Computer Aided Geometric Design 43 (2016) 82–94

Contents lists available at ScienceDirect

Computer Aided Geometric Design

www.elsevier.com/locate/cagd

Cluttered indoor scene modeling via functional part-guided graph matching

Jun Wang, Qian Xie, Yabin Xu, Laishui Zhou, Nan Ye *

College of Mechanical Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

ARTICLE INFO

Article history:
Available online 10 March 2016

Keywords:
Cluttered indoor scenes
Scene modeling
Graph matching
Raw point clouds

ABSTRACT

We propose an automatic method for fast reconstruction of indoor scenes from raw point scans, which is a fairly challenging problem due to the restricted accessibility and the cluttered space for indoor environment. We first detect and remove points representing the ground, walls and ceiling from the input data and cluster the remaining points into different groups, referred to as sub-scenes. Our approach abstracts the sub-scenes with geometric primitives, and accordingly constructs the topology graphs with structural attributes based on the functional parts of objects (namely, *anchors*). To decompose sub-scenes into individual indoor objects, we devise an *anchor-guided* subgraph matching algorithm which leverages template graphs to partition the graphs into subgraphs (i.e., individual objects), which is capable of handling arbitrarily oriented objects within scenes. Subsequently, we present a data-driven approach to model individual objects, which is particularly formulated as a model instance recognition problem. A Randomized Decision Forest (RDF) is introduced to achieve robust recognition on decomposed indoor objects with raw point data. We further exploit template fitting to generate the geometrically faithful model to the input indoor scene. We visually and quantitatively evaluate the performance of our framework on a variety of synthetic and raw scans, which comprehensively demonstrates the efficiency and robustness of our reconstruction method on raw scanned point clouds, even in the presence of noise and heavy occlusions.

1. Introduction

3D scene processing for both indoor and outdoor environments has been an important research problem in computer vision and graphics communities. Meanwhile, recent advances in scanning technology greatly improve the acquisition of point clouds both in speed and accuracy, which also renders point cloud processing receive increasing attention recently. Despite the advances in acquisition technology, the captured point cloud often suffers from severe noise and outliers, making the reconstruction of indoor models with faithful geometry and topology from such data rather arduous. In addition, the significant difficulty in indoor scans is the presence of heavy occlusions as the interior environments are usually relatively narrow and cluttered, even when multiple scan stations are set. Consequently, to automatically and efficiently reconstruct indoor scenes is particularly challenging, especially for the cluttered indoor environments with defect-laden, raw point clouds.

* Corresponding author.
E-mail address: isaac_yn@126.com (N. Ye).

http://dx.doi.org/10.1016/j.cagd.2016.02.012

Recent works (Nan et al., 2012; Kim et al., 2012; Shao et al., 2012) take advantage of the learning-based technique to infer scene segmentation, and then detect and replace multiple instances of the object within the indoor scene using for instance partial matching so as to achieve scene reconstruction. These methods are able to obtain promising modeling results even in the presence of cluttered scenes thanks to the data-driven characteristics. However, they mostly assume the indoor objects are placed always with the upward direction in terms of the ground floor. Once the assumption is invalid, they may fail in segmentation and thus reconstruction. In addition, a certain level of interactions are required during reconstruction. Moreover, when the level of data imperfection becomes high, they would perform relatively poor on the indoor scenes, due to typical clutter, missing regions and noise. The goal of our work is to automatically reconstruct cluttered indoor scenes with arbitrarily oriented objects from raw point scans.

We propose an automatic method for fast modeling raw point data captured from cluttered indoor scenes. As observed, most man-made objects of indoor scenes are assembled by parts corresponding to primitive shapes. Accordingly, we fit the point data of sub-scenes with primitive shapes to obtain a concise representation. Moreover, we notice that a man-made object generally contains at least one functional structure (namely, *anchor*), which topologically relates to the other structures (i.e., primitives) of the object. Therefore, it is reasonable to construct a topology graph, formed by connecting the anchor to the other primitives, to represent the sub-scene. Thus, we abstract sub-scenes with the topology graphs with attributes, which adequately convey the geometrical and structural information of the sub-scenes. To analyze the graphs, rather than the original point data, we render an efficient and effective way to decompose individual objects from sub-scenes.

A sub-scene usually consists of several indoor objects. Analogously, the attributed graph comprises several subgraphs. Accordingly, we formulate sub-scene decomposition as a graph matching problem. Collecting a database of man-made indoor shapes, we construct the topology graphs for them, referred to as graph templates. We use the graph templates to partition the topology graphs into subgraphs, each of which corresponds to an individual indoor object. By constructing a matching similarity function, we find the correspondences between graphs by solving a maximization problem. By reducing the computation complexity during optimization, we minimize the number of similarity comparisons between graphs to sparsify the similarity matrix. Particularly, we only measure the similarity between graph nodes sharing the same primitive type. Moreover, we establish candidate graph matches only starting from anchor nodes, and then restrict the comparisons only between edges induced from the corresponding anchor nodes. In return, the number of similarity comparisons is significantly decreased, and our graph matching can be accomplished efficiently. As a result, the individual objects are decomposed from the sub-scene, while the category of each object is determined as well.

To reconstruct individual indoor objects, we present a data-driven modeling method based on the shape database. Specifically, we formulate object modeling as a model instance recognition problem. To this end, a Randomized Decision Forest (RDF) is introduced to solve this recognition problem. We define a set of shape features for learning of the RDF classifier. The features are discriminative and insensitive to noise, outliers and data sparsity. We then exploit template fitting to compute the transformations from database models to scanned objects, which are applied to achieve geometrically faithful reconstruction from the input indoor scene.

Overall, our contributions are as follows:

1. We propose a functional part-guided modeling method for cluttered indoor scenes with raw scans. It proceeds automatically and results in high fidelity to input scenes.
2. We design an anchor-guided graph matching algorithm for scene decomposition, which is capable of handling scenes with objects arbitrarily oriented.
3. We devise a data-driven approach for object modeling based on randomized decision forest, which is robust to data imperfections.

1.1. Related work

There is an extensive amount of literature on scene modeling, ranging from image-based (Saxena et al., 2009; Xiao et al., 2010; Quattoni and Torralba, 2009), RGBD-based (Izadi et al., 2011; Bo et al., 2013) to 3D point-based approaches (Frome et al., 2004; Rusu et al., 2008; Schnabel et al., 2008; Nan et al., 2010; Shen et al., 2011; Koppula et al., 2011; Kim et al., 2012). Here, we mainly focus on the most work to ours, particularly for those regarding scene modeling, scene reconstruction and object matching.

Scene modeling. The procedural modeling of large-scale scenes has gained much attention in recent years (Parish and Müller, 2001; Wonka et al., 2003; Müller et al., 2006; Musialski et al., 2013). With the significant advances in 3D scanning recently, increasing research work has been focusing on scene reconstruction directly from 3D scan data (Shao et al., 2012; Nan et al., 2012; Lin et al., 2013; Arikan et al., 2013; Mattausch et al., 2014). Nan et al. (2012) used the repetition characteristic to model urban facades. It requires a moderate amount of user interactions to reveal the architectural structures as repetitive patterns. Kim et al. (2012) utilized object repeatability to reconstruct indoor scenes with basic primitives. Outdoor scenes, e.g. building facades, usually exhibit symmetry and repetitions, while indoor scenes are generally cluttered and objects are arranged randomly. We concentrate on cluttered indoor scenes without any assumption of repeatability and regularity.

Indoor scene reconstruction. Indoor scene reconstruction has also attracted plenty of research interest recently (Du et al., 2011; Izadi et al., 2011; Ren et al., 2012; Henry et al., 2014). From a scene database, Chen et al. (2014) learned the contextual