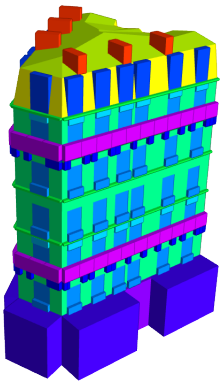
PhD Thesis Proposal at CVC/IMAGINE (2012)

# Shape grammar learning and application to building model reconstruction


## Summary

Techniques are being developed for the automatic production of semantized model (e.g., building information models) from real data (e.g., photographs and/or laser scans). One of the most promising techniques is based primarily on shape grammars. These grammars currently have to be handwritten by experts, which has a number of disadvantages: long specification time, incompleteness and inaccuracy. The aim of this thesis is to study the automatic and semi-automatic learning of grammars from examples. This will reduce the need for experts to manually design grammars. In the case of buildings, it will help in rapidly constituting a database of grammars capable of handling a wide variety of building types.

## Context

Building information modeling is revolutionizing the practices of the construction industry: precise cost evaluation, thermal performance simulation and optimization, virtual presentation of solutions to the public and policy makers, planning and monitoring of work, etc. The stakes are huge for the renovation market, particularly with regard to energy efficiency (often coupled to acoustic comfort). But most buildings do not have models, in particular older buildings, even though they have the greatest need for renovation and though they represent the vast majority of the housing stock.

To avoid the expensive manual creation of these digital models from measurements made on site, techniques are being developed for an automatic reconstruction of 3D scenes from photographs and/or laser scans, followed by regularization and semantization stages. In particular, techniques based on shape grammars, in this case building grammars, are particularly promising.

However these grammars, that are specific to given architectural types, must be presently handwritten by experts, which can be lengthy and costly. Furthermore, although careful manual specifications are a guarantee of quality for some structural aspects, it can also be a limitation. For example, even if an expert knows typical window sizes and spacings, s/he cannot provide a reliable estimate of the actual statistical distribution of these sizes and spacings, for a given type of buildings. Yet this is a valuable information to be used by a grammar-based facade semantizer. Similarly, some information may be missing from the input data, for instance regarding the roof and what is invisible from the ground, where picture shooting and laser scanning are usually made. Numerical clues that could help filling these information gaps are difficult for the expert to express.

## Objective

The goal of the proposed thesis is to study the automatic and semi-automatic learning of grammars from examples. This will reduce the need for experts to manually design grammars, and thus help in rapidly constituting a database of grammars capable of handling a wide variety of object types. Learning will address different aspects of a grammar:

- the choice of structural rules, which break down a complex element in more simple sub-elements that are related,
- the statistical distribution of values that parameterize the terminal and non-terminal elements of a grammar (e.g., geometric information, photometric clues),
- the statistical distribution of the relationships between grammatical elements, which can predict the nature and parameters of invisible elements in the data (e.g., roofs).

## Work program

The work will consist first in comparing the kind of grammar rules and parameters that are presently used in existing grammars. A database will be constructed with more or less structured and correlated examples: photographs, laser scans, models, high-level information (e.g., for a building, date of construction and number of floors, as provided by the property tax files), etc.

Then machine-learning techniques will be developed, adapted to the problem. Several avenues will be considered such the specialization of a generic grammar with respect to an observed structural regularity, or inductive logic programming. These techniques will be validated using the constructed database. This study will be linked to ongoing work on grammar-based semantization in the IMAGINE group (http://imagine.enpc.fr/).

Expected results, besides publications, include a prototype for learning grammars and a database linking measurements on buildings (photos, laser scans) and high-level information (architectural information, model), which could also be reused in other contexts. These results will help evaluate the contribution of machine learning for grammar definition, compared handwritten specification. (Data and techniques developed in this thesis could also be the basis of another line of work on the automatic inference of architectural and structural information of a building: time, style, techniques, probable materials, etc.)

## Profile and application

Required skills:

- MSc in computer vision and machine learning,
- good background in mathematical methods for numerical optimization
- proficiency in C++.

To apply, please email:

- your CV,
- a transcript of your MSc grades/marks (even if incomplete),
- the report you wrote for your MSc thesis or for a previous internship,
- reference letters of previous supervisors or teachers,
- a brief description of your research interests highlighting the links between your education/training/experience and the thesis topic,

to

- Pr. Nikos Paragios (nikos.paragios@enpc.fr) and
- Dr. Renaud Marlet (renaud.marlet@enpc.fr) and
- Dr. Olivier Tournaire (olivier.tournaire@cstb.fr).

## The IMAGINE Group

The IMAGINE Group is a joint project of the École des Ponts ParisTech (ENPC) and the French Scientific and Technical Centre for Building (CSTB), now part of the Center for Visual Computing (CVC), in association with the École Centrale de Paris (ECP), and it is part as well of the Computer Science lab (LIGM) of University Paris Est (UPE).

The domains of research of IMAGINE are computer vision, mesh processing, machine learning, statistics, optimization and constraint programming. In particular, IMAGINE has been working for several years on dense multi-view stereovision. One of the main focuses of the group has been on high precision 3D surface reconstruction from images, targeting large-scale data sets taken under uncontrolled conditions. Part of the group now also works on the interpretation and semantization of images and 3D models.

IMAGINE has got the best results worldwide on the Strecha et al. reference benchmark, with the most complete and the most precise reconstructions. This expertise and software has been transferred in 2011 to the startup company Acute3D, powering Autodesk's 123D Catch (formerly project Photofly), a web service to create 3D models from photographs. In November 2011, IMAGINE also won the PRoVisG Mars 3D Challenge, consisting in image calibration and 3D model reconstruction tasks on Mars surface.