Automatic Room Detection and Reconstruction in Cluttered Indoor Environments with Complex Room Layouts

Claudio Mura[1] Oliver Mattausch[1,2] Alberto Jaspe Villanueva[3] Enrico Gobbetti[3] Renato Pajarola[1]

[1] Visualization and MultiMedia Lab, Department of Informatics, University of Zürich [2] Computer Vision Lab, ETH Zürich
[3] Visual Computing Group, CRS4, Sardinia, Italy

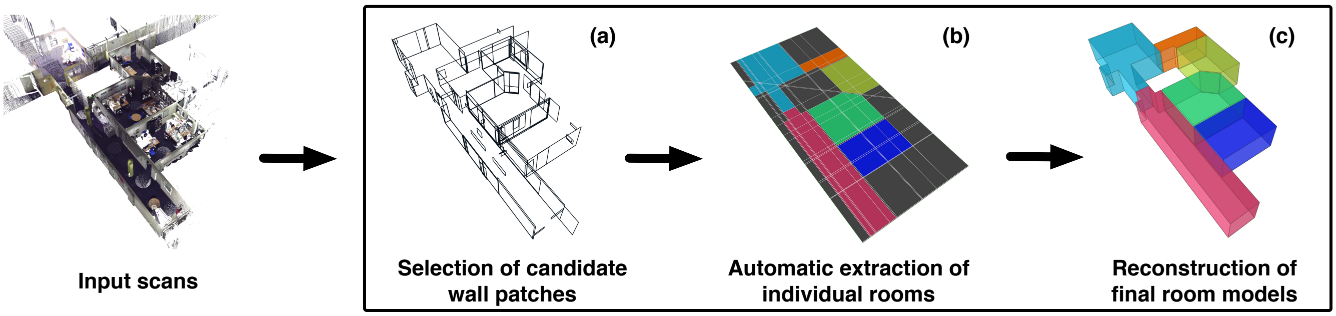

Figure 1: Overview of our pipeline: from the input scans we detect planar patches that are likely to belong to walls (a); from these we build a 2D cell complex of the floorplan, which is then segmented into individual rooms (b); finally, the 3D model of each room is reconstructed (c).

The need for methods to extract semantically rich 3D models of building interiors is growing stronger in the architecture and engineering domain. Particularly interesting is the problem of reconstructing the architectural structure of the environments, focusing on permanent structures like walls, ceilings and floor. Recent advances in 3D acquisition technologies (e.g. laser scanning) allow for fast and accurate production of raw measurements, but the reconstruction of higher-level models is made difficult by a number of factors. First of all, indoor environments contain high levels of clutter, resulting in heavy occlusions of walls and other structures of interest during the acquisition. Secondly, scanned models typically contain large-scale artifacts that originate from highly reflective surfaces (e.g. windows) often present in such scenes. Finally, creating structured 3D models of typical interior environments, such as apartments and office buildings, poses the challenge of recognizing their structure in terms of a set of individual rooms and corridors.

To address these points, we introduced a pipeline [Mura et al. 2014] for reconstructing the architectural model of an indoor environment from a set of 3D input scans. The method is robust to clutter and is able to extract the shape of the individual rooms of the environment, representing each of them as a polyhedron. The assumptions made are that the scanner positions are known and that the environment is bound by vertical walls and by horizontal ceilings of the same height.

Our pipeline (visually summarized in Fig.1) is composed of three main processing stages, which can be summarized as follows.

(a) Occlusion-aware selection of candidate wall patches. From the set of input point clouds we extract planar patches by applying a simple region growing procedure. We estimate the height of the floor and of the ceiling by finding the planar patches that have, respectively, minimum and maximum vertical coordinates. Good candidates for wall structures are computed by selecting the patches that are vertical and span a vertical extent almost equal to the height difference between ceiling and floor. To achieve robustness to viewpoint occlusions, when computing the vertical extent of a patch we also take into account the shadows cast onto it by possible occluders, that is, other patches placed between the patch considered and the scan position from which it was acquired. This allows to recover the *unoccluded* vertical extent of the vertical patches.

(b) Automatic extraction of individual rooms. The (3D) candidate wall patches selected in **(a)** are projected onto the horizontal plane, obtaining a set of 2D line segments. We merge segments that are close using the *mean-shift* clustering algorithm; the set of representative lines corresponding to the clusters obtained is used to build a 2D cell complex (i.e., an arrangement of lines) corresponding to the floor plan of the environment. The set of cells of the complex is then clustered by iteratively applying a binary version of the k-medoids algorithm (with $k = 2$). Taking advantage of the known positions of the scanpoints, this process extracts one room after another and automatically stops when all rooms have been found. The clustering is driven by *diffusion distances* between the cells of the complex, which are computed by applying a robust heat diffusion process to the complex itself.

(c) Reconstruction of the final room models. The clusters of 2D cells obtained in **(b)** correspond to the rooms of the environment; in this step, each cluster is converted into a 3D polyhedron. To do this, the edges of the boundary of each room cluster are extracted; then, the 3D wall planes corresponding to such boundary edges are computed, by robustly fitting a plane to the 3D points of the candidate wall patches that originated each segment. For each room, the final polyhedron is obtained by intersecting the 3D boundary wall planes and the ceiling and floor planes (extracted in **(a)**).

Our pipeline has been successfully validated by testing it on a number of real-world and synthetic datasets. We regard our work as a first step towards going beyond the simple geometric reconstruction to extract higher-level information about indoor environments.

References

MURA, C., MATTAUSCH, O., JASPE, A. V., GOBBETTI, E., AND PAJAROLA, R. 2014. Automatic room detection and reconstruction in cluttered indoor environments with complex room layouts. *Computers & Graphics 44*, 20–32.