ISPRS Journal of Photogrammetry and Remote Sensing 126 (2017) 180–194

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Computing multiple aggregation levels and contextual features for road facilities recognition using mobile laser scanning data

Bisheng Yang [a,*], Zhen Dong [a,*], Yuan Liu [a], Fuxun Liang [a], Yongjun Wang [b,c,d,*]

[a] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China
[b] Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing 210023, China
[c] Key Laboratory of Virtual Geographic Environment, Ministry of Education, Nanjing Normal University, Nanjing 210093, China
[d] State Key Laboratory Cultivation Base of Geographical Environment Evolution, Nanjing 210093, China

ARTICLE INFO

Article history:
Received 3 June 2016
Received in revised form 20 December 2016
Accepted 21 February 2017

Keywords:
Point clouds processing
Object recognition
Contextual features
Multiple aggregation levels features
Semantic labeling

ABSTRACT

In recent years, updating the inventory of road infrastructures based on field work is labor intensive, time consuming, and costly. Fortunately, vehicle-based mobile laser scanning (MLS) systems provide an efficient solution to rapidly capture three-dimensional (3D) point clouds of road environments with high flexibility and precision. However, robust recognition of road facilities from huge volumes of 3D point clouds is still a challenging issue because of complicated and incomplete structures, occlusions and varied point densities. Most existing methods utilize point or object based features to recognize object candidates, and can only extract limited types of objects with a relatively low recognition rate, especially for incomplete and small objects. To overcome these drawbacks, this paper proposes a semantic labeling framework by combing multiple aggregation levels (point-segment-object) of features and contextual features to recognize road facilities, such as road surfaces, road boundaries, buildings, guardrails, street lamps, traffic signs, roadside-trees, power lines, and cars, for highway infrastructure inventory. The proposed method first identifies ground and non-ground points, and extracts road surfaces facilities from ground points. Non-ground points are segmented into individual candidate objects based on the proposed multi-rule region growing method. Then, the multiple aggregation levels of features and the contextual features (relative positions, relative directions, and spatial patterns) associated with each candidate object are calculated and fed into a SVM classifier to label the corresponding candidate object. The recognition performance of combining multiple aggregation levels and contextual features was compared with single level (point, segment, or object) based features using large-scale highway scene point clouds. Comparative studies demonstrated that the proposed semantic labeling framework significantly improves road facilities recognition precision (90.6%) and recall (91.2%), particularly for incomplete and small objects.

1. Introduction

Rapidly updating the inventory of highway infrastructures is of great importance for transportation infrastructure management and intelligent transportation related applications, including intelligent drive assistant systems (Choi et al., 2012; Broggi et al., 2013; Zhu and Hyyppä, 2014; Luo et al., 2015; Seo et al., 2015; Wen et al., 2016), and traffic flow monitoring and prediction (Abadi et al., 2015; Lv et al., 2015). Techniques such as field surveying, photo/video log (Wang et al., 2010), integrated GPS/GIS mapping systems (Caddell et al., 2009; Tang et al., 2016), and aerial/satellite remote sensing (Ravani et al., 2009), have been used for highway inventory data collection. However, many limiting factors (e.g. shadows, complex illumination and low spatial accuracy) make these approaches labor intensive, time consuming, and costly. MLS systems can rapidly capture three-dimensional (3D) point clouds of road scenes with high flexibility and precision, providing a promising and feasible method for rapidly highway inventory data collection. Extensive studies have suggested methods to extract road facilities from MLS point clouds. Existing methods can be classified into pointwise labeling (Munoz, 2008; Babahajiani et al., 2014) and object based classification (Yokoyama et al., 2010; Dohan et al., 2015; Lehtomäki et al., 2015; Yang et al., 2015; Yan et al., 2016).

* Corresponding authors at: Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing 210023, China (Y. Wang).

E-mail addresses: bshyang@whu.edu.cn (B. Yang), dongzhenwhu@whu.edu.cn (Z. Dong), wangyongjun@njnu.edu.cn (Y. Wang).

http://dx.doi.org/10.1016/j.isprsjprs.2017.02.014
0924-2716/

Munoz (2008) classified the points into several classes (e.g., ground, facade, scatter, pole, trunk, and wire) by combining local feature descriptors and context based features, and achieved overall classification accuracy 91.66%. Brodu and Lague (2012) calculate multi-scale local dimensionality feature of 3D points for the classification of natural scenes. Babahajiani et al. (2014) classified non-ground points into several classes based on local feature descriptors (e.g., geometrical shape, height above ground, density, intensity, normal angle, etc.), and achieved classification accuracies of different objects from 72% to 95%.

Compared with pointwise labeling methods, object based classification methods gain more useful features from each candidate object, and are more widely used to extract candidate objects. Object based classification methods involve two major stages: segmentation and recognition. A lot of methods based on Euclidean distance clustering have been reported to segment the point clouds into individual object candidates, for example, connected k-nearest neighbors (Yokoyama et al., 2010), connected components analysis (Lehtomäki et al., 2015), and density based spatial clustering of applications with noise (Yan et al., 2016). However, Euclidean distance clustering can lead to under-segmentation, especially in areas where objects are adjacent or overlapped. To overcome this partial segmentation, advanced segmentation methods were proposed. Yang et al. (2015) first over-segmented the point clouds into supervoxels, and then merged the adjacent supervoxels into meaningful units by encoding semantic object knowledge as merging rules. Dohan et al. (2015) also first over-segmented the point clouds into supervoxels, then merged supervoxels by integrating an object classifier into a hierarchical segmentation algorithm. Yu et al. (2015a) first under-segmented the point clouds using Euclidean distance clustering approach, then developed a normalized cut segmentation method to further segment clusters containing more than one object.

After segmentation, object candidates are recognized as two or several classes based on a set of features. Existing approaches for recognizing object candidates can be classified into prior information or semantic rules based methods (Teo and Chiu, 2015; Yang et al., 2015), 3D object matching based methods (Yu et al., 2015a, b), and machine learning based methods (Golovinskiy et al., 2009; Himmelsbach et al., 2009; Lehtomäki et al., 2015). The design and construction manuals for roadside infrastructure facilities mean some types of roadside objects have predefined shapes, heights, and sizes that provide essential prior knowledge for object recognition. Teo and Chiu (2015) recognized traffic lights, traffic signs, and street trees using semantic rules for pole-like objects, such as heights, sizes, shapes, positions, etc. Yang et al. (2015) formed prior object knowledge into rules for segmenting and classifying multiple urban objects, and improved the accuracy of object extraction, particularly in the cluttered situation of occlusion and overlap. 3D object matching frameworks, benefiting from a locally affine-invariant geometric constraint (Yu et al., 2015a) and 3D shape context (Yu et al., 2015b), were developed to achieve recognition of 3D objects. Machine learning based classification uses adaptive methods to learn the mapping from feature space to the class labels using training data, and then predicts class labels of new examples (Lehtomäki et al., 2015). Several geometric features have been suggested for the machine learning based classification, such as spin images (Golovinskiy et al., 2009), point feature histograms (Himmelsbach et al., 2009), and general features (Lehtomäki et al., 2015). Lehtomäki et al. (2015) also compared the performance of three feature types (local descriptor histograms, spin images, and general features) for classification of roadside objects.

Other methods (Fischer et al., 1998; Xu et al., 2014; Niemeyer et al., 2015) computed different levels of features to improve the recognition rate for some interested objects. Fischer et al. (1998) combined feature level, feature aggregate level, building part level, and building level semantic features to automated 3D extraction of buildings from aerial images. Xu et al. (2014) proposed a multiple-entity strategy by combining individual points, planar segments, and mean shift segments based features to recognize ground, water, vegetation, roof, wall, roof element, and undefined object from airborne laser scanning data. Niemeyer et al. (2015) applied point-based Conditional Random Field (CRF) and segment-based CRF to recognize natural soil, road, gable roof, low vegetable, car, flat roof, façade, and tree from airborne laser scanning data. The above related studies show the advantages of features aggregation of different levels.

Although numbers methods for road facilities extraction have been reported, MLS software and automated algorithms for extracting them are still in relatively progress compared to the advancement of MLS hardware. Most existing road facilities extraction methods use local (point) or global (object) based features to recognize object candidates, and so can only extract limited types of objects with relatively low recognition rate, particularly for incomplete and small objects. To overcome these drawbacks, this paper proposes a semantic labeling framework by calculating multiple aggregation levels (point-segment-object) of features and contextual features to recognize road facilities from large-scale road scene point clouds. The main contributions of the proposed method are as follows:

(1) Compute multiple aggregation levels (point-segment-object) of features for road facilities recognition to improve precision and recall of road facilities recognition; and
(2) Design a series of contextual features (e.g. relative positions, relative directions, and spatial patterns) to improve the recognition performance of inter-class similar objects and incomplete objects.

Following this introduction, the key components of the proposed method are elaborated. Then the proposed method is validated in experimental studies before conclusion is drawn at the end of the paper.

2. Methodology

Fig. 1 shows the workflow of the proposed method. The proposed method first identifies ground and non-ground points, and extracts road surfaces facilities from ground points. Then, the non-ground points are then segmented into individual object candidates with a successive scheme that includes pointwise classification, multi-rule segmentation, and adjacent segments merging. Finally, multiple aggregation levels of features and contextual features are computed to facilitate recognition of each object candidate.

2.1. Road surface facilities extraction

The proposed method first identifies ground and non-ground points of each road segment using the method of Hernández and Matcotegui (2009). Then the ground points are segmented into several planes using the random sample consensus (RANSAC) algorithm (Fischler and Bolles, 1981), and road surface planes are identified according the following rules:

- Road surface segments are large planes at a certain distance below the 3D trajectory of the vehicle;
- The normal vectors of road surface segments are approximately parallel to the Z-direction.