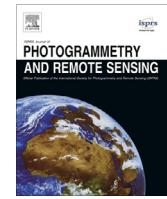




Contents lists available at SciVerse ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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



A shape-based segmentation method for mobile laser scanning point clouds

Bisheng Yang*, Zhen Dong*

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

ARTICLE INFO

Article history:

Received 25 December 2012

Received in revised form 4 April 2013

Accepted 8 April 2013

Available online 9 May 2013

Keywords:

Point classification

Object segmentation

Mobile laser scanning

Object extraction

ABSTRACT

Segmentation of mobile laser point clouds of urban scenes into objects is an important step for post-processing (e.g., interpretation) of point clouds. Point clouds of urban scenes contain numerous objects with significant size variability, complex and incomplete structures, and holes or variable point densities, raising great challenges for the segmentation of mobile laser point clouds. This paper addresses these challenges by proposing a shape-based segmentation method. The proposed method first calculates the optimal neighborhood size of each point to derive the geometric features associated with it, and then classifies the point clouds according to geometric features using support vector machines (SVMs). Second, a set of rules are defined to segment the classified point clouds, and a similarity criterion for segments is proposed to overcome over-segmentation. Finally, the segmentation output is merged based on topological connectivity into a meaningful geometrical abstraction. The proposed method has been tested on point clouds of two urban scenes obtained by different mobile laser scanners. The results show that the proposed method segments large-scale mobile laser point clouds with good accuracy and computationally effective time cost, and that it segments pole-like objects particularly well.

© 2013 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS) Published by Elsevier B.V. All rights reserved.

1. Introduction and related work

Within the last two decades, airborne, terrestrial, and mobile laser scanning have become widely used to capture geospatial information for a variety of applications, including mobile mapping, cultural heritage management, reverse engineering, building reconstruction, DEM generation, and urban planning. Airborne and terrestrial laser scanning clearly differ in terms of their capture modes, typical project sizes, scanning mechanisms, and obtainable accuracy and resolution (Vosselman and Mass, 2010). Filtering, classification, and segmentation using airborne laser scanning have recently become an attractive topic. Extensive research has explored how to deal with airborne laser-scanning point clouds in various applications, including DEM generation (Kraus and Pfeifer, 1998; Axelsson, 2000), 3D building reconstruction (Maas and Vosselman, 1999; Haala and Brenner, 1999; Overby et al., 2004), and tree detection (Bienert et al., 2007).

Methods which deal with airborne laser-scanning data are only partly adaptable for processing terrestrial and mobile laser-scanning data due to differences in scanning patterns, point densities, and geometric characteristics. Although terrestrial laser scanning can provide enough detail (e.g., facades), the efficiency of

large-site (e.g., street-scene) data acquisition is generally not satisfactory because the method only captures the local part of a scene each time. In the last decade, many mobile laser scanning systems using terrestrial laser scanners mounted on vans have been widely used for various purposes (e.g., LYNX and StreetMapper). Compared with the rapid advances in mobile laser scanning system hardware, effective methods for automated information extraction from mobile laser scanning point clouds have progressed relatively slowly. The main challenges in processing mobile laser scanning point clouds are the disparity in the sizes of geometric features, occlusion of features by moving objects, complex and incomplete scene structures, variable point densities, and huge point data volumes (billions).

The segmentation of airborne laser scanning point clouds has been extensively studied, and a large number of proposed methods have been published, including edge-based segmentation (Bhanu et al., 1986; Sappa and Devy, 2001; Wani and Arabnia, 2003) and region-growing segmentation (Jiang et al., 2000; Sithole and Vosselman, 2003). Due to the non-unique correspondence between (X, Y) coordinate pairs and Z-coordinates and the disparity in the size of geometric features, existing segmentation and classification methods designed for airborne laser-scanning points have difficulties in dealing with mobile laser-scanning point clouds. These clouds are usually captured by a van-based platform as on-the-fly observations along a trajectory. There are usually many objects, including buildings, pedestrians, moving cars, and trees, in the mo-

* Corresponding authors. Tel.: +86 27 6877 9699; fax: +86 27 6877 8043.

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

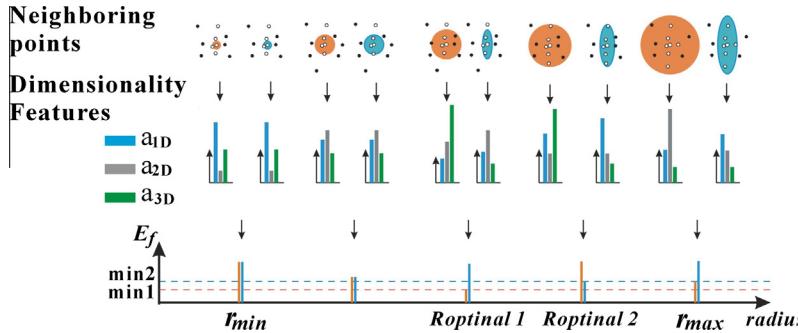


Fig. 1. Selection of optimal point neighborhood size.

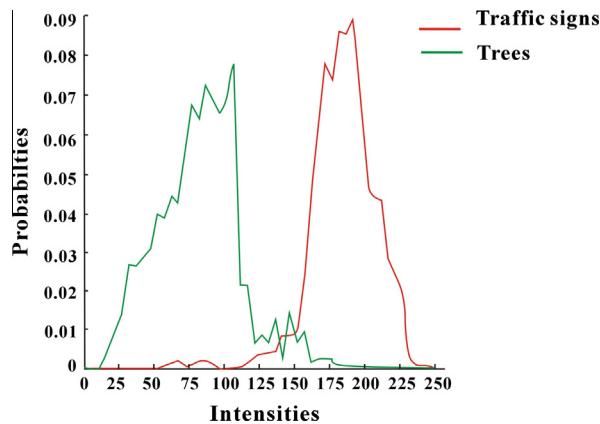


Fig. 2. Probability distribution of the intensities of different kinds of objects.

bile laser scanning point clouds. Efficiently segmenting these point clouds for further object extraction and reconstruction poses a challenge.

Existing segmentation methods for mobile laser scanning point clouds can be classified into three main categories. The first class of approaches, which have been proposed for segmenting buildings and trees, is based on scanning-line information. Manandhar and Shibasaki (2001) relied on information from individual scanning lines to classify points into roads and buildings. Abuhadrous et al. (2004) classified points into buildings, roads, and trees by analyzing two histograms in the Z- and Y-directions in one profile

to reflect the horizontal or vertical nature of urban targets. Nevertheless, the use of scanning-line information is not applicable to unordered point clouds, particularly when point clouds from more than two laser scanners are mixed. The second class of approaches is to extract planar patches, specific features, or objects (e.g., roads and pole-like objects) to segment mobile laser scanning point clouds. Examples of these approaches include clustering-based classification by normal vectors or texture measures (Flinn, 2002; Biosca and Lerma, 2008), region-growing-based classification with planar patches (e.g., Dold and Brenner, 2004), and 3D Hough-transformation-based classification (Overby et al., 2004; Vosselman et al., 2004). Rabbani et al. (2007) presented a model-driven segmentation approach for extracting a range of geometric primitives, including planes, cylinders, and spheres. Boyko and Funkhouser (2011) investigated the extraction of roads from mobile laser scanning point clouds incorporating large-scale vector map data. Lehtomaki et al. (2010) detected vertical pole-like objects from vehicle-based laser scanning points. Pu et al. (2011) recognized basic structures like poles in mobile laser scanning data and partitioned the point clouds into broad parts, which were then divided into three main classes: ground surface, objects on the ground, and objects above the ground. They focused mainly on detecting poles near roads according to their geometric attributes (e.g., size, orientation, and percentile height). Becker (2009) proposed a model-driven method to extract windows from building façades using a formal grammar. This kind of method is mainly focused on planar patches (e.g., building facades and poles) and has difficulties in dealing with complex natural scenes. Another approach is to classify mobile laser scanning point clouds using image segmenta-

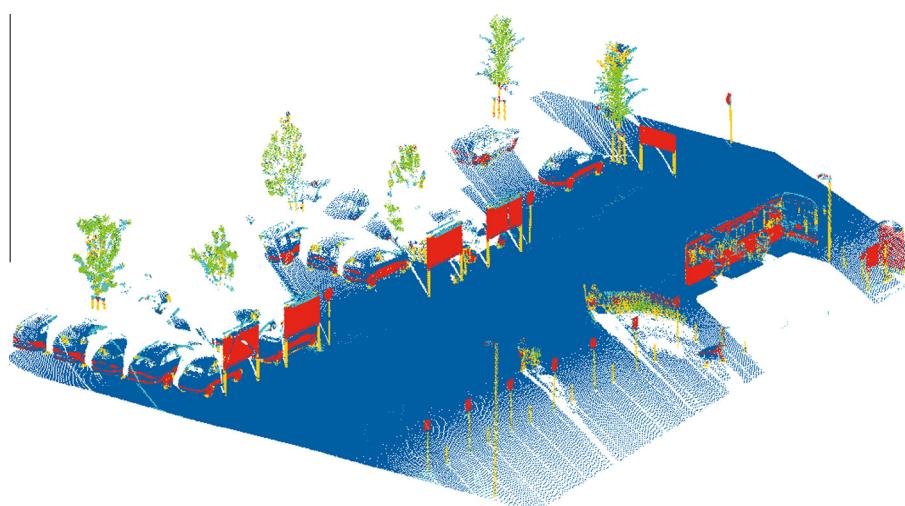


Fig. 3. Classification of point clouds by SVMs.

Table 1

Parameters and corresponding values for merging points.

Parameters	Values
R	$r_{optimal}$
P_θ	10°
N_θ	10°
I_A	10

tion or fusion of auxiliary data (e.g., vector data, imagery, and intensity). Barnea and Filin (2013) proposed a solution for segmentation of terrestrial laser scanning point clouds incorporating colors from imagery. However, such auxiliary data are not always available, and registration between point clouds and imagery is still a non-trivial issue. Golovinskiy et al. (2009) designed a system to recognize objects in urban scenes from point clouds and integrated a number of techniques to localize, segment, and classify objects using support vector machines (SVMs). Hernández and Marcotegui (2009) proposed a method for generating range images from point clouds and extracted features from the connected components detected, which were then classified into four categories using SVMs. Yang et al. (2012a) proposed a method to generate feature imagery from point clouds for extracting buildings and trees. The generated feature image provides an alternative solution for extracting road markings and facade footprints (Yang et al., 2012b, 2013). Nevertheless, the parameters of feature-imagery generation require good tuning to extract different kinds of objects (e.g., buildings and power lines).

Although the techniques described above are generally able to extract a specific kind of object (e.g., windows and road markings) and either planar patches or linear features of interest, they require further improvement to deal with large-scale urban scenes containing multiple objects with varied and complex shapes, including incomplete structures and variable point densities. Moreover, existing techniques have paid little attention to the effects of variable point densities and neighborhood sizes on point-cloud segmentation and classification.

This paper therefore proposes a shape-based method for the segmentation of large-scale urban scene point clouds into a meaningful segmented geometric abstraction. The procedure of the proposed method for segmenting mobile laser scanning point clouds consists of three sequential steps: the classification of points based on their geometric features using SVMs; the segmentation of the classified points and further refinement of the segmentation; and

merging of the segmentation output for geometric abstraction of objects. The main contributions of the proposed method are as follows:

- adaptive determination of point neighborhood size for accurate geometric feature estimation (e.g., surface normals) of point clouds;
- robust refinement of the segments of linear, planar, and spherical patches using the proposed similarity criteria to overcome over-segmentation; and
- efficient and accurate extraction of pole-like objects from the segmentation output.

The segmentation output of the proposed method is examined and validated by extracting pole-like objects, which are a widespread type of objects in urban scenes.

Following the introduction, the key steps of the proposed method will be elaborated. Before conclusions are drawn, the proposed method will be tested and validated by experimental studies.

2. Classification of mobile laser scanning point clouds

This step aims to classify mobile laser scanning point clouds into several categories (e.g., planar and spherical) according to the geometric features (e.g., eigenvalues and principal directions) of the points. Efforts have been made to address the problem of computing point geometric features for classification and feature extraction in the domain of point-sampled data (Tang and Medioni, 2002; Pauly et al., 2003). Finding the optimal neighborhood of each point is important for computing each point's local covariance matrix and obtaining an accurate shape estimate. In particular, the distribution of mobile laser point clouds has variable point densities because of occlusion, varying scanning angles, and varying distances to the laser scanner. Using a fixed neighborhood size leads to poor shape estimation and segmentation (Demantke et al., 2011). Therefore, the optimal neighborhood size for each point should be selected carefully to derive an accurate local covariance matrix. It is clear that the local covariance matrix of one point is constructed using the points within its neighborhood. Once the local covariance matrix of one point has been determined, the eigenvalues of the local covariance matrix are calculated to derive various geometric features.

Principal components analysis (PCA) has been used to synthesize the distribution of points in three dimensions (Tang and Medioni, 2002) and thus to model the principal directions and

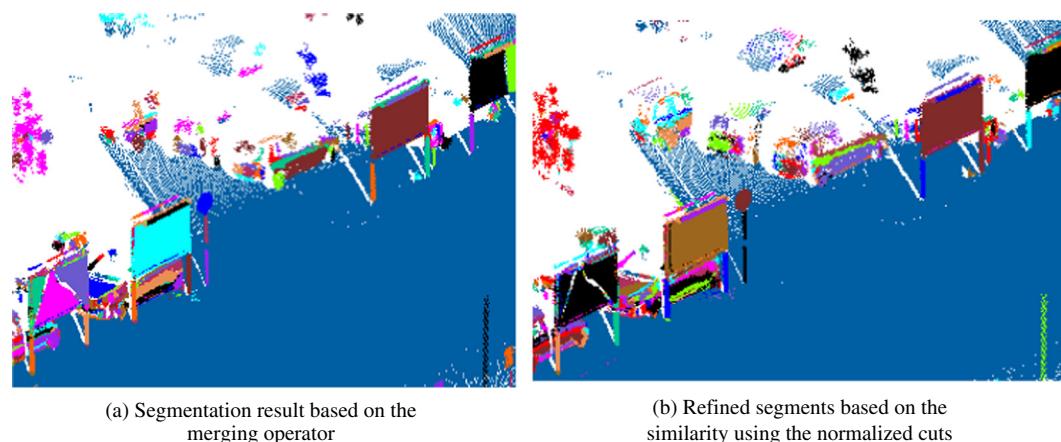


Fig. 4. Shaped based segmentation of point clouds.

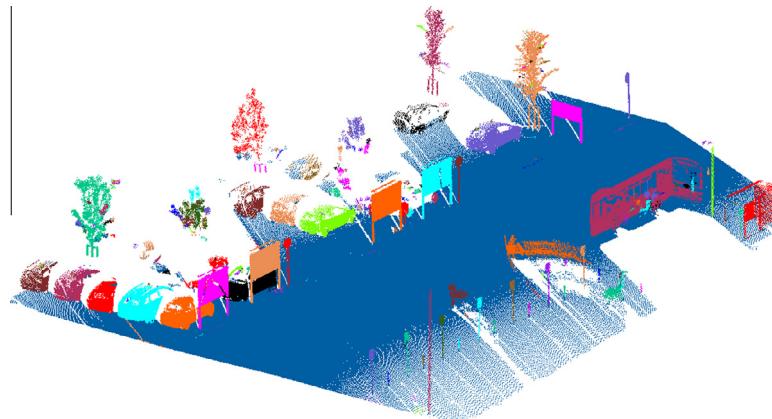


Fig. 5. Merging of segmentation output.

magnitudes of variation of the point distribution around the center of gravity. Let the neighboring points of p_i be $p_m = (x_m, y_m, z_m)$, $m = (1, 2, \dots, k)$.

The local covariance matrix of p_i is formatted according to PCA and can be written as:

$$M = \frac{1}{k} \sum_{i=1}^k (p_i - \bar{p})(p_i - \bar{p})^T, \quad (1)$$

where $\bar{p} = \frac{1}{k} \sum_{i=1}^k p_i$.

Note that M is always symmetric positive semi-definite, therefore has non-negative eigenvalues, and can be expressed as:

$$M = [\vec{e}_1 \quad \vec{e}_2 \quad \vec{e}_3] \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \begin{bmatrix} \vec{e}_1 \\ \vec{e}_2 \\ \vec{e}_3 \end{bmatrix}. \quad (2)$$

The eigenvalues are positive and ordered, so $\lambda_1 > \lambda_2 > \lambda_3 > 0$. For the 3D scattered points, no dominant direction can be found with $\lambda_1 \approx \lambda_2 \approx \lambda_3$. In the case of a linear structure, the principal direction will be the tangent to the curve with $\lambda_1 \gg \lambda_2 \approx \lambda_3$. In the case of a planar surface, the principal direction is parallel to the surface normal with $\lambda_1 \approx \lambda_2 \gg \lambda_3$. The dimensionality features of $a_{1D} = \frac{\sqrt{\lambda_1} - \sqrt{\lambda_2}}{\sqrt{\lambda_1}}$, $a_{2D} = \frac{\sqrt{\lambda_2} - \sqrt{\lambda_3}}{\sqrt{\lambda_1}}$, $a_{3D} = \frac{\sqrt{\lambda_3}}{\sqrt{\lambda_1}}$ as described in Demantke et al. (2011) are derived to indicate the likelihood that the shape behavior of the neighboring points of point p_m is linear, planar, or spherical respectively. The dimensionality features ($\alpha_{1D}, \alpha_{2D}, \alpha_{3D}$) are different from the curvature-based stick voting fields of Tang and Medioni (2002), which label voxels as hyperbolic, planar, and parabolic or elliptic. $\vec{e}_1 = (p_x, p_y, p_z)$ and $\vec{e}_3 = (n_x, n_y, n_z)$ are the principal direction and the normal vector of point p_m .

2.1. Selection of laser point neighborhood size

The dimensionality features ($\alpha_{1D}, \alpha_{2D}, \alpha_{3D}$) derived from the eigenvalues of the local covariance matrix of a point can be formulated using the entropy function proposed by Demantke et al. (2011) as:

$$E_f = -\alpha_{1D} \ln(\alpha_{1D}) - \alpha_{2D} \ln(\alpha_{2D}) - \alpha_{3D} \ln(\alpha_{3D}). \quad (3)$$

The optimal neighborhood size is obtained when:

$$r_{optimized} = \operatorname{argmin}(E_f). \quad (4)$$

For each laser point p_k , the steps for calculating $r_{optimized}$ for p_k are described below.

- Step 1: Let the size of neighborhood r_c of p_k be between r_{min} and r_{max} , and let the increment of r_c be r_A . Initialize the neighborhood size r_c to r_{min} .
- Step 2: Select the neighboring points of p_k according to the neighborhood size r_c : the intensity difference between each neighboring point and p_k is less than a threshold I_A until all the neighboring points of p_k have been traversed.
- Step 3: Calculate the eigenvectors and eigenvalues of the local covariance matrix of p_k using PCA.
- Step 4: Calculate the value E_i of the entropy function E_f .
- Step 5: Set $r_c = r_c + r_A$. If $r_c \leq r_{max}$, go to Step 2. Otherwise, go to Step 6.
- Step 6: For each neighborhood size r_c , the corresponding values E_i are compared; the neighborhood size r_c with minimum value E_{min} is selected as the optimal neighborhood size $r_{optimized}$.

Fig. 1 illustrates the procedure for point neighborhood size selection according to the calculated entropy function E_f . The neighboring points in the orange sphere and the light blue ellipsoid are the selections of points outside and inside the intensity difference of the points. It can be seen from Fig. 1 that the shape of a point neighborhood is always a sphere, according to the method proposed by Demantke et al. (2011). Nevertheless, the intensity difference of the points has an effect on the selection of neighboring points (Step 2). As illustrated in Fig. 1, the white and black points represent objects of different kinds and intensities. A significant difference from the work of Demantke et al. (2011) is that not all the points in the neighborhood r_c are involved in the local covariance matrix calculation, resulting in an accurate estimation of shape features. As shown in Fig. 1, the orange spherical neighborhood is selected without considering point intensities, but the corresponding shape is classified as spherical by mistake (α_{3D} has the maximum value). However, the light blue neighborhood is selected and the corresponding shape is classified as a linear label (α_{1D} has the maximum value) when neighborhood-point intensity is integrated into the selection of neighborhood points (Step 2). Fig. 2 depicts the probability distribution of point intensities in a local area that mixes different kinds of objects. It can be seen that the points representing different objects in a local area exhibit distinct peaks. Although point intensities are closely related to the distances from the laser scanner to the objects, the incidence angles of the laser beam, and the materials of the scanned objects, the standard deviation of point intensities in a small local area can be calculated to distinguish the points of the various scanned objects. This proves that point intensity has a crucial effect on the selection of the optimal neighborhood size, thus resulting in

Table 2

Descriptions of the two test data sets.

	Number of points (million)	Point density (points/m ²)	Data length (m)	Data width (m)
Lynx data set	8.5	512	220	89
VMX-450 data set	105	120	5100	121

more accurate shape estimation, particularly in areas with mixed multiple objects.

Once the optimal neighborhood size $r_{optimized}$ of one point has been determined, the local covariance matrix of the point is formulated and its eigenvalues calculated to derive the geometric features of the point. These features are used for point classification by the SVMs.

2.2. Classification of laser scanning point clouds based on SVMs

The local covariance matrix of point p_i is formulated using its neighboring points according to PCA, and the following geometric features associated with p_i are thus derived:

- the dimensionality feature ($\alpha_{1D}, \alpha_{2D}, \alpha_{3D}$) indicates the likelihood that the neighboring points of p_i represent a linear, planar, or spherical patch respectively;
- the normal vector (n_x, n_y, n_z) indicates the direction of the planar patch constituted by the neighboring points of p_i ;
- the principal direction (p_x, p_y, p_z) is the eigenvector corresponding to the maximum eigenvalue and indicates the direction of the linear patch constituted by the neighboring points of p_i ; and
- the eigenvalues (n_{vx}, n_{vy}, n_{vz}) of the normal vector indicate the smoothness of the neighboring points of p_i .

Each point associated with the above geometric features is labeled as a linear, planar, or spherical point by LIBSVM, which is an integrated software tool for support vector classification, regression, and distribution estimation and supports multi-class classification (Chang and Lin, 2011). To obtain an optimal classification model, a training set and a validation set are generated manually within a real mobile laser scanning point cloud, and the training set is trained using LIBSVM with a radial basis function (RBF) kernel. Finally, the mobile laser point clouds are classified point by point into three types of shapes (linear, planar, and spherical) using LIBSVM with the optimal classification model, as shown by the yellow, blue, and green dots respectively (Fig. 3). The classified point clouds are used to segment objects in mobile laser scanning point clouds.

Table 3

Points classification performance of the proposed method.

	The proposed method without intensities integration		The proposed method with intensities integration	
	Precision (%)	Recall (%)	Precision (%)	Recall (%)
Linear points	84.8	90.7	95.0	93.5
Planar points	92.8	91.2	96.8	93.4
Sphere points	73.4	83.7	80.2	92.3
Overall accuracy	89.0		92.9	

3. Segmentation of mobile laser scanning point clouds

The mobile laser scanning point clouds, each of which has been classified into one of the three types of shapes, are segmented according to the associated attributes of their points. A set of rules are defined to merge points with identical labels (e.g., planar and spherical).

3.1. Merging points with identical labels

The principal direction, the normal vector, and the intensity of each labeled point are used for merging points into segments according to the principle of region growing. First, randomly select one point from the labeled points as a seed point. Let the principal direction, the normal vector, and the intensity of the seed point be $\vec{P}_{seed} = (p_{seedx}, p_{seedy}, p_{seedz})$, $\vec{N}_{seed} = (n_{seedx}, n_{seedy}, n_{seedz})$, and I_{seed} respectively. The corresponding rules for merging linear, planar, and spherical points are defined as follows:

- Merge rule for linear points:

$$\vec{P}_{seed} \cdot \vec{P}_{candidate} \leq P_\theta \wedge |I_{seed} - I_{candidate}| \leq I_\Delta \wedge \|V_{seed} - V_{candidate}\| \leq r.$$

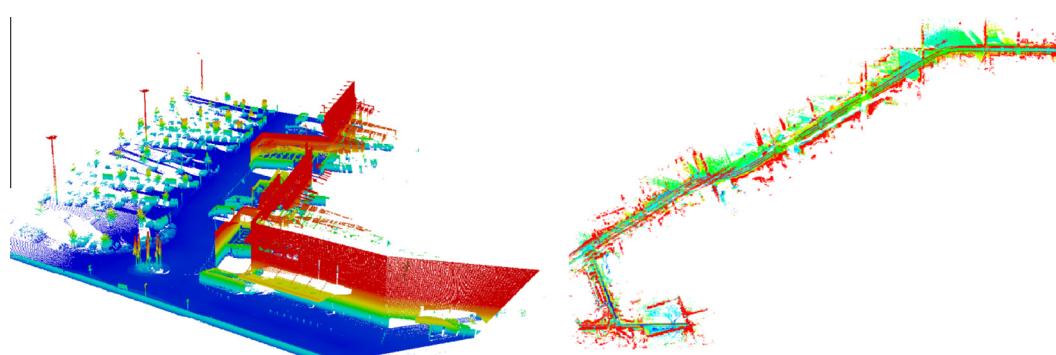
- Merge rule for planar points:

$$\vec{N}_{seed} \cdot \vec{N}_{candidate} \leq N_\theta \wedge |I_{seed} - I_{candidate}| \leq I_\Delta \wedge \|V_{seed} - V_{candidate}\| \leq r.$$

- Merge rule for spherical points:

$$\|V_{seed} - V_{candidate}\| \leq r,$$

where $\vec{P}_{candidate}$, $\vec{N}_{candidate}$, and $I_{candidate}$ are the principal direction, the normal vector, and the intensity of the candidate point; V_{seed} , $V_{candidate}$ are the coordinates of the seed point and the candidate point; P_θ , N_θ , I_Δ , r are the thresholds for principal direction differences, normal vector differences, intensity differences, and the search radius of the seed point.

**Fig. 6.** Overview of the test data sets.

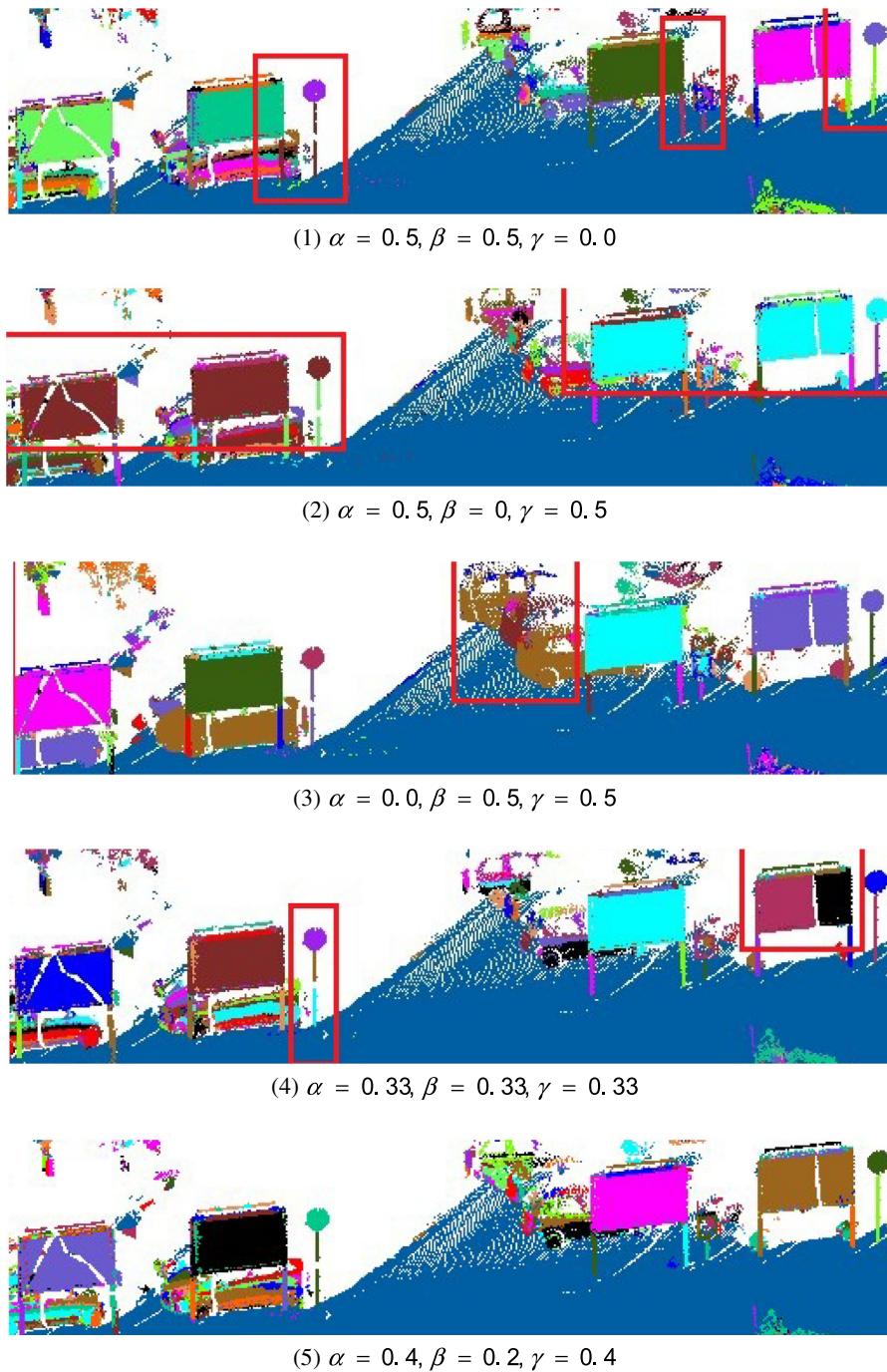


Fig. 7. Comparison of segmentation results with different thresholds of α, β, γ .

The parameter values for merging points with identical labels are listed in Table 1.

The values of the parameters P_θ, N_θ are sensitive to the results of merging points into segments. Larger values of P_θ, N_θ may result in under-segmentation of the point clouds, while smaller parameter values may lead to over-segmentation of the mobile laser points. In the method proposed here, smaller values (10°) were used for the parameters P_θ, N_θ to overcome under-segmentation. The points representing cars and traffic signs are over-segmented because of complex surfaces or occlusion (Fig. 4a). To overcome the limitations of over-segmentation, the over-segmented parts were refined to obtain a global optimal segmentation using normalized cuts (Shi and Malik, 2000).

3.2. Refining segments using normalized cuts based on similarity measurements

Normalized cuts (Shi and Malik, 2000) have proven effective for overcoming over-segmentation. The key aspect of normalized cuts is measuring the similarity criteria of the segments. These similarity criteria are different in different applications. The mobile laser point clouds are segmented into numerous segments by the merge operator described in Section 3.1. Each segment is regarded as a vertex of an undirected weighted graph according to the principles of normalized cuts. Hence, an undirected weighted graph can be generated and represented as $G = (V, E)$, where the weight of each edge is assigned according to the similarity between its two corre-

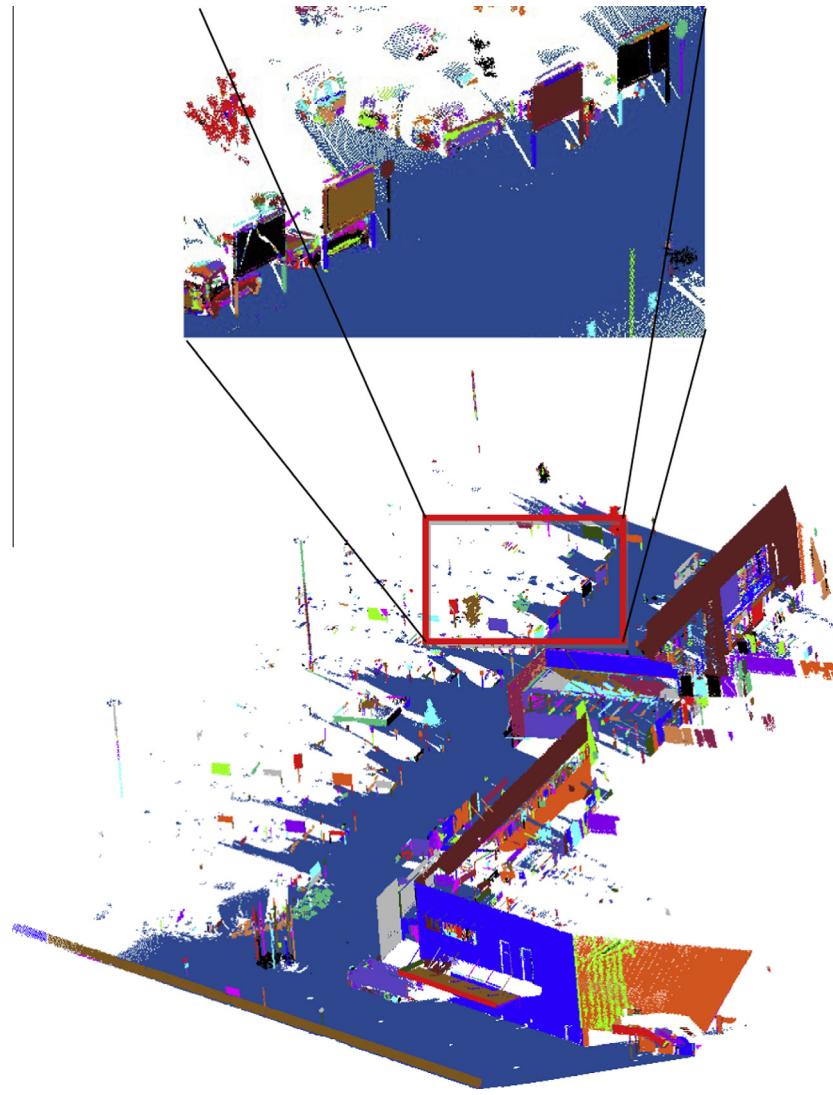


Fig. 8. Segmentation of the Lynx data set.

sponding vertices (segments). In the proposed method, the similarity between homogeneous segments is measured by the angle between the normal vectors of the segments, the angle between the principal directions of the segments, and the distance between the segments. The similarity between linear patches is calculated as:

$$WL(i,j) = e^{-\alpha \times \Delta p_{ij}} \times e^{-\beta \times ld_{ij}} \times e^{-\gamma \times ct_{ij}} (\alpha > 0, \beta > 0, \gamma > 0, \alpha + \beta + \gamma = 1.0), \quad (5)$$

where Δp_{ij} is the angle between the principal directions of two linear-shaped patches, ld_{ij} is the minimum Euclidean distance between two linear patches, ct_{ij} is the Euclidean distance between the centers of two linear patches, and α, β, γ are the weights associated with each factor.

The calculation for planar patches is:

$$WP(i,j) = e^{-\alpha \times \Delta n_{ij}} \times e^{-\beta \times pd_{ij}} \times e^{-\gamma \times cop_{ij}} (\alpha > 0, \beta > 0, \gamma > 0, \alpha + \beta + \gamma = 1.0), \quad (6)$$

where Δn_{ij} is the angle between the normal vectors of two planar patches, pd_{ij} is the minimum Euclidean distance between two planar patches, cop_{ij} is the Euclidean distance between the centers of two planar patches, and α, β, γ are the weights associated with each factor.

The calculation for spherical patches is:

$$WS(i,j) = e^{-\beta \times sd_{ij}} (\beta > 0), \quad (7)$$

where sd_{ij} is the minimum Euclidean distance between two spherical patches and β is the set of weights associated with the factor.

The value of α is responsible for varying the weight of the angle between the principal directions or normal vectors of neighboring patches. The value of β is responsible for varying the weight of the distances between neighboring patches. The value of γ is responsible for varying the weight of the distance between points and lines and between points and planes.

The normalized cuts measure the total similarity within patches as well as the overall dissimilarity between different patches. As far as the undirected weighted graph $G = (V, E)$ is concerned, the goal is iteratively to find an edge cut that divides G into k disjoint subsets such that $V = \bigcup_{i=1}^k V_i$ and $\forall (i, j), V_i \cap V_j = \emptyset$. Moreover, the cost of the edge cut should have a minimum value. According to the principles of normalized cuts, the cost of an edge cut for linear patches can be calculated as:

$$NLinearCut(A, B) = \frac{\sum_{i \in A, j \in B} WL(i,j)}{\sum_{i \in A, j \in V} WL(i,j)} + \frac{\sum_{i \in A, j \in B} WL(i,j)}{\sum_{i \in B, j \in V} WL(i,j)}, \quad (8)$$

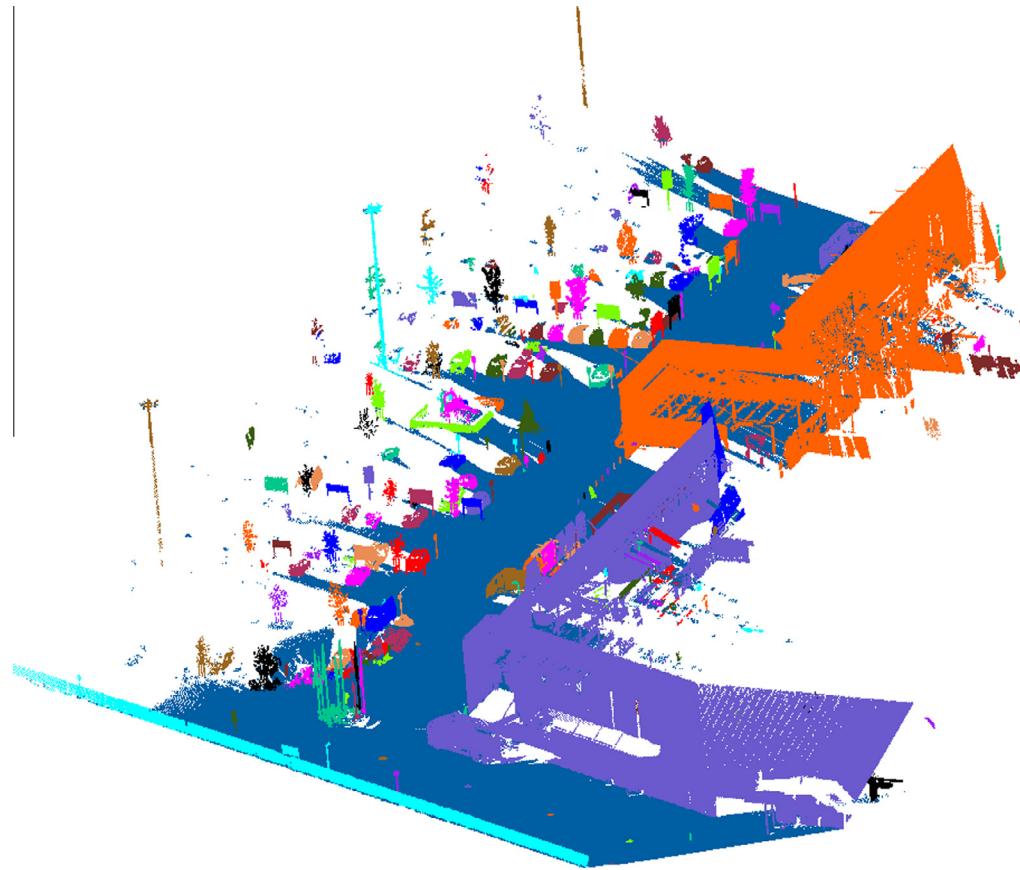


Fig. 9. Merged segments for the Lynx data set.

where $\sum_{i \in A, j \in B} WL(i, j)$ is the sum of the weights of edges connecting vertices within subsets A and B , $\sum_{i \in A, j \in V} WL(i, j)$ is the sum of the weights of edges from vertices in A to all vertices in graph G , and $\sum_{i \in B, j \in V} WL(i, j)$ is defined similarly.

The costs of the edge cuts of planar patches, $NPlanarCut(A, B)$, and of spherical patches, $NSphericalCut(A, B)$, are defined similarly. The process of generating normalized cuts will continue until the cost of an edge cut is less than a specified threshold N_T .

Three undirected weighted graphs (for linear, planar, and spherical patches) are generated for the corresponding patches. For each type of segment, the normalized cuts are used with the corresponding edge-cut cost. When the value of the edge-cut cost is larger, the associated patches have a higher probability of being merged. The procedure is executed iteratively until the edge-cut cost value is less than a specified threshold N_T . Fig. 4b shows the refined segments created by merging patches based on the similarity measurement. Comparison of Fig. 4a and b shows that the over-segmentation of point clouds representing traffic signs has been overcome.

4. Merging neighboring segments for object extraction

The mobile laser point clouds are divided into numerous segments at the point level using the procedure described in Sections 2 and 3. Each segment may belong to a part of a certain object. The segments should be merged at the object level to provide helpful hints for object detection and basic structure recognition (Monnier et al., 2012; Pu et al., 2011). In urban scenes, mobile laser scanning point clouds consist of numerous objects, such as roads, buildings, trees, poles, traffic signs, pedestrians, and cars. The road surface is associated mainly with planar segments at lower elevations whose

normal vectors are approximately parallel to the z -axis. Hence, their ground segments are easily detected from the segmentation output of the proposed method and classified as roads. Generally, building facades are combinations of segments at higher elevations, such as planar segments, and traffic signs are combinations of linear and planar segments. Trees are combinations of segments at higher elevations as well, including linear segments (tree trunks) and spherical segments (tree crowns). Lamps are combinations of linear segments. Hence, the segments need to be further merged to classify them as meaningful geometrical abstractions of objects such as poles and trees. Such further abstraction would be beneficial and meaningful for detection and recognition of objects from mobile laser scanning point clouds of urban scenes.

To merge segments at higher elevations, an operator is used that checks whether these segments are adjoining. The operator calculates the minimum hulls of the segments and the adjacency relationships between the minimum hulls. Suppose that two segments at higher elevations are adjacent. The two segments at higher elevations are then merged as a geometrical abstraction of an object. Segments at higher elevations are merged by the operator until all the segments at higher elevations have been checked. Fig. 5 shows that the linear and planar segments representing traffic signs have been merged well and include all the segments in the objects. Moreover, the ground segments have also been accurately merged.

5. Experimental analysis

Two test data sets were acquired using the Optech Lynx mobile mapping system and the Riegl VMX-450 mobile mapping system. The Lynx is equipped with two laser scanners, which are mounted

Table 4

Time performance of the proposed method.

Data	Steps						Total costs (min)	
	Point classification (s)		Point segmentation (s)		Segmentation classification (s)			
	Features calculation	SVM classification	Region growing	Region merging	Segmentation classification	Object extraction		
Lynx dataset	171	58	44	18	39	11	5.5	
VMX-450 dataset	1891	629	432	289	223	89	59.2	

perpendicularly to each other. The two laser scanners are oriented at an angle of 45° from the driving direction and collect data at a rate of 500,000 measurements per second with a field of view (FOV) of 360°. The hardware of the VMX-450 consists of two Riegl VQ-450 laser scanners, a camera system, a portable control unit box, and a GNSS/INS unit. The effective measurement rate of the VMX-450 can be up to 1,100,000 points/s. **Table 2** gives a description of the two test data sets, and **Fig. 6** shows an overview of the two scenes.

5.1. Classification results based on SVM

Determining the optimal neighborhood size of points is the first step in the proposed method. Because the minimum point density is approximately 120 points/m², at least two points are included within a circle with radius 0.08 m. The objects to be detected should have a minimum area of 0.6 m². The standard deviation of the intensity in a local neighborhood is approximately 10. Hence, the minimum neighborhood size r_{\min} , the maximum neighborhood size r_{\max} , the increment r_A , and the intensity difference I_A were specified as 0.08 m, 0.48 m, 0.08 m, and 10 m respectively. Then the dimensionality features, the principal directions, the normal vectors, and the eigenvalues of the normal vectors associated with each point of the two test data sets were calculated. To use SVMs to classify the two data sets, 100 training targets and 100 validation targets were manually selected from the Lynx data set. The training targets included planar areas, pole-like areas, and spherical areas. Each target consisted of approximately 100 points. The selected training targets were trained using LIBSVM to obtain an optimal classification model that was used to label the point clouds of the two data sets as linear, planar, and spherical points using SVMs. **Table 3** lists the classification accuracies of the proposed method evaluated according to the validation targets. It can be seen from **Table 3** that the precision, recall, and overall accuracy for linear, planar, and spherical points have been improved by 5%, resulting in better data for further segmentation. This demonstrates that a significant improvement in point classification accuracy can be achieved by integrating point intensities in a small local area.

5.2. Segmentation results based on the proposed method

The points in the two test data sets were labeled as linear, planar, and spherical points plus unclassified points. The merge operator was used to segment the mobile laser scanning point clouds into homogeneous patches. The merge operator has the property that points that share similar shape features and are adjoining in space have a high probability of being assigned to the same patch. In the experiment, the classified linear, planar, and spherical patches were segmented by the merge operator according to the defined rules and the corresponding values of the parameters as listed in **Table 1**.

Following the application of the merge operator, the segmented linear, planar, and spherical patches were refined by normalized cuts using the similarity measurement formulations defined in Section 3.2. To test the effects of α , β , γ on the segmentation results, the results were visually inspected with different thresholds of α , β , γ (**Fig. 7**). **Fig. 7-1–7-5** shows that the best segmentation result (**Fig. 7-5**) was obtained with thresholds of $\alpha = 0.4$, $\beta = 0.2$, and $\gamma = 0.4$, which overcame the unreasonable merging tendencies between linear (**Fig. 7-1**), planar (**Fig. 7-2**), and spherical patches (**Fig. 7-3**). In general, the value of the edge-cut cost in the normalized cuts is assigned a value between 0.01 and 0.2. Here, the value of the edge-cut cost was specified as 0.1.

The segmentation results for the Lynx data set are illustrated in **Fig. 8**. It can be seen that the facades and traffic signs have been segmented into different patches. Moreover, the incomplete parts of the traffic sign have been merged into a single planar patch, and the poles with traffic signs attached have been segmented into a linear patch.

5.3. Results of merging neighboring segments for object extraction

Because many objects have multiple parts of different shapes, for example, a lamp has a pole-like part and a sphere-like part which generate different segments (**Fig. 8**), once the point clouds of many objects have been classified into different segments, these segments should be merged. To fulfill these requirements, the segments of the two test data sets were further merged according to

Table 5

The precisions, recalls, and overall accuracy of the proposed method for pole-like object extraction.

Data	Object	Lamps	Traffic signs	Trees	Others	Classification overall	Precision	Recall
Lynx data set	Lamps	15	0	0	1	16	93.7	83.3
	Traffic signs	0	62	1	2	65	95.4	96.8
	Trees	2	0	41	1	44	93.2	91.1
	Others	1	2	3	129	135	95.5	96.9
	True overall	18	64	45	133	260	Overall accuracy: 95.0	
VMX-450 data set	Lamps	257	1	3	11	272	94.5	97.3
	Traffic signs	0	11	0	2	13	84.6	84.6
	Trees	5	1	1729	78	1813	95.4	97.6
	Others	2	0	39	609	650	95.4	95.0
	True overall	264	13	1771	700	2748	Overall accuracy: 94.8	

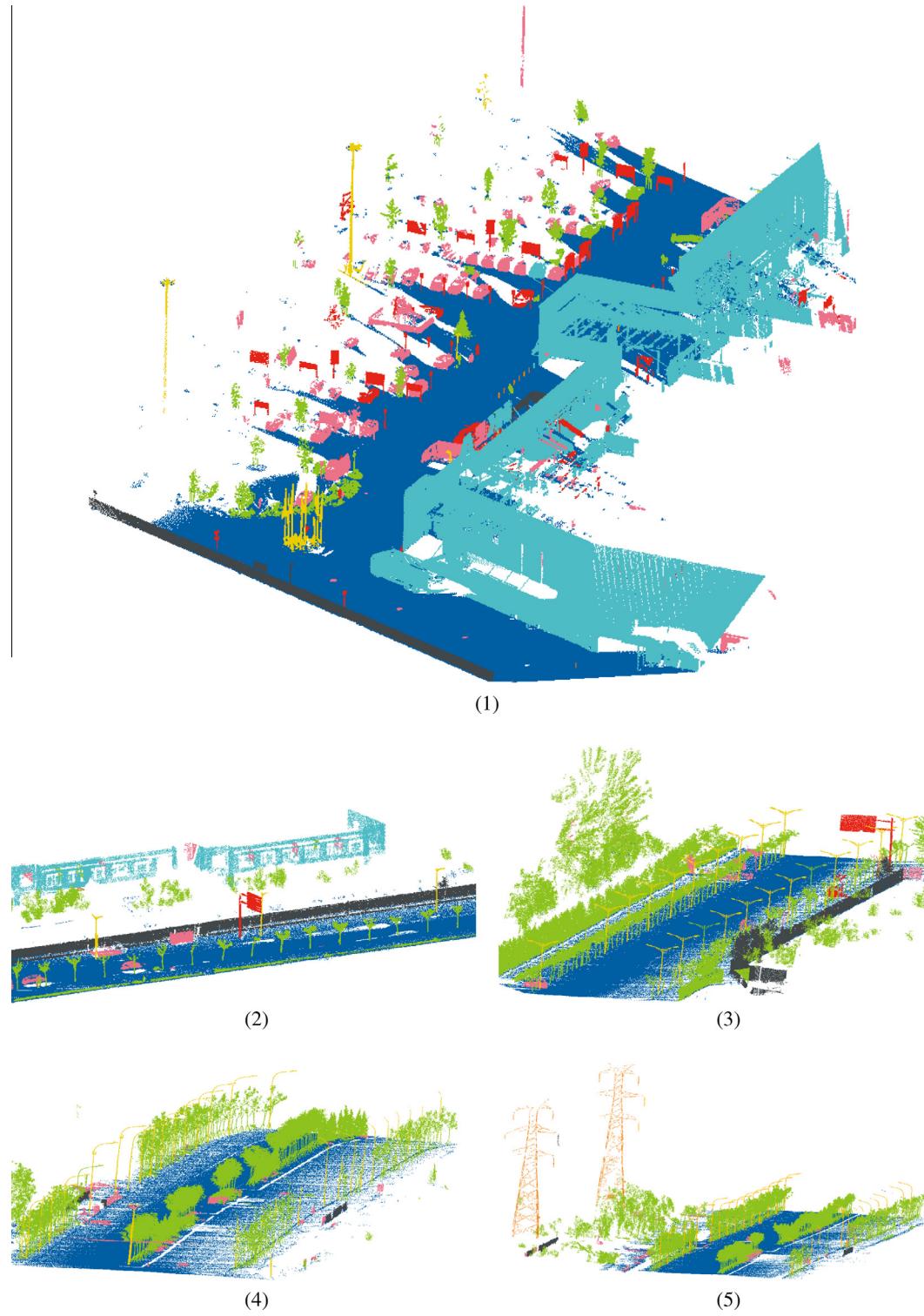


Fig. 10. Pole-like object extraction from the two data sets.

the adjacent topology between the segmented patches. This merging of segmented patches serves to combine the patches into objects. Fig. 9 illustrates the final segmentation results for the two test data sets. It can be seen that the points of one object are classified into one segment. Moreover, each combined segment in the final segmentation result records the sizes, the normal vectors, and the principal directions of all its sub-segments. The recorded data are used for further object extraction (e.g., poles, traffic signs, and trees).

5.4. Pole-like object extraction performance of the proposed method

The procedure to implement the proposed method encompasses several steps and entails heavy computing requirements. The experiment was performed using a computer with 8 GB RAM and an Intel®Core™i3 540@ 3.07 GHz CPU. Table 4 lists the time performance of the proposed method for segmenting mobile laser point clouds, including the time costs of each step. It takes about 1 h to segment the data for 105 million points and to extract

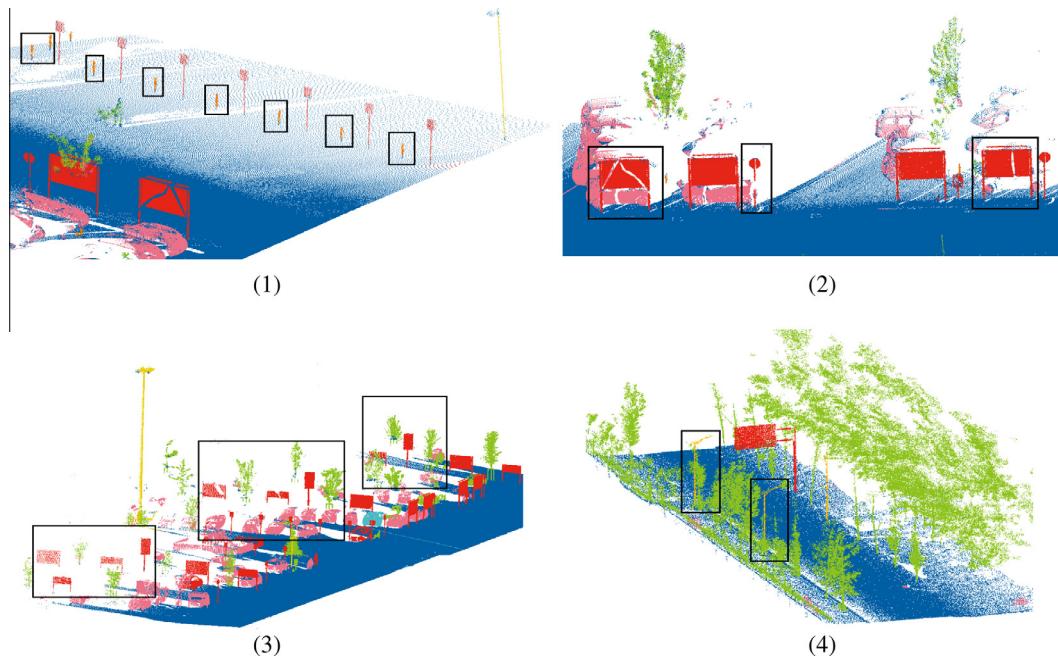


Fig. 11. Typical pole-like object extraction.

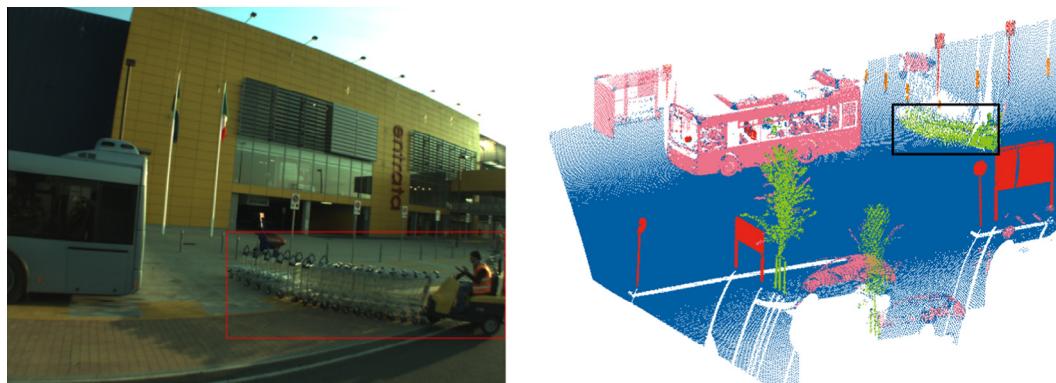


Fig. 12. Misclassification scenarios for pole-like objects.

pole-like objects from the segmentation output. The time costs of segmenting the two data sets show that the proposed method offers high efficiency in segmenting mobile laser point clouds, particularly for the large data volumes in mobile laser point clouds.

The segmentation output of the proposed method was checked by extracting pole-like objects (e.g., lamps, traffic signs, and trees) from the final segments of the two test data sets according to a defined pole-like model. The pole-like model should have a linear patch that at least touches the ground surface; this linear feature may be attached to planar or spherical patches of certain sizes (e.g., 2 m^2). The final segments of the two test data sets were matched with the pole-like model to extract pole-like objects. The pole-like objects extracted were further classified into three kinds of objects (trees, lamps, and traffic signs) according to the size of the attached planar or spherical patches. Table 5 lists the accuracy and recall values for extracting pole-like objects from the test data sets. The true numbers of trees, traffic signs, and lamps in the two test data sets were counted manually to evaluate the performance of pole-like object extraction.

The pole-like object extraction results show that the proposed method achieved an overall accuracy of 95% and a precision of 93% for extracting trees, lamps, and traffic signs. In particular,

VMX-450 data were captured in a complex urban scene, including numerous objects such as trees, traffic signs, lamps, and cars. The proposed method still achieves good performance in extracting pole-like objects from the segmented point clouds and classifies the pole-like objects extracted into trees, traffic signs, and lamps (Fig. 10). Fig. 10-1 shows the results of pole-like object extraction from the whole scene of the Lynx data set. Parts of the pole-like object extraction results from the VMX-450 data are illustrated in Fig. 10-2–10-5. The extracted trees, lamps, and traffic signs are indicated by green, yellow, and red dots respectively.

The extraction results show that the proposed method extracts the lamps successfully, whether they are close to or far from the road. In particular, it extracts the short lamp posts (Fig. 11-1), poles with incomplete structures (Fig. 11-2), and parking meters far from the road (Fig. 11-3). Moreover, the proposed method extracts and classifies pole-like objects in mixed areas correctly. For example, it extracts traffic signs and lamp posts from among the trees (Fig. 11-4).

However, a few misclassifications still occurred in pole-like object extraction. As illustrated in Fig. 12, carts were misclassified as trees. It can be seen from the image that the carts have shapes very similar to trees.

6. Conclusions

Efficiently and robustly segmenting mobile laser scanning point clouds of urban scenes into objects is of crucial importance for object detection and reconstruction, in particular when processing large point-cloud data volumes. In this paper, a shape-based method is proposed for efficiently segmenting mobile laser scanning point clouds of large-scale urban scenes into objects. The proposed method achieves accurate estimation of point geometric features by incorporating point intensity and adaptive neighborhood size, resulting in a classification accuracy improvement for point clouds. A merge operator is then used to partition the classified point clouds into segments of linear, planar, and spherical patches that are then refined by normalized cuts using proposed similarity criteria for shape patches to overcome over-segmentation. Finally, the segmented patches are further merged into segments of objects.

Experimental results show that the proposed method robustly partitions mobile laser point clouds of urban scenes into distinct patches and then combines the classified patches into objects, as demonstrated by extracting and classifying pole-like objects with an overall accuracy of 95.0%. The time costs of partitioning point clouds and classifying patches reveal that the proposed method efficiently segments mobile laser point clouds, in particular for large-volume datasets. It takes about 59 min to segment 2769 patches embedded in an urban scene of 105 million points. Calculation of point features consumes approximately 60% of the total time costs. On the other hand, the segmentation output of the proposed method shows promise for object extraction and recognition, as proved by the performance of pole-like object extraction. Although there is still room for improving the segments generated by the proposed method and for reducing the computing costs of point features, future work will focus on incorporating additional features such as colors and contextual features into object segmentation and extraction.

Acknowledgements

Work described in this paper was jointly supported by National Basic Research Program of China (No. 2012CB725301), the NSFC Project (No. 41071268), and Ph.D. Programs Foundation of Ministry of Education of China (No. 20120141110035). Special thanks go to anonymous reviewers for their constructive comments that substantially improve quality of the paper.

References

- Axelsson, P., 2000. DEM generation from laser scanner data using adaptive TIN models. *International Archives of Photogrammetry and Remote Sensing* 33 (Part B4), 110–117.
- Abuhadrous, I., Ammoun, S., Nashashibi, F., Goulette, F., Laurgeau, C., 2004. Digitizing and 3D modeling of urban environments and roads using vehicle-borne laser scanner system. In: 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems, Sendai, Japan, 28 September–2 October, pp. 76–81.
- Barnea, S., Filin, S., 2013. Segmentation of terrestrial laser scanning data using geometry and image information. *ISPRS Journal Photogrammetry and Remote Sensing* 76, 33–48.
- Biosca, J.M., Lerma, J.L., 2008. Unsupervised robust planar segmentation of terrestrial laser scanner point clouds based on fuzzy clustering methods. *ISPRS Journal of Photogrammetry and Remote Sensing* 63 (1), 84–98.
- Bienert, A., Scheller, S., Keane, E., Mohan, F., Nugent, C., 2007. Tree detection and diameter estimations by analysis of forest terrestrial laser scanner point clouds. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* XXXVI-3/W52, 50–55.
- Becker, S., 2009. Generation and application of rules for quality dependent façade reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing* 64 (6), 640–653.
- Bhanu, B., Lee, S., Ho, C., Henderson, T., 1986. Range data processing: representation of surfaces by edges. In: Proceedings of the Eighth International Conference on Pattern Recognition, Paris, France, pp. 236–238.
- Boyko, A., Funkhouser, T., 2011. Extracting roads from dense point clouds in large scale urban environment. *ISPRS Journal of Photogrammetry and Remote Sensing* 66 (6), S2–S12.
- Chang, C.-C., Lin, C.-J., 2011. LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology* 2 (27), 1–27.
- Dold, C., Brenner, C., 2004. Automatic matching of terrestrial scan data as a basis for the generation of detailed 3D city models. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 35 (Part B3), 1091–1096.
- Demantke, J., Mallet, C., David, N., Vallet, B., 2011. Dimensionality based scale selection in 3D lidar point clouds. *ISPRS Workshop on Laser Scanning 2011*, Calgary, Canada, 29–31 August, 6p (on CDROM).
- Filin, S., 2002. Surface clustering from airborne laser scanning data. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 34 (Part 3A), 117–124.
- Golovinskiy, A., Kim, V., Funkhouser, T., 2009. Shape-based recognition of 3D point clouds in urban environments. In: International Conference on Computer Vision, Kyoto, Japan, 27 September–4 October, pp. 2154–2161.
- Haala, N., Brenner, C., 1999. Extraction of buildings and trees in urban environments. *ISPRS Journal of Photogrammetry and Remote Sensing* 54 (2–3), 130–137.
- Hernández, J., Marcotegui, B., 2009. Point cloud segmentation towards urban ground modeling. *URBAN'09*, 5th GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas, Shanghai, China, 20–22 May 2009, 5p (on CDROM).
- Jiang, X., Bowyer, K., Morioka, Y., Hiura, S., Sato, K., Inokuchi, S., Bock, M., Guerra, C., Loke, R.E., du Buf, J.M.H., 2000. Some further results of experimental comparison of range image segmentation algorithms. In: 15th International Conference on Pattern Recognition, ICPR'00, Barcelona, Spain, 3–8 September, pp. 877–881.
- Kraus, K., Pfeifer, N., 1998. Derivation of digital terrain models in wooded areas. *ISPRS Journal of Photogrammetry and Remote Sensing* 53 (4), 193–203.
- Lehtomaki, M., Jaakkola, A., Hyppä, J., Kukko, A., Kaartinen, H., 2010. Detection of vertical pole-like objects in a road environment using vehicle-based laser scanning data. *Remote Sensing* 2, 641–664.
- Maas, H.-G., Vosselman, G., 1999. Two algorithms for extracting building models from raw laser altimetry data. *ISPRS Journal of Photogrammetry and Remote Sensing* 54 (2–3), 153–163.
- Manandhar, D., Shibasaki, R., 2001. Feature extraction from range data. In: Proceedings of the 22nd Asian Conference on Remote Sensing, Singapore, 5–9 November, 6p (on CDROM).
- Monnier, F., Vallet, B., Soheilian, B., 2012. Trees detection from laser point clouds acquired in dense urban areas by a mobile mapping system. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* I-3, 245–250.
- Overby, J., Bodum, L., Kjems, E., Ilsøe, P., 2004. Automatic 3D building reconstruction from airborne laser scanning and cadastral data using Hough transform. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 34 (Part B3), 296–301.
- Pauly, M., Keiser, R., Gross, M., 2003. Multi-scale feature extraction on point-sampled surfaces. *Computer Graphics Forum* 22 (3), 281–289.
- Pu, S., Rutzinger, M., Vosselman, G., Oude Elberink, S., 2011. Recognizing basic structures from mobile laser scanning data for road inventory studies. *ISPRS Journal of Photogrammetry and Remote Sensing* 66 (6), S28–S39.
- Rabbani, T., Dijkman, S., van den Heuvel, F., Vosselman, G., 2007. An integrated approach for modelling and global registration of point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing* 61 (6), 355–370.
- Sithole, G., Vosselman, G., 2003. Automatic structure detection in a point cloud of an urban landscape. In: Proceedings 2nd GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas: URBAN2003, Berlin, Germany, 22–23 May, pp. 67–71.
- Sappa, A., Devy, M., 2001. Fast range image segmentation by an edge detection strategy. In: IEEE Third International Conference on 3D Digital Imaging and Modeling, Quebec, Canada, 28 May–1 June, 8p (on CDROM).
- Shi, J., Malik, J., 2000. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22 (8), 888–905.
- Tang, C.-K., Medioni, G., 2002. Curvature-augmented tensor voting for shape inference from noisy 3D data. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24 (6), 858–864.
- Vosselman, G., Mass, H.-G., 2010. Airborne and Terrestrial Laser Scanning. Whittles Publishing, Scotland, UK.
- Vosselman, G., Gorte, B.G.H., Sithole, G., Rabbani, T., 2004. Recognising structure in laser scanner point clouds. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 46 (part 8/W2), 33–38, Freiburg, Germany, 4–6 October.
- Wani, M.A., Arabnia, H.R., 2003. Parallel edge-region-based segmentation algorithm targeted at re-configurable multiring network. *The Journal of Supercomputing* 25 (1), 43–62.
- Yang, B.S., Wei, Z., Li, Q., Li, J., 2013. Semi-automated building facade footprint extraction from mobile lidar point clouds. *IEEE Geoscience and Remote Sensing Letters* 10 (4), 766–770.
- Yang, B.S., Wei, Z., Li, Q., Li, J., 2012a. Automated extraction of street-scene objects from mobile lidar point clouds. *International Journal of Remote Sensing* 33 (18), 5839–5861.
- Yang, B.S., Fang, L.N., Li, Q., Li, J., 2012b. Automated extraction of road markings from mobile lidar point clouds. *Photogrammetric Engineering & Remote Sensing* 78 (4), 331–338.