Combining meshes and geometric primitives for accurate and semantic modeling

Florent Lafarge[1]
http://imagine.enpc.fr/~lafargef/
Renaud Keriven[1]
http://imagine.enpc.fr/~keriven/
Mathieu Brédif[2]
http://recherche.ign.fr/labos/matis/cv.php?nom=Bredif

[1] IMAGINE, ENPC/CSTB, LabIGM
Université Paris est
Paris, France
[2] MATIS, IGN
Saint-Mandé, France

We propose an original hybrid modeling process that represents three-dimensional (3D) models as a combination of meshes and 3D-primitives. Meshes describe details such as ornaments or statues, whereas 3D-objects code for regular shapes such as walls or columns. Starting from a mesh obtained by multi-view stereo techniques, these primitives are substituted to the mesh where they are detected. This strategy allows the introduction of semantic knowledge, the simplification of the modeling, and even corrections of errors generated by the acquisition process.

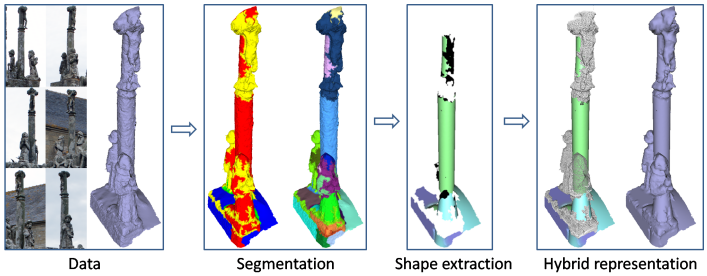

Figure 1: Overview of the method.

3D-models of urban scenes are very useful for many applications such as urban planning, virtual reality, disaster recovery or computer games [8]. The reconstruction of such scenes is a well known computer vision problem which has been addressed by various approaches providing integral building representations such as [2, 3, 4]. With the new perspectives offered for the aid to navigation by general public softwares such as *Street View* (Google) or *GeoSynth* (Microsoft), 3D building modeling is a topic of growing interest. Two main families of approaches may be distinguished in the literature.

The first family represents buildings as 3D-object layouts. The associated works efficiently detect and insert various urban objects such as windows or doors in 3D building models. However, these limited parametric descriptions fail to model fine details. On the contrary, the reconstruction of buildings with high order details, such as ornament, statues and other irregular shapes, is mainly addressed by mesh generation techniques using Laser scanning or multi-view stereo processes. Multi-view stereo techniques have significantly progressed during recent years as underlined in the comparative studies [5, 6]. However, buildings are man made objects containing many regular components such as planar or cylindrical shapes. Such a mesh representation gives a large amount of useless information concerning these regular elements which could be more relevantly described by parametric objects (*e.g.* wall facets by planes or columns by cylinders).

The two families have complementary advantages : semantic knowledge and model compaction for the former, detail modeling and non-restricted use for the latter. A natural idea, but still weakly unexplored, would consist in merging both the families in order to propose a **hybrid modeling**. Regular elements would be representing by 3D primitives whereas irregular structures would be described by meshes. In this paper, we propose a process for substituting regular mesh patches by 3D-objects (see Figure 2). This is of interest for several reasons: *(i)* the introduction of semantic knowledge in the mesh; *(ii)* the simplification of the modeling while preserving details; and *(iii)* the corrections of some errors generated by the multi-view stereo processes.

We adopt a two step strategy consisting in *(i)* segmenting the mesh, and *(ii)* fitting 3D-primitives on the obtained partition where it is relevant (see Figure 1). We first present the segmentation process based on a curvature analysis of the mesh. A multi-label energy taking topological smoothness constraints into account is formulated. The optimal labeling is estimated by α-expansion [1]. The 3D-primitive extraction from the obtained partition is then described. An error parameter controls the fitting quality and decides whether a mesh cluster has to be substituted by a plane, sphere, cylinder, cone or torus. In addition, a refinement process corrects the eventual errors generated during the segmentation step. Experimental results on real building meshes and also on synthetic data show the interesting potential of the proposed approach.

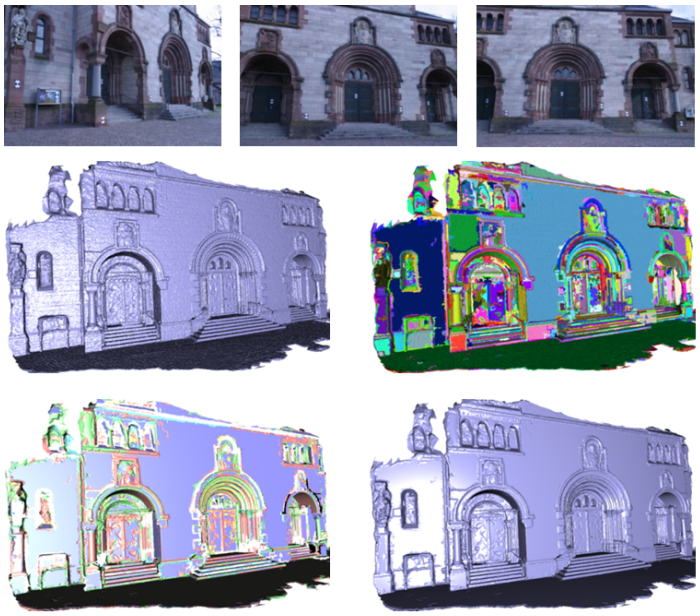

Figure 2: Hybrid modeling - *from top to bottom, and left to right:* some multiview stereo images from [6], mesh obtained from [7], segmented mesh, extracted 3D-primitives *(purple=plane, pink=cylinder, blue=cone, yellow=sphere, green=torus)*, our hybrid representation.

[1] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *PAMI*, 23(11):1222 – 1239, 2001.

[2] A.R. Dick, P.H.S. Torr, and R. Cipolla. Modelling and interpretation of architecture from several images. *IJCV*, 60(2):111–134, 2004.

[3] C. Fruh and A. Zakhor. An automated method for large-scale, ground-based city model acquisition. *IJCV*, 60(1):5–24, 2004.

[4] F. Lafarge, X. Descombes, J. Zerubia, and M. Pierrot-Deseilligny. Building reconstruction from a single DEM. In *CVPR*, Anchorage, United States, 2008.

[5] S. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. In *CVPR*, New York, United States, 2006.

[6] C. Strecha, W. Von Hansen, L. Van Gool, P. Fua, and U. Thoennessen. On benchmarking camera calibration and multi-view stereo for high resolution imagery. In *CVPR*, Anchorage, United States, 2008.

[7] H. Vu, R. Keriven, P. Labatut, and J.-P Pons. Towards high-resolution large-scale multi-view. In *CVPR*, Miami, United States, 2009.

[8] Z. Zhu and T. Kanade. Special issue on modeling and representations of large-scale 3d scenes. *International Journal of Computer Vision*, 78(2-3), 2008.