ISPRS Journal of Photogrammetry and Remote Sensing 99 (2015) 45–57

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Hierarchical extraction of urban objects from mobile laser scanning data

Bisheng Yang *, Zhen Dong *, Gang Zhao, Wenxia Dai

State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China

ARTICLE INFO

Article history:
Received 16 August 2014
Received in revised form 24 October 2014
Accepted 27 October 2014
Available online 20 November 2014

Keywords:
Mobile laser scanning
Multi-scale supervoxel
Segmentation
Object extraction
Classification
Filtering

ABSTRACT

Point clouds collected in urban scenes contain a huge number of points (e.g., billions), numerous objects with significant size variability, complex and incomplete structures, and variable point densities, raising great challenges for the automated extraction of urban objects in the field of photogrammetry, computer vision, and robotics. This paper addresses these challenges by proposing an automated method to extract urban objects robustly and efficiently. The proposed method generates multi-scale supervoxels from 3D point clouds using the point attributes (e.g., colors, intensities) and spatial distances between points, and then segments the supervoxels rather than individual points by combining graph based segmentation with multiple cues (e.g., principal direction, colors) of the supervoxels. The proposed method defines a set of rules for merging segments into meaningful units according to types of urban objects and forms the semantic knowledge of urban objects for the classification of objects. Finally, the proposed method extracts and classifies urban objects in a hierarchical order ranked by the saliency of the segments. Experiments show that the proposed method is efficient and robust for extracting buildings, streetlamps, trees, telegraph poles, traffic signs, cars, and enclosures from mobile laser scanning (MLS) point clouds, with an overall accuracy of 92.3%.

1. Introduction

Mobile laser scanning (MLS) systems capture three-dimensional (3D) point clouds with high flexibility and precision, thus are widely used for various applications (e.g., transportation, forestry). However, MLS systems, unlike traditional surveying equipment, collect huge data volumes, resulting in urgent demands for efficient and effective processing of the point clouds, or the productivity gained in the data collection phase may be lost during processing. Points collected by MLS systems from urban scenes contain numerous objects with significant disparities in size, complicated and incomplete structures, holes, varied point densities, and huge data volumes, raising great challenges for automated point segmentation and object extraction. In recent years, there have been many scientific contributions aiming to process mobile laser scanning point clouds from urban scenes, focusing on segmentation (Aijazi et al., 2013; Barnea and Filin, 2013; Lari and Habib, 2013; Serna and Marcotegui, 2014; Yang and Dong, 2013; Yao et al., 2009), roads extraction and modelling (Boyko and Funkhouser, 2011; Guan et al., 2014; Hernández and Matcotegui, 2009a,b; Yang et al., 2012, 2013a; Zhu and Hyyppa, 2014), pole-like objects extraction (Cabo et al., 2014; Lehtomäki et al., 2010; Li and Elberink, 2013; Monnier et al., 2012; Pu et al., 2011; Yang and Dong, 2013), and building extraction and reconstruction (Jochem et al., 2011; Pu and Vosselman, 2009; Yang et al., 2013b).

Segmentation is the fundamental step for extracting objects from MLS point clouds. Barnea and Filin (2013) proposed a segmentation method for terrestrial laser scanning data by integrating ranges, normals, and colors in a panoramic representation. The proposed segmentation method yielded more physically meaningful segments. Objects were extracted by forming the segments as meaningful units according to predefined rules (Pu et al., 2011; Yang and Dong, 2013). Pu et al. (2011) presented a framework for structured recognition from MLS point clouds and classified the objects as traffic signs, trees, building walls, and barriers based on the characteristics of points segments, such as size, shape, orientation, and topological relationship. Yang and Dong (2013) proposed a shape-based segmentation method that classified points according to the geometric features derived from support vector machines (SVMs) for objects extraction. All of the above methods have difficulties extracting individual objects from areas of dense mixed objects, and have heavy computing costs because the local geometric features of each point must be calculated.

* Corresponding authors.
E-mail addresses: bshyang@whu.edu.cn (B. Yang), dongzhenwhu@whu.edu.cn (Z. Dong).

http://dx.doi.org/10.1016/j.isprsjprs.2014.10.005
0924-2716/

Lehtomäki et al. (2010) developed an automated method for the detection of vertical pole-like structures (e.g., traffic signs, lamp posts) in road environments with the support of profile information from the scanner. Yang et al. (2013a) presented a method for extracting and delineating roads from large-scale MLS point clouds by detecting curb points from a set of consecutive scanning lines. These methods are suitable for the rapid classification of MLS data and the extraction of specific kinds of objects, but they are difficult to deal with unordered laser scanning data (particularly when point clouds from more than two laser scanners are mixed).

Hernández and Matcotegui (2009a) presented a method to generate range images from point clouds for extracting features using morphological operators. Serna and Marcotegui (2014) proposed an automatic and robust approach to classify objects from the points of urban scenes using elevation images. The result is re-projected onto the 3D point cloud, to detect, segment and classify urban objects. These methods accelerate the speed of operation and improve the efficiency, but at the cost of loss of accuracy in the process of generating images. In addition, pixel size requires careful tuning to extract different kinds of objects. The quality of objects extraction depends on the generation of images from the points.

To reduce the computing cost of point clouds, the voxels based segmentation and classification method was developed (Aijazi et al., 2013; Lim and Suter, 2008, 2009). Lim and Suter (2008, 2009) first over-segmented the points into 3D voxels, then calculated the local and regional features of voxels, and finally used multi-scale Conditional Random Fields to classify 3D outdoor terrestrial laser scanning data. Aijazi et al. (2013) presented a method to classify urban scenes based on voxels segmentation of sparse 3D data obtained from LiDAR sensors. The method first partitioned 3D point cloud into voxels, then joined voxels by a link-chain method to create objects, and finally classified these objects using geometrical models and local descriptors. These voxels based methods accelerate the computing speed. Nevertheless, the segmentation quality is subject to the size of the voxels. Voxels of a fixed size lead incorrect results, particularly in the areas of dense mixed objects.

Although the reported methods are generally able to extract specific kinds of objects based on segmentation or with the support of scanning lines, they suffer from the quality of segmentation and there remains much room for improvement. On the one hand, the semantic knowledge of urban objects should be formed as rules for extracting and classifying urban objects. On the other hand, the time efficiency of extracting objects should be improved. This paper proposes an automated method to extract and classify urban objects following the pipeline of segmentation.

The contributions of the proposed method have three aspects:

- Generate multi-scale supervoxels from scattered MLS points to improve the estimation of local geometric structures of neighboring points and the time efficiencies of segmentation.
- Form semantic knowledge of urban objects into rules for merging adjacent segments into meaningful units, resulting in the meaningful units having better consistency with physical objects.
- Define a hierarchical strategy to extract and classify objects from urban scenes in the order of the saliency of the segments, resulting in the improved accuracy of object extraction, especially in the cluttered situation of occlusion and overlapping between closely neighboring objects.

Following the introduction to the subject, we elaborate the proposed method and test and validate the method with two MLS datasets before drawing our conclusions.

2. Methodology

The point clouds of urban scenes usually contain a huge number of points with varied point densities and occlusions, including ground points and non-ground points. The proposed method aims to extract urban objects from non-ground points of urban scenes robustly and efficiently. It removes ground points using the approach of Hernández and Matcotegui (2009b) before extraction of urban objects. The proposed method firstly generates the supervoxels of two different sizes according to the attributes (e.g., colors, intensities) of non-ground points and spatial relations between the points, resulting in multi-scale supervoxels by integrating the generated supervoxels. Each supervoxel contains the points of a certain kind of objects and has unique geometric property, and then the proposed method segments the supervoxels according to the principles of graph-based segmentation. Secondly, the proposed method defines the formula of the saliency of the segments using several factors and models the sematic knowledge of urban objects as formal representation. Finally, the proposed method forms the segments as physical urban objects according to a set of predefined rules in a hierarchical order ranked by the saliency of the segments, and classifies the objects according to the formal representation of urban objects, resulting in robust and efficient extraction and classification of urban objects.

2.1. Generating multi-scale supervoxels of non-ground points

Point-wise segmentation processing has a heavy computing cost and may lead to over or under segmentation. To reduce the computing costs of the point clouds of large scale urban scenes, the scene space is partitioned into 3D voxels and the points are allocated into the corresponding 3D voxels according to their coordinates. Then, we utilize a weighted distance measurement to reallocate the voxels of each point by its color or intensity and the spatial distance between itself and other close points. Inspired by the work of Achanta et al. (2012) for 2D image segmentation, we construct supervoxels of the point clouds by the weighted distance. The key steps of generating 3D supervoxels are:

Step 1: Initialize the size, S, of the 3D voxels, and partition the scene of point clouds into 3D voxels according to their associated coordinates.

Step 2: Set the centroid of each 3D voxel k, with the attribute $C_k = (x_k, y_k, z_k, L_k, a_k, b_k$, where $x_k, y_k, z_k, L_k, a_k, b_k$ are the coordinates and the color values in CIE Lab, respectively, of the closest point to the centroid of the 3D voxel. The mathematical formulations of the descriptive about transforming RGB values to CIE Lab colors are given in Appendix A.

Step 3: Initialize the distances between each point $p_i(x_i, y_i, z_i, L_i, a_i, b_i)$ in each 3D voxel and its associated centroid as infinitely large $d_i = \infty$.

Step 4: Traverse the 3D voxels one by one, search the neighboring points of the centroid C_k, of each voxel within a sphere radius S, the voxel size, and calculate the associated distances between each neighboring point and the centroid by:

$$d_{ik} = \sqrt{\left(\frac{d_{ik}^c}{N_c}\right)^2 + \left(\frac{d_{ik}^s}{N_s}\right)^2} \tag{1}$$

$$d_{ik}^s = \sqrt{(x_i - x_k)^2 + (y_i - y_k)^2 + (z_i - z_k)^2}$$

$$d_{ik}^c = \sqrt{(L_i - L_k)^2 + (a_i - a_k)^2 + (b_i - b_k)^2}$$