ISPRS Journal of Photogrammetry and Remote Sensing 94 (2014) 183–193

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

A global optimization approach to roof segmentation from airborne lidar point clouds

Jixing Yan [a], Jie Shan [a,b,*], Wanshou Jiang [c]

[a] School of Remote Sensing and Information Engineering, Wuhan University, 129 Luoyu Rd., 430079 Wuhan, China
[b] School of Civil Engineering, Purdue University, West Lafayette, IN 47907, USA
[c] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 129 Luoyu Rd., 430079 Wuhan, China

ARTICLE INFO

Article history:
Received 6 August 2013
Received in revised form 27 April 2014
Accepted 29 April 2014

Keywords:
Segmentation
City modeling Buildings
Lidar
Point clouds
Global optimization

ABSTRACT

This paper presents a global plane fitting approach for roof segmentation from lidar point clouds. Starting with a conventional plane fitting approach (e.g., plane fitting based on region growing), an initial segmentation is first derived from roof lidar points. Such initial segmentation is then optimized by minimizing a global energy function consisting of the distances of lidar points to initial planes (labels), spatial smoothness between data points, and the number of planes. As a global solution, the proposed approach can determine multiple roof planes simultaneously. Two lidar data sets of Indianapolis (USA) and Vaihingen (Germany) are used in the study. Experimental results show that the completeness and correctness are increased from 80.1% to 92.3%, and 93.0% to 100%, respectively; and the detection cross-lap rate and reference cross-lap rate are reduced from 11.9% to 2.2%, and 24.6% to 5.8%, respectively. As a result, the incorrect segmentation that often occurs at plane transitions is satisfactorily resolved; and the topological consistency among segmented planes is correctly retained even for complex roof structures.

1. Introduction

Building roof reconstruction is of a current research interest in 3D city modeling. Because of being able to directly collect dense, accurate 3D point clouds of urban objects, lidar (Light Detection and Ranging) technology provides an efficient solution to this need. Reported methods for building roof reconstruction mostly fall into two categories: data-driven (bottom–up) and model-driven (top–down). In terms of data-driven methods, a common assumption is that a building is a polyhedron consisting of planes and edges. As a crucial step, the point clouds of a building are usually segmented into disjointed planar regions. Subsequent tasks, including roof edge extraction and topologic reconstruction, are dependent on the quality of segmentation. A poor segmentation may make these tasks fail. As for model-driven methods, a building is assumed to be an assembly of roof primitives (e.g., gable roof and hipped roof), which and whose topology are predefined in a model library (Tarsha-Kurdi et al., 2007a; Huang et al., 2013). However, roof segmentation is still a required step in many model-driven methods, such as the graph matching approach (Verma et al., 2006; Oude Elberink and Vosselman, 2009). A poor segmentation may alter the topology among roof planes and make the matching task fail (Oude Elberink and Vosselman, 2009).

Building roof segmentation can be accomplished via various approaches, such as data clustering, region growing, energy minimization, and model fitting. A review of these approaches can be found in (Awwad et al., 2010; Sampath and Shan, 2010). Data clustering is basically a statistical technique that classifies the point clouds into primitives based on certain pre-calculated local surface properties or features. Filin and Pfeifer (2006) propose a slope adaptive neighborhood for such calculation. Considering the variations in local point density, Sampath and Shan (2010) use the Voronoi neighborhood to estimate the local surface properties, whereas Lari et al. (2011) use a cylindrical neighborhood for this purpose. As for clustering the feature vectors, mode-seeking (Filin and Pfeifer, 2006), conventional mean-shift (Melzer, 2007), and fuzzy k-means (Sampath and Shan, 2010) are applied. In spite of the popularity and efficiency of this approach, it suffers the difficulty in neighborhood definition and is sensitive to noise and outliers.

Region growing is a region-based segmentation method that partitions point clouds into disjoint homogenous regions. It starts with a selected seed point or region and expands to its neighboring points. Gorte (2002) selects the triangles in triangulated irregular

* Corresponding author. Address: Jie Shan School of Civil Engineering, Purdue University, West Lafayette, IN 47907, USA.

E-mail addresses: yanjx@whu.edu.cn (J. Yan), jshan@ecn.purdue.edu (J. Shan), jws@whu.edu.cn (W. Jiang).

http://dx.doi.org/10.1016/j.isprsjprs.2014.04.022
0924-2716/

networks (TINs) as seed regions and extends them to neighboring triangles. Zhang et al. (2006) perform a local plane fitting at points and select the points with good planarity as seed points. To obtain robust seed points, Chauve et al. (2010) develop an iterative Principal Component Analysis (PCA) to estimate local planarity. You and Lin (2011) present a non-iterative approach using tensor voting for this purpose. Unlike the aforementioned approaches, Dorninger and Pfeifer (2008) determine seed clusters (regions) by a hierarchical clustering approach. As for expanding seed regions, similarity measures such as distances of points to planes (Zhang et al., 2006; Dorninger and Pfeifer, 2008; Chauve et al., 2010) and angle differences between normal vectors (Dorninger and Pfeifer, 2008; Chauve et al., 2010; You and Lin, 2011) are used. Nevertheless, region growing is susceptible to the selection of seed regions (Awwad et al., 2010) and difficult to stop when transitions between two regions are smooth (Sampath and Shan, 2010).

The energy minimization approach is a global solution that formulates the segmentation as an optimization problem. Its objective function may consist of fidelity to data, continuity of feature values and compactness of segment boundaries (Kim and Shan, 2011; Vitti, 2012). A widespread application of this approach to image segmentation can be found in (Vitti, 2012). As for the segmentation of lidar data, multiphase level set approach is adopted to segment planar roof primitives under an energy minimization formulation (Kim and Shan, 2011). Compared to the RANSAC (RANdom SAmple Consensus, Fischler and Bolles, 1981) based approaches, it is global and multiple roof planes can be segmented at one time. However, a common shortcoming of this approach is that poor segmentation may occur when the energy function converges to a local minimum.

Since the reconstructed models are dependent on the robust estimate at planar primitives, robust model fitting methods such as RANSAC and Hough transform (Duda and Hart, 1972) are also applied to roof segmentation. Lidar points that fit a mathematical plane with most inliers are first extracted and regarded as a planar segment. This approach is robust to noise and outliers, but it tends to result in spurious planes (Tarsha-Kurdi et al., 2007b; Yan et al., 2012). With the help of available building ground plans, Vosselman and Dijkman (2001) split the dataset into small parts before applying Hough transform to prevent the detection of spurious planes. Some extended RANSAC considering local surface normals (Bretar and Roux, 2005; Schnabel et al., 2007; Awwad et al., 2010; Chen et al., 2012) are also developed for this purpose. Considering spatial connectivity, Zhang et al. (2006) and Chauve et al. (2010) combine model fitting and region growing. Nevertheless, most of the model fitting approaches are order-dependent and based on a single-model. Segmentation results are dependent on the order in which the planes are extracted. When multiple planes are present, each plane instance needs to be sequentially extracted. As a result, points at transitions between roof faces will be assigned to the first extracted planes. In most cases, this approach performs well with some additional constraints. However, for complex roof structures it tends to result in mistakes, such as spurious planes (segments that do not exist in reality), over-segmented planes (one actual plane is segmented into multiple segments), and under-segmented planes (multiple actual planes are segmented into one segment).

To overcome these problems, this paper seeks a global optimization solution to the problem of roof segmentation from airborne lidar point clouds. A multi-label (plane) optimization approach is introduced for this purpose. It intends to reduce the number of mistakes derived from a plane fitting based on region growing and to improve the topological consistency among segmented planes. In our study, the point clouds of building roofs are first extracted from airborne lidar data. Starting with initial planes derived from a plane fitting approach, a global energy function consisting of fidelity to data, spatial smoothness, and the number of models (i.e., the number of planes) is constructed to optimize the segmented planes. Comparing to existing approaches, the proposed method incorporates spatial smoothness between data points into plane fitting. It can produce spatially coherent segments and improve the segmentation quality. Additionally, the proposed approach is global, i.e., multiple roof planes are determined at the same time and their corresponding model parameters can be refined when minimizing the energy function. It yields both high completeness and high correctness rates. More noticeably, the incorrect segmentation that often occurs at plane transitions is satisfactorily avoided and the topological consistency among segmented planes is correctly retained.

The remainder of the paper is structured as follows. Section 2 formulates the segmentation task as a multi-label (plane) optimization problem and presents a graph cuts based global solution. Section 3 starts with the test data description, followed by a presentation of individual and overall test results to demonstrate the solution procedure. Assessment and discussion constitute Section 4, where we define our quality metrics, examine the sensitivity of the solution to relevant parameters, and assess the metric quality and topologic quality of the segmentation outcome. Both quantitative and qualitative evaluations are presented. Section 5 consists of our concluding remarks on the properties of the method and future effort.

2. Multi-label optimization

2.1. Formulation

The segmentation task can be noted as a labeling problem and formulated in terms of energy minimization. Eq. (1) provides such a formulation (Delong et al., 2012; Isack and Boykov, 2012)

$$E(L) = \overbrace{\sum_{p\in P} D_p(L_p)}^{data\ \ cost} + \overbrace{\sum_{p,q\ \in N} w_{pq}\cdot\delta(L_p \neq L_q)}^{smooth\ \ cost} + \overbrace{\sum_{l\ \in\ \mathcal{L}} h_l\cdot\delta(l\in\ \mathcal{L}')}^{label\ \ cost} \tag{1}$$

where $\mathcal{L}$ is a given set of labels (planes) and $\delta(.)$ is an indicator function. Let P represent a set of data points, the multiple labeling task is to assign a point $p \in P$ a label $L_p \in \mathcal{L}$ such that the labeling L minimizes the energy $E(L)$, where $\mathcal{L}'$ is the set of labels appearing in L and N is an assumed neighborhood for data points. Three energy terms are considered in the energy formula. The data cost term (the first term in Eq. (1)) measures the discrepancy between data points and labels. It is the sum of the distances of points to their assigned labels. The smooth cost term (the second term in Eq. (1)) measures the label inconsistency between neighboring points. It is the sum of weight w_{pq} of each pair of neighboring points p and q that are assigned to different labels. The label cost term (the third term in Eq. (1)) measures the number of labels appearing in L. It is the sum of the label cost h_l of each label $l \in \mathcal{L}'$. Fig. 1 illustrates a labeling of data points and their fitted lines. Two lines A and B are fitted to

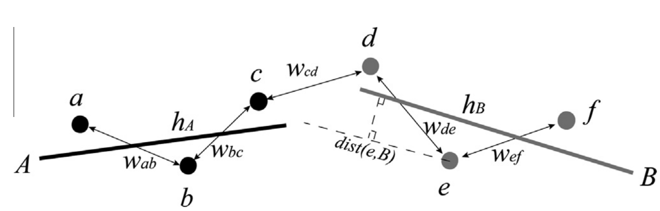

Fig. 1. Labeling of data points and their fitted lines. The double-arrowed lines link each pair of neighboring points. Data points and their corresponding fitted lines are shown with the same shade.