

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

Available online 3 March 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 combining 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.

© 2017 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

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 Hyppä, 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/

* 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).

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 (Muñoz, 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).

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. Babahajani 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.

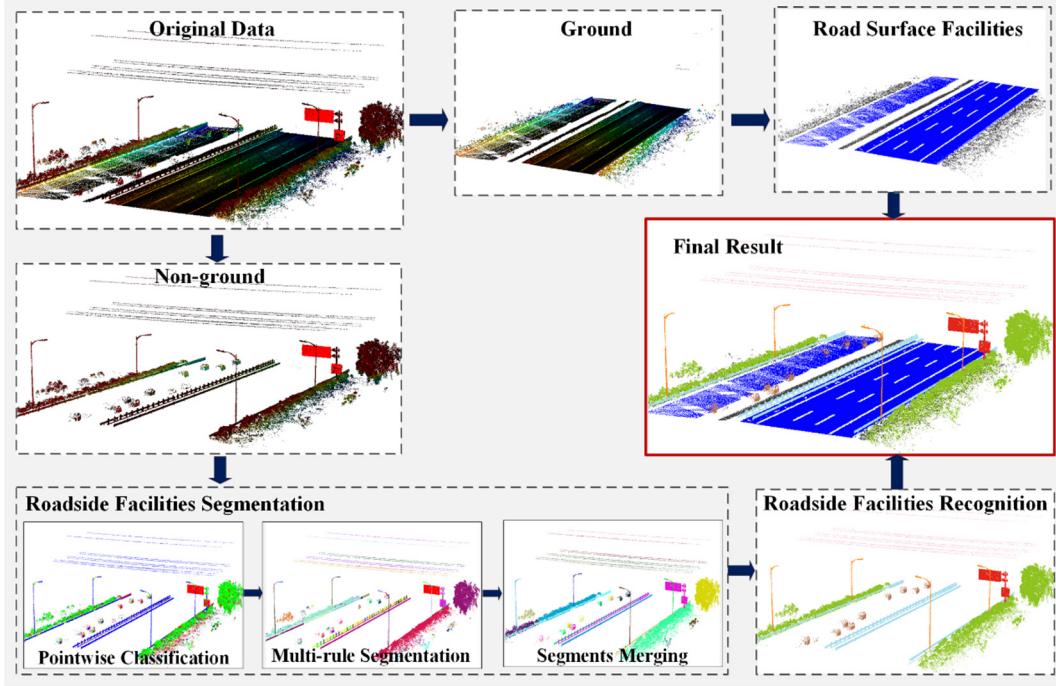


Fig. 1. Overview of the proposed framework.

The α -shape algorithm (Edelsbrunner and Mücke, 1994) is adopted to extract the boundaries of each road surface, and road markings are extracted from the identified road surface points using the method of Yang et al. (2012).

2.2. Roadside facilities extraction

2.2.1. Multi-rule region growing for roadside facilities segmentation

The proposed method segments the non-ground points into individual object candidates with a successive scheme that includes non-ground points classification, multi-rule segmentation, and adjacent segment merging. First, the non-ground points are classified into three categories (linear, planar, and spherical points) according to local dimensional features (Demantke et al., 2011) of each point. Dimensional features are estimated for each non-ground point p_j , using Eqs. (1) and (2), and it is placed into one of the primitive types (linear: $Label_j = 1$, planar: $Label_j = 2$, or spherical: $Label_j = 3$) using Eq. (3). The normal Nd_j , and principal direction Pd_j of p_j are then obtained, corresponding to the eigenvectors of the smallest eigenvalue λ_3 and the largest eigenvalue λ_1 , respectively.

$$f_j^{1D} = \frac{\sqrt{\lambda_1} - \sqrt{\lambda_2}}{\sqrt{\lambda_1}}, f_j^{2D} = \frac{\sqrt{\lambda_2} - \sqrt{\lambda_3}}{\sqrt{\lambda_1}}, f_j^{3D} = \frac{\sqrt{\lambda_3}}{\sqrt{\lambda_1}}, \quad (1)$$

where f_j^{1D} , f_j^{2D} , and f_j^{3D} are the linear, planar, and spherical dimensional features of p_j ; and $\lambda_1, \lambda_2, \lambda_3$ ($\lambda_1 \geq \lambda_2 \geq \lambda_3 > 0$) are the eigenvalues of the covariance matrix M . The covariance matrix M is written as:

$$M_{3 \times 3} = \frac{1}{k_p} \sum_{i=1}^{k_p} (\vec{p}_i - \vec{p}_c)(\vec{p}_i - \vec{p}_c)^T \quad (2)$$

where p_i ($i = 1, 2, \dots, k_p$) are the adjacent points of p_j , and $\vec{p}_c = \frac{1}{k_p} \sum_{i=1}^{k_p} p_i$ is the average of the adjacent points, k_p is the number of adjacent points.

Each point is labeled by

$$Label_j = \arg \max_{d \in \{1,2,3\}} (f_j^{dD}), \quad (3)$$

where $Label_j$ is the label of point p_j .

Approaching the segmentation of non-ground points by seeking consistency around a single cue or parameter is likely to generate partial results (Barnea and Filin, 2013). Hence, the proposed method segments the non-ground points that share similar labels into connected components using the multi-rule region growing algorithm that exploits labels, directions, and spatial connectivity attributes of neighboring points. A point p_s is chosen randomly (the seed point), and a new empty segment is created. Then, the multi-rule region growing algorithm retrieves the k -nearest neighbor points p_i ($i = 1, 2, \dots, k_p$) of p_s using the kd-tree. Neighbors that satisfy one of the following conditions are added to the segment and to a temporary queue.

- $Label_i = Label_s = 1$, the angle between their principal directions is less than a fixed threshold σ_a ;
- $Label_i = Label_s = 2$, the angle between their normal directions is less than a fixed threshold σ_a ;
- $Label_i = Label_s = 3$.

The algorithm then loops over all the points in the temporary queue. Once the queue is empty, a point that has not yet been visited is randomly selected as a new seed, and operations are repeated until all points are visited. Each segment formed by the multi-rule region growing may be a complete object, but most of them are a part of a certain object. Finally, adjacent segments (where the minimum distance between the segments is less a threshold, r_s) are merged as meaningful geometric abstractions of objects. Calculating the adjacency relationships between the segments is time consuming. Hence the proposed method calculates the minimum hull of each segments to derive the adjacency relationships between the hulls as the adjacency relationships between the segments. The segmentation algorithm may cause under-segmentation for highly overlapped objects but is feasible

and efficient for the objects scattering away from each other in highway scenes.

2.2.2. Computing multiple aggregation levels of features and contextual features

After roadside facilities segmentation, non-ground points are segmented into individual object candidates, and the object candidates are further classified into several classes based on the multiple aggregation levels features and contextual features associated with each candidate object.

• Point based features

Point based features describe the geometric information of the local neighborhood around specific points, and are more robust to occlusion and clutter compared to object based features (Guo et al., 2014). Existing point based features such as normal based signature (Li and Guskov, 2005), exponential map (Novatnack and Nishino, 2008), rotational projection statistics (Guo et al., 2013), and 3D shape context (Frome et al., 2004) are sensitive to noise and point density of the point cloud (Wang et al., 2015). Hence, the fast point feature histogram (FPFH) (Rusu et al., 2009) is computed as point based features because of its efficiency and discrimination. According to Rusu et al. (2009), the FPFH of each point p_s is calculated as follows: (1) search the k-nearest points of p_s ; (2) construct a local reference frame (LRF) for each of the k-nearest points p_k and p_s . Fig. 2a shows the constructed LRF, where N_s, N_t are the normal vectors of p_s, p_t respectively, $p_t - p_s$ is the direction of the line connecting the two points; (3) calculate the three angle differences (α, θ, ϕ) of each pair of points p_s and p_k , as shown in Fig. 2a; (4) generate the simplified point feature histogram (SPFH) of p_s with the quantization of the three angle differences (α, θ, ϕ) of all the pairs into a histogram; (5) calculate the FPFH of p_s by Eq. (4), as shown in Fig. 2b

$$\text{FPFH}(p_s) = \text{SPFH}(p_s) + \frac{1}{k} \sum_{i=1}^k \frac{1}{\omega_k} \cdot \text{SPFH}(p_k) \quad (4)$$

where the weight ω_k is the distance between the query point p_s and a neighboring point p_k . The implementation of FPFH with the optimum parameters (feature dimension $F_d = 33$) is available in the Point Cloud Library (PCL) Version 1.7.1 (Rusu and Cousins, 2011). The average of the FPFHs of all points of each object candidate is computed to describe its shape features. Typical FPFHs of different objects are shown in Fig. 3.

• Segment based features

Segment based features describe the shapes, sizes, and directions of object components. For example, most traffic signs are composed of vertical linear segments and vertical planar segments, whereas trees are composed of vertical linear segments (trunk) and spherical segments (crown) (Yang and Dong, 2013). Hence, the dimensional features, minimum bounding box, normal and principal directions of each segment are computed to describe their shapes (linear, planar, or spherical segments), sizes (length, width, and height) and directions (vertical or horizontal segments), respectively. Typical segment based features for different objects are shown in Fig. 4.

• Object based features

Object based features effectively and concisely describe the entire 3D object (Castellani et al., 2008), and are more robust to noise and varying point density compared to point based features (Guo et al., 2014). Existing object based features such as geometric 3D moments (Paquet et al., 2000), shape distribution (Osada et al., 2002), viewpoint feature histogram (Rusu et al., 2010), and potential well space embedding (Shang and Greenspan, 2010) ignore shape details and require priori segmentation of the object from the scene. Therefore, object based features have difficulty in recognizing partially visible or incomplete objects from cluttered scenes. In this paper, the viewpoint feature histogram (VFH) and minimum bounding box are computed to describe the shapes and sizes (length, width, and height) of each object candidate. According to Rusu et al. (2010), the VFH of each object is calculated as follows: (1) Calculate the centroid p_c of the object and the distance d from each point of the object to the centroid, as shown in Fig. 5a. (2) construct a local reference frame (LRF) for each point p_t and centroid p_c . As illustrated in Fig. 5b, N_c, N_t are the normal vectors of p_c, p_t respectively, $p_t - p_c$ is the direction of the line connecting the two points. (3) Calculate the three angle differences (α, θ, ϕ) of each pair of points p_c and p_t , as shown in Fig. 5b. (4) Calculate the include angles β between the viewpoint direction and the normal of each point. As shown in Fig. 5c, V_p is the viewpoint, and $p_t - V_p$ is the viewpoint direction. (5) Generate the VFH of the candidate object with the quantization of the four angle differences and distances ($\beta, \alpha, \theta, \phi, d$) of all the pairs into a histogram. The default VFH implementation uses 128 binning subdivisions for the viewpoint component β , 45 binning subdivisions for each of the three angles (α, θ, ϕ), and another 45 binning subdivisions for the distances d between

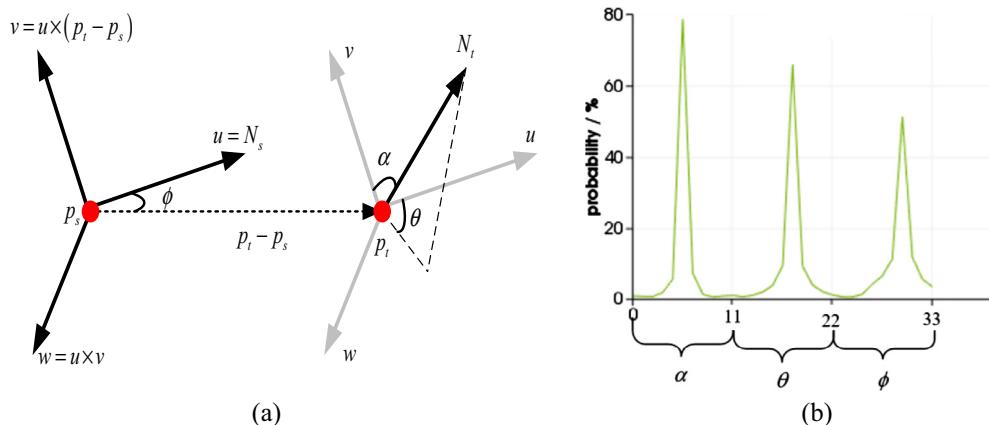


Fig. 2. Example of Fast Point Feature Histograms calculation: (a) local reference frame construction, (b) typical Fast Point Feature Histogram.

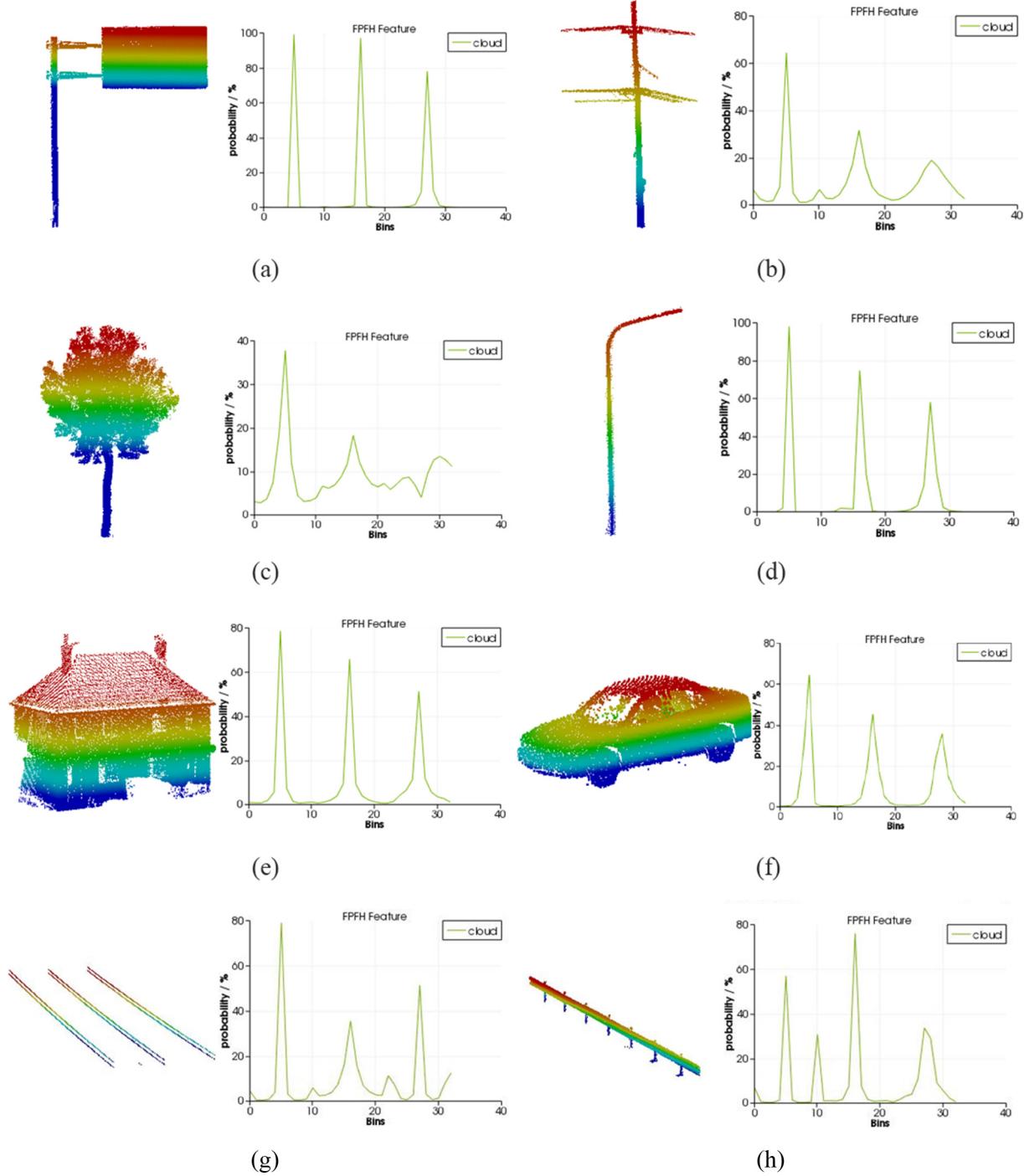


Fig. 3. Typical FPFHs of different types of objects: (a) traffic sign, (b) utility pole, (c) tree, (d) street lamp, (e) building, (f) car, (g) power line, and (h) guardrail.

each point and the centroid p_C , resulting in a 308-byte array of float values (Rusu and Cousins, 2011), as shown in Fig. 5d. Typical VFHs of different objects are shown in Fig. 6.

• Contextual features

In some cases, it is difficult to accurately recognize multiple types of object candidates using only geometric features due to similarities between specific inter-class objects and incomplete objects. To overcome the above problems, we compute three types of context features: relative position, relative direction, and spatial distribution pattern.

Relative position: Based on contextual knowledge of the road environment, relative position is an important cue in object recognition (Teo and Chiu, 2015). For example, roadside objects are placed along the boundaries of roads and on the ground rather than floating in air. Therefore, we utilize 2D distances from the center of an object candidate to road boundaries to describe the relative horizontal position. And we compute height differences between the lowest elevation point of an object candidate and the road surface to describe the relative vertical position. Fig. 7a shows an example where $D_{\text{lamp}1}$, $D_{\text{lamp}2}$, $D_{\text{car}1}$, and $D_{\text{car}2}$ are the 2D distances of the street lamp and car to the road boundaries, respectively; and H_{lamp} and H_{car} are the height differences from their lowest elevation points to the road surface.

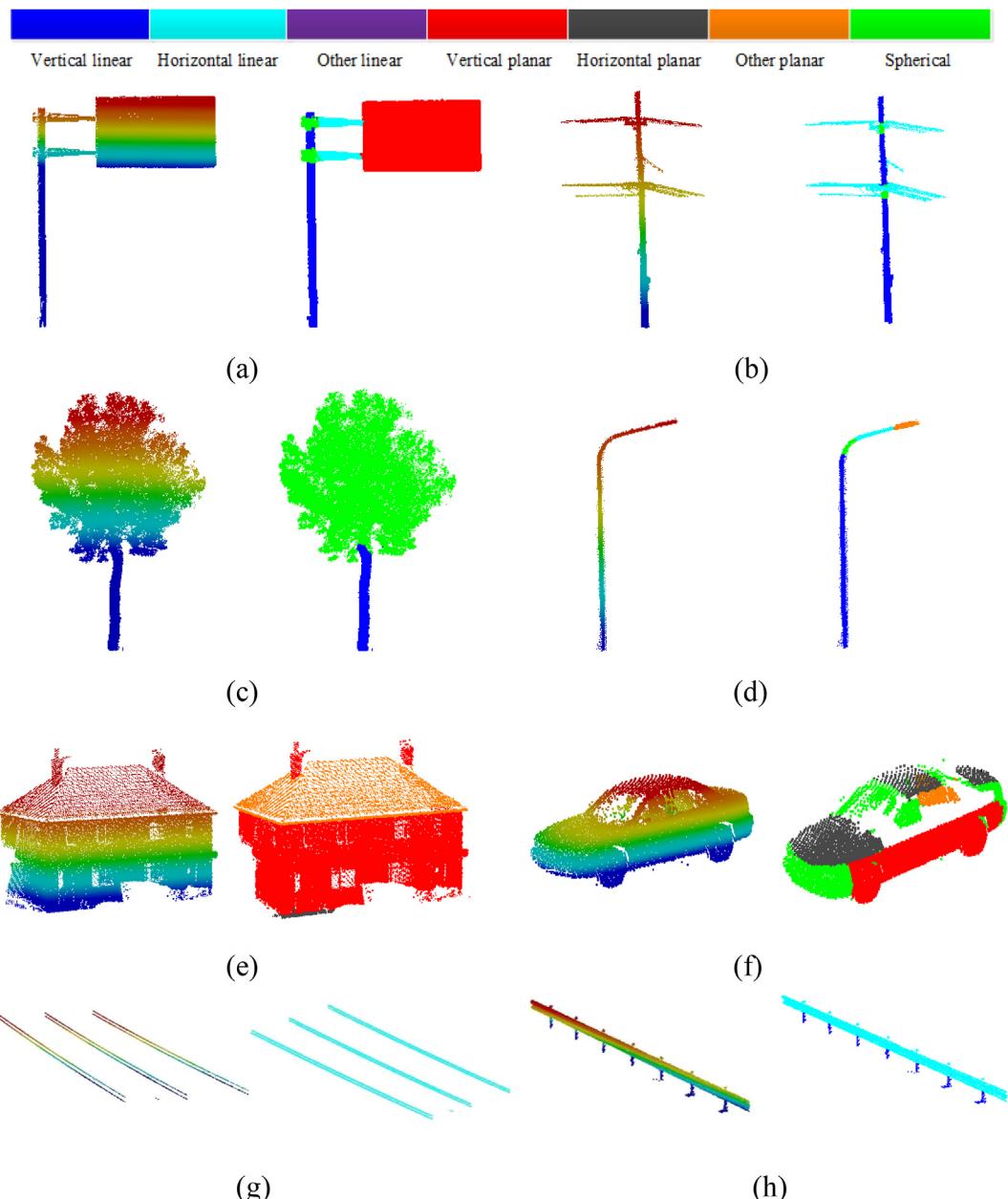


Fig. 4. Typical segment based features of different types of objects: (a) traffic sign, (b) utility pole, (c) tree, (d) street lamp, (e) building, (f) car, (g) power line, and (h) guardrail.

Relative direction relation: The relative direction relation describes the relative direction between object candidates and the principal direction of the corresponding road segment. For example, the principal directions of guardrails and traffic lane lines are approximately parallel to the principal direction of the corresponding road segment, and the normal directions of traffic signs are approximately parallel to the principal direction of the corresponding road segment. Hence, the angles between the principal direction of the road segment with the normal and principal direction of each object candidate are computed to describe the relative direction relation between the object candidates and the corresponding road segment. An example is shown in Fig. 7b, where P_r , P_g and N_t are the principal directions of the road segment and guardrail, and the normal direction of the traffic sign, respectively, α is the angle between the normal direction of traffic sign and the principal direction of road segment, and β is the angle between the principal directions of the guardrail and the road segment.

Spatial distribution pattern: Spatial distribution pattern describes the regularity of object candidates. For example, street lamps are often found at regular intervals along the roadside, and the spatial distributions of road markings also have regular patterns, especially in highway scenes. Hence the 2D distances between the object candidate and its O_k nearest neighbor objects are computed to describe the surrounding spatial distribution. In this paper, O_k was set to 10. An example is shown in Fig. 7c, where D_i ($i = 1, 2, \dots, O_k$) is the 2D distance between the current candidate object and its adjacent objects.

2.2.3. Candidate objects recognition using support vector machine (SVM)

The calculated multiple aggregation levels of features and contextual features are scaled into the range [0, 1]. For recognizing candidate objects into multiple classes (e.g. buildings, trees, traffic signs, street lamps, cars, guardrails, utility poles, power lines, etc.),

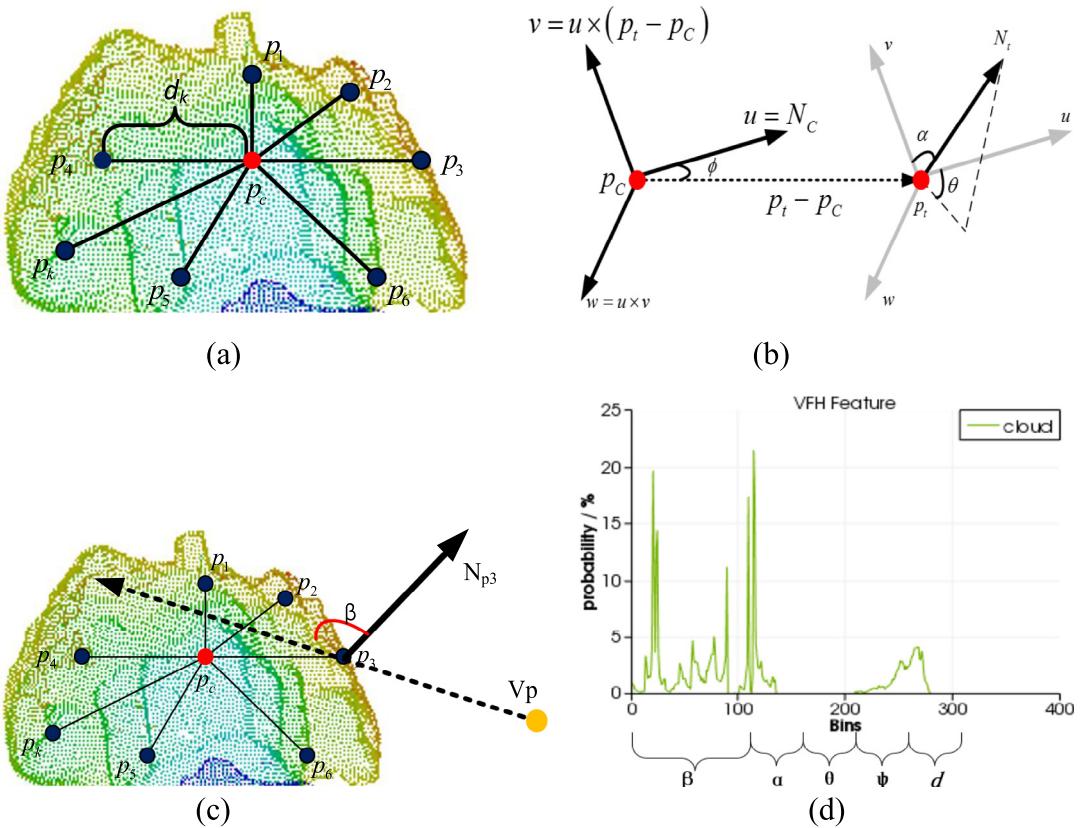


Fig. 5. Example of Viewpoint Feature Histograms calculation: (a) centroid p_c and the distance d from each point to the centroid, (b) the local reference frame (LRF) and the triple angle differences (α, θ, ϕ), (c) include angle β between the viewpoint direction and the normal of each point, (d) typical Viewpoint Feature Histograms.

A SVM classifier was trained on the manually labeled training data set. In order to increase the sigma-completeness and representativeness of the training data set, we labeled candidate objects from datasets with varying point densities, point position precisions, and percentage of outliers as the training data set. More specifically, the candidate objects belonging to uninterested classes (e.g. the grass, bush, billboard and so on) in the training data set were labeled as “others”.

The C-SVM version with one-versus-one classification strategy (Hsu and Lin, 2002) and the radial basis function (RBF) kernel in LIBSVM package (Chang and Lin, 2011) were used to perform training and prediction. The C-SVM version with an RBF kernel has two hyperparameters: complexity parameter C and inverse kernel width γ . The cross validation strategy was adopted to determine the optimal choice of the parameters. More specifically, the training data set was randomly divided into two folds of equal size. One fold was used for training the SVM classifier using the current parameter choices and the other fold was used for evaluating the accuracy of the SVM classifier. After the training process, the remaining unrecognized candidate objects were classified by the trained SVM classifier.

3. Results and analysis

3.1. Data description

The proposed method performance was checked using three highway scene datasets with different average point densities, point position precisions, and percentage of outliers. The average point density is calculated as follows: First, we randomly selected 1000 regions with a certain bound. Then, we calculate the point

density of each region (points/m²). Finally, we calculate the average point density of the 1000 regions as the average point density of the point clouds. Dataset A was captured using a Riegl VMX-450 MLS system on the Beijing-Chengde expressway, covering road length of approximately 6.0 km. The VMX-450 MLS system integrates two Riegl VQ-450 laser scanners with 360° field of view (FOV), four high resolution digital cameras, and a position and orientation system. Dataset B was captured using a SSW-MMTS MLS system on the Shenyang-Haerbin expressway, covering road length of approximately 50.8 km. The SSW-MMTS MLS system is equipped with one laser scanner with 360° FOV, a navigation and positioning system, and six digital cameras (22 million pixels). Dataset C was captured using a Trimble MX2 MLS system on the Beijing-Miyun expressway, covering road length approximately 79.8 km. The Trimble MX2 MLS system is composed of two laser scanners with 360° FOV, a combined Trimble Applanix GNSS, and inertial geo-referencing module for precise positioning. Table 1 shows the detailed description of the MLS systems and point clouds, and Fig. 8 shows an overview of the three MLS systems and point cloud scenes.

3.2. Road facilities extraction

Table 2 shows the parameters of road surface features and roadside objects extraction, set by trial and error. It was found in trials that the object recognition results were insensitive to parameter values thresholds in a certain range. More specifically, the desired object recognition results can be achieved with the values of parameters $k_p, \sigma_a, r_s, F_r, O_k$ ranging from 10 to 40, from 5° to 15°, from 0.2 m to 0.5 m, from 0.3 m to 0.7 m, and from 5 to 20, respectively. And the performance of object recognition deteriorates

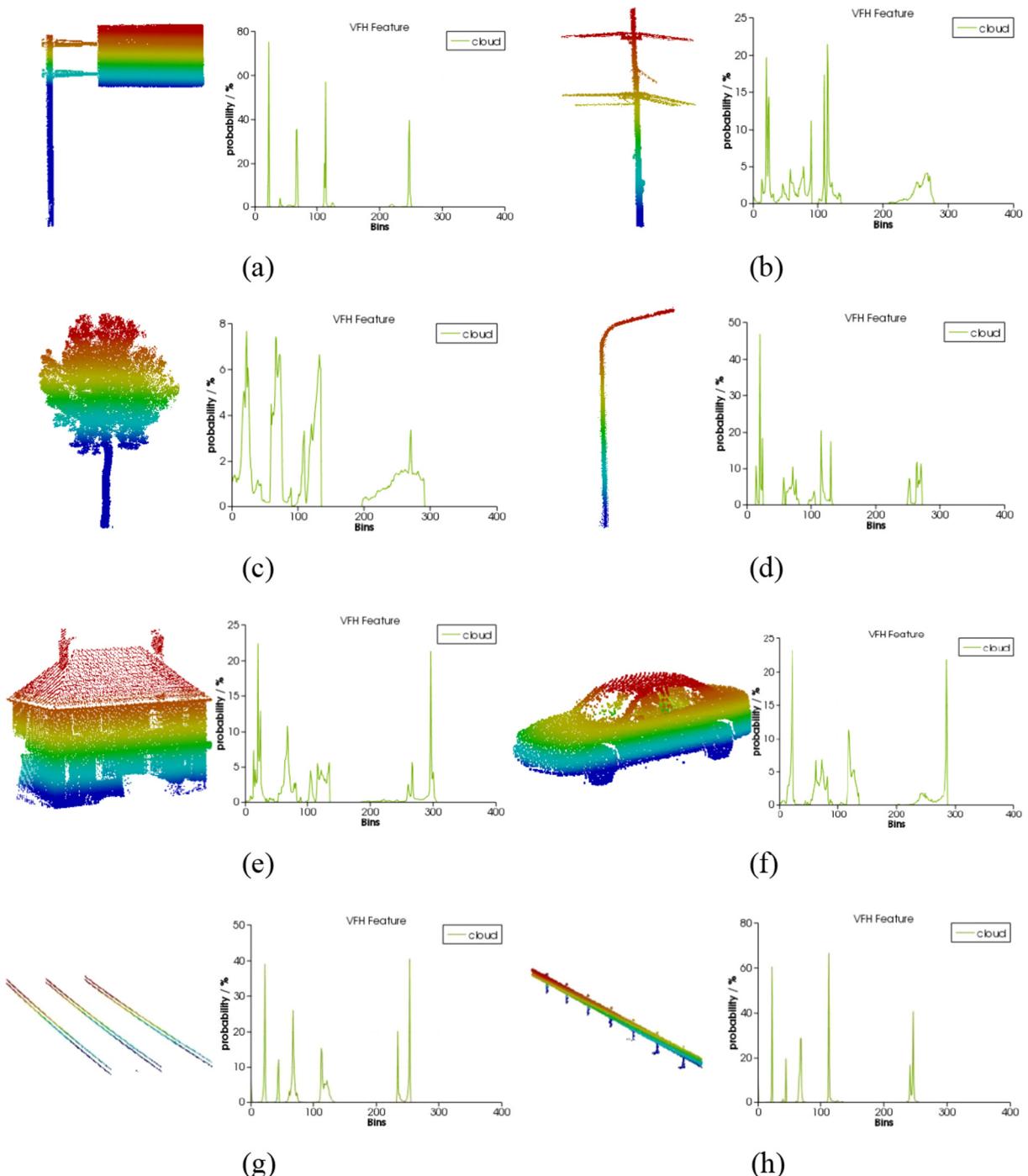


Fig. 6. Typical VFHs of different kinds of objects: (a) traffic sign, (b) utility pole, (c) tree, (d) street lamp, (e) building, (f) car, (g) power line, and (h) guardrail.

sharply if the parameter values change over the certain range. Figs. 9–11 show the road facilities detection outcomes for selected point clouds testdata 1, testdata 2, testdata 3 from Datasets A, B, and C, respectively. Figs. 9a, 10a, and 11a show the three selected point clouds, colored by the elevation of each point. After ground and non-ground points identified by the method of Hernández and Matcotegui (2009), the road surfaces facilities were extracted from the ground points. Figs. 9b, 10b, and 11b show road surface extraction outcomes, where gray, blue and white points represent points from ground, road surface, and road markings respectively. Figs. 9c, 10c, and 11c show roadside facilities segmentation out-

comes, where each object is drawn in one color. Figs. 9d, 10d, and 11d show object candidate recognition outcomes, dotted in different colors. Fig. 12 shows more details of the roadside objects recognition outcomes. Fig. 12a shows the recognition outcomes for small traffic signs, Fig. 12b shows recognition outcomes for moving cars, and Fig. 12c shows recognition outcomes for incomplete objects, including traffic signs, street lamps, and cars. The recognition results show that the proposed method provides good performance in recognizing roadside objects, even small traffic signs, moving cars, and incomplete traffic signs cars, and street lamps. However, there remain a few wrong cases of roadside

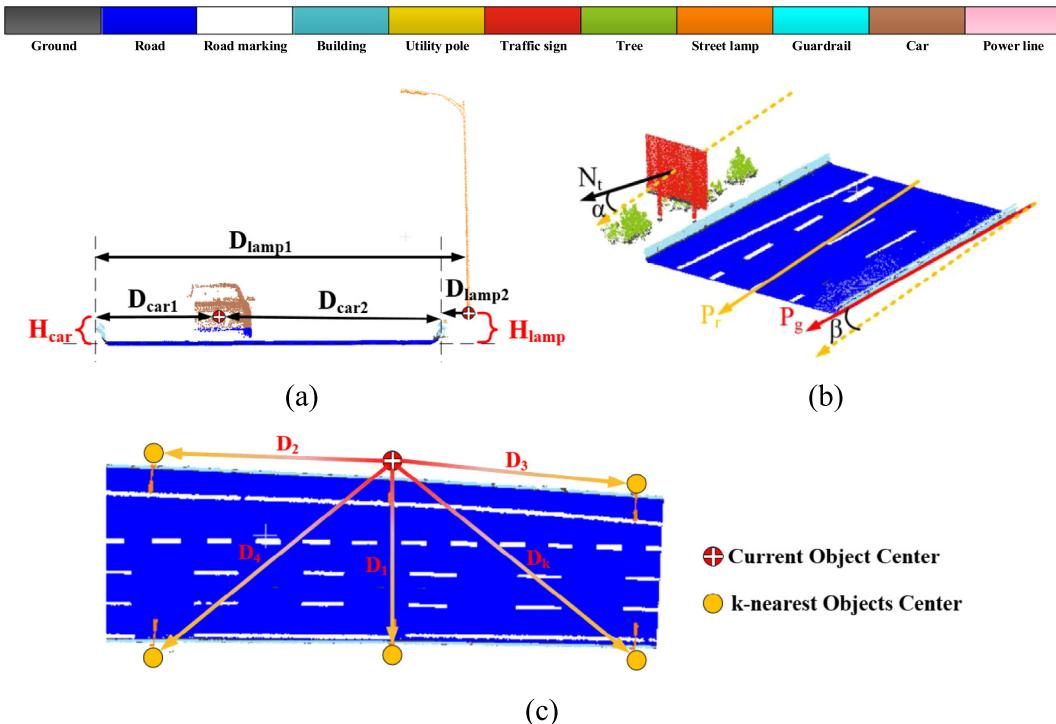


Fig. 7. Schematic diagrams of context based features: (a) Relative position, (b) Relative direction, and (c) Spatial distribution pattern.

Table 1
MLS systems and point clouds description.

	Dataset A	Dataset B	Dataset C
MLS system	VMX-450	SSW-MMTS	Trimble MX2
Number of laser scanners	2	1	2
Maximum valid range (m)	800	300	250
Maximum measurement rate	1100 K	200 K	36 K
Maximum line scan rate	400	100	20
Number of points (million)	167.8	825.4	564.2
Length (km)	6.0	50.8	79.8
Average point density (points/m ²)	204	82	33
Point position accuracy (cm)	0.5–0.8	1–2	2–3
Percentage of outliers (%)	0.5	1.5	3.2

facilities recognition. Most of the wrong cases were caused by under-segmentation that mixed with attached objects (Fig. 13), and incomplete objects (Fig. 14). Fig. 13 shows the wrong case of a utility pole mixed with tree is misclassified as a tree. Fig. 14 shows the wrong case of a tree without canopy is misclassified as others.

3.3. Evaluation of the proposed method

The proposed method was implemented on a computer with 16.0 GB RAM and an Intel (R) Core (TM) i7-6700HQ @ 2.60 GHz CPU. Table 3 lists time performance of the proposed method for road facilities extraction, including time costs for each step. It took approximately 108 min to recognize road facilities from MLS point clouds covering 79.8 km, and containing 564.2 million points. The proposed method has high time efficiency in road facilities recognition, particularly for the large data volume MLS point clouds. The high time efficiency of the proposed method in road facilities extraction owes to the multi-threading parallel processing strategy, which partitions point clouds into several road segments along the road centerlines, and distributes them within a multi-thread computing environment.

Precision and recall were calculated to evaluate the performance of the proposed method for recognizing roadside facilities on these three data sets,

$$\text{Precision} = \frac{TP}{TP + FP} * 100\%, \quad (5)$$

$$\text{Recall} = \frac{TP}{TP + FN} * 100\%, \quad (6)$$

where TP is the number of true positives, i.e., the number of detected objects with correct classes; FP is the number of false positives, i.e., the number of detected objects with incorrect classes; and FN is the number of false negatives, i.e., the number of undetected objects.

Table 4 shows the precision and recall of road facilities recognition from these three data sets. The proposed method achieves good performance in recognizing road facilities with an average precision and recall of (90.6%, 91.2%), (89.7%, 90.1%) and (86.3%, 87.5%) for the three data sets, respectively. Precision and recall decline slightly with the sharp decreasing point position accuracy, and point density. Thus, the proposed method is robust to varying point density, point position accuracy, and percentage of outliers in the point cloud.

3.4. Comparative studies

3.4.1. Single and multiple aggregation levels of features

Single level of features and the proposed multiple aggregation levels of features were compared to further verify the effectiveness of the proposed method for object recognition, as shown in Table 5.

Among single level of features, object based features is the most distinctive with precision and recall 76.0% and 77.4%, respectively, and the highest precision and recall for buildings, 90.3% and 91.6%, respectively. Point based features is the most effective features for power lines classification with precision and recall 85.2% and 87.1%, respectively. Segment based features is also distinctive for

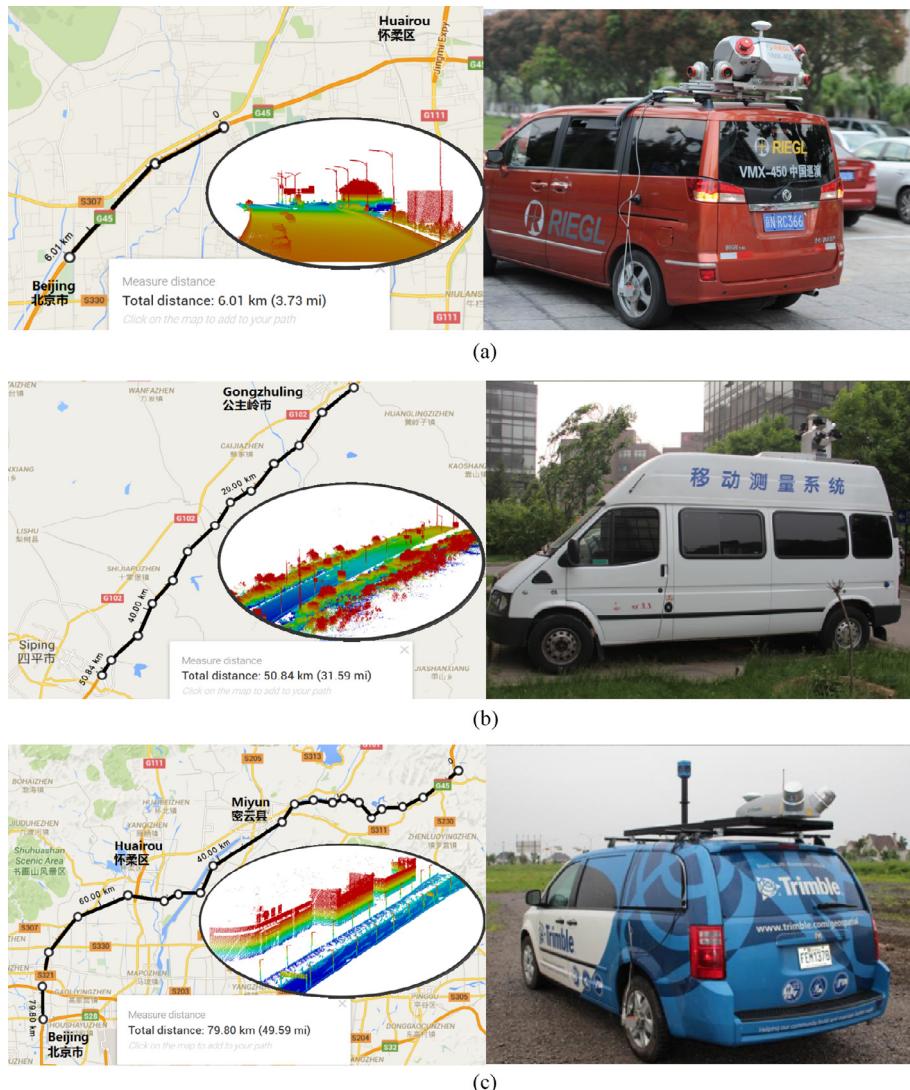


Fig. 8. Overview of the MLS systems and point clouds: (a) Dataset A and the VMX-450 MLS system, (b) Dataset B and SSW-MMTS MLS system, and (c) Dataset C and Trimble MX2 MLS system.

classification of utility poles (precision and recall 86.7% and 87.4%) and traffic signs (precision and recall 84.1% and 86.3%), respectively. Contextual features were the most indistinctive features for roadside objects classification with lowest average precision and recall of 24.7% and 31.9%, respectively. However, contextual features were very important for classification of incomplete objects and objects with similarities between specific inter-classes.

The combination of point and segment based features significantly improved classification precision and recall for utility poles, traffic signs, and trees, compared to point based features. The addition of object based features evidently raised the precision and recall for buildings from 80.3% and 82.5% to 95.4% and 96.2%, respectively. Adding contextual features significantly improved classification precision and recall of cars from 67.8% and 69.4% to 86.4% and 87.2%, traffic signs from 87.2% and 87.8% to 92.5% and 92.6%, and street lamps from 83.2% and 85.4% to 90.8% and 91.7%, respectively. The improvement of the classification precision and recall for cars, traffic signs and street lamps owes to the contextual features that are beneficial for the recognition of moving cars, small size traffic signs and incomplete street lamps. The proposed method achieved the highest precision and recall, with precision and recall 90.7% and 91.8%, respectively.

3.4.2. Comparison with other methods

The proposed method was compared with the methods of Yang et al. (2015) and Lehtomäki et al. (2015) in terms of precision and recall for recognizing roadside facilities, as listed in Table 6. In this paper, we use the same training data with different features to train the classifiers and same testing data to evaluate the proposed method and other methods. Precision and recall numbers missing for other methods is caused by the limitation of the compared methods instead of datasets. More specifically, the method Yang et al. (2015) is designed to extract urban objects including buildings, streetlamps, trees, telegraph poles, traffic signs, cars, and enclosures, it cannot recognize guardrails and power lines. Hence the precision and recall numbers of guardrails and power lines are invalid. The method Lehtomäki et al. (2015) is designed to classify and recognize objects in road environment including trees, lamp post, traffic poles, car, hoarding and pedestrians, it cannot recognize buildings, utility pole, guardrails and power lines. Hence the precision and recall numbers of the corresponding objects are invalid.

Yang et al. (2015) mainly employs segment based features (e.g. dimensional features, direction and size of segments, and relationship between adjacent segments) and object based features (e.g.

Table 2

Parameters for experimental datasets.

Procedure	Parameter	Descriptor	Value
Road facilities segmentation	k_p	Number of adjacent points for dimensional feature calculation	25
	σ_a	Angular separation threshold between normal directions or principal directions of adjacent points for multi-rules region growing	10°
	r_s	Minimum distance threshold between candidate segments for segments merging	0.3 m
Road facilities recognition	F_r	Radius for FPFH calculation	0.5 m
	F_d	Feature dimension of FPFH	33
	V_d	Feature dimension of VFH	308
	O_k	Number of adjacent objects for spatial distribution pattern estimation	10

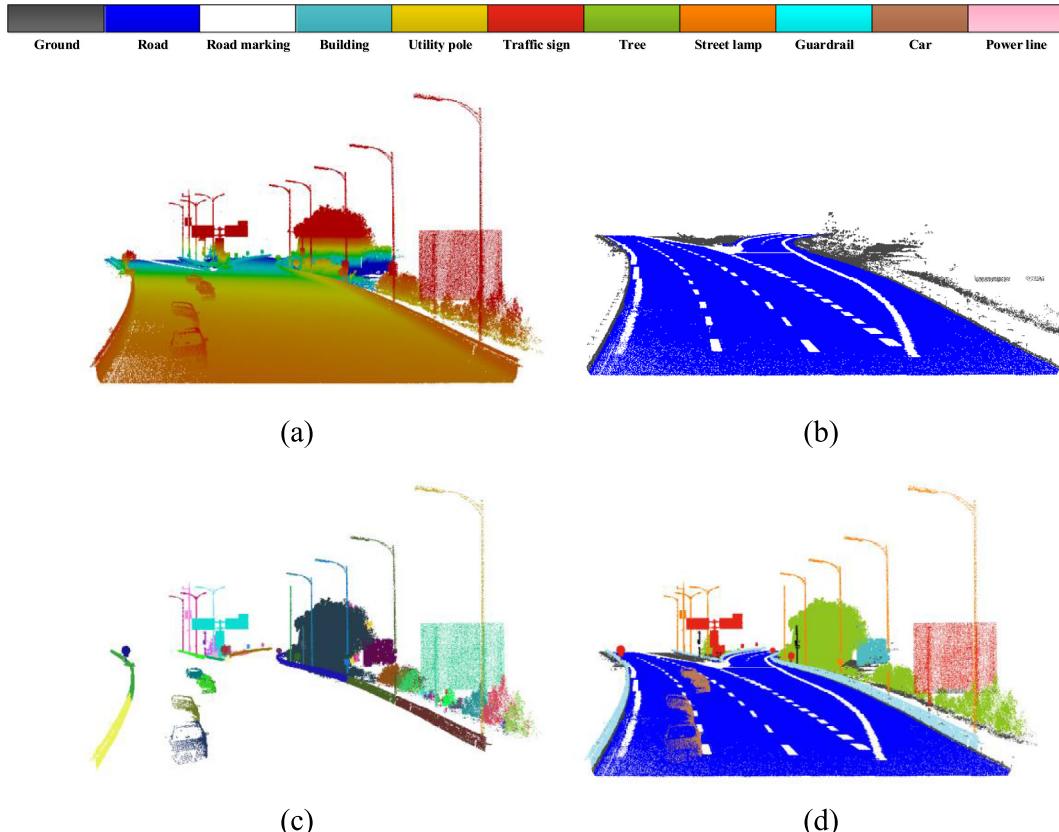


Fig. 9. Road facilities extraction outcomes from testdata 1: (a) original point clouds, (b) road surface facilities recognition, (c) roadside facilities segmentation, and (d) road facilities recognition.

size and number of segments comprising the object candidate) to extract objects from MLS point clouds. Although the method demonstrates good performance in extracting structural objects (e.g. buildings (90.1%, 90.4%), utility poles (85.7%, 86.8%), and traffic signs (82.9%, 83.6%)), it has difficulties in dealing with unstructured objects (e.g. cars (62.7%, 64.5%) and trees (79.4%, 81.6%)) and incomplete objects. The precision values (94.4% utility pole, 94.6% traffic sign, 91.0% trees, 93.5% street lamp) reported in Yang et al. (2015) are better than those of the proposed method (91.7% utility pole, 92.5% traffic sign, 84.9% trees, 90.8% street lamp). The reasons are caused by two aspects. Firstly, the datasets in Yang et al., 2015 were collected in early morning when there were few cars in the roads, and the datasets in this paper were collected in day time when there were many cars in the roads, resulting in more occlusions and varying point densities. Secondly, the small size traffic signs (as shown in Fig. 12a) and the utility pole, traffic sign, trees, and street lamp far away from the roadside are manually labeled as

“others” in the ground truth of Yang et al., 2015, instead they are labeled as their corresponding categories in the ground truth in this paper.

Lehtomäki et al. (2015) uses point based features (local descriptor histograms) and object based features (e.g. spin images and general features) to extract roadside objects from MLS point clouds. The method demonstrates good performance in extracting most roadside objects in terms of precision and recall. However it is difficult to deal with small objects and incomplete objects. The proposed method employs multiple aggregation levels of features derived from points, segments, and objects to improve feature distinctiveness. The proposed method also combines contextual features with multiple aggregation levels of features, improving recognition performance for incomplete and small objects. Thus, the proposed method has better precision and recall of object recognition than those of Yang et al. (2015), and Lehtomäki et al. (2015).

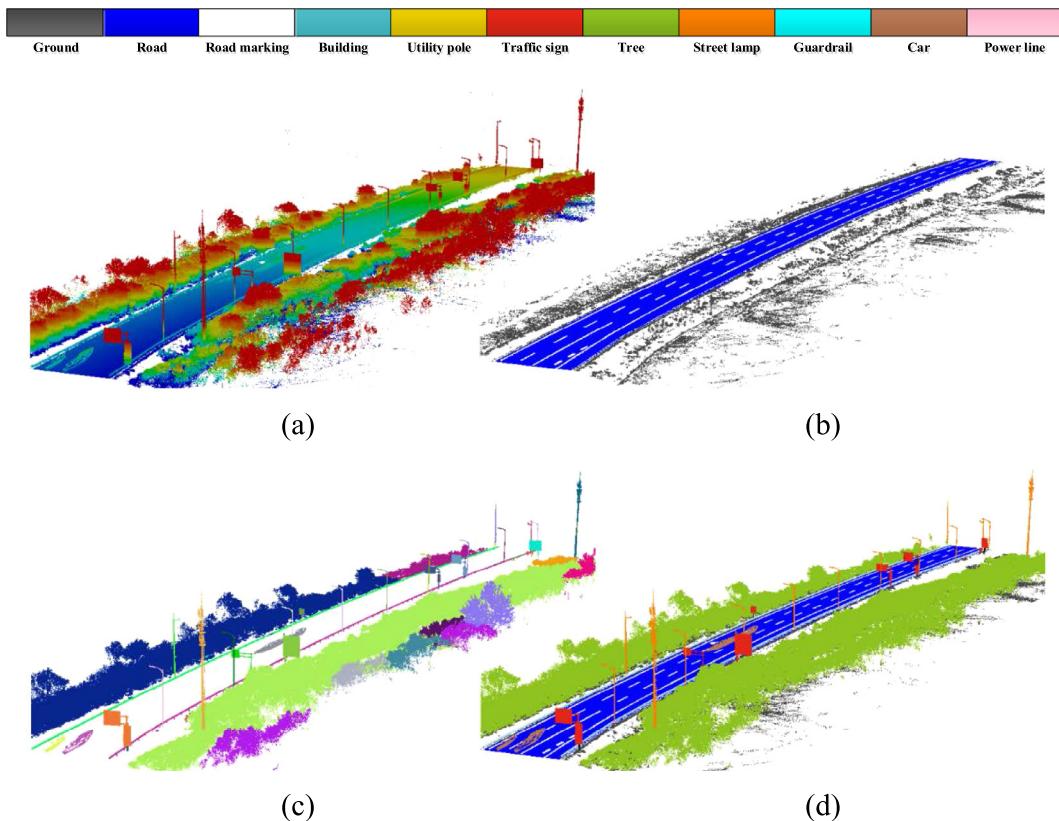


Fig. 10. Road facilities extraction outcomes from testdata 2: (a) original point clouds, (b) road surface facilities recognition, (c) roadside facilities segmentation, and (d) road facilities recognition.

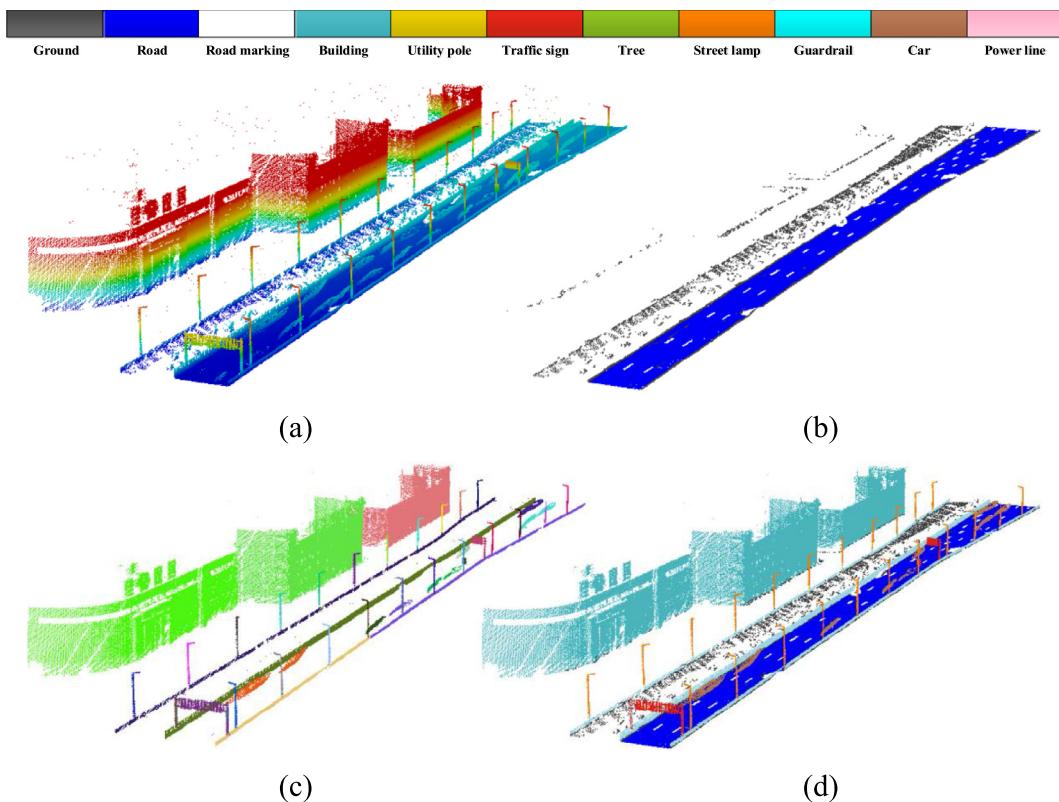


Fig. 11. Road facilities extraction result from testdata 3: (a) original point clouds, (b) road surface facilities recognition, (c) roadside facilities segmentation, and (d) road facilities recognition.

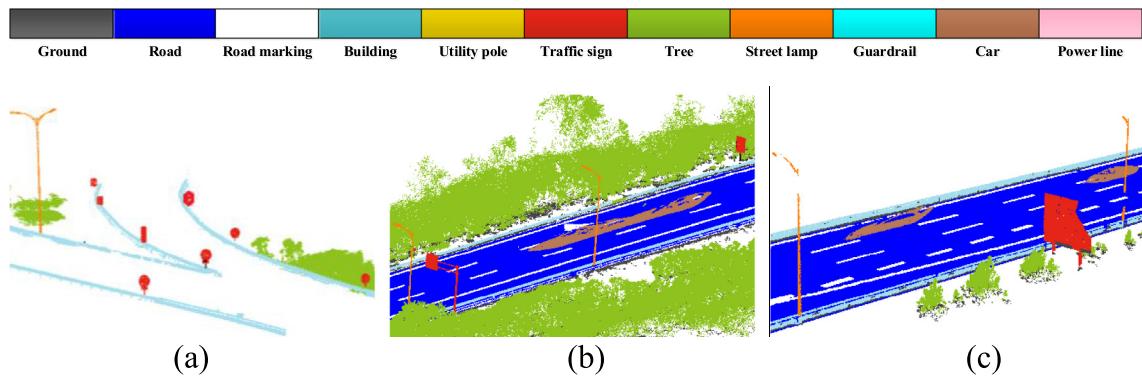


Fig. 12. Details of the roadside facilities recognition outcomes: (a) small traffic signs, (b) moving cars, and (c) incomplete street lamps, traffic signs, and cars.

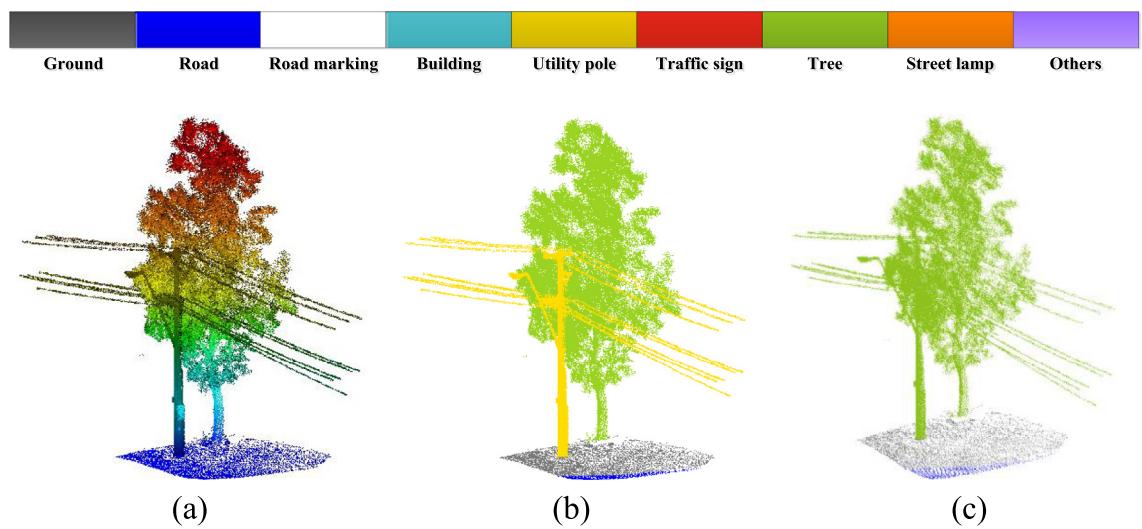


Fig. 13. Error case caused by under-segmentation: (a) original point clouds; (b) ground truth; (c) utility pole mixed with tree is misclassified as trees.

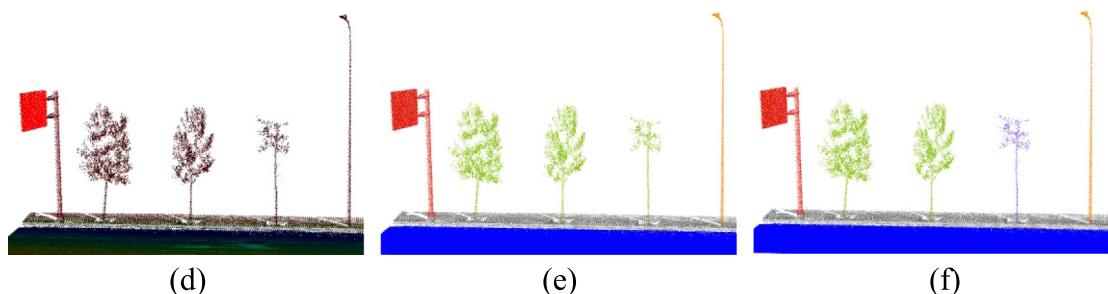


Fig. 14. Error case caused by incomplete object: (a) original point clouds; (b) ground truth; (c) trees without canopy is misclassified as others.

Table 3
Time performance of the proposed method.

Dataset	Road surface facilities extraction (s)		Roadside facilities extraction (s)			Total time cost (min)
	Road surface extraction	Road markings extraction	Roadside facilities segmentation	Features computation	Object recognition	
A	116	63	407	158	12	12.6
B	983	566	3652	1248	84	108.9
C	554	345	2108	732	53	63.2

Table 4

Precision and Recall of the proposed method for road facilities recognition.

	Dataset A		Dataset B		Dataset C	
	Precision %	Recall %	Precision %	Recall %	Precision %	Recall %
Building	96.2	96.8	95.9	96.4	94.3	95.5
Utility pole	90.7	91.3	90.4	91.1	88.6	88.9
Traffic sign	92.5	92.9	92.3	92.7	86.2	87
Tree	83.2	83.6	81.4	82.0	79.4	80.6
Streetlamp	90.6	91.3	89.2	91.3	87.7	88.9
Guardrail	93.2	94.5	93.1	93.8	85.8	87.3
Power line	90.1	90.7	89.9	89.6	84.3	86.9
Car	86.1	86.5	85.9	88.1	83.1	83.7
Average	90.6	91.2	89.7	90.5	86.3	87.5

Table 5

Recognition performance of single and multiple aggregation levels of features.

	PF		SF		OF		CF		PF + SF		PF + SF + OF		PF + SF + OF + CF	
	Precision %	Recall %	Precision %	Recall %	Precision %	Recall %								
Building	73.6	75.2	43.6	45.2	90.3	91.6	25.6	32.7	80.3	82.5	95.4	96.2	95.9	96.4
Utility pole	66.8	67.9	86.7	87.4	78.5	79.7	9.4	20.3	88.2	89.3	90.5	91.2	91.7	91.8
Traffic sign	58.5	59.4	84.1	86.3	82.4	84.5	8.9	23.8	86.1	87.4	87.2	87.8	92.5	92.6
Tree	65.3	67.4	60.5	63.4	74.3	76.7	12.7	21.2	76.3	77.2	83.4	87.5	84.9	88.6
Streetlamp	75.5	76.1	52.5	56.9	76.2	78.6	38.4	39.5	82.0	83.1	83.2	85.4	90.8	91.7
Guardrail	74.3	76.1	54.8	56.3	72.8	74.1	16.7	27.3	80.3	82.1	90.6	91.3	93.7	94.3
Power line	85.2	87.1	75.6	78.4	80.2	81.4	32.4	35.3	87.3	87.7	87.5	88.0	89.7	91.8
Car	52.4	56.8	41.2	45.1	53.1	52.4	53.6	55.1	63.2	65.1	67.8	69.4	86.4	87.2
Average	68.9	70.7	62.4	64.9	76.0	77.4	24.7	31.9	80.5	81.8	85.7	87.1	90.7	91.8

PF: Point based feature, SF: Segment based feature, OF: Object based feature, CF: Context based feature, PF+SF: combination of PF and SF, PF + SF + OF: combination of PF, SF and OF, PF + SF + OF + CF: combination of PF, SF, OF and CF.

Table 6

Performance comparison between the proposed method and others.

	Yang et al. (2015)		Lehtomäki et al. (2015)		Proposed method	
	Precision %	Recall %	Precision %	Recall %	Precision %	Recall %
Building	90.1	90.4	\	\	95.9	96.4
Utility pole	85.7	86.8	\	\	91.7	91.8
Traffic sign	82.9	83.6	91.4	91.8	92.5	92.6
Tree	79.4	81.6	83.1	84.3	84.9	88.6
Streetlamp	85.9	86.3	90.1	91.3	90.8	91.7
Guardrail	\	\	\	\	93.7	94.3
Power line	\	\	\	\	89.7	91.8
Car	62.7	64.5	79.2	84.5	86.4	87.2
Average	81.1	82.2	86.0	88.1	90.7	91.8

4. Conclusions

Automatic and accurate recognition of road facilities is a requirement for road infrastructures inventory and intelligent transportation related applications. This paper proposes a semantic labeling framework by integrating multiple aggregation levels (point-segment-object) of features and contextual features for road facilities recognition from large scale highway scene point clouds, resulting in improved recognition accuracy for candidate objects, particularly for incomplete and small objects. The accuracy of the proposed method was evaluated experimentally, validated that the proposed method improved recognition precision and recall for road facilities. Future work will focus on incorporating additional features such as spectral information and crowd sourced data into the object segmentation and recognition scheme.

Acknowledgements

The work described in this paper was jointly supported by the NSFC projects (No. 41531177, No. 41071268), and the project (No. 2016YFB0502300).

References

- Abadi, A., Rajabioun, T., Ioannou, P.A., 2015. Traffic flow prediction for road transportation networks with limited traffic data. *Intell. Transport. Syst. IEEE Trans.* 16 (2), 653–662.
- Babahajani, P., Fan, L., Gabbouj, M., 2014. Object recognition in 3D point cloud of urban street scene. In: *Computer Vision-ACCV 2014 Workshops*. Springer International Publishing, pp. 177–190.
- Barnea, S., Filin, S., 2013. Segmentation of terrestrial laser scanning data using geometry and image information. *ISPRS J. Photogram. Remote Sens.* 76, 33–48.
- Brodu, N., Lague, D., 2012. 3D terrestrial lidar data classification of complex natural scenes using a multi-scale dimensionality criterion: applications in geomorphology. *ISPRS J. Photogram. Remote Sens.* 68, 121–134.
- Broggi, A., Buzzoni, M., Debattista, S., et al., 2013. Extensive tests of autonomous driving technologies. *Intell. Transport. Syst. IEEE Trans.* 14 (3), 1403–1415.
- Caddell, R., Hammond, P., Reinmuth, S., 2009. Roadside Features Inventory Program. Washington State Department of Transportation.
- Castellani, U., Cristani, M., Fantoni, S., Murino, V., 2008. Sparse points matching by combining 3D mesh saliency with statistical descriptors. *Comput. Graph. Forum* 27 (2), 643–652. Blackwell Publishing Ltd.
- Chang, C.C., Lin, C.J., 2011. LIBSVM: a library for support vector machines. *ACM Trans. Intell. Syst. Technol. (TIST)* 2 (3), 27–39.
- Choi, J., Lee, J., Kim, D., et al., 2012. Environment-detection-and-mapping algorithm for autonomous driving in rural or off-road environment. *Intell. Transport. Syst. IEEE Trans.* 13 (2), 974–982.
- Munoz, Daniel, 2008. Directional associative markov network for 3-D point cloud classification. In: *Fourth International Symposium on 3D Data Processing, Visualization and Transmission*.

- Demantke, J., Mallet, C., David, N., Vallet, B., 2011. Dimensionality based scale selection in 3D lidar point clouds. *Int. Arch. Photogram. Remote Sens. Spatial Inform. Sci.* 38 (Part 5/W12) (on CDROM).
- Dohan, D., Matejek, B., Funkhouser, T., 2015. Learning hierarchical semantic segmentations of LIDAR data. In: 3D Vision (3DV), 2015 International Conference on. IEEE, pp. 273–281.
- Edelsbrunner, H., Mücke, E.P., 1994. Three-dimensional alpha shapes. *ACM Trans. Graph. (TOG)* 13 (1), 43–72.
- Fischler, M.A., Bolles, R.C., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* 24 (6), 381–395.
- Fischer, A., Kolbe, T.H., Lang, F., et al., 1998. Extracting buildings from aerial images using hierarchical aggregation in 2D and 3D. *Comput. Vis. Image Underst.* 72 (2), 185–203.
- Frome, A., Huber, D., Kolluri, R., et al., 2004. Recognizing objects in range data using regional point descriptors. *Computer Vision-ECCV 2004*. Springer, Berlin Heidelberg, pp. 224–237.
- Golovinskiy, A., Kim, V.G., Funkhouser, T., 2009. Shape-based recognition of 3D point clouds in urban environments. In: Proc. IEEE 12th International Conference on Computer Vision. ICCV2009, pp. 2154–2161.
- Guo, Y., Bennamoun, M., Sohel, F., Lu, M., Wan, J., 2014. 3D object recognition in cluttered scenes with local surface features: a survey. *Pattern Anal. Mach. Intell. IEEE Trans.* 36 (11), 2270–2287.
- Guo, Y., Sohel, F., Bennamoun, M., et al., 2013. Rotational projection statistics for 3D local surface description and object recognition. *Int. J. Comput. Vision* 105 (1), 63–86.
- Hernández, J., Matcotegui, B., 2009. Filtering of artifacts and pavement segmentation from mobile Lidar data. *Int. Arch. Photogram. Remote Sens. Spatial Inform. Sci.* 38 (Part 3/W8), 329–333.
- Himmelsbach, M., Luettel, T., Wueensche, H.J., 2009. Real-time object classification in 3D point clouds using point feature histograms. In: Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on. IEEE, pp. 994–1000.
- Hsu, C.W., Lin, C.J., 2002. A comparison of methods for multiclass support vector machines. *Neural Netw. IEEE Trans.* 13 (2), 415–425.
- Lehtomäki, M., Jaakkola, A., Hyppä, J., et al., 2015. Object classification and recognition from mobile laser scanning point clouds in a road environment. *IEEE Trans. Geosci. Remote Sens.* 54 (2), 1226–1239.
- Li, X., Guskov, I., 2005. Multiscale features for approximate alignment of point-based surfaces. *Symp. Geom. Process.* 2005 (255), 217–226.
- Luo, H., Wang, C., Wen, C., et al., 2015. Patch-based semantic labeling of road scene using colorized mobile LiDAR point clouds. *Intell. Transport. Syst. IEEE Trans.* 17 (5), 1286–1297.
- Lv, Y., Duan, Y., Kang, W., et al., 2015. Traffic flow prediction with big data: a deep learning approach. *Intell. Transport. Syst. IEEE Trans.* 16 (2), 865–873.
- Niemeyer, J., Rottensteiner, F., Soergel, U., et al., 2015. Contextual classification of point clouds using a two-stage Crf. *Int. Arch. Photogram. Remote Sens. Spatial Inform. Sci.* 40 (3), 141–148.
- Novatnack, J., Nishino, K., 2008. Scale-dependent/invariant local 3D shape descriptors for fully automatic registration of multiple sets of range images. *Computer Vision-ECCV 2008*. Springer, Berlin Heidelberg, pp. 440–453.
- Osada, R., Funkhouser, T., Chazelle, B., et al., 2002. Shape distributions. *ACM Trans. Graph. (TOG)* 21 (4), 807–832.
- Paquet, E., Rioux, M., Murching, A., et al., 2000. Description of shape information for 2-D and 3-D objects. *Signal Process.: Image Commun.* 16 (1), 103–122.
- Ravani, B., Dart, M., Hiremagalur, J., Lasky, T.A., Tabib, S., 2009. Advanced Highway Maintenance and Construction Technology Research Center. CALTRANS, Sacramento, CA, p. 2009.
- Rusu, R.B., Blodow, N., Beetz, M., 2009. Fast point feature histograms (FPFH) for 3d registration. In: Robotics and Automation, 2009. ICRA'09. IEEE International Conference on. IEEE, pp. 3212–3217.
- Rusu, R.B., Bradski, G., Thibaux, R., Hsu, J., 2010. Fast 3d recognition and pose using the viewpoint feature histogram. In: Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on, pp. 2155–2162.
- Rusu, R.B., Cousins, S., 2011. 3d is here: Point cloud library (pcl). In: Robotics and Automation (ICRA), 2011 IEEE International Conference on. IEEE, pp. 1–4.
- Seo, Y.W., Lee, J., Zhang, W., et al., 2015. Recognition of highway workzones for reliable autonomous driving. *Intell. Transport. Syst. IEEE Trans.* 16 (2), 708–718.
- Shang, L., Greenspan, M., 2010. Real-time object recognition in sparse range images using error surface embedding. *Int. J. Comput. Vision* 89 (2–3), 211–228.
- Tang, L., Yang, X., Dong, Z., et al., 2016. CLRIC: collecting lane-based road information via crowdsourcing. *Intell. Transport. Syst. IEEE Trans.* <http://dx.doi.org/10.1109/TITS.2016.2521482>.
- Teo, T.A., Chiu, C.M., 2015. Pole-like road object detection from mobile lidar system using a coarse-to-fine approach. *IEEE J. Select. Top. Appl. Earth Observ. Remote Sens.* 8, 4805–4818.
- Yan, W.Y., Morsy, S., Shaker, A., et al., 2016. Automatic extraction of highway light poles and towers from mobile LiDAR data. *Opt. Laser Technol.* 77, 162–168.
- Wang, K.C.P., Hou, Z., Gong, W., 2010. Automated road sign inventory system based on stereo vision and tracking. *Comput.-Aid. Civ. Infrastruct. Eng.* 25 (6), 468–477.
- Wang, Z., Zhang, L., Fang, T., Mathiopoulos, P.T., Tong, X., Qu, H., Xiao, Z., Li, F., Chen, D., 2015. A multiscale and hierarchical feature extraction method for terrestrial laser scanning point cloud classification. *Geosci. Remote Sens. IEEE Trans.* 53 (5), 2409–2425.
- Wen, C., Li, J., Luo, H., et al., 2016. Spatial-related traffic sign inspection for inventory purposes using mobile laser scanning data. *Intell. Transport. Syst. IEEE Trans.* 17 (1), 27–37.
- Xu, S., Vosselman, G., Elberink, S.O., 2014. Multiple-entity based classification of airborne laser scanning data in urban areas. *ISPRS J. Photogram. Remote Sens.* 88, 1–15.
- Yang, B., Dong, Z., 2013. A shape-based segmentation method for mobile laser scanning point clouds. *ISPRS J. Photogram. Remote Sens.* 81, 19–30.
- Yang, B., Dong, Z., Zhao, G., Dai, W., 2015. Hierarchical extraction of urban objects from mobile laser scanning data. *ISPRS J. Photogram. Remote Sens.* 99, 45–57.
- Yang, B., Fang, L., Li, Q., Li, J., 2012. Automated extraction of road markings from mobile lidar point clouds. *Photogram. Eng. Remote Sens.* 78 (4), 331–338.
- Yokoyama, H., Date, H., Kanai, S., et al., 2010. Detection and classification of pole-like objects from mobile laser scanning data of urban environments. *Int. J. CAD/CAM* 13 (2).
- Yu, Y., Li, J., Guan, H., Wang, C., 2015a. Automated extraction of urban road facilities using mobile laser scanning data. *Intell. Transport. Syst. IEEE Trans.* 16 (4), 2167–2181.
- Yu, Y., Li, J., Guan, H., Wang, C., 2015b. Semiautomated extraction of street light poles from mobile lidar point-clouds. *Geosci. Remote Sens. IEEE Trans.* 53 (3), 1374–1386.
- Zhu, L., Hyppä, J., 2014. The use of airborne and mobile laser scanning for modelling railway environments in 3D. *Remote Sens.* 6 (4), 3075–3100.