PhD Thesis Proposal at CVC/IMAGINE (2012)

Global modeling of geometry and semantics, and application to the reconstruction of technical equipment

Context

Building information models (BIM) and 3D technologies are revolutionizing the construction domain. However, 3D data acquisition techniques (laser scans, photogrammetry) have a number of limitations:

- The data consists of points or mesh surfaces representing the observed scene, not abstract geometric models labeled with high-level semantics.

- The data is often partial due to areas that are difficult to access or unprofitable to complete (rear of pipes or columns), or because of self-occlusion (portion of an object masking another one).

- The data can be noisy or even erroneous due to acquisition or reconstruction errors. These defects are of different orders depending on the techniques but none is immune. Errors occur for instance with lasers as well as cameras because of reflections on stainless steel pipes or glass surfaces.
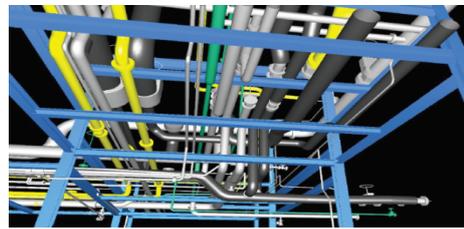
Techniques such as shape grammars are being studied to automate the reconstruction of semantized geometric models, including for noisy or incomplete data. But parsing algorithms generally have a fragmented view of the problem, which is treated as a succession of tasks rather than as a whole. This may be sufficient to treat the envelope of a building, where the forms are relatively few, simple and linked with uncomplicated combinations, but it is not well suited for complex technical equipment, e.g., a pipe network.

Objective

The objective of the proposed thesis is to study a new framework for the automatic interpretation of complex 3D scenes. The hypothesis is that the 3D data can be approximated by a combination of geometric primitives (cylinders, cubes, spheres, etc.) that are constrained by the type of the observed scene, given domain-specific rules. The idea here is to jointly address the recognition of geometric primitives and their optimal combination in accordance with the semantic composition rules. The originality lies in the global and simultaneous processing of geometric and semantic estimation of the entire scene. Solving this large-scale optimization problem on discrete and continuous variables can borrow some techniques used for shape grammar parsing, including "top-down" statistical inference as well as "bottom-up" detection-based analysis. Expected benefits are an improved treatment of complex scene and a greater flexibility thanks to an easy adaptation of the primitive dictionary to handle new types of scenes or objects.

The research will be guided by a case study concerning the reconstruction of the technical equipment in a building.

Profile and application

Required skills:

- MSc in computer vision,
- excellent background in mathematical methods for numerical optimization,
- proficiency in C++,
- some knowledge in computational geometry.

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).

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.