



**HAL**  
open science

## **CSTB - Projet Imagine : bilan 2011-2013, rapport final**

Olivier Tournaire, Souheil Soubra

► **To cite this version:**

Olivier Tournaire, Souheil Soubra. CSTB - Projet Imagine : bilan 2011-2013, rapport final. CSTB - Centre scientifique et technique du bâtiment. 2013. hal-01044737

**HAL Id: hal-01044737**

**<https://cstb.hal.science/hal-01044737>**

Submitted on 24 Jul 2014

**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.

# Simulation et outils numériques

## Projet Imagine Bilan 2011-2013

Rapport final

TOURNAIRE Olivier

**Novembre 2013**

**CENTRE SCIENTIFIQUE ET TECHNIQUE DU BATIMENT**

SIÈGE SOCIAL > 84 AVENUE JEAN JAURÈS | CHAMPS-SUR-MARNE | 77447 MARNE-LA-VALLÉE CEDEX 2

TÉL. (33) 01 64 68 82 82 | FAX. (33) 01 60 05 70 37 | SIRET 775 688 229 000 27 | [www.cstb.fr](http://www.cstb.fr)

ÉTABLISSEMENT PUBLIC À CARACTÈRE INDUSTRIEL ET COMMERCIAL | RCS MEAUX 775 688 229 | TVA FR 70 775 688 229

MARNE-LA-VALLÉE | PARIS | GRENOBLE | NANTES | SOPHIA-ANTIPOLIS



# **Simulation et outils numériques**

## **Projet IMAGINE**

### **Bilan 2011-2013**

TOURNAIRE Olivier

Université Paris-Est, Centre scientifique et technique du bâtiment, Direction  
technologies de l'information et diffusion du savoir, MODEVE

SOUHEIL Soubra

**Novembre 2013**

© 2014 CSTB



Ce texte est distribué sous les termes de la licence Creative Commons Attribution 3.0 non transposé (CC BY 3.0).

Le texte complet de la licence est disponible à l'adresse suivante : <http://creativecommons.org/licenses/by/3.0/legalcode>

Résumé des droits et conditions de la licence :

⇒ **Vous êtes libre de :**

- partager (reproduire, distribuer et communiquer) l'œuvre ;
- remixer, adapter l'œuvre ;
- d'utiliser cette œuvre à des fins commerciales.

⇒ **Selon les conditions suivantes :**

- Attribution (paternité, crédit) : vous devez attribuer l'œuvre de la manière indiquée par l'auteur de l'œuvre ou le titulaire des droits (mais pas d'une manière qui suggérerait qu'ils vous approuvent, vous ou votre utilisation de l'œuvre).

Toute citation d'extraits, reproduction ou utilisation doit obligatoirement faire apparaître la référence de ce document sous la forme : **TOURNAIRE O., SOUBRA S., Projet Imagine : bilan 2011-2013, rapport final, CSTB, novembre 2013, 37 p.**

⇒ **Comprenant bien que**

- les droits suivants ne sont en aucune manière affectés par la licence :
  - Vos prérogatives issues des exceptions et limitations aux droits exclusifs ou *fair use* ;
  - Les droits moraux de l'auteur que rien dans ce contrat ne diminue ou ne restreint.
- A chaque réutilisation ou distribution de cette œuvre, vous devez faire apparaître clairement au public la licence selon laquelle elle est mise à disposition. La meilleure manière de l'indiquer est un lien vers cette page web : <http://creativecommons.org/licenses/by/3.0/deed.fr>.

## **Résumé**

Ce document dresse un panorama des activités de recherche et développement menées au sein de l'équipe d'Imagine, pôle de recherche commun entre le CSTB et l'École des Ponts ParisTech. L'ensemble des résultats y sont présentés, en portant une attention particulière aux réalisations dédiées à la ville et au bâtiment, qui ont pour objectif de permettre d'obtenir une maquette numérique sémantisée, à ces deux échelles, de façon automatique, à partir de données facilement accessibles.

**Mots clés** : Vision par ordinateur, photogrammétrie, sémantisation, maquette numérique.

## **Abstract**

This document provides a global overview of research and development activities conducted within the Imagine team, joint research pole between CSTB and Ecole des Ponts ParisTech. All the results are presented, with particular attention to achievements dedicated to the city and building, which are intended to allow for a digital model semanticized, at these two scales, automatically, from data easily accessible.

**Keywords**: Computer vision, photogrammetry, semantization, digital mock-up

## Table des matières

<b>INTRODUCTION</b> .....	<b>8</b>
<b>1. CONTEXTE ET ENJEUX</b> .....	<b>9</b>
<b>2. PRINCIPAUX RÉSULTATS</b> .....	<b>11</b>
<b>2.1 Calibration et reconstruction 3D</b> .....	<b>11</b>
2.1.1 Calibration .....	11
2.1.2 Reconstruction 3D .....	13
<b>2.2 Sémantisation 2D</b> .....	<b>14</b>
2.2.1 Détection de fenêtres dans les images de façades .....	14
2.2.2 Grammaires de formes .....	16
2.2.3 Grammaires de graphes.....	17
2.2.4 Apprentissage de grammaires.....	19
<b>2.3 Sémantisation 3D</b> .....	<b>19</b>
2.3.1 Analyse de géométrie 3D brute .....	20
2.3.1.1 Simplification de maillages .....	20
2.3.1.2 Interprétation de nuages de points.....	22
2.3.2 Grammaires attributaires contraintes.....	23
2.3.2.1 Interprétation sémantique de modèles CAD.....	23
2.3.2.2 Détection d’erreurs dans les modèles et contrôle qualité .....	26
<b>3. PERSPECTIVES</b> .....	<b>27</b>
<b>3.1 Reconstruction terrestre et conjointe aérien / terrestre</b> .....	<b>27</b>
<b>3.2 Sémantisation 3D</b> .....	<b>29</b>
<b>3.3 Intégration de bases de données externes</b> .....	<b>31</b>
<b>3.4 Intégration et gestion des standards et modèles de données</b> .....	<b>31</b>
<b>3.5 Applications métiers</b> .....	<b>32</b>
<b>CONCLUSION</b> .....	<b>33</b>
<b>PUBLICATIONS</b> .....	<b>35</b>
<b>Articles de revue avec comité de lecture et d’audience internationale</b> .....	<b>35</b>
<b>Conférences avec actes</b> .....	<b>35</b>
<b>Thèses</b> .....	<b>36</b>

## Liste des figures

Figure 1 – Principe de la cohérence géométrique pour la mise en correspondance robuste .....	12
Figure 2 – Calibration par la méthode <i>a contrario</i> .....	12
Figure 3 - Quelques exemples de reconstructions .....	13
Figure 4 – Illustration de la reconstruction 3D sur la ville de Chamonix par sous-ensembles .....	14
Figure 5 – Résultats de la détection de fenêtres sur des images rectifiées et non rectifiées .....	15
Figure 6 - Principe des grammaires de formes : l'image de départ est récursivement segmentée en sous-ensembles cohérents. ....	16
Figure 7 – Amélioration de la carte de mérite à partir des détections de fenêtres : (a) avant, (b) après. Amélioration des hypothèses de découpes en utilisant les gradients proches des fenêtres : (c) hypothèses verticales initiales, (d) nouvelles hypothèses.....	17
Figure 8 –Résultats de l'état de l'art(a) et (c) et nos résultats (b) et (d), pour lesquels la position et la taille des fenêtres sont notablement améliorés .....	17
Figure 9 - Le principe des grammaires de graphes .....	18
Figure 10 - Comparaison entre les résultats obtenus par grammaires de formes (a) et (c) et par grammaires de graphes (b) et (d) .....	18
Figure 11 - Grammaire simple divisant une façade en étage, et chaque étage en travées et baies. A droite, l'arbre d'interprétation associé. ....	19
Figure 12 - Les différentes étapes de la simplification planaire .....	20
Figure 13 - Simplification planaire d'un modèle 3D à partir d'images multivues. (De gauche à droite) Exemples d'images utilisées en entrée ; Le nuage de point reconstruit ; Résultats de la simplification planaire. (de haut en bas) Cluny ; Marseille .....	21
Figure 14 - Simplification d'un maillage photogrammétrique par détection de primitives géométriques simples .....	21
Figure 15 – De gauche à droite et de haut en bas : Normales reconstruites sur un angle droit par minimisation par moindres carrés ; Normales obtenues par notre méthode ; Segmentation d'un nuage de points par cohérence locale des normales (chaque objet, quel que soit sa géométrie est extrait correctement : colonne, chanfrein, mobilier ...). ..	22
Figure 16 - Les différents éléments constitutifs d'un escalier .....	23
Figure 17 - Modélisation d'un escalier par une grammaire attributaire contrainte.....	24
Figure 18 - Résultats de la détection des escaliers dans des modèles CAO avec le formalisme de grammaires attribuées contraintes.....	25
Figure 19 - Résultats de la sémantisation globale sur un modèle CAO avec le formalisme de grammaires attribuées contraintes .....	25
Figure 20 – a) Résultat de sémantisation ; b) et d) Primitives extraites ; c) et e) Texturation automatique en fonction de la sémantisation .....	26
Figure 21 - Erreurs et incohérences dans les modèles CAD mis en évidence lors de l'étape de sémantisation. ....	27
Figure 22 - De gauche à droite : Une image acquise par <i>Stéréopolis</i> , système de cartographie mobile développé au laboratoire MATIS de l'IGN ; Détails sur des reconstructions 3D obtenues à partir de ces images.....	28
Figure 23 – En haut : reconstruction à partir d'images aériennes à 10cm (gauche) et terrestres (droite). En bas : zone commune aux deux reconstructions (gauche) et fusion des deux maillages (droite).....	29
Figure 24 – Fusion des emprises de bâtiments issues de la BD Topo sur un modèles 3D (Mandelieu). ....	30



## **INTRODUCTION**

Le présent rapport dresse un bilan du projet Imagine, pôle de recherche commun entre le CSTB et l'École des Ponts ParisTech pour la période 2011-2013. Ce pôle de recherche a été créé en 2009, et la collaboration a été renouvelée entre les deux organismes au début de l'année 2013. L'objectif majeur est l'étude du bâtiment et des ensembles urbains, via des techniques d'apprentissage, de modélisation statistique et de vision par ordinateur.

Dans un premier temps, sont rappelés le contexte et les enjeux des recherches conduites au sein du pôle Imagine. Sont ensuite présentés dans une seconde partie les principaux résultats et avancées scientifiques obtenus sur la période écoulée. La troisième et dernière partie évoque les pistes de recherche qu'il convient de suivre afin de poursuivre les travaux entrepris.

## 1. CONTEXTE ET ENJEUX

Photogrammétrie et vision par ordinateur sont devenues en quelques années des outils incontournables pour la production de données 3D, aussi bien à l'échelle du bâtiment qu'à celle du quartier ou de la ville entière. En effet, ces techniques permettent, en disposant de plusieurs vues prises sous des angles différents d'un même objet ou territoire, d'en obtenir une reconstruction géométrique tridimensionnelle. Les intérêts majeurs de ces approches sont leur automaticité totale et leur faible coût, ce dernier se limitant au coût d'acquisition des données qui dépend du vecteur utilisé (images terrestres ou aéro/héli-portées) pour leur acquisition.

Dans le même temps, que ce soit à l'échelle du bâtiment isolé de son environnement, du quartier ou de la ville, la maquette numérique est devenue un outil fondamental à plusieurs titres, et s'adresse aux différents acteurs de la ville : monde de la construction, urbanistes, architectes, décideurs publics, usagers, ... Aussi, il est nécessaire, afin de toucher ces sensibilités variées, de disposer d'outils adaptés aux diverses situations rencontrées. Une modélisation fine et précise pour l'étude et la représentation d'un bâtiment seul peut être nécessaire, alors que pour la réalisation de simulations, une représentation simplifiée ne conservant que les grandes lignes de la construction sera plus adaptée. L'enjeu consiste donc à trouver le bon compromis entre simplification et qualité, et à adapter la représentation des modèles aux échelles d'étude.

Par ailleurs, des contraintes spécifiques peuvent s'ajouter aux types de reconstructions numériques que l'on souhaite obtenir. En effet, modéliser l'extérieur d'un bâtiment est une problématique bien différente de la modélisation de son intérieur. Lorsque le bâtiment est occupé, de nombreux objets viennent « polluer » la scène ou en cacher une partie (mobilier, équipements, ...). En outre, la reconstruction d'intérieurs est un champ de recherche à part entière, et les méthodes traditionnelles de photogrammétrie et de vision par ordinateur, qui opèrent à partir d'images, sont, dans l'immense majorité des cas, inaptés à reconstruire de tels environnements. En effet, l'absence de textures des parois principales (murs homogènes), ou les variations brutales de points de vue lorsque l'on passe d'une pièce à l'autre, amènent à devoir imaginer des méthodes nouvelles, introduisant plus de contraintes, ou des acquisitions avec des matériels spécifiques, notamment avec des scanners laser qui permettent d'obtenir un nuage de points 3D de l'environnement.

Enfin, et c'est l'un des enjeux majeurs à l'heure actuelle, les reconstructions tridimensionnelles obtenues de manière automatique ne représentent que la géométrie de la scène acquise. Pour disposer d'une représentation et d'une compréhension globale d'une scène, il est nécessaire d'enrichir les modèles produits en leur ajoutant une information de haut niveau sur son contenu. C'est la sémantisation, qui, à l'échelle du bâtiment va s'attacher à reconnaître les différents éléments architecturaux d'une façade (découpage en étages, nombre, tailles et positions des fenêtres, ...), et, à l'échelle urbaine, va s'intéresser à l'extraction des différents composants de la ville (bâtiments, végétation, ...). De nombreuses applications sont aujourd'hui tributaires de cette information riche. Aussi, il est important de pouvoir l'extraire de façon automatique.

Les recherches pour avancer dans ces domaines apparaissent donc multiples et à la croisée de plusieurs domaines thématiques (vision par ordinateur, algorithmie géométrique, apprentissage et optimisation, ...). Elles mêlent par ailleurs les échelles, passant du bâtiment isolé à la ville dans son ensemble. Répondre à ces différentes problématiques et satisfaire l'ensemble des acteurs de la ville est un enjeu crucial qui ne peut aboutir que par des avancées scientifiques aussi bien théoriques que pratiques. Il convient ainsi de développer les méthodes existantes pour les améliorer, mais également d'explorer de nouvelles pistes afin d'ouvrir les verrous aujourd'hui existants. Ce dernier point passe par des avancées théoriques qui sont fondamentales pour poser les problèmes de la façon la plus juste qui soit.

La diversité des sources de données pour l'acquisition de l'existant est une variable à prendre en compte. En effet, il est possible de s'appuyer sur des images simples ou multiples, sur des relevés laser qui acquièrent directement la géométrie tridimensionnelle ou sur des modèles construits manuellement (type CAD/CAO). Cet éventail de sources permet d'adapter les méthodes à mettre en œuvre aux besoins (précision, texture, ...), aux moyens (différences de coûts) et aux contraintes d'acquisition, mais aussi à la disponibilité de données existantes.

Sur la base de ces constats et afin de répondre aux problèmes et aux interrogations soulevées dans le monde du bâtiment et de la construction par la vision par ordinateur, les recherches du pôle Imagine, se sont orientées autour de deux axes majeurs : la reconstruction 3D à partir d'images multivues, et la sémantisation, aussi bien d'images 2D que d'images de géométrie tridimensionnelle, ce dernier point permettant de faire le lien avec la reconstruction de la géométrie.

Concernant la reconstruction 3D, un thème phare et bien maîtrisé d'Imagine, il était important, en plus de disposer d'un outil spécifiquement dédié à cette tâche, de prendre en compte des critères d'exploitation dans un cadre applicatif. Aussi, l'amélioration de la qualité de la reconstruction a été explorée. Des gains qualitatifs pouvaient être générés en accroissant la qualité de la calibration, processus consistant à déterminer la géométrie de la prise de vue, mais aussi en utilisant au mieux l'information image dans le processus de reconstruction. Le passage à l'échelle des algorithmes est aussi un point crucial qui ne doit pas être négligé. En effet, reconstruire un bâtiment isolé ou une petite zone constituée de quelques dizaines ou centaines d'images est aisément réalisable, mais traiter des volumes importants de données reste problématique. Par ailleurs, les besoins en précision s'accroissant, les prises de vue aériennes tendent à être de plus en plus résolues (de l'ordre de quelques centimètres). Couvrir un territoire, même à échelle moyenne, requiert d'immenses quantités de données. Les approches développées à l'heure actuelle ne permettent pas leur traitement. Il convient donc de s'approprier cette problématique afin de la traiter efficacement, sans dégrader la performance des résultats.

L'autre grand défi actuel reste la sémantisation automatique. L'engouement ressenti il y a quelques années pour la 3D s'est aujourd'hui quelque peu émoussé. Les acteurs de la construction et de la ville s'entendent cependant sur le fait que la représentation numérique est source de richesses. Cela dit, une simple représentation des territoires, sans information de haut niveau, n'est

aujourd'hui plus satisfaisante. Bien que la seule géométrie permette déjà de conduire un certain nombre d'études (dispersion de fluides, études de visibilité, propagation des ondes, ...), les besoins applicatifs récents sont plus exigeants sur la donnée d'entrée. Les études se tournent en effet vers des considérations sur l'efficacité des ensembles bâtis et de la ville, ou leurs potentialités. Pour étudier cette efficacité, les outils numériques de simulation nécessitent en entrée des informations que la géométrie seule ne peut apporter. Par exemple, avoir une première approche de l'efficacité énergétique d'un bâtiment requiert de qualifier son enveloppe : baies et parois opaques, matériaux des façades et toitures, ... Evaluer le potentiel photovoltaïque d'un quartier nécessite de disposer des surfaces sur lesquelles il est possible d'installer des panneaux. A ce titre, la modélisation du bâtiment en tant que tel ne suffit pas, il est nécessaire de pouvoir en identifier les différents composants : par exemple, fenêtres, toits et superstructures pour le calcul de masques fins.

## **2. PRINCIPAUX RÉSULTATS**

Sur la période écoulée, les recherches menées au sein du pôle de recherche IMAGINE se sont principalement concentrées sur la sémantisation, aussi bien à partir d'images que de géométries tridimensionnelles. Néanmoins, de substantielles avancées ont également été obtenues sur le passage à l'échelle des algorithmes de reconstruction, ainsi que sur la calibration des caméras, préalable nécessaire à toute reconstruction géométrique à partir d'images multivues. Ce sont parfois de petits détails, mais, pour obtenir une reconstruction, précise, fine et détaillée, tous les détails sont importants.

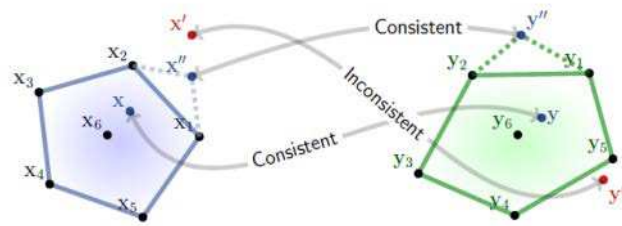
### **2.1 CALIBRATION ET RECONSTRUCTION 3D**

Une étape fondamentale de la chaîne de reconstruction 3D à partir d'images multivues est la calibration. Cette étape permet de reconstruire la géométrie de la prise de vue, c'est-à-dire de replacer dans un système cohérent les points d'acquisition et les orientations des caméras dans ce repère.

#### **2.1.1 CALIBRATION**

Deux nouvelles approches ont été développées pour la calibration. Toutes deux visent à obtenir une plus grande précision et apportent une plus grande souplesse d'utilisation ainsi qu'une meilleure robustesse par rapport aux méthodes de l'état de l'art.

La première approche, utilisée pour la calibration mais aussi pour la reconnaissance d'objets (cf. 2.2.1), consiste à aborder le problème de mise en correspondance entre images comme un problème de cohérence locale entre groupes de points d'intérêts (OK, et al., 2012). En effet, en ajoutant une contrainte d'appariement de régions (et donc de cohérence géométrique) plutôt que de points isolés, la robustesse s'en trouve accrue : il est plus facile de rejeter de faux appariement dans ce cadre que si l'on se base uniquement sur une information très locale et purement radiométrique.

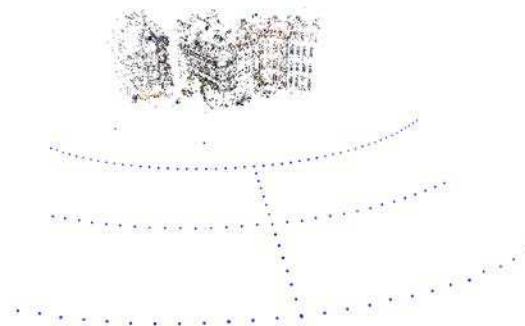


**Figure 1 – Principe de la cohérence géométrique pour la mise en correspondance robuste**

Cette méthode, qui peut multiplier les points d'intérêts potentiels en entrée sans crainte de contamination massive par les *outliers*, a permis d'améliorer significativement les méthodes de l'état de l'art en permettant de calibrer plus de caméras que les méthodes traditionnelles, notamment dans des conditions difficiles.

Une autre méthode de calibration (Moulon, et al., 2012), basée sur une méthodologie *a contrario* a également été développée. Cette approche a prouvé ses performances à calibrer finement un ensemble de caméras, grâce à des seuils adaptatifs, contrairement aux méthodes classiques qui utilisent des seuils « en dur », et généralement choisis pour l'ensemble du jeu de données (la méthode *a contrario* permet de déterminer des seuils spécifiques à chaque image, ce qui en fait également un facteur clé de robustesse).

Cette adaptation automatique des seuils dans la recherche de correspondances entre les images pour la calibration est particulièrement importante. En effet, étant données les conditions d'acquisition, les variations d'éclairage que peuvent présenter les images doivent être prises en compte. La méthode *a contrario* utilisée permet de gérer aisément le problème, en évitant d'avoir à fixer des valeurs fixes sur les seuils. Cette méthode permet aussi d'éviter les dérives accumulées en cas de boucle, en prenant en compte un modèle global.



**Figure 2 – Calibration par la méthode *a contrario***

Les positions des caméras sont représentées par les points bleus.

### 2.1.2 RECONSTRUCTION 3D

Les technologies de reconstruction automatique de modèles 3D à partir d'images ont été améliorées, tant en termes de qualité que de capacité à traiter de grands volumes de données en des temps raisonnables. L'approche retenue a consisté à utiliser au mieux l'information image pour la précision, les capacités de calcul des cartes graphiques pour la rapidité et un découpage du problème en sous-problèmes pour le passage à l'échelle.

Le maillage est maintenant raffiné en utilisant une information photométrique consistante avec les images, c'est-à-dire en utilisant la surface pour la reprojection dans les images. Cela permet d'assurer l'exactitude de la correspondance et une meilleure qualité du maillage, notamment sur les petits détails. Quelques exemples de reconstructions sont présentés sur la Figure 3.



**Figure 3 - Quelques exemples de reconstructions**

A gauche, Cluny à partir de 26 images ; à droite, un calvaire

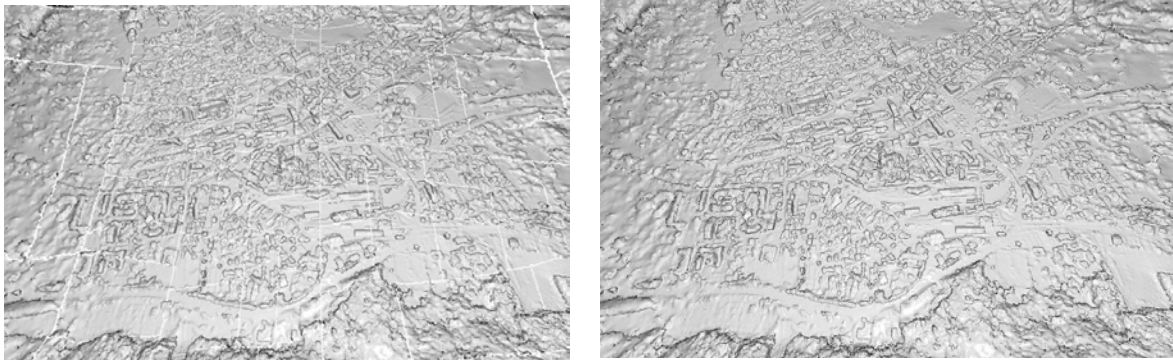
Un problème majeur concernait le passage à l'échelle des algorithmes. En effet, pour traiter une acquisition sur une ville entière, il n'est pas pensable de stocker simultanément en mémoire la totalité du chantier. Pour résoudre ce problème, la reconstruction peut maintenant être effectuée par sous-ensembles, qui sont par la suite fusionnés. La cohérence globale du maillage est respectée, ainsi que sa topologie. Couplée à une implémentation GPGPU<sup>1</sup>, cette approche permet de traiter en un temps raisonnable (quelques heures de calculs), des chantiers sur des villes entières, à haute résolution (cf. Figure 4).

Cette technologie de reconstruction 3D a fait l'objet en janvier 2011 de la création de la start-up Acute3D montée par deux membres de l'équipe Imagine, Renaud Keriven (ENPC) et Jean-Philippe Pons (CSTB).

Il est aussi à noter que ces techniques de mise en correspondance, de calibration et de reconstruction 3D ont été utilisées avec succès sur de nombreux jeux de données de l'état de l'art. Elles ont par ailleurs permis de remporter le *PRoVisG Mars 3D Challenge*, un concours international qui mettait à l'épreuve des systèmes de calibration et de reconstruction 3D à partir d'images

<sup>1</sup> *General-Purpose Computation on Graphics Processing Unit*

stéréoscopiques dans la perspective de futures explorations planétaires robotisées.



**Figure 4 – Illustration de la reconstruction 3D sur la ville de Chamonix par sous-ensembles**

A gauche, les reconstructions par dalles ; à droite, la reconstruction globale après fusion des différents maillages

## 2.2 SÉMANTISATION 2D

La sémantisation a été l'une des activités phare du pôle IMAGINE ces dernières années. Les premiers travaux ont porté sur de la détection d'objets répétitifs, avec une application aux bâtiments, pour notamment chiffrer la part de surface vitrée d'une façade. Des algorithmes existants ont également été étendus et améliorés pour une interprétation sémantiquement plus riche.

### 2.2.1 DÉTECTION DE FENÊTRES DANS LES IMAGES DE FAÇADES

En utilisant les travaux d'un doctorant (Ok, 2013), il est possible de repérer dans une image les objets répétitifs. En effet, la méthode d'appariement de régions entre images peut être appliquée à une image unique pour en extraire les régions, et donc les objets similaires. Ces travaux ont été naturellement appliqués à la détection de fenêtre dans des images de façades. Ainsi, si l'on dispose d'une instance de fenêtre à détecter (obtenue automatiquement ou saisie manuellement), un algorithme itératif permet d'extraire toutes les instances de cet objet. La prise en compte, dans l'algorithme d'appariement, de régions de déformation affine permet d'utiliser l'algorithme sur des images non rectifiées, ce qui constitue une avancée de l'état de l'art (cf. Figure 5).



**Figure 5 – Résultats de la détection de fenêtres sur des images rectifiées et non rectifiées**

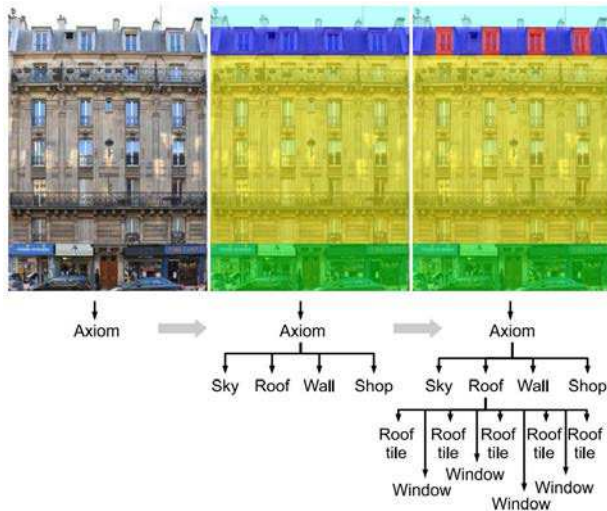
Première ligne : résultats sur images rectifiées de la base de données ECP Paris<sup>2</sup> ; seconde ligne : résultats sur images non rectifiées de la base de données eTRIMS<sup>3</sup>

<sup>2</sup> <http://vision.mas.ecp.fr/Personnel/teboul/data.php>

<sup>3</sup> [http://www.ipb.uni-bonn.de/projects/etrimis\\_db/](http://www.ipb.uni-bonn.de/projects/etrimis_db/)



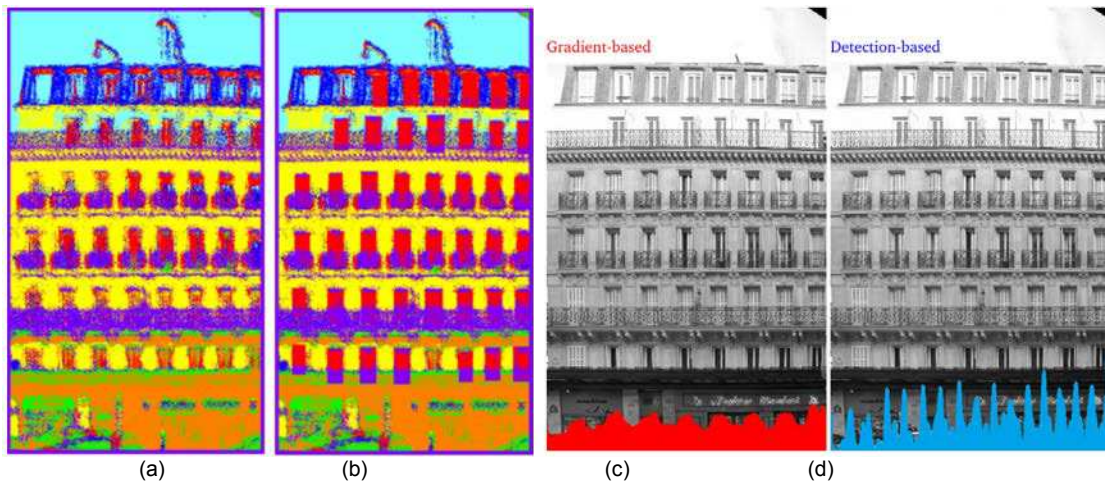
### 2.2.2 GRAMMAIRES DE FORMES



**Figure 6 - Principe des grammaires de formes : l'image de départ est récursivement segmentée en sous-ensembles cohérents.**

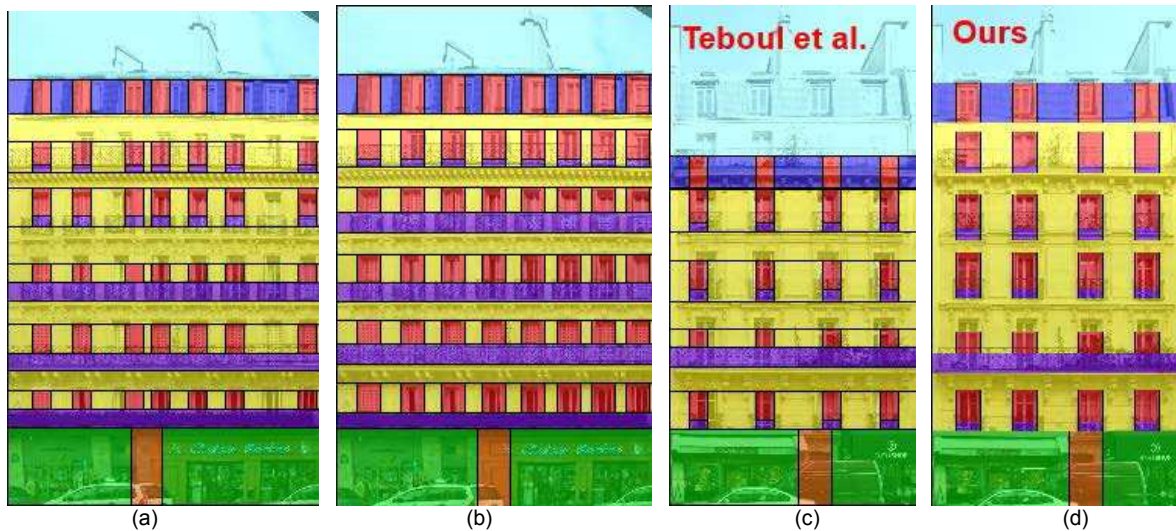
Cet algorithme de détection de fenêtre a été utilisé pour améliorer les résultats de l'état de l'art (Ok, et al., 2012) reposant sur une approche par grammaire de forme. Le principe de cet algorithme consiste à définir de façon formelle les relations qui existent entre les différents objets d'une image, et ce, de façon récursive (cf. Figure 6). La façade se trouve alors récursivement décomposée en ses différents éléments constitutifs, ce qui forme un arbre, la racine étant l'image entière, les feuilles représentant les objets détectés. La grammaire définie comporte 7 classes (ciel, toit, fenêtre, balcon, mur, porte et vitrine) et s'applique à l'architecture de style haussmannien.

L'approche originale reposait sur le calcul d'une carte de mérite qui classe les pixels en fonction de leur probabilité d'appartenance à une classe, ainsi que sur des hypothèses de découpes horizontales et verticales de l'image, basés sur les gradients. Or, cette carte était relativement bruitée, et les hypothèses de découpes peu claires. La carte de mérite a pu être améliorée en utilisant l'algorithme de détection de fenêtre présenté précédemment (cf. 2.2.1). Par ailleurs, en utilisant la détection de cet objet extrêmement structurant, la position des hypothèses de découpe obtenues via les gradients est remplacée par des positions dérivées de celles des fenêtres (cf. Figure 7).



**Figure 7 – Amélioration de la carte de mérite à partir des détections de fenêtres : (a) avant, (b) après. Amélioration des hypothèses de coupes en utilisant les gradients proches des fenêtres : (c) hypothèses verticales initiales, (d) nouvelles hypothèses**

La combinaison de ces deux éléments permet aujourd'hui de sensiblement améliorer les résultats de l'état de l'art, et la rapidité de calcul s'en trouve favorisée (cf. Figure 8).

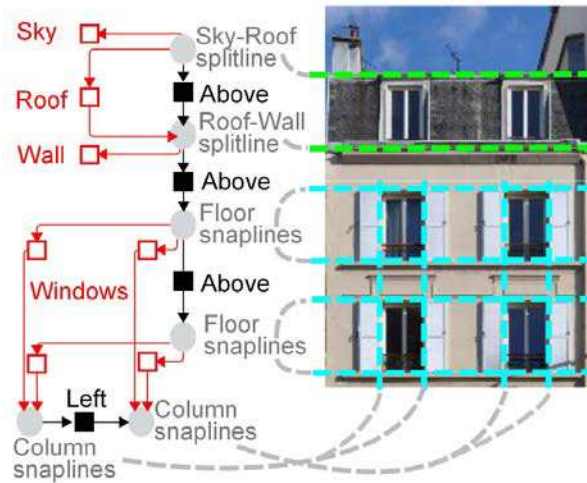


**Figure 8 – Résultats de l'état de l'art (a) et (c) et nos résultats (b) et (d), pour lesquels la position et la taille des fenêtres sont notablement améliorés**

### 2.2.3 GRAMMAIRES DE GRAPHES

La précédente approche repose sur une exploration d'un grand espace, combinant la structure de l'objet et les positions de ses composants. La détermination des positions est conduite par des procédés stochastiques ou « gloutons », ces derniers dépendant fortement d'une bonne solution initiale.

Pour traiter ces problèmes, une approche par grammaire de graphes a été proposée. Elle permet de découpler le problème en optimisant la structure de l'objet et ses positions séparément.



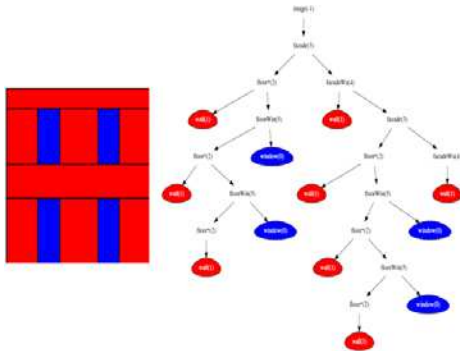
**Figure 9 - Le principe des grammaires de graphes**

Les différents nœuds représentent des objets architecturaux de la façade, et les relations de position sont encodées



**Figure 10 - Comparaison entre les résultats obtenus par grammaires de formes (a) et (c) et par grammaires de graphes (b) et (d)**

### 2.2.4 APPRENTISSAGE DE GRAMMAIRES



**Figure 11 - Grammaire simple divisant une façade en étage, et chaque étage en travées et baies. A droite, l'arbre d'interprétation associé.**

Les méthodes précédentes à base de grammaires ont prouvé leur robustesse et leur fiabilité. Cependant, l'écriture d'une grammaire est une étape lourde et fastidieuse, réalisée par un expert, qui doit être répétée pour chaque type d'architecture (cf. Figure 11). L'objectif de ces travaux est d'apprendre les règles spécifiques d'une grammaire, à partir d'une grammaire de formes générique, en utilisant une "vérité terrain" propre à chaque classe architecturale.

L'algorithme développé consiste à commencer par générer un arbre de dérivation à partir d'une vérité terrain, ce qui permet d'en déduire une grammaire. Une étape retour est ensuite réalisée afin de valider les règles de la grammaire en interprétant la vérité terrain, ce qui permet de valider ou non les règles. La grammaire est finalement optimisée en recherchant dans l'arbre de dérivation obtenu les sous-graphes identiques.

### 2.3 SÉMANTISATION 3D

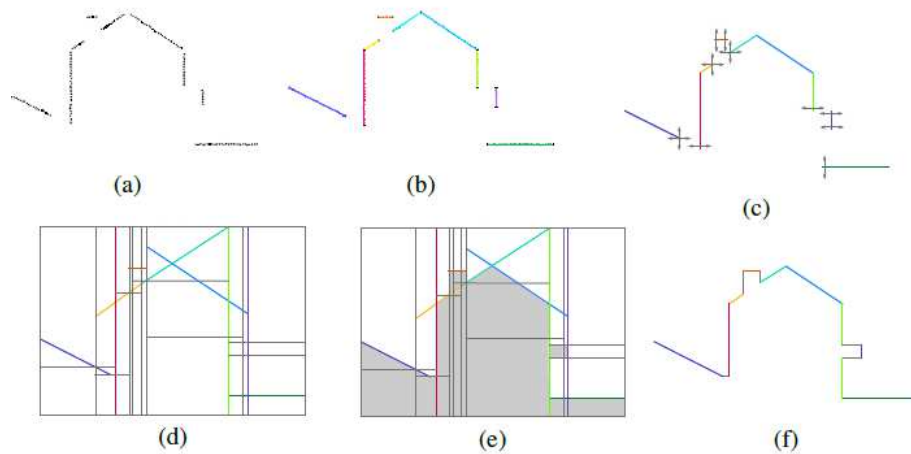
La sémantisation est un processus qui peut aussi être conduit directement à partir de la géométrie 3D. Les méthodes sont cependant différentes, la donnée d'entrée ne portant pas les mêmes informations. En effet, l'information photométrique est généralement absente (il est cependant possible d'y revenir si la géométrie a été obtenue par des procédés photogrammétriques, où une information d'intensité est généralement acquise dans le cas d'un scan laser).

A partir de géométrie 3D, la sémantisation peut s'entendre sous plusieurs sens. Le premier, consiste à obtenir à partir d'une donnée complexe (type nuage de points 3D), une représentation simplifiée prenant en compte les contraintes géométrique liées, par exemple, aux espaces bâtis. Une autre approche de plus haut niveau consiste à identifier, comme dans le cas de l'interprétation des images de façades, les objets qui composent la scène observée. Des avancées significatives ont été obtenues dans ces deux domaines.

## 2.3.1 ANALYSE DE GÉOMÉTRIE 3D BRUTE

### 2.3.1.1 Simplification de maillages

La simplification de maillages reconstruits à partir d'images multivues, avec la prise en compte des contraintes spécifiques aux bâtiments ou aux environnements urbains, constitue le premier niveau de sémantisation. C'est par ailleurs un outil important pour l'analyse de scène qui permet l'exploitation des maquettes numériques pour des applications nécessitant une représentation simplifiée et cohérente. Enfin, c'est un préalable pour la sémantisation riche (cf. 2.3.2) qui nécessite de disposer en entrée de polygones simples et des relations d'adjacences. Aussi, un formalisme a été développé pour répondre à ces problématiques. L'idée générale est résumée par la figure suivante :



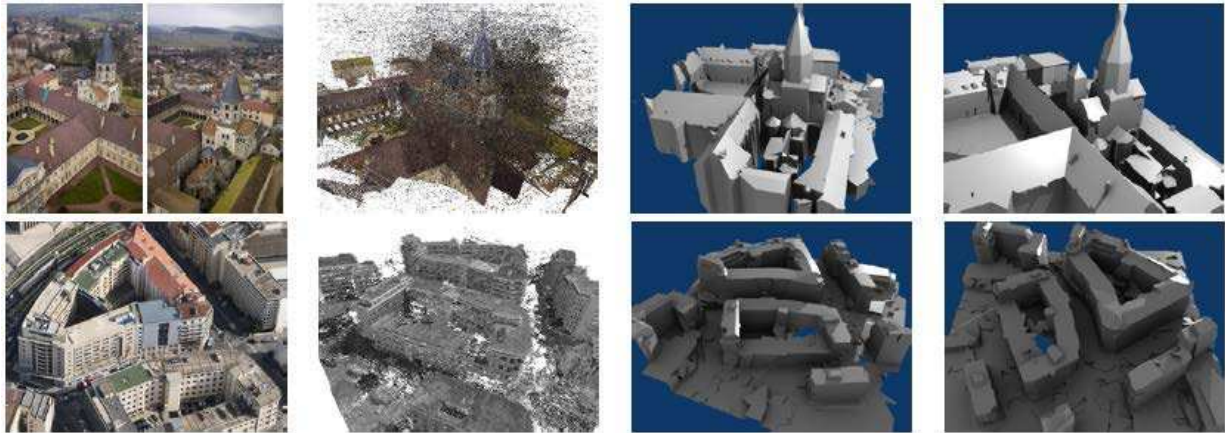
**Figure 12 - Les différentes étapes de la simplification planaire**

Cf. le corps du texte pour les détails

Lorsque l'on dispose d'une donnée d'entrée incomplète (ce qui est le cas lorsque l'on reconstruit la géométrie d'un bâtiment avec des images acquises d'un point de vue aérien, ou plus généralement dès lors qu'il y a des parties cachées), le modèle obtenu est troué. L'approche développée a consisté à obtenir des représentations planaires des environnements urbains tout en prenant en compte la complétion des parties cachées. Dans un premier temps, les primitives planaires sont détectées (b). L'inférence des parties manquantes est réalisée en créant aux extrémités des plans détectés des hypothèses de plans horizontaux et verticaux (c). Ces hypothèses, combinées aux détections permettent de construire une partition de l'espace via un arrangement 3D (d). La surface réelle est alors à rechercher dans cet arrangement comme celle délimitant au mieux l'objet (e). La recherche de cette surface est réalisée par un algorithme de *graph-cut*, appliqué sur le graphe dual de l'arrangement (chaque nœud du graphe est une cellule, chaque arête une relation d'adjacence entre cellules). La reconstruction finale est alors une surface complète, cohérente et étanche, ce qui est indispensable dans de nombreux cas.

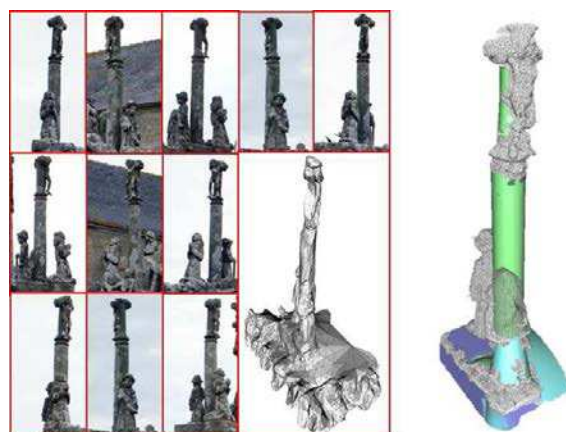
Cette méthodologie permet par ailleurs d'obtenir des modèles simples pour lesquels le nombre de triangles est drastiquement réduit par rapport à une

reconstruction photogrammétrique. Quelques résultats sur des données réelles sont présentés ci-dessous.



**Figure 13 - Simplification planaire d'un modèle 3D à partir d'images multivues. (De gauche à droite) Exemples d'images utilisées en entrée ; Le nuage de point reconstruit ; Résultats de la simplification planaire. (de haut en bas) Cluny ; Marseille**

Cette méthode très efficace contraint par sa formulation à n'utiliser que des primitives planaires. Une autre approche visant à étendre l'analyse à d'autres primitives (sphères, cônes, cylindres ...) a également été développée. Cette seconde approche permet par ailleurs de ne remplacer que certaines parties du maillage, et préserve donc les détails là où l'appariement avec des primitives simples engendrerait une simplification trop grande du maillage. Fondée sur une approche de géométrie stochastique, elle consiste à utiliser les courbures locales du maillage pour en déduire quel type de primitive peut s'accrocher sur un ensemble de points du maillage :



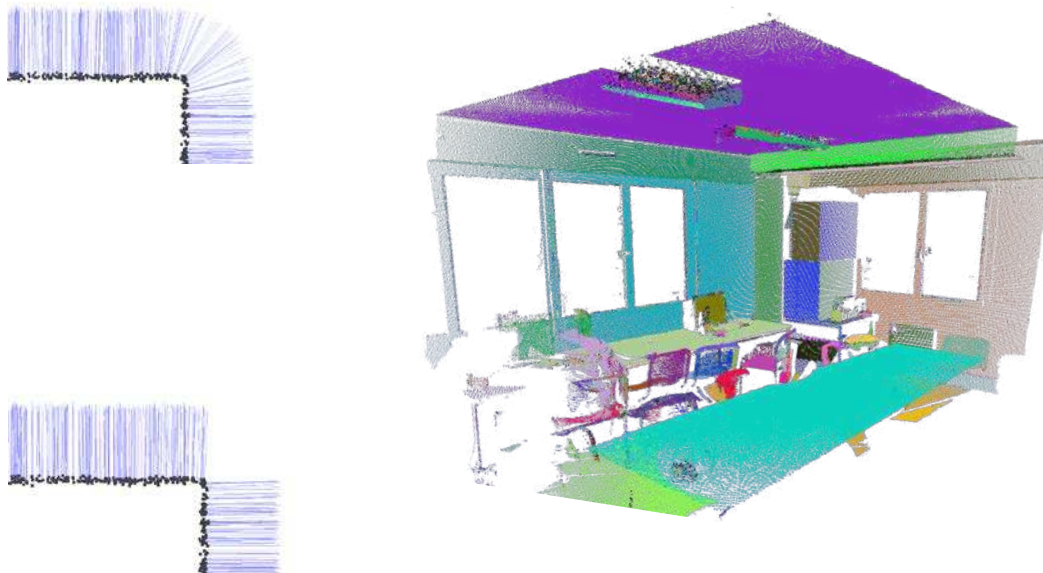
**Figure 14 - Simplification d'un maillage photogrammétrique par détection de primitives géométriques simples**

Une fois la détection effectuée, le maillage peut alors être simplifié en remplaçant les triangles par les primitives. Cela conduit à une représentation à la fois compacte et lissée de l'objet reconstruit.

### 2.3.1.2 Interprétation de nuages de points

Des travaux dans la même direction ont été menés, mais, pour cette fois travailler sur des scans laser qui permettent d'obtenir de façon directe une information 3D peu bruitée, extrêmement dense et sur tout type de surface (hormis les surfaces spéculaires, types vitres). De même que pour les travaux précédents, l'idée est d'obtenir ici une segmentation du nuage de points, localement cohérente, en vue de procéder à des traitements de plus haut niveau par la suite (cf. 2.3.2).

Dans un premier temps, il est nécessaire de calculer les normales en chaque point. Les approches de l'état de l'art sont soit rapides, soit robustes. Une approche combinant rapidité et robustesse (notamment aux discontinuités) a été développée, reposant sur une *Randomized Hough Transform*. Elle permet par ailleurs de gérer les problèmes d'anisotropie et ainsi d'obtenir des résultats précis en tous les points sur lesquels peuvent reposer la segmentation.



**Figure 15 – De gauche à droite et de haut en bas : Normales reconstruites sur un angle droit par minimisation par moindres carrés ; Normales obtenues par notre méthode ; Segmentation d'un nuage de points par cohérence locale des normales (chaque objet, quel que soit sa géométrie est extrait correctement : colonne, chanfrein, mobilier ...).**

L'ensemble de ces trois approches forme un corpus de méthodes applicables à différentes sources, à différents niveaux de qualité de la donnée d'entrée, et avec des objectifs variables en fonctions du niveau de représentation souhaité en sortie. Ces méthodes constituent par ailleurs un préalable à une sémantisation de haut niveau ayant pour but de classifier nommément les objets d'une scène.

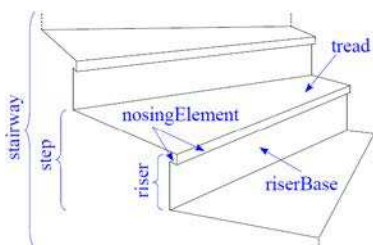
## 2.3.2 GRAMMAIRES ATTRIBUTAIRES CONTRAINTES

Dans cette optique de sémantisation de géométrie 3D, le pôle Imagine, partant de son expertise sur les grammaires, a développé une approche utilisant des grammaires attribuées contraintes. Ce type de formalisme est pour l'instant à même de travailler sur des modèles CAD, et est en cours d'extension à des données acquises par laser (dans premier temps) et à des données issues de reconstructions photogrammétriques (dans un second temps).

### 2.3.2.1 Interprétation sémantique de modèles CAD

Pour une première approche, le parti a été pris d'utiliser une géométrie d'entrée aussi « propre » que possible. Pour cela, des modèles IFC ont été utilisés après conversion en modèles purement géométriques (maillages de triangles). La sémantique contenue dans les modèles originaux a ainsi été enlevée, mais, est utilisée en sortie de l'algorithme pour valider l'approche et comparer les résultats obtenus avec une vérité terrain.

Pour illustrer le principe de la méthode, les figures suivantes présentent comment formaliser sous forme de grammaire attributive contrainte un escalier, aussi générique soit-il (taille des marches, tournant ou droit ...). Le principe consiste à décomposer l'objet escalier en ses différents composants : marches et contremarches, chaque marche pouvant ou non avoir un nez de marche.



**Figure 16 - Les différents éléments constitutifs d'un escalier**

La grammaire se contente ensuite de modéliser la façon dont sont assemblés les différents éléments : une marche est horizontale et perpendiculaire à une contremarche, une contremarche est adjacente par un côté à une marche et est verticale, un escalier est une séquence maximale de marches et contremarches,...

Les contraintes géométriques augmentent ensuite le modèle en réduisant les valeurs possibles des différents attributs : tailles minimales et maximales des marches...



```

// A riser base is a vertical polygon of constrained size.
riserBase b → polygon p
    ( vertical(p),
      0.1 <= p.breadth <= 0.3,
      0.5 <= p.length <= 2 )

// A nosing element is a thin polygon whose length direction is horizontal.
nosingElement e → polygon p
    ( horizontal(p.lengthVector),
      p.breadth <= 0.08,
      0.5 <= p.length <= 2 )

// A riser with optional nose is a riser base topped with a possibly empty sequence
// of nosing elements (see Figure 2).
riserOptNose ron → riserBase b, maxseq(nosingElement, edgeAdj) n
    ( edgeAdj(b, n),
      above(b, n) )

// A tread is a horizontal polygon of constrained size.
tread t → polygon p
    ( horizontal(p),
      0.1 <= p.breadth <= 2,
      0.5 <= p.length <= 2 )

// A riser topped with a tread underside and nose (see Figure 3) — useful for CAD models.
riserBottomTreadBase bt → riserBase b, tread t
    ( adj(b, t),
      above(b, t) )
riserBottomTreadNose rbtn → riserBottomTreadBase bt,
    maxseq(nosingElement, edgeAdj) n
    ( 1 <= n.size,
      edgeAdj(bt, n), above(bt, n) )

// A riser is any kind of riser.
riser r → riserOptNose ron
    | riserBottomTreadNose rbtn

// A step is a riser topped by a tread.
step s → riser r, tread t
    ( edgeAdj(r, t),
      above(r, t) )

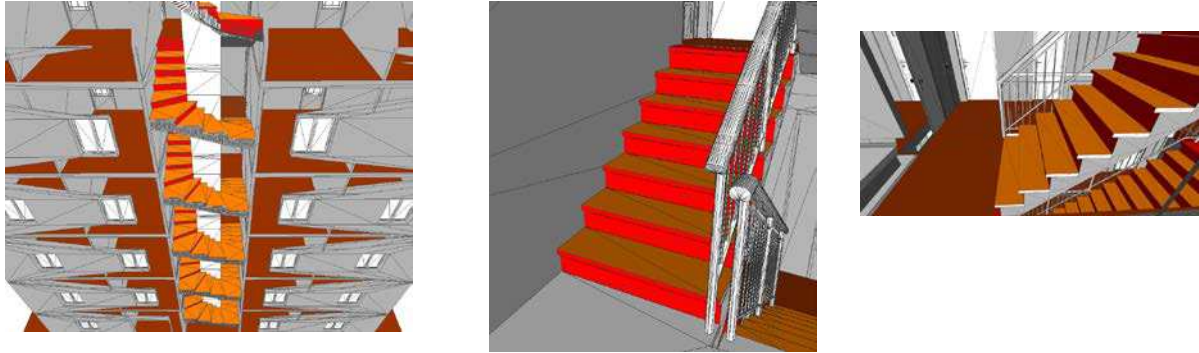
// A stairway is a sequence of steps topped by an optional riser.
stairway w → maxseq(step, edgeAdj) s, optional riser r
    ( edgeAdj(s, r),
      above(s, r) )

```

**Figure 17 - Modélisation d'un escalier par une grammaire attributive contrainte**

Pour appliquer les règles de la grammaire, l'algorithme repose sur des polygones en entrée<sup>4</sup>. Ils peuvent être obtenus par des approches comme celle présentées en 2.3.1, ou obtenus par un algorithme simple de fusion lorsque la géométrie est parfaite (fusion des faces adjacentes et coplanaires). Un graphe encode ensuite les relations de voisinages entre les polygones reconstruits. La recherche des objets définis dans la grammaire se résout comme un problème de recherche de sous-graphes. L'espace d'exploration est réduit par l'introduction progressive des contraintes (de la plus forte à la moins forte). Ainsi, l'algorithme se déroule en temps quasi linéaire.

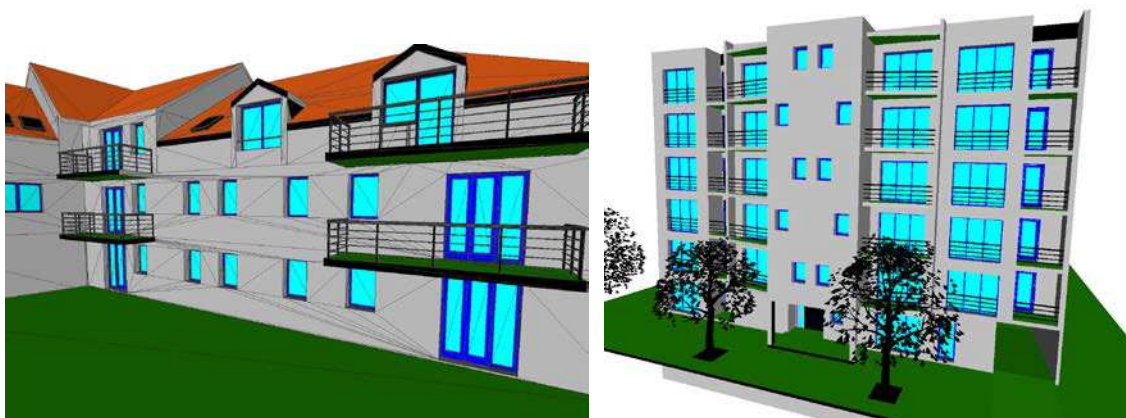
<sup>4</sup> C'est en quoi les méthodes présentées en 2.3.1 constituent un préalable requis pour mettre en œuvre cette approche.



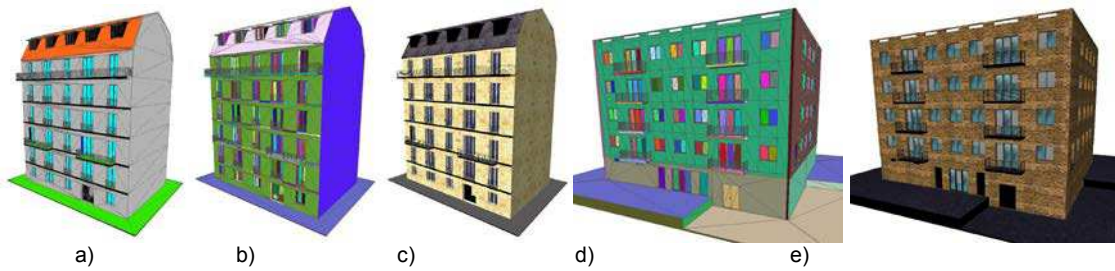
**Figure 18 - Résultats de la détection des escaliers dans des modèles CAO avec le formalisme de grammaires attribuées contraintes.**

Dans la grammaire définie pour traiter les modèles CAO, les objets détectables sont les toits, les murs, les dalles / sols, les ouvertures et les escaliers. Les figures ci-dessus et ci-dessous montrent des exemples de résultats détaillés (escaliers) et globaux sur l'ensemble de la maquette d'entrée. L'évaluation quantitative et qualitative montre que les résultats sont fiables et permettent de reconstruire la sémantique de la quasi-totalité des objets présents dans la scène (seuls les objets en noirs ne sont pas classifiés). Les objets non sémantisés sont ceux pour lesquels il n'existe pas de définition dans la grammaire (gouttières, fenêtres de toit et chiens assis, rebords de dalles ...) ou pour lesquels des erreurs de modélisation existent. Il est également à noter que la différence entre porte et fenêtre n'a pas pu être faite, notamment à cause de l'objet « hybride » porte-fenêtre qui crée une ambiguïté dans la définition.

La méthodologie ayant pu être validée, ces recherches sont à l'heure actuelle en cours d'extension afin de pouvoir traiter des géométries réelles, soit celles obtenues par scans laser ou reconstruction à partir d'images multivues. Pour traiter ce type de données, il est nécessaire d'adapter le modèle pour prendre en compte le bruit ou les reconstructions incomplètes, et une transposition immédiate de la méthode n'est pas possible.



**Figure 19 - Résultats de la sémantisation globale sur un modèle CAO avec le formalisme de grammaires attribuées contraintes**



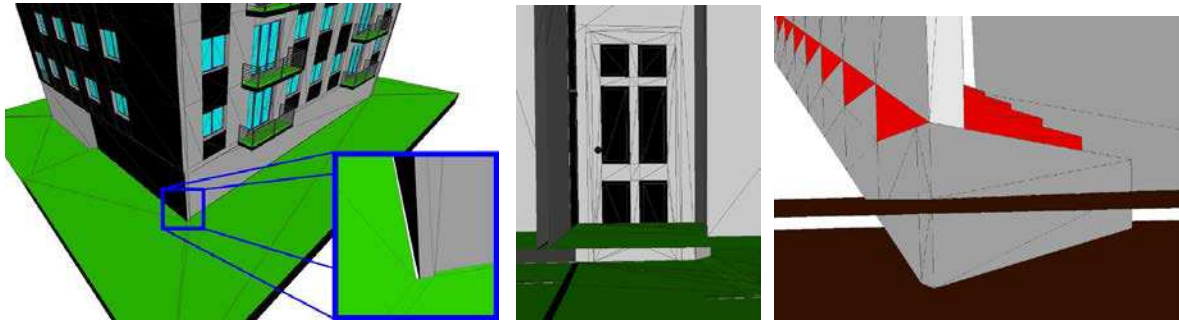
**Figure 20 – a) Résultat de sémantisation ; b) et d) Primitives extraites ; c) et e) Texturation automatique en fonction de la sémantisation**

### 2.3.2.2 Détection d'erreurs dans les modèles et contrôle qualité

Les recherches menées sur la sémantisation 3D, avec les approches par grammaires attributaires présentées ci-dessus, ont abouti à un effet collatéral et inattendu. Dans certains des modèles CAD utilisés, et selon la définition de la grammaire, certains objets n'étaient pas détectés, alors que leur modélisation géométrique aurait dû répondre aux attributs spécifiés. Ces sous-détections sont apparues lors de l'évaluation quantitative de la méthode, et ont, par exemple, été rencontrées pour certains murs, des baies ou des escaliers.

En se penchant sur ces erreurs, on a pu constater que le modèle utilisé n'était pas en cause, mais qu'elles provenaient de la modélisation géométrique. En effet, comme l'illustrent les images ci-dessus, lorsque l'on regarde de près la géométrie, des approximations non décelables au premier abord deviennent évidentes. Dans certains cas, comme dans celui d'un mur non connecté au sol, cela ne porte pas préjudice. D'autres cas sont cependant plus problématiques : portes rentrant dans le sol, ou première marche d'un escalier partiellement dans la dalle de son étage de départ. Dimensionner les objets en utilisant directement les informations de la maquette numérique aurait pu conduire à des conséquences rendant incohérente la construction.

Aussi, cette méthode de sémantisation par grammaires attributaires, en mettant en exergue ces erreurs de modélisation pourrait être envisagée pour effectuer du contrôle qualité des maquettes, et permettrait ainsi d'en détecter les incohérences.



**Figure 21 - Erreurs et incohérences dans les modèles CAD mis en évidence lors de l'étape de sémantisation.**

De gauche à droite : Le mur pignon n'est pas reconnu, car il n'est pas connecté directement au sol ; La porte n'est pas détectée car elle rentre dans le sol ; La première marche, et donc l'objet escalier dans sa totalité, n'est pas sémantisée comme telle puisqu'elle est partiellement incluse dans la dalle (les marches sont cependant détectées).

### 3. PERSPECTIVES

Les travaux de recherche menés au sein du pôle de recherche Imagine durant la période 2011-2013 font clairement apparaître une orientation vers les problématiques liées au bâtiment et à la ville. La reconstruction de géométrie tridimensionnelle, compétence majeure et reconnue d'Imagine, a été étendue pour être plus précise et pour permettre de traiter de grands volumes de données. Cela a permis d'ouvrir la voie à une industrialisation de la technologie, aujourd'hui valorisée dans la start-up Acute3D créée en 2011.

La simple reconstruction 3D n'étant plus à l'heure actuelle suffisante, les objectifs du laboratoire se sont en partie tournés vers la sémantisation et l'interprétation haut niveau de modèles 3D ou d'images. Il convient cependant de continuer l'effort dans ce domaine de la sémantisation, une partie des verrous n'étant pas levés à l'heure actuelle. La reconstruction 3D reste aussi un sujet pour lequel des avancées doivent être faites, notamment concernant la fusion de sources multiples de données (données aériennes et terrestre par exemple).

Enfin, il conviendra pour tirer parti de l'ensemble des récentes avancées d'intégrer les standards actuels aux processus afin de proposer des applicatifs métiers.

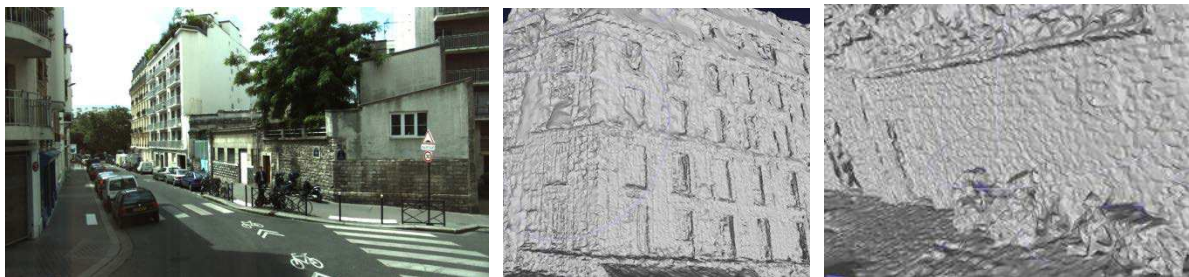
#### 3.1 RECONSTRUCTION TERRESTRE ET CONJOINTE AÉRIEN / TERRESTRE

La technologie actuellement utilisée pour la reconstruction de géométrie 3D à partir d'images multivues est extrêmement efficace à partir d'images aériennes. Cependant, avec le développement depuis quelques années des systèmes de cartographie mobile, des véhicules embarquant de nombreux capteurs, de nouvelles sources de données sont disponibles. Ainsi, pour acquérir des images depuis le sol, il n'est plus nécessaire de se déplacer pour faire soi-même les

images, ce qui permet d'acquérir de vastes zones d'images, de la rue à la ville, en passant par le quartier et l'îlot. Ces systèmes disposent par ailleurs de capteurs de positionnement qui offrent ainsi des images géoréférencées. Il est à noter que les coûts d'acquisition sont bien inférieurs à ceux d'une campagne aérienne.

A l'échelle du bâtiment, la reconstruction est relativement bien maîtrisée, dès lors que l'on dispose de suffisamment d'images, présentant des variations d'angles faibles. En utilisant les images issues de systèmes de cartographie mobile, ces contraintes ne sont pas respectées, et les reconstructions obtenues sont relativement bruitées (cf. Figure 22). Par ailleurs, quelle que soit la configuration de prise de vue, les surfaces spéculaires (vitres) sont mal reconstruites, et, les objets mobiles (piétons, véhicules, ...) posent des problèmes.

Deux pistes principales pour l'amélioration des résultats sont à envisager : d'une part, l'amélioration de la calibration, et, d'autre part, la prise en compte de contraintes spécifiques pour la reconstruction.

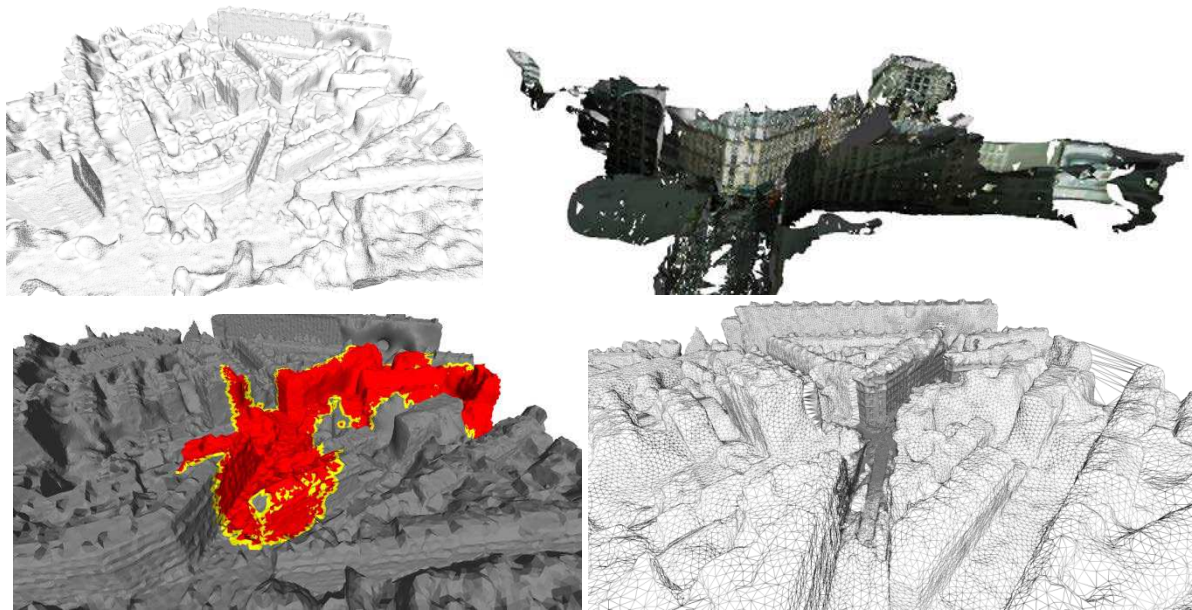


**Figure 22 - De gauche à droite : Une image acquise par *Stéréopolis*, système de cartographie mobile développé au laboratoire MATIS de l'IGN ; Détails sur des reconstructions 3D obtenues à partir de ces images**

En particulier, faire des hypothèses de planarité par morceaux de la scène semble raisonnable et prometteur. Il apparaît en effet plus simple (et pertinent) de détecter des plans *a priori* plutôt que de chercher à les repérer en post-traitement de la reconstruction. L'intégration de résultats de sémantisation serait aussi de nature à aider à la construction de ces hypothèses et aiderait également à contraindre le problème de la reconstruction, notamment sur les surfaces vitrées. Une telle reconstruction ouvrirait la voie à une modélisation simulable de la ville.

Enfin, il est important de préciser que les reconstructions obtenues à partir de données aériennes et terrestres sont complémentaires : depuis le sol, une plus grande résolution est possible, et permet de voir précisément les façades ; depuis les images aériennes, les toits et cours intérieures sont visibles. Des reconstructions issues de ces deux sources de données permettraient une fusion des sources, et ce, afin d'obtenir à grande échelle une modélisation détaillée, avec une précision dépendant des objets et de leur points d'observation. Des premiers tests ont été menés en ce sens (cf. Figure 23) et méritent d'être étendus afin de traiter les problèmes spécifiques qu'ils ont mis à jour : recalage

entre des données de résolution et points très différents, fusion, choix des meilleures surfaces locales ...



**Figure 23 – En haut : reconstruction à partir d’images aériennes à 10cm (gauche) et terrestres (droite). En bas : zone commune aux deux reconstructions (gauche) et fusion des deux maillages (droite)**

### 3.2 SÉMANTISATION 3D

La sémantisation reste l’un des enjeux majeurs des prochaines années. De nombreuses avancées ont été réalisées, mais, concernant la sémantisation 3D, de nombreux points restent à traiter pour parvenir à la construction, automatique, d’une maquette numérique exploitable. Il est important de garder à l’esprit que l’objectif ambitieux est de pouvoir sémantiser des modèles obtenus par les techniques de reconstruction 3D à partir d’images.

De ce fait, plusieurs pistes progressives sont envisageables. Elles permettront de lever un à un les verrous, tout en offrant à chaque étape des réponses à des besoins spécifiques. Dans un premier temps, il apparaît pertinent de passer de la sémantisation des modèles CAO (cf. 2.3.2.1) à la sémantisation de nuages de points issus de relevés laser. En effet, dans le cas d’intérieurs, c’est la seule technique qui permet aujourd’hui d’obtenir de la 3D. Par ailleurs, les données sont relativement peu bruitées, ce qui peut permettre de transposer la technique développée pour les modèles CAO. Ainsi, il pourrait être envisagé d’utiliser une grammaire attribuée contrainte sur les primitives détectées dans le nuage de points (cf. 2.3.1.2). Une première passe de sémantisation permettrait de détecter les objets encodés dans la grammaire : murs, dalles, baies. Le mobilier serait ainsi éliminé. La construction d’un arrangement tridimensionnel avec les primitives restantes permettrait ensuite, via une phase d’optimisation, de retrouver l’ensemble de la surface composant la structure du bâtiment, et ainsi

de compléter les parties cachées. Cette étape pourrait s'appuyer sur les travaux déjà réalisés (cf. 2.3.1.1).

Dans un second temps, l'intégration de la sémantisation obtenue à partir d'images pourrait être efficacement utilisée dans les reconstructions afin d'apporter des contraintes. Cela permettrait, d'une part, de guider le processus de reconstruction, et, d'autre part, de le contraindre. En effet, le niveau de sémantisation accessible à l'heure actuelle permet d'extraire les éléments architecturaux principaux de la façade. Cela donne une indication sur les zones que les algorithmes de reconstruction ne peuvent traiter de façon robuste, et également sur des zones de discontinuité géométrique. Par exemple, intégrer dans les algorithmes des connaissances *a priori* sur la position des surfaces vitrées éviterait d'obtenir des reconstructions bruitées. Cela passerait bien entendu par une étape préalable de sémantisation multi-images, puis de mise en correspondance de ces sémantisations. Il pourrait également être pertinent d'utiliser ce contexte multivues pour la sémantisation, en utilisant l'information de l'ensemble des images pour en accroître la robustesse. Au final, l'objectif serait d'automatiser la sémantisation conjointement avec la reconstruction, l'un guidant l'autre, à travers un processus itératif de raffinement.

Les objectifs précédents s'entendent au niveau du bâtiment isolé. Afin de pouvoir disposer de maquettes numériques à l'échelle urbaine, il est indispensable de savoir sémantiser à cette échelle. La donnée la plus commune étant le maillage de l'ensemble de l'espace urbain, c'est vers des méthodologies aptes à traiter ce type de données qu'il convient de s'orienter. Les méthodes à base de grammaire ayant des difficultés à passer à l'échelle, s'orienter sur des approches d'apprentissage statistique semble pertinent. Dans un premier temps, l'intérêt pourrait se porter sur les objets les plus communs du paysage urbain : bâtiments, sols, routes et végétation. Une base de données d'apprentissage pourra être construite en utilisant des bases de données vectorielles externes, type BD Topo de l'IGN, qu'il faudra aligner avec le modèle. Il faudra pour la construction de cette base être vigilant aux écarts entre les données vectorielles et géométriques (généralisation ou incohérences temporelles : cf. Figure 24).



**Figure 24 – Fusion des emprises de bâtiments issues de la BD Topo sur un modèles 3D (Mandelieu).**

Noter les bâtiments absents de la base de données vectorielle et l'imprécision de la géométrie causant des problèmes de recalage.

Une voie prometteuse pour cette approche consisterait à utiliser un modèle graphique, modèle permettant d'encoder les informations et les relations locales entre des éléments de la scène : géométrie, photométrie, voisinages ... De ce point de vue, la sémantisation se définit comme un problème d'optimisation discret.

### **3.3 INTÉGRATION DE BASES DE DONNÉES EXTERNES**

L'intégration des bases de données externes est un véritable enjeu. En effet, ces sources de données apportent une information complémentaire, qui parfois peut être extrêmement riche et bénéfique pour la modélisation. Outre les bases de données vectorielles de l'IGN, des bases de données plus traditionnelles (« textuelles ») sont disponibles. En particulier, la base de données Majic de la DGFIP<sup>5</sup> qui décrit très précisément les ensembles bâtis. A partir des informations qu'elle contient, plusieurs perspectives sont ouvertes. En effet, bien qu'un certain nombre de données soient probablement obsolètes (nombre de salles de bain et points d'eau par appartement par exemple), d'autres sont plus pérennes : matériaux des murs et toitures, nombre d'étages, nombre d'appartements, année de construction, ... A partir de ces informations, il est possible de dériver des hypothèses qui peuvent être utilisées dans les différents processus, de la reconstruction à la sémantisation. Ce sont également des informations riches pour des applications métier (cf. 3.5).

Par exemple, pour la sémantisation d'images avec des grammaires, nous avons vu qu'il était nécessaire d'instancier la grammaire en fonction d'un type architectural. La connaissance de l'année de construction peut permettre d'avoir une première idée de ce style. Couplée aux matériaux des façades et au nombre d'étages, il serait également possible d'en déduire un système constructif, ce qui pourrait automatiser une partie du processus de sélection de la grammaire adaptée. Pour permettre d'exploiter la richesse des données et obtenir des solutions robustes aux erreurs et fiables (prise en compte de la non mise à jour), des approches par apprentissage statistiques, telles que celles qui pourraient être mises en place pour la sémantisation 3D (cf. 3.2) semblent pertinentes. Une telle formalisation serait à même d'intégrer des données venant de sources variées, accroissant ainsi la connaissance de l'environnement.

### **3.4 INTÉGRATION ET GESTION DES STANDARDS ET MODÈLES DE DONNÉES**

Pour pouvoir être interopérable, les modèles produits se doivent de répondre à des formats et modèles standards et reconnus. Aussi, il est important de garder à l'esprit cette contrainte, qui permet de construire, dans un référentiel admis par la communauté visée, les résultats des algorithmes. L'adéquation des données produites aux standards est particulièrement importante pour les applications métiers qui peuvent intégrer des données issues de différents intervenants. Il est donc primordial de disposer d'un « langage » commun que chacun soit à même de comprendre et d'interpréter.

---

<sup>5</sup> Direction générale des finances publiques



A l'échelle du bâtiment, le standard aujourd'hui admis est l'IFC. A l'heure actuelle, les résultats des algorithmes ne sont pas formalisés dans ce format. Une attention particulière devra être portée sur cette action, d'autant plus que les résultats de sémantisation arriveront bientôt à maturité.

A l'échelle urbaine, le pendant de l'IFC est le CityGML (cf. projet de recherche *Ville numérique* – Tâche 2 – Données et couplages moteurs externes – CityGML). Il s'agit à la fois d'un format d'échange et d'un modèle de données, ce qui lui donne tout son avantage : il est possible de se reposer sur le modèle pour bâtir des applications, tout en conservant l'interopérabilité. Cela permet d'intégrer tout type de sources et de les faire travailler ensemble, ce qui est la plupart du temps nécessaire à l'échelle urbaine. Aussi, la sémantisation à l'échelle urbaine est nécessaire pour pouvoir structurer les données dans ce formalisme.

### **3.5 APPLICATIONS MÉTIERS**

Le développement d'applicatifs métiers utilisant les résultats issus de ces recherches est un point fondamental. En effet, c'est à la fois un guide qui oriente les besoins (en tant que consommateur de données), et un objectif (en tant que recherche opérationnelle). Les applications de ces recherches sont variées et nombreuses. Aussi, ne sont pointés ci-dessous que quelques exemples pour amorcer la réflexion.

L'une des premières applications visée concerne l'étude du potentiel photovoltaïque. Un modèle pour ce type d'analyse nécessite en entrée une géométrie 3D afin de calculer les masques. Elle peut également tirer parti des résultats de l'analyse sémantique, et doit l'intégrer pour être complète. En effet, la définition des surfaces d'intérêt sur lesquelles, par exemple, il est possible de disposer des panneaux nécessite ce niveau d'information. L'étude d'impact d'un nouveau projet est aussi un cas très intéressant qui permettrait d'enrichir et d'automatiser les études actuellement menées. Par exemple, si l'ensemble de l'environnement (3D) autour d'un projet de construction est sémantisé (la position des baies peut dans un premier temps suffire), il sera possible d'obtenir pour chaque bâtiment des données quantitatives, par exemple, sur l'ombrage causé par le nouveau projet. La sémantisation s'avèrerait ici intéressante puisqu'elle donnerait des informations fines et détaillées : plutôt que de spécifier simplement quels seront les bâtiments impactés, il deviendrait possible de qualifier et de quantifier l'impact du projet sur chacun des bâtiments : telle baie, de tel étage, de tel bâtiment, sur telle période temporelle.

Le diagnostic thermique automatisé semble également être une application prometteuse. Réaliser un diagnostic précis et complet apparaît difficile, mais, obtenir un premier aperçu de l'efficacité énergétique des bâtiments paraît réaliste. Pour y parvenir, il s'agirait d'acquérir depuis un véhicule des données images thermiques sur les bâtiments. Ces données, croisées avec des bases de données sur la géométrie des bâtiments et sur leurs attributs (position des baies, matériaux, nombre de logements, système constructif, moyen de chauffage, ...) permettrait de qualifier les bâtiments, si ce n'est en valeur absolue, au moins en valeur relative. L'efficacité thermique d'un bâtiment pourrait alors être classée par rapport à sa rue, une rue par rapport à un quartier, et un quartier par

rapport à une ville. Une telle approche fournirait des éléments de décision pour orienter des travaux de rénovation, en focalisant les efforts sur des bâtiments ou ensembles spécifiques, ou, pour un bâtiment donné, sur une intervention particulière : remplacement des fenêtres, isolation, ... Un tel outil pourrait être un support inédit d'aide à la décision à destination des autorités locales ou des producteurs d'énergie. Pour compléter cette étude, la reconstruction des environnements intérieurs est une piste à suivre. En effet, une simulation thermique correcte ne saurait s'appuyer uniquement sur des données de l'enveloppe du bâtiment. Aussi, étudier par inférence les liens existants entre l'intérieur et l'extérieur du bâtiment constitue un objet d'étude pertinent.

Ces recherches pourraient être menées en collaboration avec d'autres organismes, notamment le LNE ou l'IGN. Concernant ce dernier, c'est d'ailleurs l'un des axes de collaboration avec Imagine qui a été identifié.

## **CONCLUSION**

L'ensemble des résultats de recherches présenté dans ce rapport forme un tout cohérent. Les différentes sources de données, l'interprétation des images et de la géométrie 3D, les approches grammaire et l'apprentissage, sont en effet complémentaires. Pour relever le défi des années à venir de la reconstruction totalement automatique de maquettes numériques sémantisées, la prise en compte de l'ensemble des informations semble incontournable. Les enjeux sont très importants et répondent à des exigences formulées par les industriels du bâtiment et de la construction, investis notamment dans la rénovation du parc des bâtiments anciens dans une perspective de développement durable en accord avec les objectifs du Grenelle de l'environnement. Hormis les questions de sémantique pure, de nombreux défis subsistent, en calibration, visions par ordinateur et apprentissage statistique. Il conviendra d'étudier de manière concourante ces domaines afin d'aboutir à une approche cohérente répondant aux objectifs définis. Le développement de ces objectifs passera par des collaborations académiques et industrielles.

De nombreuses applications telles que celles présentées ci-dessus sont envisageables. Afin de valoriser les résultats de la recherche, il est important de consacrer une part non négligeable des moyens pour les développer et chercher les moyens de les déployer efficacement. Ces recherches, ainsi que ses applications opérationnelles auront tout intérêt à être capitalisées, en interne, dans une plateforme commune. eveCity3 est apte à héberger les résultats des algorithmes, et, grâce à son système modulaire, permet de les utiliser dans le cadre de simulations et d'études interopérables. Outre l'aspect de mise en commun des résultats, une telle intégration fournirait un support réaliste et opérationnel de navigation dans un espace virtuel. Son développement et ses liens étroits avec les équipes d'expertise est aussi un atout puisque cela permet de faire remonter des besoins en relations avec les usages recherchés.

Par ailleurs, le développement de ces points permettra d'enrichir l'offre Decision3D, consortium regroupant le CSTB, Acute3d et Technicom et qui a pour objectif de réaliser, à destination des collectivités locales et des agglomérations,

une maquette numérique de référence, en assurant toutes les phases de production. Enrichir cette offre de possibilités d'applications métiers ajouterait au support numérique des outils de grand intérêt pour l'ensemble des acteurs de la ville.

## **PUBLICATIONS**

### **ARTICLES DE REVUE AVEC COMITÉ DE LECTURE ET D'AUDIENCE INTERNATIONALE**

C. Allène, J.Y. Audibert, M. Courpie, R. Keriven, Some links between extremum spanning forests, watersheds and min-cuts, *Image and Vision Computing*, 28(10): 1460-1471, 2010

A. Boulch, S. Houillier, R. Marlet, O. Tournaire, Semantizing Complex 3D Scenes using Constrained Attributes Grammars, *Computer Graphics Forum*, 32(5): 33-42, 2013

A. Boulch, R. Marlet, Fast Normals Estimation for Point Clouds with Sharp Features using a Robust Randomized Hough Transform, *Computer Graphics Forum*, 31(5): 1765-1774, 2012

P. Labatut, J.P. Pons, R. Keriven, Robust and Efficient Surface Reconstruction from Range Data, *Computer Graphics Forum*, 28(8): 2275-2290, 2009

F. Lafarge, R. Keriven, M. Brédif, Insertion of 3D-Primitives in Mesh-Based Representation: Towards Compact Models Preserving the Details, *IEEE Transactions on Image Processing*, 19(7): 1683-1694, 2010

G. Peyré, M. Péchaud, R. Keriven, L. Cohen, Geodesic Methods in Computer Vision and Graphics, *Foundations and Trends in Computer Graphics and Vision*, 5(3-4): 197-397, 2010

H.H. Vu, P. Labattut, J.P. Pons, R. Keriven : High Accuracy and Visibility-Consistent Dense Multiview Stereo, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(5): 889-901, 2012

### **CONFÉRENCES AVEC ACTES**

A. Chariot, R. Keriven, GPU-boosted online image matching, *ICPR - 19<sup>th</sup> International conference on Pattern Recognition*, p. 1-4, 2008

A.L. Chauve, P. Labatut, J.P. Pons, Robust piecewise-planar 3D reconstruction and completion from large-scale unstructured point data, *CVPR 2010, 23<sup>rd</sup> IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, p. 1261-1268, 2010

A.L. Chauve, J.P. Pons, J.Y. Audibert, R. Keriven, Transductive Segmentation of Textured Meshes, *9th Asian conference on Computer Vision (ACCV 2009), Lecture notes in Computer Science*, vol. 5995, p. 1261-513, 2010

A. Dalalyan, R. Keriven, Sparse learning approach to the problem of robust estimation of camera locations. *ICCV - IEEE 12<sup>th</sup> International Conference on Computer Vision Workshops*, p. 436-443, 2009

Z. Jank, J.P. Pons, Spatio-Temporal Image-Based Texture Atlases for Dynamic 3D Models, *IEEE International Workshop on 3D Digital Imaging and Modeling*, 2009

P. Labatut, J.P. Pons, R. Keriven, Hierarchical shape-based surface reconstruction for dense multi-view stereo, *IEEE 12<sup>th</sup> International Conference on Computer Vision Workshops (ICCV Workshops)*, p. 1598-1605, 2009

F. Lafarge, R. Keriven, M. Brédif, H.H. Vu, Hybrid multi-view reconstruction by Jump-Diffusion, *CVPR 2010, 23<sup>rd</sup> IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, p. 350-357, 2010

F. Lafarge, R. Keriven, M. Brédif, Combining meshes and geometric primitives for accurate and semantic modeling., *BMVC – British Machine Vision Conference*, p. 38.1-38.11, 2009

P. Moulon, P. Monasse, R. Marlet, Adaptive Structure from Motion with a contrario model estimation. *ACCV 2012 – Proceedings of the 11th Asian Conference on Computer Vision, Lectures Notes in Computer Science*, vol. 7727, p. 257-270, 2012

D. Ok, M. Kozinski, R. Marlet, N. Paragios, High-Level Bottom-Up Cues for Top-Down Parsing of Façade Images, *3DIMPVT – Proceedings of the 2<sup>nd</sup> Joint 3DIM/3DPVT Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission*, 2012

D. OK, R. Marlet, J.Y. Audibert, Efficient and Scalable 4<sup>th</sup> Order Match Propagation. *ACCV 2012 – Proceedings of the 11<sup>th</sup> Asian conference on Computer Vision, Lectures Notes on Computer Science*, vol. 7724, p. 460-473, 2012

J.P. Pons, S. Soubra, 3D Reconstruction of large scale city models as a support to sustainable development. *Cooperative Design, Visualization and Engineering*, p. 172-175, 2008

O. Tournaire, J.P. Pons, R. Keriven, S. Soubra, Modélisation 3D Urbaine Automatique par Photogrammétrie Haute Résolution. *Colloque du GIS Modélisation Urbaine « La Modélisation de la Ville : du Modèle au Projet Urbain »*, 2011

H.H. Vu, R. Keriven, P. Labatut, J.P. Pons, Towards a high-resolution large-scale multi-view stereo. *CVPR – Proceedings of Conference on Computer Vision and Pattern Recognition*, p. 1430-1437, 2009

## **THÈSES**

D. Ok, *Mise en correspondance robuste et détection d'éléments visuel appliquées à l'analyse de façades*. Thèse de doctorat, Université Paris-Est, 2011

H.H. Vu, *Large-scale and high-quality multi-view stereo*. Thèse de doctorat, Université Paris-Est, 2011

**SIEGE SOCIAL**

84, AVENUE JEAN JAURÈS | CHAMPS-SUR-MARNE | 77447 MARNE-LA-VALLÉE CEDEX 2  
TÉL. (33) 01 64 68 82 82 | FAX (33) 01 60 05 70 37 | [www.cstb.fr](http://www.cstb.fr)

**CSTB**  
*le futur en construction*

CENTRE SCIENTIFIQUE ET TECHNIQUE DU BATIMENT | MARNE-LA-VALLÉE | PARIS | GRENOBLE | NANTES | SOPHIA-ANTIPOLIS