Generation of Simple 3D Building Models from a Template and Possibly Inconsistent Photos

Master 2 Internship at Dassault Systèmes Provence,
in collaboration with Inria Sophia Antipolis - Méditerranée

Abstract. *The objective of this Master internship is the design and implementation of an algorithm for the semi-automated reconstruction of textured 3D building models. The algorithm will take as input a series of street view photos of a building, shot by professional photographers or drones. It will generate as output a 3D model of the building, with high resolution textures derived from the photos and conforming to the LOD1 level of detail of the CityGML open data model (see figure below). This internship should be considered as the first steps of a Ph.D. thesis.*

Figure. *From a series of multi-view photos of a building and coarse camera calibrations (left), the objective is to generate a simple 3D model with high resolution texture (right). The output textured model should be photo-consistent.*

Context. The recent technological advances make it possible to generate in routine large series of overlapping high resolution photos of buildings, either from street-view photographers or from instrumented unmanned aerial vehicles (drones), with high coverage and capability to overcome occlusion and poor vantage points. Despite these advances and mature approaches for multi-view stereo, the automated reconstruction of 3D building models from high resolution photos (referred to as photo-modeling) is still a scientific challenge, due to (1) imperfect camera calibration, i.e. uncertain recovery of the 3D location and orientation of the cameras, (2) inconsistent lightings between the views, (3) repeated patterns that are common on building facades, and (4) reflections owing to highly reflective surfaces such as glass or polished metal.

Imperfect calibration translates into uncertainties when matching pixels of overlapping images. *Inconsistent lightings* also challenge the matching processes and hamper the direct generation of textures from the input photos. *Repeated patterns* further challenge the matching processes and often yield outliers, i.e. erroneous 3D geometry in the reconstructed surfaces. *Reflections*, be they spurious due to specular surfaces, or large scale due to reflective glasses or mirrors, are sources of complex optical illusions (seeing behind

walls) that turn into highly inconsistent matchings. Similar types of hurdles occur for transparent surfaces.

Objective. The objective of the internship is to devise a template-based approach, restricted to generate *simple* 3D models in the form of 3D right prisms, i.e. n-sided polygons extruded in the vertical direction. Each polygon, corresponding to the *coarse* 2D footprint of a building, is assumed to be given as input and either selected in a library of template models [1] or generated via procedural modeling [3]. The photos given as input are assumed to be already geo-localized and coarsely calibrated using an existing multi-view stereo algorithm.

Given as input an initial template 2D building model and the set of calibrated photos, the goal is to find the corresponding textured 3D model that best matches some multi-view *photo-consistency* criteria. A model is considered photo-consistent when its color at all points appears to be similar to all the cameras that can see it.

Challenges:

- Devise a method with high resilience to potentially inconsistent input photos, with changes in exposure, shadows, reflections, transparent surfaces and even presence of transient outliers such as cars.
- Devise a numerical method that optimizes for photo-consistency while dealing with inaccurate models that are restricted to be 3D right prisms, i.e. LOD1 levels of detail of the CityGML open data model [2, 5].
- Generate a high resolution texture, in accordance to the photo-consistent criteria used for optimization. More specifically, the input photos must be optimally blended [4] into a texture map so that the resulting textured model appear as similar as possible to all the cameras that shot the building.

Evaluation criteria:

- Visual fidelity. When rendered in 3D from similar virtual cameras, the output 3D textured model should be visually faithful to the input photos.
- Geometric fidelity. The output model should be close to the physical building, in the sense of a geometric error metric such as two-sided Hausdorff distance. We will use as ground truth either CAD models of existing buildings, or metrology measurement data.
- Robustness. We will evaluate the stability of the method against several criteria: sparsity of input (low number of input images), inconsistencies (shadows, reflections) and density of outliers.

References.

[1] F. Lafarge, X. Descombes, J. Zerubia and M. Pierrot Deseilligny. *Structural Approach for Building Reconstruction from a Single DSM*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010.

[2] Y. Verdie, F. Lafarge and P. Alliez: *LOD generation for Urban Scenes*. ACM Transactions on Graphics, 2015.

[3] P. Muller, P. Wonka, S. Haegler, A. Ulmer and L. Van Gool. *Procedural Modeling of Buildings*. Proceedings of ACM SIGGRAPH, 2006.

[4] M. Callieri, P. Cignoni, M. Corsini and R. Scopigno. *Masked Photo Blending: mapping dense photographic dataset on high-resolution 3D models*. Computer & Graphics, 2008.

[5] CityGML: http://www.citygml.org/

Location. The internship will be hosted at Dassault Systèmes Provence, in Aix-en-Provence.

Duration: 6 months starting from spring 2017.

Required qualification. The required qualification for this internship is a MS degree in Computer Science, or equivalent, and motived for a Ph.D. program after this internship. The basic requirement is a solid understanding of geometric data structures and algorithms, computer vision, applied mathematics and C++ programming. Familiarity with common computer vision and geometry processing algorithms such as photogrammetry and surface reconstruction is a real plus. Previous experience with image processing as well as numerical optimization techniques is highly relevant.

Contacts. André Lieutier, Vincent Guitteny and Pierre Alliez.

Partners.

- **Dassault Systèmes** develops and markets PLM software and services that support industrial processes by providing a 3D vision of the entire lifecycle of products from conception to maintenance. Dassault Systèmes Provence, created in 1999, contributes to the development of Dassault Systèmes core technologies. In particular, the hosting team focus is 3D modelling, vision and machine learning.
- **Inria**. The TITANE team focuses on geometric modeling of 3D environments, and analysis of satellite images, see https://team.inria.fr/titane/. Our central research themes involve methods commonly referred to as digital geometry processing, 3D urban reconstruction (both indoor and outdoor) and machine learning approaches for analysis of large-scale images. A distinctive property of the TITANE team is a synergy between geometry, computer vision and image processing. We made several advances on the computerized geometric modeling of complex scenes from physical measurements.