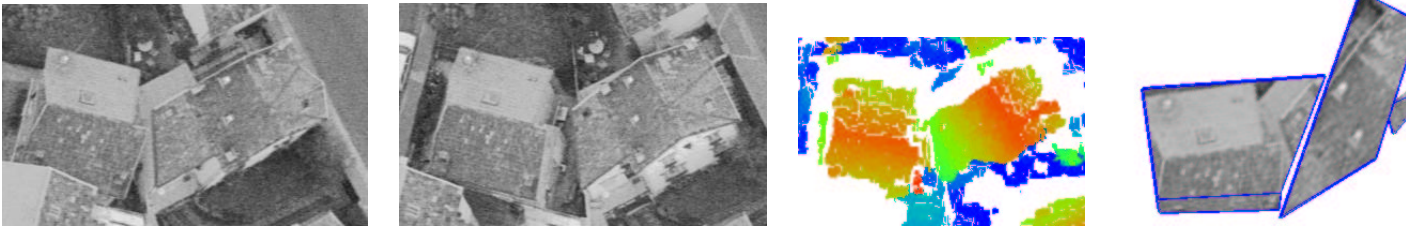


FIG. H.71 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 49

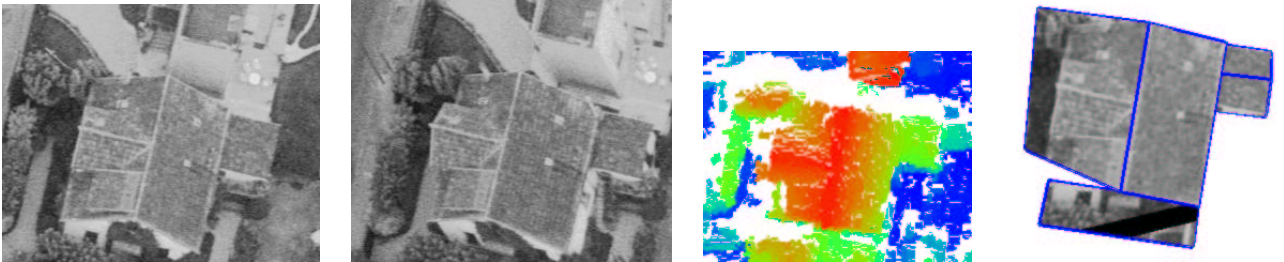


FIG. H.72 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 48

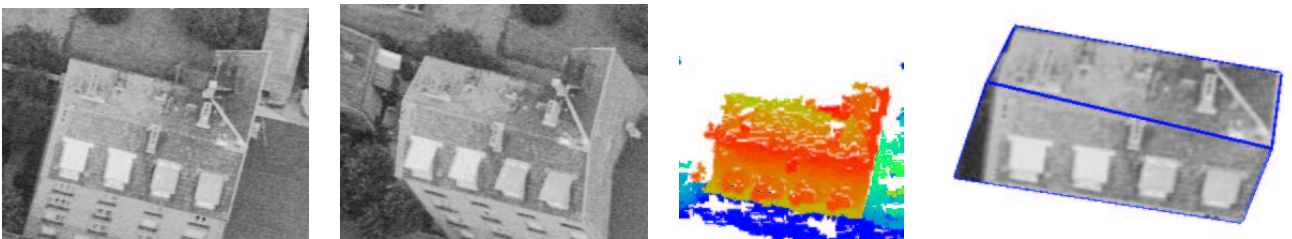


FIG. H.73 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 47

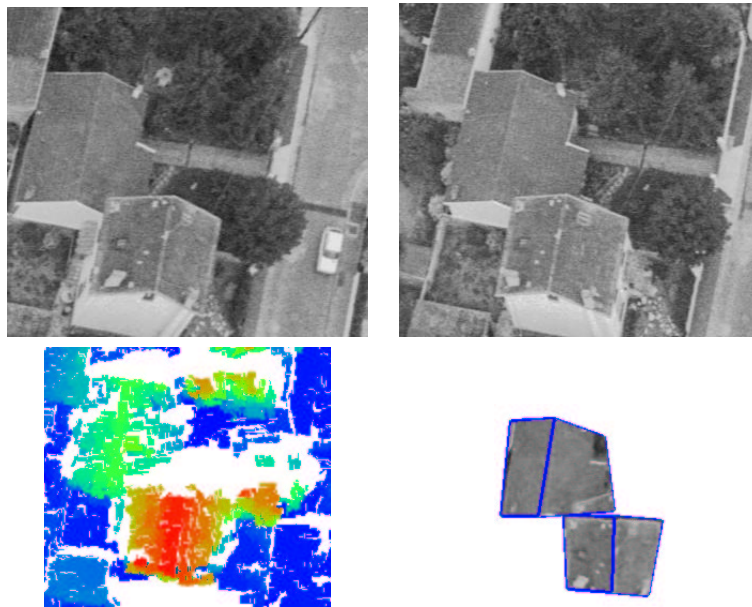


FIG. H.74 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 46

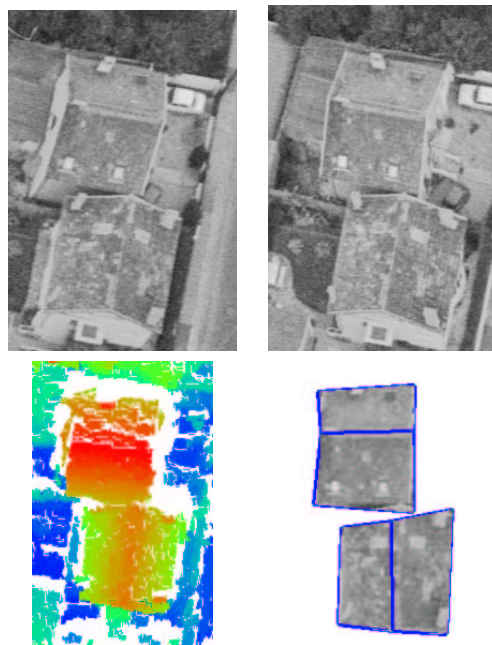


FIG. H.75 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 45

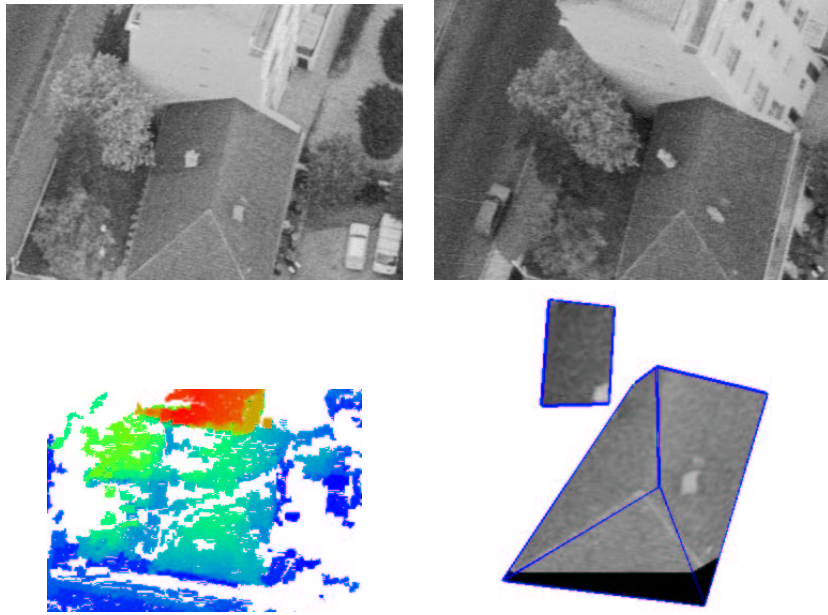


FIG. H.76 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 42

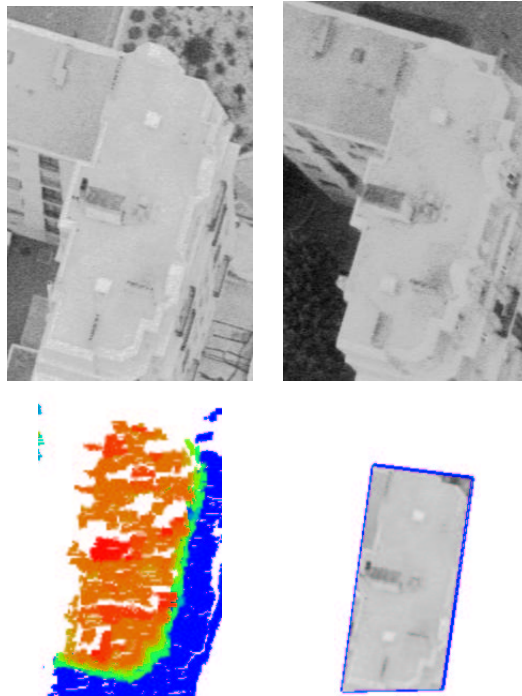


FIG. H.77 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 41

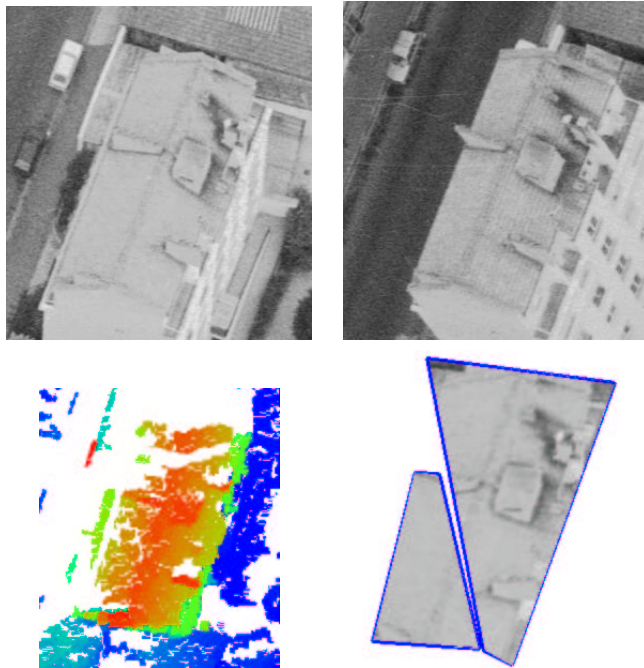


FIG. H.78 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 40

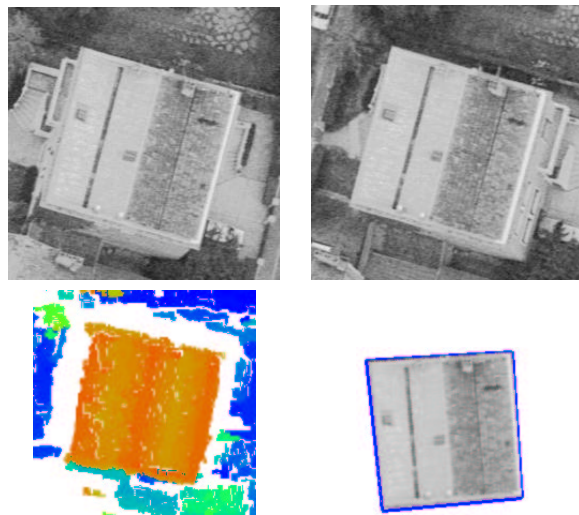


FIG. H.79 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 39

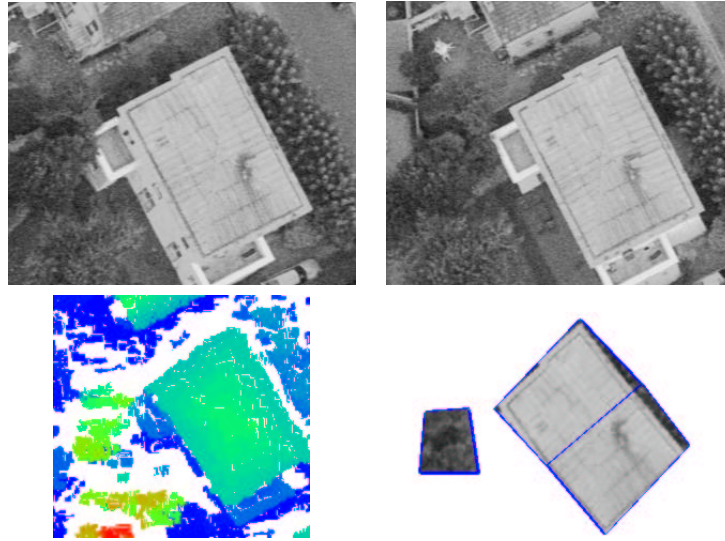


FIG. H.80 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 38

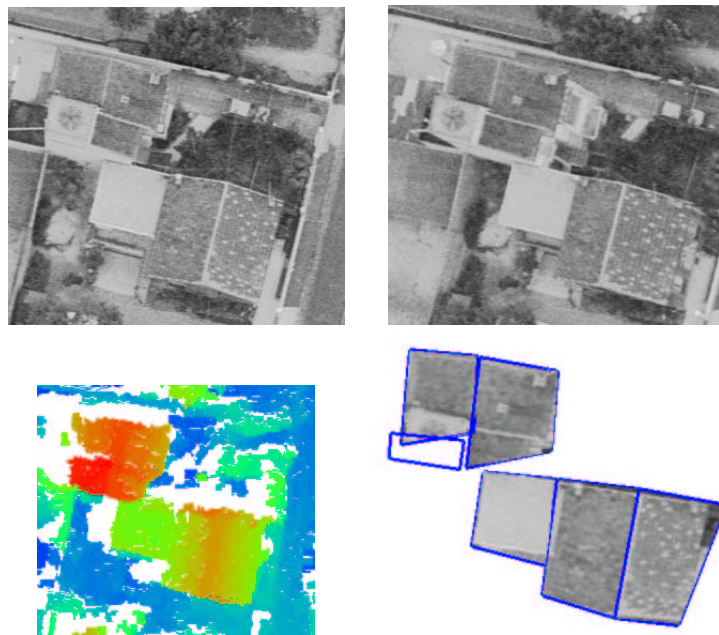


FIG. H.81 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 37

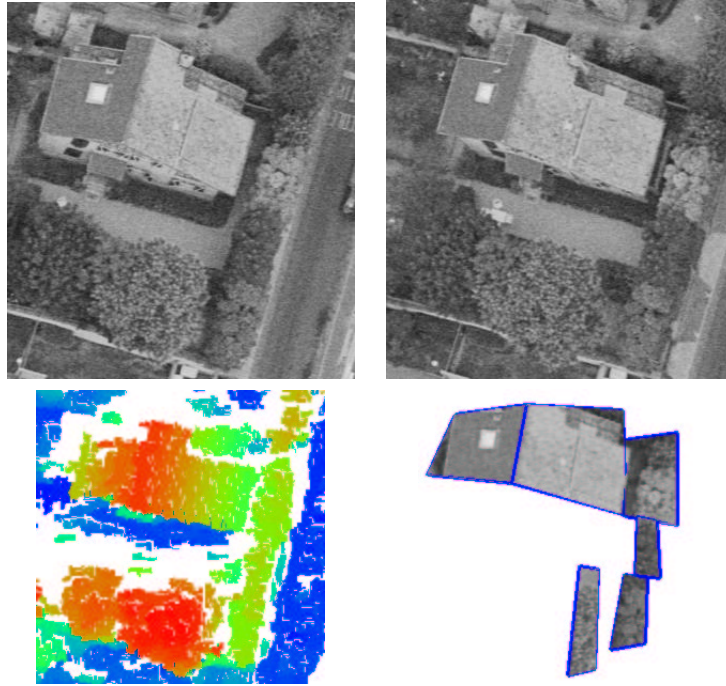


FIG. H.82 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 36

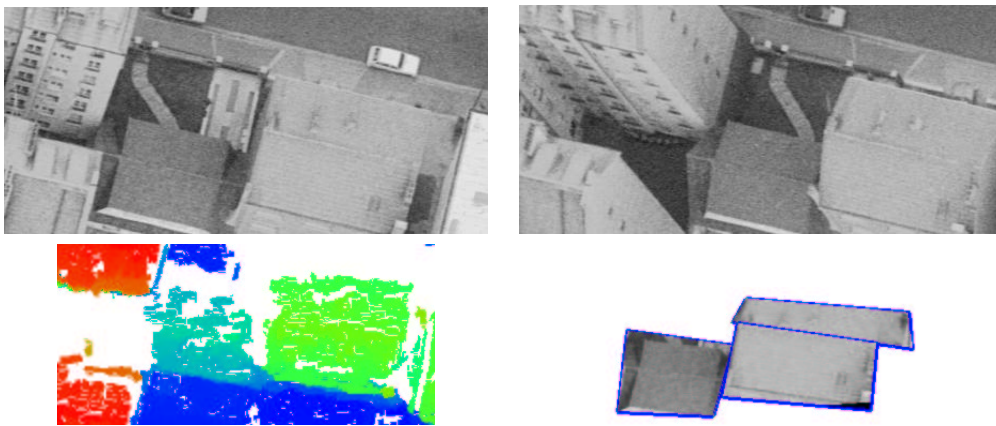


FIG. H.83 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 35

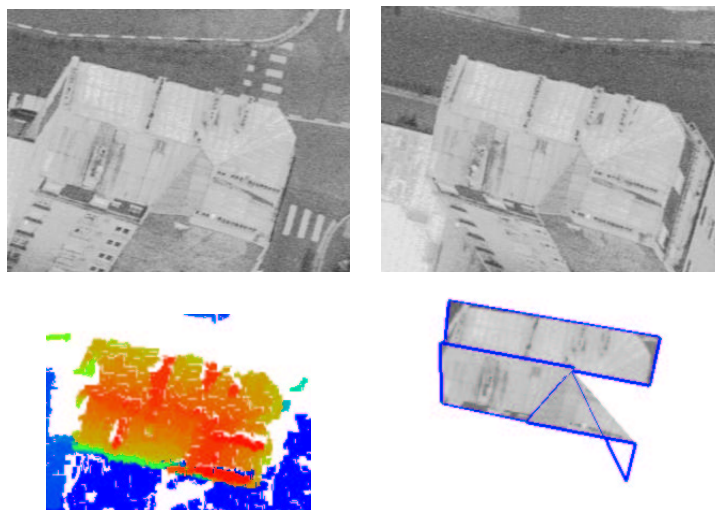


FIG. H.84 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 34

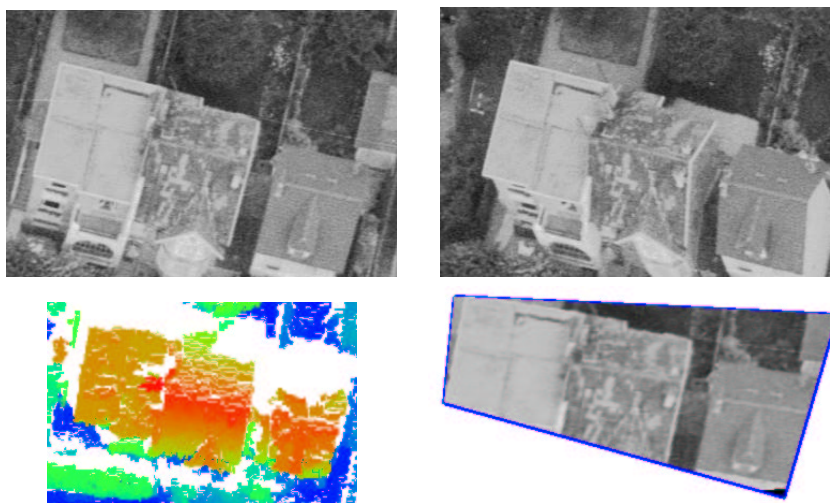


FIG. H.85 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 33

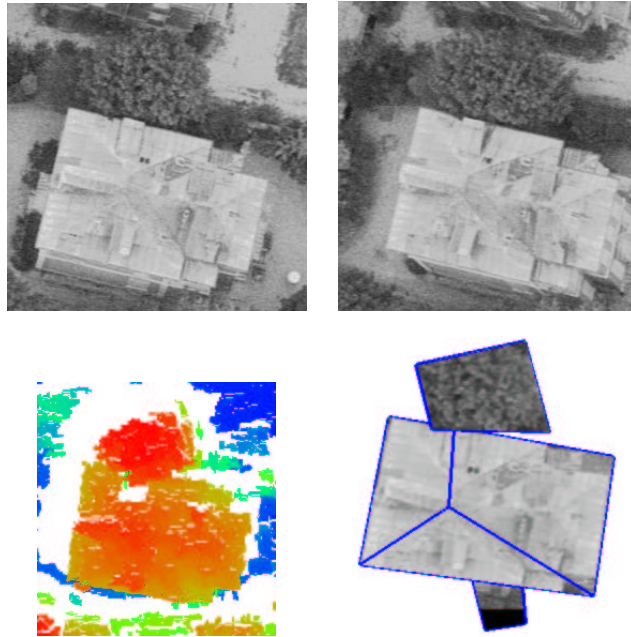


FIG. H.86 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 31

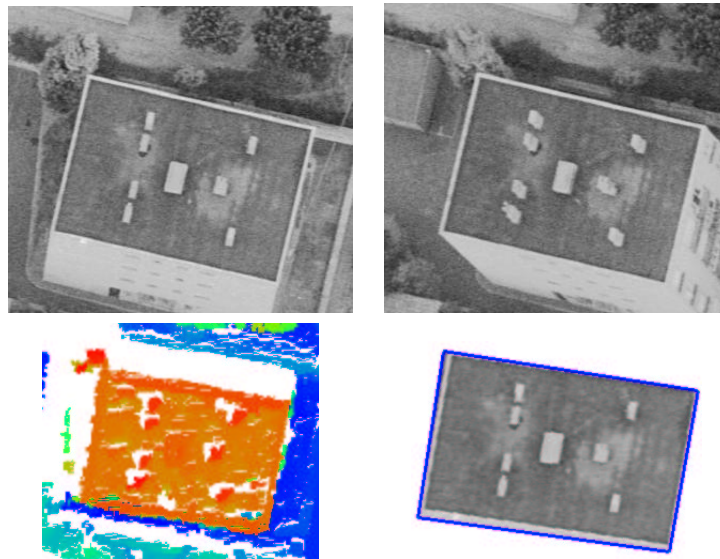


FIG. H.87 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 30

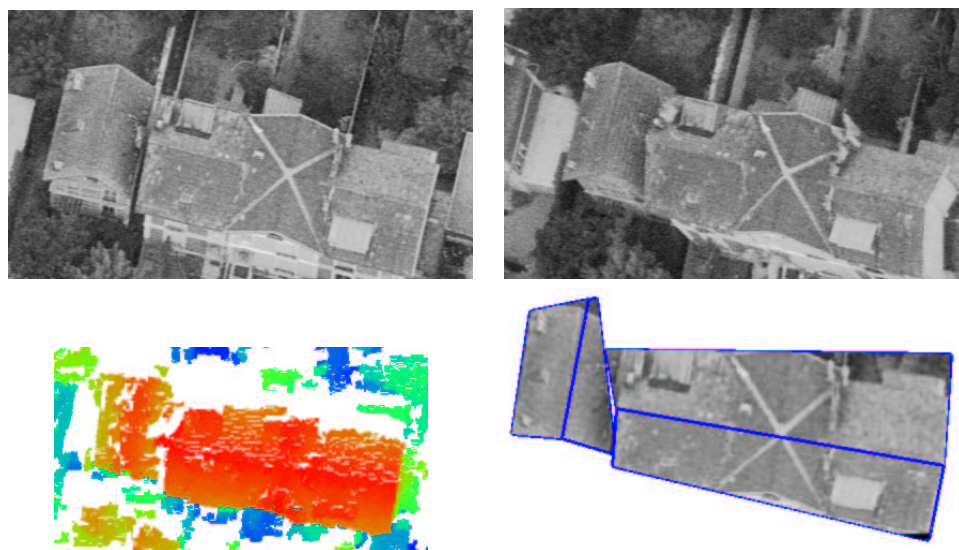


FIG. H.88 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 29

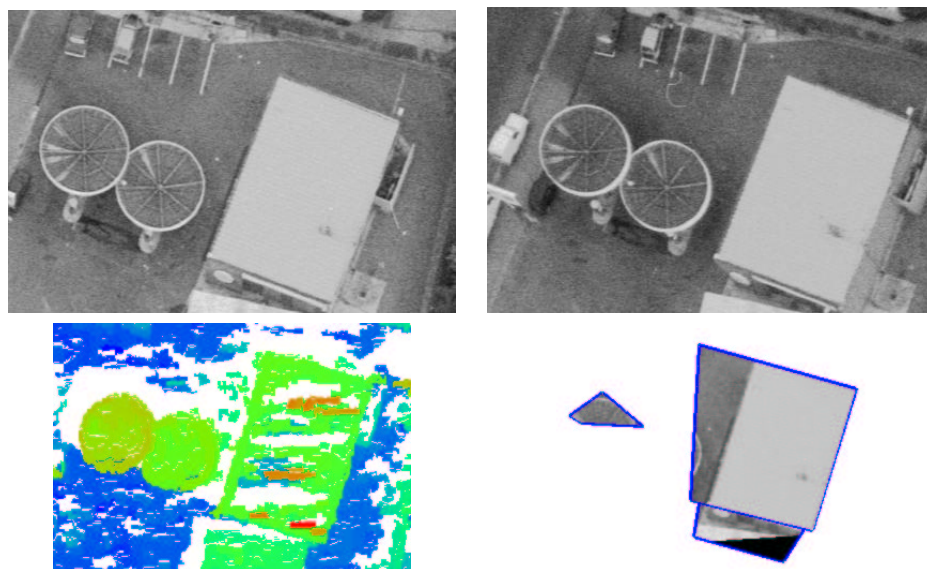


FIG. H.89 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 28

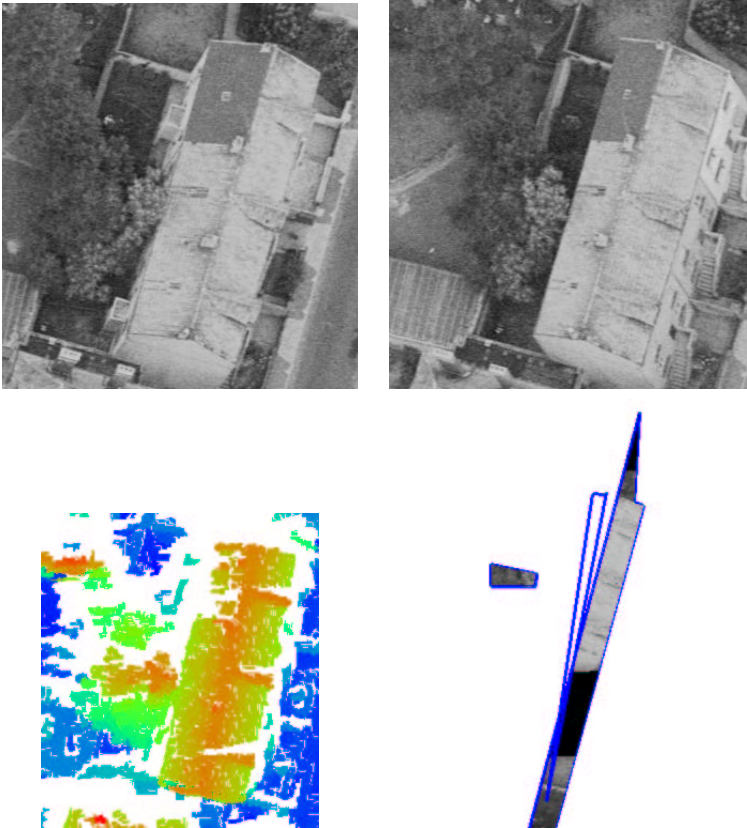


FIG. H.90 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 27

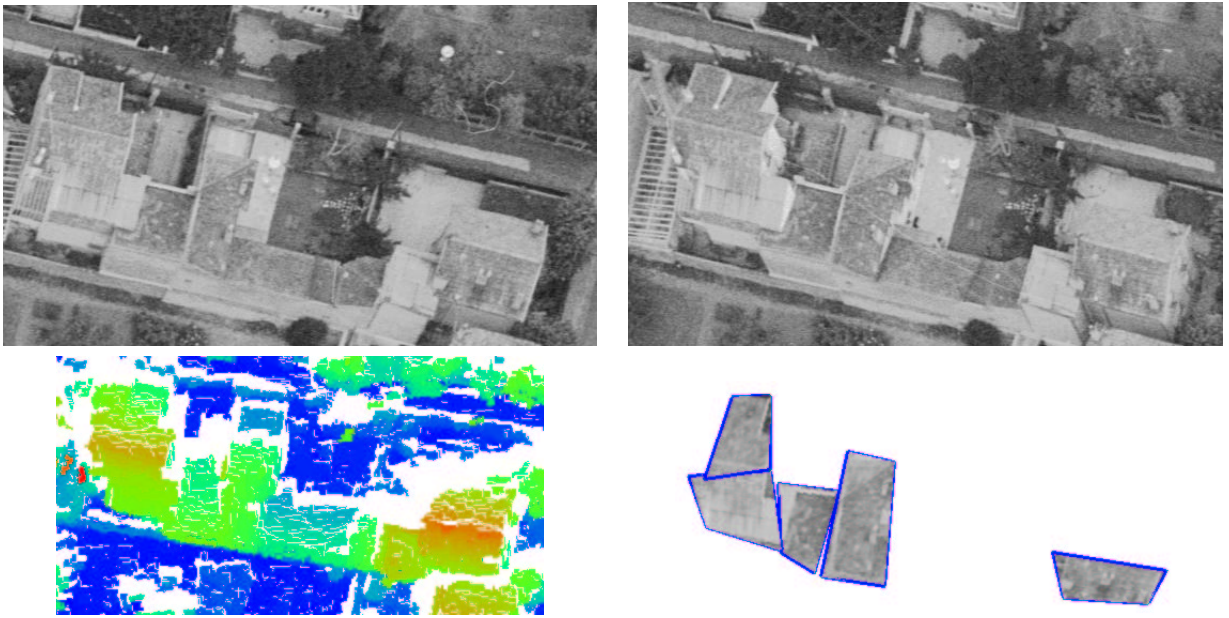


FIG. H.91 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 24

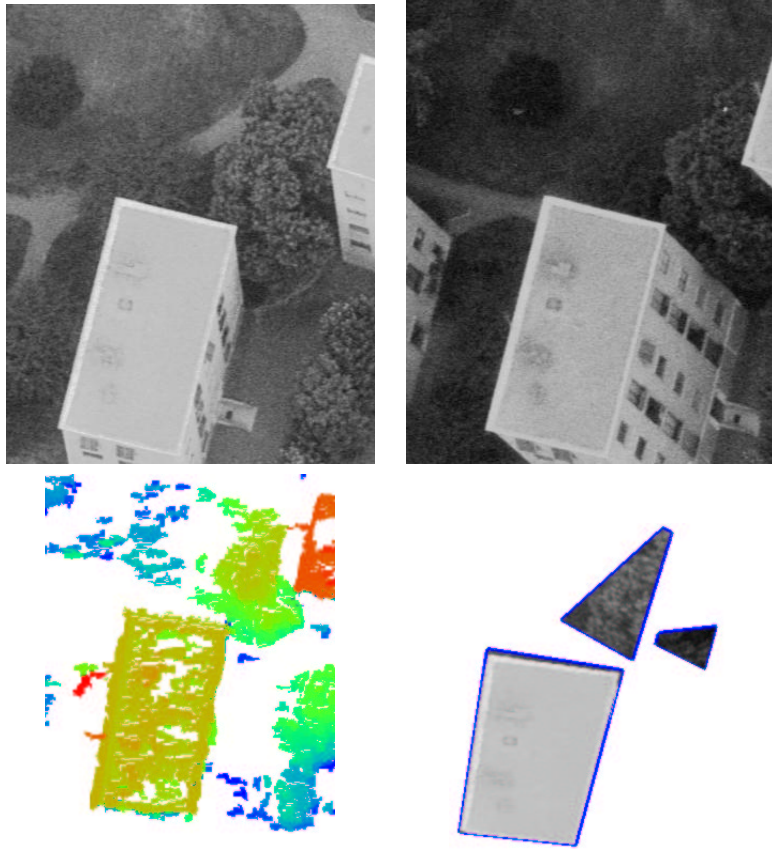


FIG. H.92 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 23

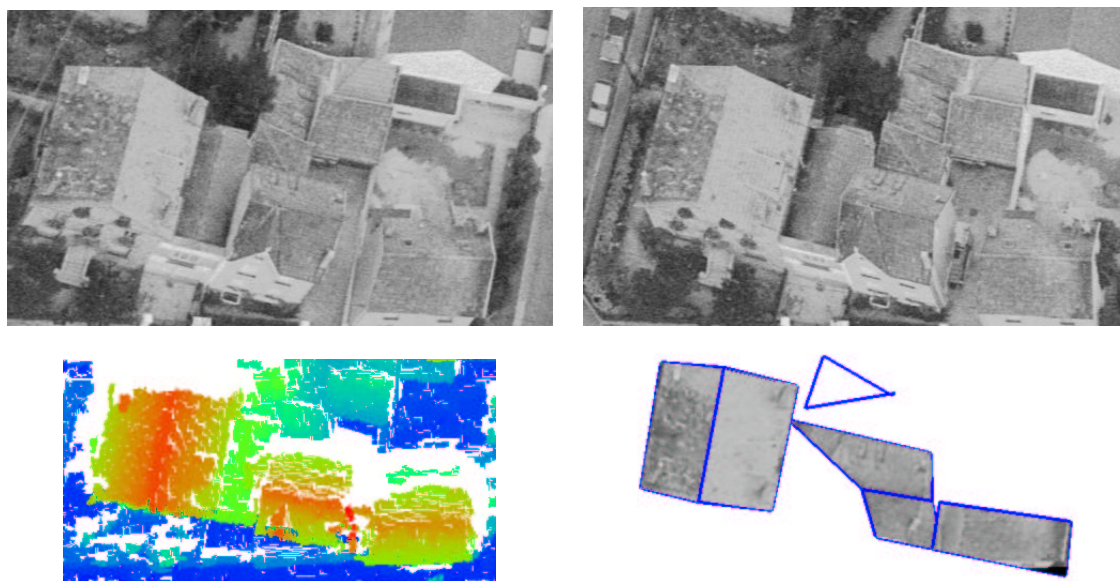


FIG. H.93 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 22

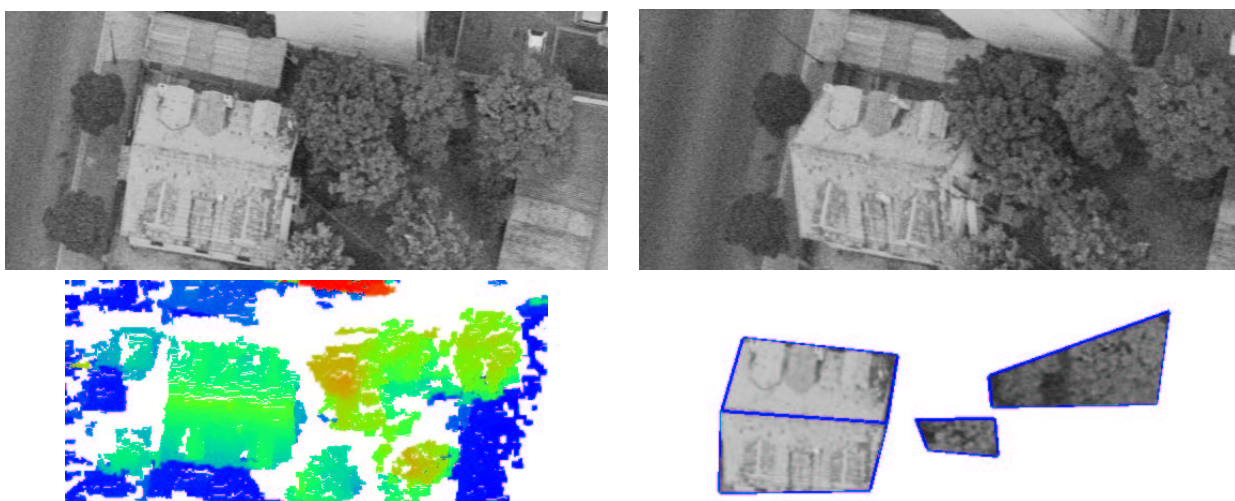


FIG. H.94 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 21

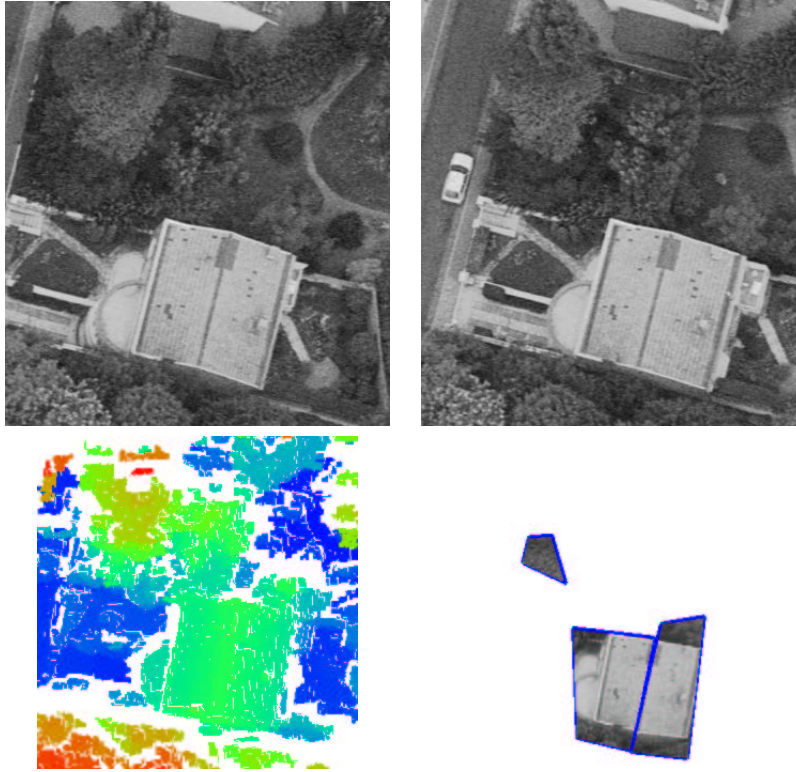


FIG. H.95 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 20

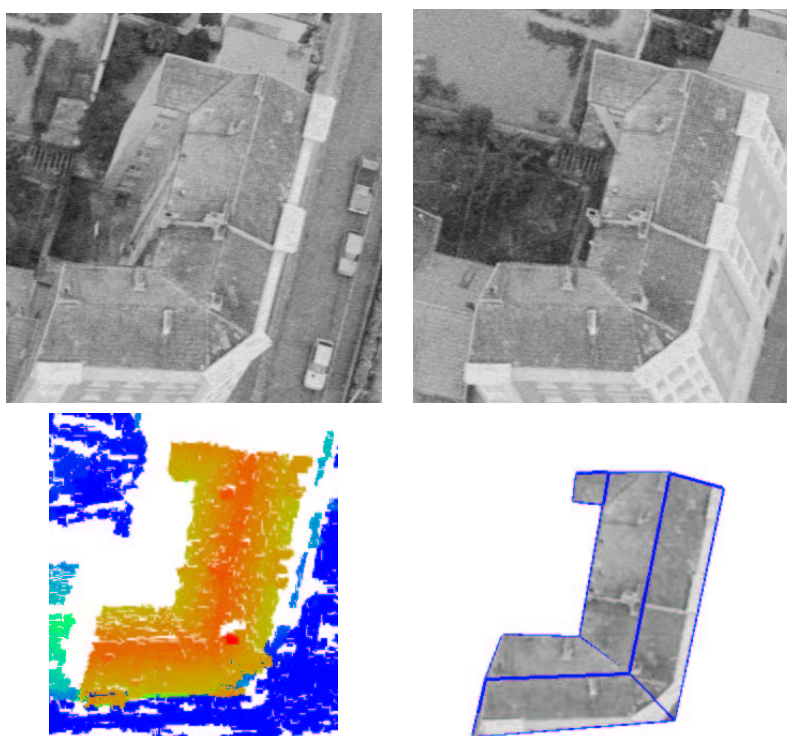


FIG. H.96 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 19

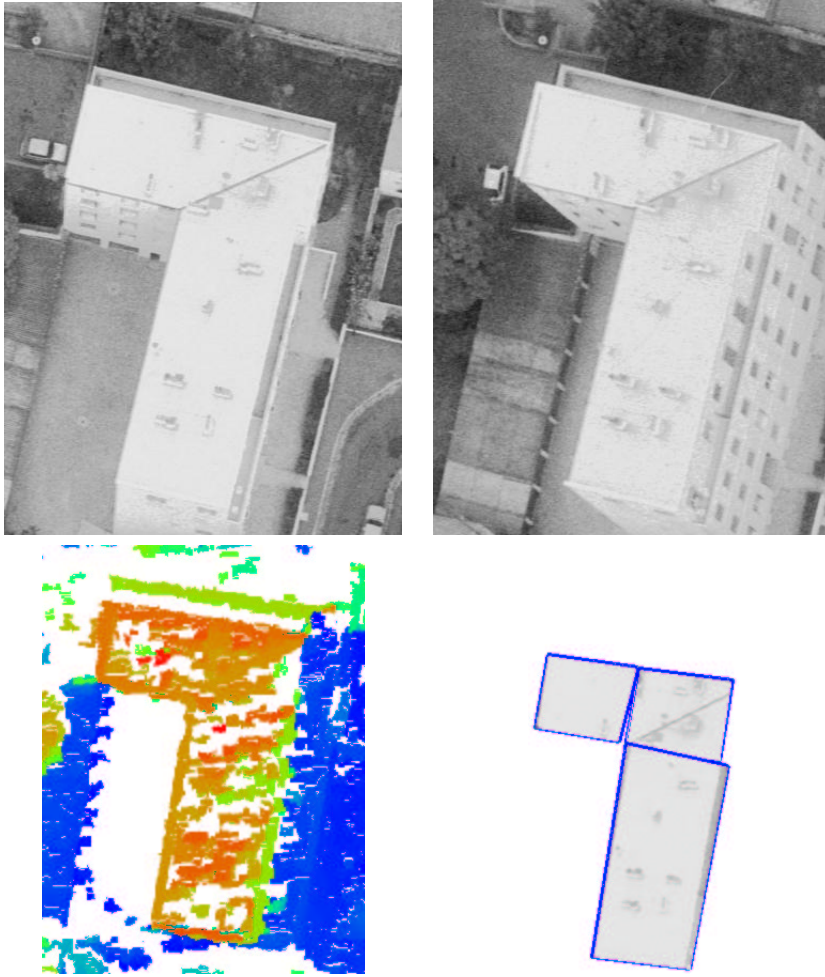


FIG. H.97 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 18

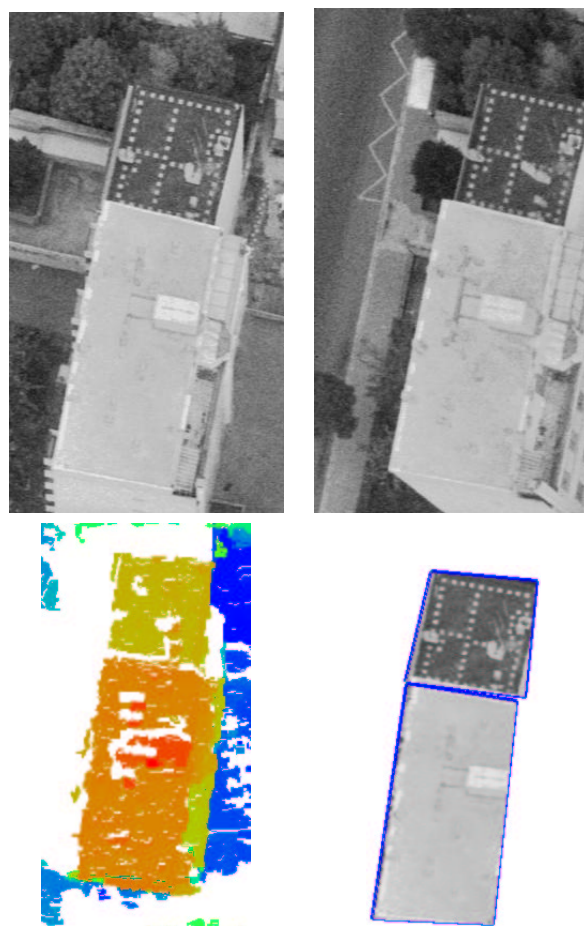


FIG. H.98 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 17

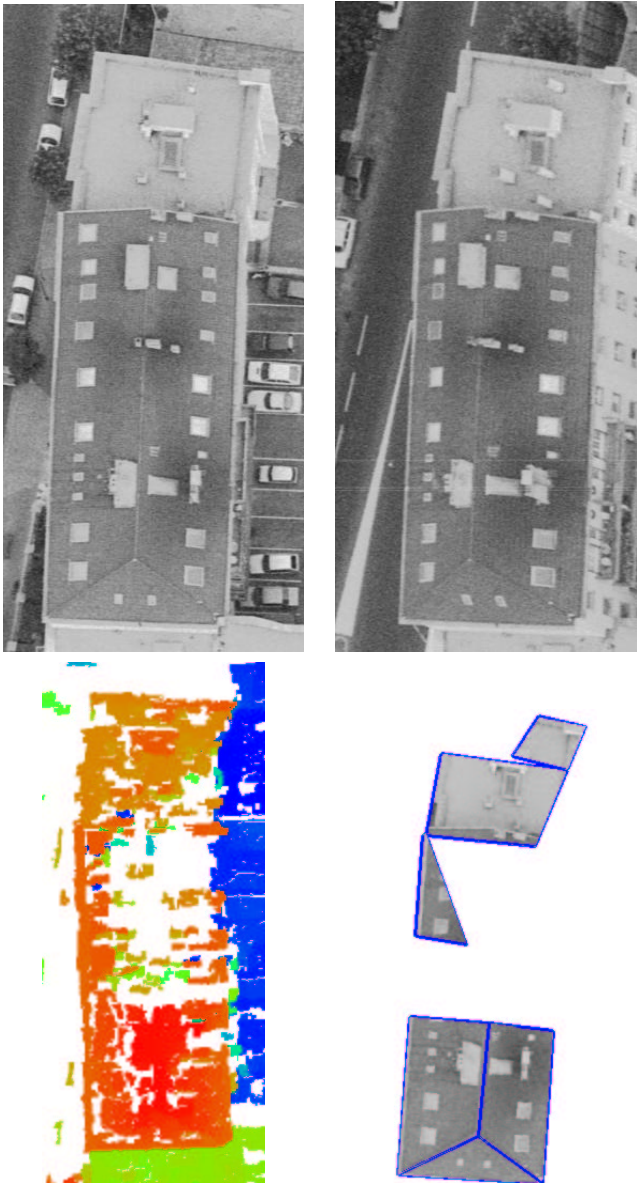


FIG. H.99 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 14

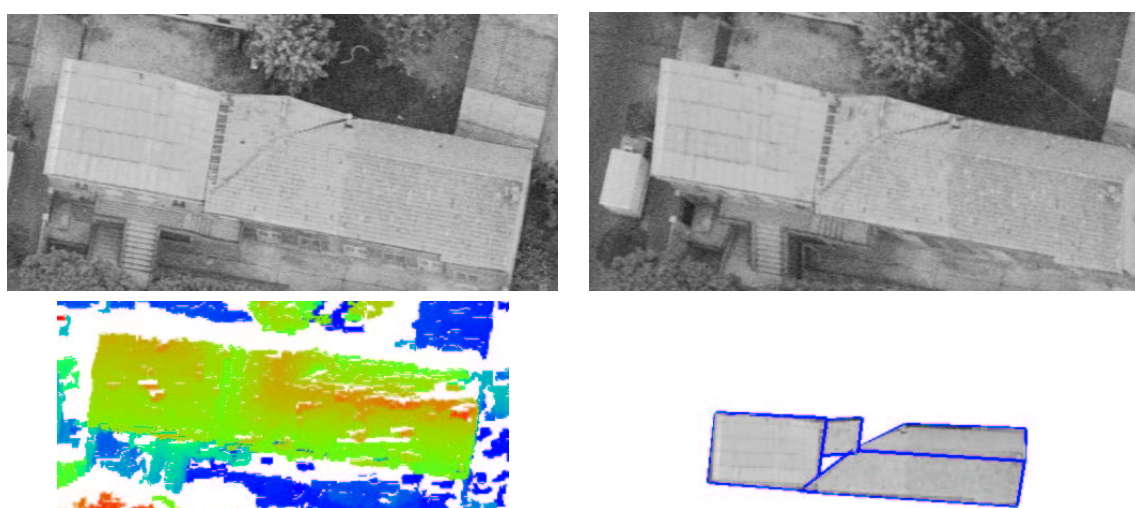


FIG. H.100 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 13

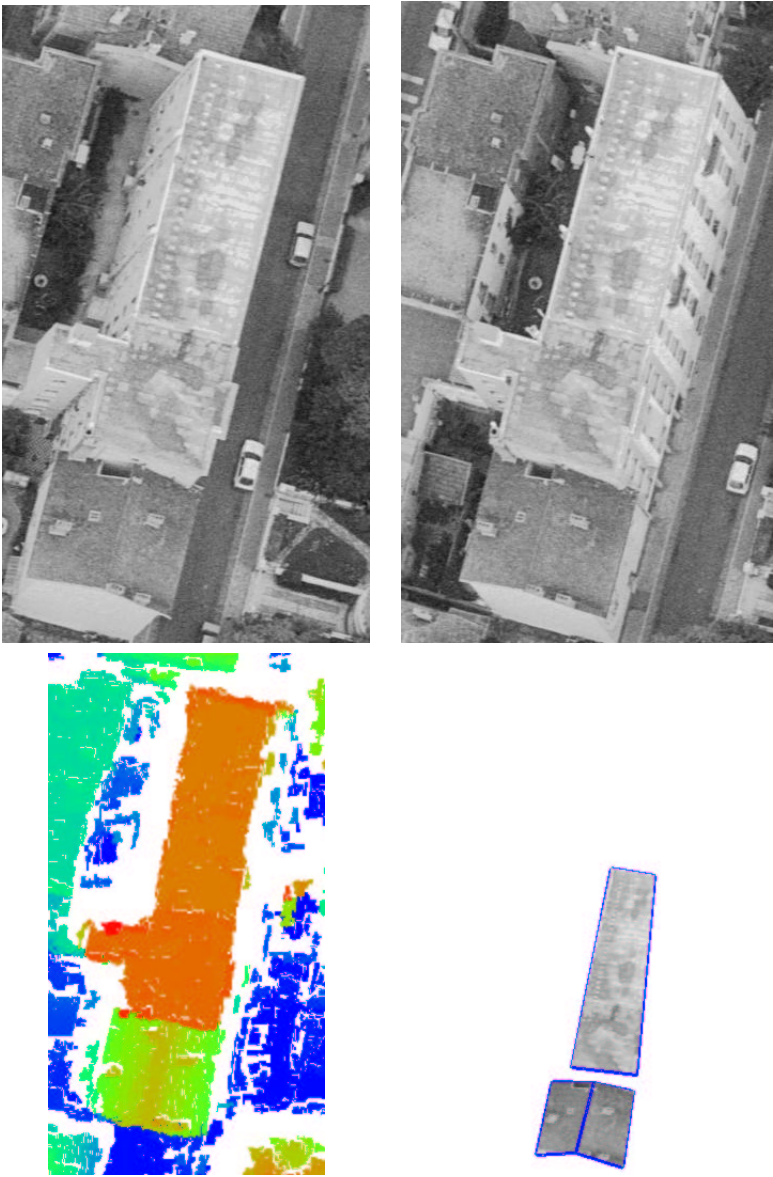


FIG. H.101 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 11

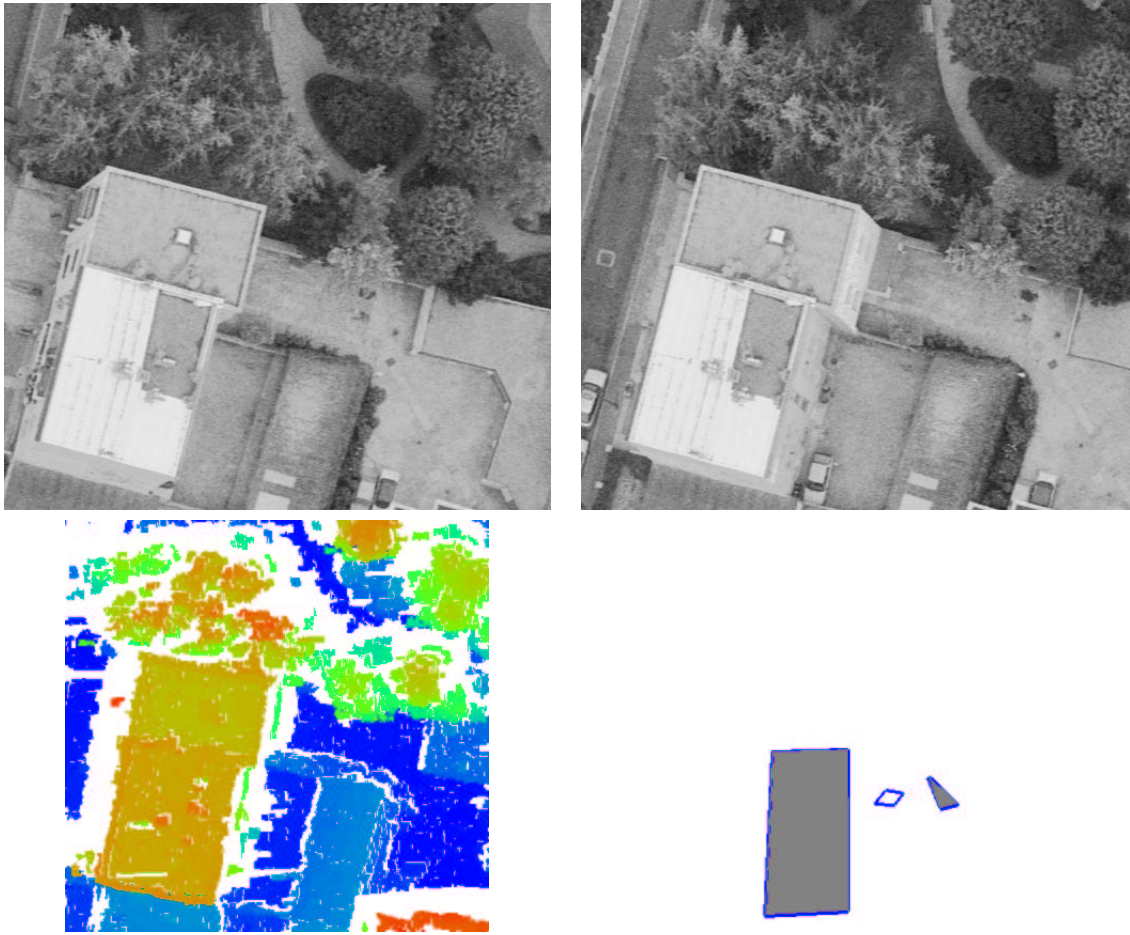


FIG. H.102 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 10

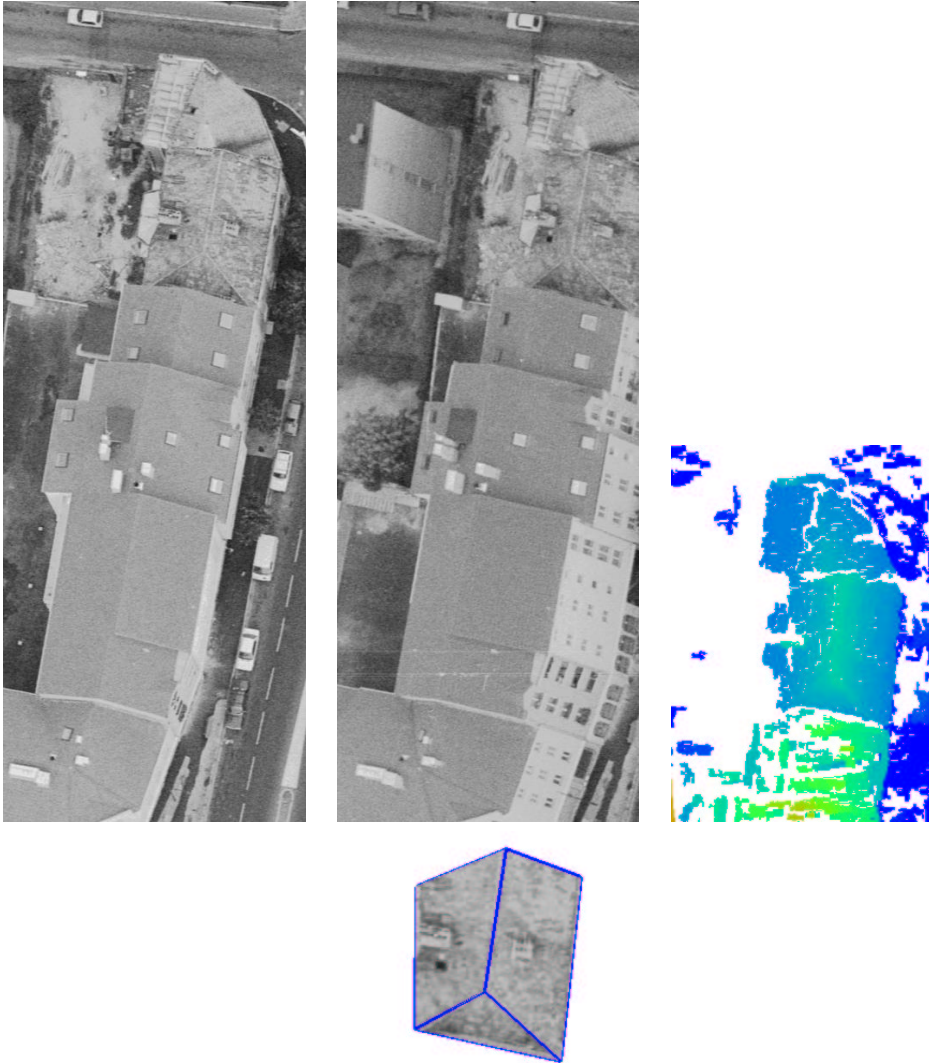


FIG. H.103 – Couple stéréoscopique, MNE, et reconstruction finale pour la zone 4. Dans cette zone, seul le tiers nord est considéré, car le MNE du reste de la zone est fortement dégradé en raison d'effets de bord, cette zone se situant en effet en limite de scène.

Bibliographie

- [asc01] *International Workshop on Automatic Extraction of Man-Made objects from Aerial and Space Images (III)*. 2001. Ascona.
- [Bai97] Caroline Baillard. *Analyse d'images aériennes stéréoscopiques pour la restitution 3-D des milieux urbains. Détection et caractérisation du sursol*. Thèse de doctorat, ENST, 1997.
- [BCG94] C. Ballester, V. Caselles, et M. González. Affine invariant segmentation by variational methods. In *Actes du 9ème Congrès RFIA*, pages 379–390, 1994.
- [BFJW01] Didier Boldo, Frank Fuchs, Franck Jung, et Yann Wehrstedt. Automating the aerial triangulation of digital frame camera images. The 3^d international Symposium on Mobile Mapping Technology, janvier 2001. Session 2.
- [BM99] Caroline Baillard et Henri Maître. 3-D Reconstruction of Urban Scenes from Aerial Stereo Imagery : A Focusing Strategy. *Computer Vision and Image Understanding*, pages 244–258, dec 1999.
- [BMS95] E. Baltsavias, S. Mason, et D. Stallmann. Use of DTMs-DSMs and Orthoimages to Support Building Extraction. In *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, pages 199–210, 1995.
- [Bre99] C. Brenner. Interactive tools for 3d building reconstruction. In *Proceedings of the Photogrammetric Week*, pages pp 23–34. Université de Stuttgart, 1999.
- [BW] R. Basri et D. Weinshall. Distance metric between 3d models and 2d images for recognition and classification.
- [BW98] Ansgar Brunn et Uwe Weidner. Hierarchical bayesian nets for building extraction using dense digital surface models. *ISPRS Journal of Photogrammetry and Remote Sensing*, 53 :296–307, 1998.
- [CC99] Tat-Jen Cham et Roberto Cipolla. Automated b-spline curve representation incorporating mdl and error-minimizing control point insertion strategies. *IEEE Trans. on PAMI*, 21(1) :49–53, 1999.
- [CH98] D. J. Cross et E. R. Hancock. Graph matching with a dual-step em algorithm. *IEEE Trans. on PAMI*, 20(11) :1236–1253, 1998.
- [Cor98] M. Cord. *Analyse d'images aériennes haute résolution : Détection et modélisation du Bâti en zone urbaine*. Thèse de doctorat, Université de Cergy Pontoise, 1998.

- [Dan94] Tuan Dang. *Interprétation et restitution automatique des bâtiments isolés à partir d'un couple stéréoscopique d'images aériennes*. Thèse de l'ENST, 1994.
- [DBMJ97] O. Dissard, C. Baillard, H. Maître, et O. Jamet. Above-ground objects in urban scenes from medium scale aerial imagery. In *Proc. Workshop on Automatic Extraction of Man-Made Objects from Aerial and Space Images (II)*, 1997.
- [Der87] R. Deriche. Using canny's criteria to derive a recursively implemented optimal edge detector. *International Journal of Computer Vision*, 1 :167–187, 1987.
- [DM] P. Dosch et G. Masini. Urban environment modelling by fusion of a cadastral map and a digital elevation model.
- [Dos00] P. Dosch. *Un environnement pour la reconstruction 3D d'édifices à partir de plans d'architecte*. Thèse de doctorat, LORIA - Université Poincaré - Nancy 1, 2000.
- [DTM00] P. E. Debevec, C. J. Taylor, et J. Malik. Modeling and rendering architecture from photographs : A hybrid geometry- and image-based approach. In *Actes SIGGRAPH*, 2000.
- [FG87] M.A. Förstner et E. Gülch. A fast operator for detection and precise location of distinct points, corners, and center of circular features. In *Proceedings of the Intercommission of the International Society for Photogrammetry and Remote Sensing*, 1987.
- [FM99] F. Fuchs et H. Le Men. Detecting planar patches in urban scenes. In *Proceedings of SPIE, Visual Information Processing VIII*, volume 3716, pages 167–176, avril 1999. Orlando.
- [FM00] F. Fuchs et H. Le Men. Efficient subgraph isomorphism with a-priori knowledge. application du building reconstruction for cartography. *Lecture Notes in Computer Science, 1876*, 2000.
- [Fri99] D. Fritsch. Virtual cities and landscape models - what has photogrammetry to offer ? In *Proceedings of the Photogrammetric Week*, pages pp 3–14. Université de Stuttgart, 1999.
- [FWH98] A. M. Finch, R. C. Wilson, et E. H. Hancock. Symbolic graph matching with the em algorithm. *Pattern Recognition*, 31(11) :pp. 1777–1790, 1998.
- [GGMR98] S. Girard, P. Guérin, H. Maître, et M. Roux. Construction d'une carte de disparité dense pour la reconnaissance de bâtiments en imagerie aérienne. In *Actes de la conférence RFIA '98, Clermont-Ferrand*, janvier 1998.
- [GM95] M. Gondran et M. Minoux. *Graphes et algorithmes*. Collection de la direction des Etudes et Recherches d'Electricité de France, 1995.
- [Grü98] A. Grün. Tobago - a semi-automated approach for the generation of 3-d building models. *ISPRS journal of Photogrammetry & Remote Sensing*, 53 :pp 108–118, 1998.
- [Gui97] Laurent Guigues. Recalage de modèles indéformables de bâtiments pour la restitution semi-automatique du bâti répétitif. Technical report, IGN, 1997.
- [GW98] A. Grün et X. Wang. Cc-modeler : a topology generator for 3-d city models. *ISPRS Journal of Photogrammetry and Remote Sensing*, 53 :pp. 286–295, 1998.
- [Haa99] N. Haala. Combining multiple data sources for urban data acquisition. In *Proceedings of the Photogrammetric Week*, pages pp 329–339. Université de Stuttgart, 1999.

- [Hab92] Annegret Habel. *Hyperedge Replacement : Grammars and Languages*, volume 643. Lecture Notes in Computer Science, 1992.
- [HB97] N. Haala et C. Brenner. Generation of 3-d city models from airborne laser scanning data. In *3rd EARSeL Workshop on LIDAR Remote Sensing of Land and Sea*, pages 105–112, juillet 1997.
- [HDRZ98] Christine Hivernat, Xavier Descombes, Sabine Randriamasy, et Josiane Zerubia. Mise en correspondance et recalage de graphes : application aux réseaux routiers extraits d'un couple carte/image. Technical Report N°3529, INRIA, 1998.
- [Hen96] Olof Henricsson. *Analysis of Image Structures using Color Attributes and Similarity Relations*. PhD Thesis No. 11663, Swiss Federal Institute of Technology, May 1996.
- [HN] S. Heuel et R. Nevatia. Including interaction in an automated modelling system.
- [HS88] C. Harris et M. Stephens. A combined edge and corner detector. In *Proceeding of the 4th Alvey Conference*, pages 147–151, 1988. Manchester.
- [JB94] X. Jiang et H. Bunke. Fast segmentation of range images into planar regions by scan line grouping. *Machine Vision and Applications*, 7 :pp 115–122, 1994.
- [JB96] X. Y. Jiang et H. Bunke. Fast segmentation of range images into planar regions by scan line grouping. Technical report, Institute of Informatics and Applied Mathematics, 1996.
- [JB99] X. Y. Jiang et H. Bunke. Edge detection in range images based on scan line approximation. *Computer Vision and Image Understanding*, 73(2) :pp 183–199, février 1999.
- [JCCP99] M. Jordan, M. Cord, J.P. Cocquerez, et N. Paparoditis. Automatic extraction and modelling of urban buildings from high-resolution aerial images. In *Automatic Extraction of GIS Objects from Digital Imagery*. IAPRS, 1999.
- [Jib00] H. Jibrini. Reconstruction 3d des bâtiments à partir de données 2d cadastrales vectorisées et d'images aériennes. In *RFIA*, février 2000.
- [KK00] Ho-Yon Kim et Jin H. Kim. Hierarchical random graph representation of handwritten characters and its application to hangul recognition. *Pattern Recognition*, 34(2) :187–201, 2000.
- [Lee99] T. C. M. Lee. A minimum description length based image segmentation procedure, and its comparison with a cross-validation based segmentation procedure. *Journal of the American Statistical Association*, 1999.
- [LF] F. Lang et W. Förstner. 3d-city modeling with a digital one-eye stereo system. In *ISPRS Commission IV, WG 4*.
- [LG98] T. Läbe et E. Gülch. Robust techniques for estimating parameters of 3d building primitives. In *Proceedings of the ISPRS Conference '98, Cambridge*, pages 169–176. ISPRS, 1998. Commission II, Working group 6.
- [Li92] M. Li. Minimum Description Length Based 2D Shape Description. Technical report, Department of Numerical Analysis and Computing Science, Suède, octobre 1992.
- [LK97] J. W. Lee et I.-S. Kweon. Extraction of line features in a noisy image. *Pattern Recognition*, Vol. 30(No 10), October 1997.

- [Loh96] U. Lohr. Imaging distance measurement by pulsed laser. In *3D-City Models Workshop*. Institut für Photogrammetrie, Universität Bonn, Octobre 1996.
- [MB96] B. T. Messmer et H. Bunke. Fast error-correcting graph isomorphism based on model precompilation. Technical report, Institut für Informatik und angewandte Mathematik, Universität Bern, Neubrückestr. 10, Bern, Suisse, septembre 1996. IAM-96-012.
- [MB98] B. T. Messmer et H. Bunke. A new algorithm for error-tolerant subgraph isomorphism detection. *IEEE Trans. on PAMI*, 20(5) :493–504, mai 1998.
- [Mes96] B. T. Messmer. *Efficient Graph Matching Algorithms for Preprocessed Model Graphs*. PhD Thesis, Institut für Informatik und angewandte Mathematik, Universität Bern, Switzerland., 1996.
- [MF99] H. Maître M. Fradkin, M. Roux. Building detection from multiple views. In *Automatic Extraction of GIS Objects from Digital Imagery*. IAPRS, 1999.
- [ML92] H. Maître et W. Luo. Using models to improve stereo reconstruction. *IEEE Trans. on PAMI*, 14(2) :269–277, 1992.
- [MM96] Henri Moissinac-Massenat. *Utilisation conjointe d'informations symboliques et de mesures numériques dans la prise de décision en traitement d'images*. Thèse de doctorat, ENST, 1996.
- [MME00] J. Marchadier, S. Michelin, et Y. Egels. A graph labelling approach for connected feature selection. In *Lecture Notes in Computer Science*, volume 1876, 2000.
- [MS89] D. Mumford et J. Shah. Optimal approximations by piecewise smooth functions and variational problems. *Communications on Pure and Applied Mathematics*, XLII(4), 1989.
- [MS94] J.C. McGlone et J.A. Shufelt. Projective and object space geometry for monocular building extraction. Technical report, Carnegie Mellon University, 1994.
- [MS98] Farzin Mokhtarian et Riku Suomela. Robust image corner detection through curvature scale space. *IEEE Trans. on PAMI*, 20(12), décembre 1998.
- [MWH00] R. Myers, R. C. Wilson, et E. R. Hancock. Bayesian edit distance. *IEEE Trans. on PAMI*, 22(6) :pp 626–635, 2000.
- [Ris83] J. Rissanen. A universal prior for integers and estimation by minimum description length. *The Annals of Statistics*, 11(2) :416–431, 1983.
- [Ris86] J. Rissanen. Stochastic complexity and modeling. *The Annals of Statistics*, 14(3) :1080–1100, 1986.
- [RM97] M. Roux et Henri Maître. Map analysis for guided interpretation of aerial images. In *GREC' 97*, Nancy, 1997. IAPR Workshop.
- [RMG97] M. Roux, H. Maître, et S. Girard. A step towards stereo reconstruction of urban aerial images. In *IAPRS*, volume 32, septembre 1997. partie 3-4W2.
- [Rou97] M. Roux. Cooperative analysis of maps and aerial images for urban scene description. In *SPIE Europto*, 1997. Londres.
- [RP90] I. Rock et S. Palmer. The legacy of gestalt psychology. In *Scientific American*, décembre 1990.

- [Rua99] Anne Ruas. *Modèle de généralisation de données géographiques à base de contraintes et d'autonomie*. Thèse de doctorat, IGN - Université de Marne La Vallée, 1999.
- [Sch96] Cordelia Schmid. *Appariement d'images par invariants locaux de niveaux de gris*. Thèse de doctorat, Institut National Polytechnique de Grenoble, 1996.
- [Shu99] J. A. Shufelt. Performance evaluation and analysis of monocular building extraction from aerial imagery. In *IEEE Trans. on PAMI*, volume 21, pages pp 329–339, 1999.
- [Ste97] V. Steinhage. On the integration of object modeling and image modeling in automated building extraction from aerial images. In *Automatic Extraction of Man-Made Objects from Aerial and Space Images (II)*, pages 139–148, 1997.
- [SVP⁺96] Michel Scholl, Agnès Voisnard, Jean-Paul Peloux, Laurent Raynal, et Philippe Rigaux. *SGBD Géographiques*. Thomson Publishing, 1996.
- [SZ97] Cordelia Schmid et Andrew Zisserman. Automatic line matching across views. *Computer Vision and Pattern Recognition*, Juin 1997.
- [Ull76] J. R. Ullman. An Algorithm for Subgraph Isomorphism. *Journal of the Association for Computing Machinery*, Vol. 23, No 1 :pp 31–42, 1976.
- [Ves00] Christophe Vestri. *Outils pour la reconstruction automatique de bâtiments à partir d'images aériennes*. Thèse de doctorat, Université de Nice - Sophia, 2000.
- [WA96] W. Willuhn et F. Ade. A rule-based system for house reconstruction from aerial images. In *Proceedings of ICPR '96*, pages 885–889, 1996.
- [WCH98] R. C. Wilson, A. D. J. Cross, et E. R. Hancock. Structural matching with active triangulations. *Computer Vision and Image Understanding*, 72(1) :PP. 21–38, octobre 1998.
- [WF95] U. Weidner et W. Förstner. Towards automatic building reconstruction from high-resolution digital elevation models. *ISPRS Journal of Photogrammetry and Remote Sensing*, 50(4) :38–49, 1995.
- [WH97] R. C. Wilson et E. R. Hancock. Structural matching by discrete relaxation. *IEEE Trans. on PAMI*, 19(6) :634–648, juin 1997.
- [ZY98] S. C. Zhu et A. Yuille. Region competition : Unifying snakes, region growing, and bayes/mdl for multiband image segmentation. *IEEE Trans. on PAMI*, 18(9), septembre 1998.