

A Three-Step Approach for TLS Point Cloud Classification

Zhuqiang Li, Liqiang Zhang, Xiaohua Tong, Bo Du, *Senior Member, IEEE*, Yuebin Wang, Liang Zhang, Zhenxin Zhang, Hao Liu, Jie Mei, Xiaoyue Xing, and P. Takis Mathiopoulos, *Senior Member, IEEE*

Abstract—The ability to classify urban objects in large urban scenes from point clouds efficiently and accurately still remains a challenging task today. A new methodology for the effective and accurate classification of terrestrial laser scanning (TLS) point clouds is presented in this paper. First, in order to efficiently obtain the complementary characteristics of each 3-D point, a set of point-based descriptors for recognizing urban point clouds is constructed. This includes the 3-D geometry captured using the spin-image descriptor computed on three different scales, the mean RGB colors of the point in the camera images, the LAB values of that mean RGB, and the normal at each 3-D point. The initial 3-D labeling of the categories in urban environments is generated by utilizing a linear support vector machine classifier on the descriptors. These initial classification results are then first globally optimized by the multilabel graph-cut approach. These results are further refined automatically by a local optimization approach based upon the object-oriented decision tree that uses weak priors among urban categories which significantly improves the final classification accuracy. The proposed method has been validated on three urban TLS point clouds, and the experimental results demonstrate that it outperforms the state-of-the-art method in classification accuracy for buildings, trees, pedestrians, and cars.

Index Terms—Discriminative feature, multilabel graph-cut, object-oriented decision tree, optimization, point cloud.

I. INTRODUCTION

DURING the few last decades, advances in laser scanning technology have led to the development of research and development (R&D) activities related to the processing of large-scale outdoor ground scenes such as building modeling or detection [1]–[3] and tree modeling [4]–[6]. For such applications, precise automated classification procedures that can preprocess complex 3-D point clouds in a variety of natural

environments are highly desirable [7]. Although in recent years many researchers have made advancements in this area of R&D activities, it still remains a challenging task to achieve high-quality classification results mainly due to the following reasons: 1) data sets of important objects, such as trees, are often incompletely captured due to the limited scanning resolution or occlusions by the front objects; 2) raw data can be typically noisy and cluttered because of the nonstationary interference created by the winds; and 3) the lack of effective methods to efficiently derive discriminative characteristics of the objects in complex urban environments.

In the past, numerous research activities have been devoted to improve the classification of point clouds acquired by laser range sensors [7]–[11]. Patterson IV *et al.* [12] have recognized objects by using the EGIs [13] and spin-image descriptors. Golovinskiy *et al.* [8] first extracted small objects from the combination of mobile laser scanning and airborne laser scanning (ALS) point clouds. Then, they used a support vector machine (SVM) classifier to classify the objects into semantic groups such as cars, streetlights, trees, and fire hydrants. Although their approach can localize and segment small objects, the classification results are not very accurate because the shape features are not discriminative enough. It is noted that, for both methods reported in [8] and [12], label objects at the object level. Clearly, if the objects cannot be separated from the point clouds well, it is difficult to achieve good classification results of these two methods.

In another approach [14], during detecting and reconstructing complex 3-D watertight buildings using ALS point clouds, vegetation, buildings, and terrains were accurately classified using the graph-cut optimization technique followed by a hierarchical Euclidean clustering method. However, because of the unreasonable constraints, which have been considered in [14], e.g., that the boundaries of all parts of a complex rooftop are assumed as rectilinear, this approach decreases the accuracy of the building models. Furthermore, for complex and cluttered scenes, it is very difficult to achieve high classification performance based only upon the shape-based features of the objects. In another approach, Mallet *et al.* [15] applied full-waveform LiDAR point clouds to classify urban areas into buildings, ground, and vegetation. In their method, waveform features, including echo amplitude and radiometric features, are extracted as input for an SVM classifier.

In [16], the reflected laser intensity is utilized to distinguish marls and lime stones. Due to the difficulty in precisely correcting for the distance and incidence effects [17], this method cannot be applied to 3-D surfaces [18]. Following another

Manuscript received December 24, 2015; revised March 12, 2016; accepted April 30, 2016. Date of publication May 24, 2016; date of current version August 2, 2016. This work was supported by the National Natural Science Foundation of China under Grant 41371324. (Corresponding author: Liqiang Zhang.)

Z. Li, L. Zhang, Y. Wang, L. Zhang, Z. Zhang, H. Liu, J. Mei, and X. Xing are with the State Key Laboratory of Remote Sensing Science, Beijing Normal University, Beijing 100875, China (e-mail: skybelongtous@foxmail.com; zhanglq@bnu.edu.cn; xxgcdxwyb@163.com; yammer_0303@126.com; zhenxin066@163.com; apprentice_g@163.com; meijie@mail.bnu.edu.cn; xyxing@mail.bnu.edu.cn).

X. Tong is with the School of Surveying and Geo-informatics, Tongji University, Shanghai 200092, China (e-mail: xhtong@tongji.edu.cn).

B. Du is with the Computer School, Wuhan University, Wuhan 430072, China (e-mail: gunsapace@163.com).

P. T. Mathiopoulos is with the Department of Informatics and Telecommunications, National and Kapodistrian University of Athens, 157 84 Athens, Greece (e-mail: mathio@di.uoa.gr).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TGRS.2016.2564501

approach, Antonarakis *et al.* [19] used elevation and LiDAR intensity data to broadly classify forest and ground types using a supervised object orientated method. Pu *et al.* [20] propose a two-step approach for recognizing urban basic structures from mobile laser scanning data. In particular, first ground, on-ground objects, and off-ground objects were largely classified. Based upon these three categories, by analyzing the geometrical attributes and topological relations among urban objects, traffic signs, trees, building walls, and barriers were further precisely classified by using decision trees. In another research effort, in order to classify urban 3-D point clouds derived from oblique aerial imagery and vertical aerial imagery, a rule-based hierarchical semantic classification scheme that utilizes spectral information and geometry- and topology-related features has been developed [21].

We know that natural surfaces are heterogeneous and their distinctive properties are hardly defined at a unique scale. To overcome the difficulty in choosing a specific spatial scale at which natural geometrical features must be analyzed, the multiscale criteria are effective for purely geometry-based classification of point cloud scenes [18]. Lim and Suter [22], [23] presented a supervoxel method for 3-D terrestrial range data classifications in which they applied multiscale conditional random field methods to classify the terrestrial laser scanning (TLS) point clouds. However, they only label some large objects such as trees, grounds, and buildings but do not try to label small objects such as cars and pedestrians. In another approach, a multiscale and hierarchical framework was presented to describe the classification of TLS point clouds of cluttered urban scenes [6]. In particular, point clouds were first resampled into different scales, and then, the resampled data set of each scale was aggregated into several hierarchical point clusters. In [24], a multiscale point cloud surface classification approach was developed. More specifically, three multiscale operators, namely, approximate shape, density of a defined neighborhood, and distance of its mean point from its geometric center, are fused into a single feature vector used to identify individual point observations corresponding to rock surfaces in TLS point clouds. However, most of these methods ignore the context information. For example, cars are found on streets, often in a line [8], or the average height of a building above the ground should be greater than a certain value [25]. Use of this type of information can help improve the classification results.

To enhance the recognition of on-ground objects from complex and cluttered scenes, we propose a new method for point cloud classification which considers point clouds of an urban environment as input. In particular, we classify each point in the point cloud into one semantic class (building, tree, pedestrian, or car), using a linear SVM classifier trained on the feature descriptor. Then, a global as well as a local optimization algorithm is proposed, with the purpose of further refining the previous labeling results. Fig. 1 illustrates the main procedure for classifying urban point clouds. More specifically, at the first step, we compute the feature descriptor per 3-D point. The descriptor is composed of a set of lightweight 3-D features. Next, a linear SVM classifier is trained on the descriptor, and the initial 3-D labeling of the urban environment is generated. As the second and third steps, the global and local optimization algorithms

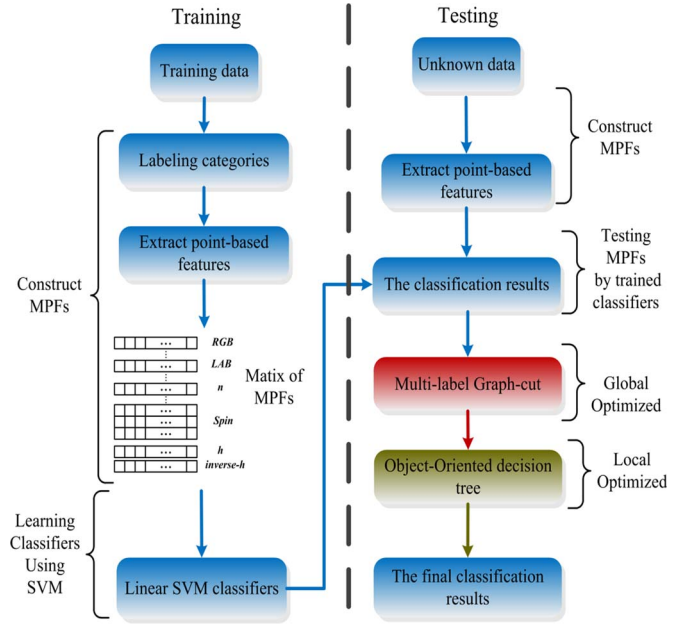


Fig. 1. Methodology for TLS point cloud classification.

are employed to further refine the labeling results. The multi-label graph-cut-based global optimization makes the labeling spatially local continuity and global optimization. Based on the global optimization, the object-oriented decision-tree-based local optimization that uses weak priors reduces the misclassification and significantly increases the classification accuracy.

In summary, the contributions of our approach are threefold.

- 1) A three-step point cloud parsing framework is presented to formulate and solve a joint classification and segmentation problem from TLS point clouds.
- 2) A set of lightweight descriptors for each 3-D point is constructed. They can effectively characterize the shape and color information of the objects in each class.
- 3) The global and local optimizations integrated with prior knowledge are performed on the initial labeling of the 3-D scene to greatly enhance the recognition accuracy of the point clouds in cluttered urban environments.

II. FEATURE CONSTRUCTION AND LEARNING FEATURES

A. Construction of Multipoint-Based Features

In real outdoor scenes, the structures of tree crowns are generally complicated, while the shapes of building facades and roofs are smooth. The point-based features can preferably discriminate the shape-specific objects. Because the local regions of tree trunks and cars are planar, the points on them are often misclassified into the building class through the point-based features. The point clusters [6] or supervoxels [22] based classification algorithms can usually overcome the drawback. However, the object shapes in a point cluster and supervoxel are often incomplete because of oversegmentation or undersegmentation, which reduce the classification performance.

Since the point-based features are sensitive to noise, for each point p_i in the point cloud, we use the spin images with three

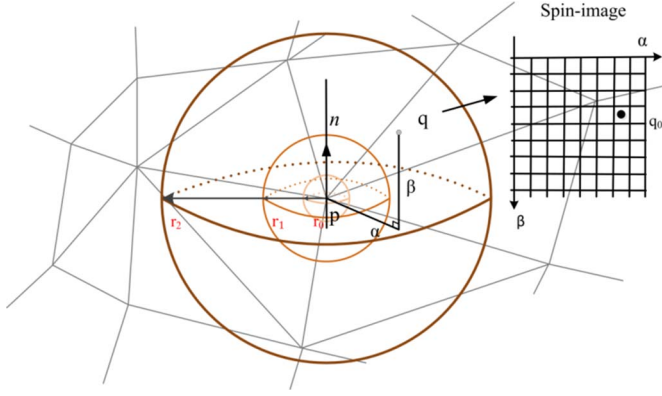


Fig. 2. Construction of multiscale spin images.

scales to capture the feature of the region around p_i . The terrain points are first removed by using the method in [26], and then, the on-ground points are organized by the k -d tree. Let $N_p = \{q | q \text{ is one of the neighboring points of } p_i, \text{ and } |p_i - q| \leq r\}$ be the point set within the sphere of which radius is r and centroid is p_i . N_p is the support region of p_i . The normal vector of a point in N_p is taken as the rotational axis of the spin image. Equations (1) and (2) are applied to compute the coordinate of p_i in the spin image

$$\alpha = \sqrt{\|q - p_i\|^2 - [n_i \cdot (q - p_i)]^2} \quad (1)$$

$$\beta = n_i \cdot (q - p_i) \quad (2)$$

where α denotes the x coordinate of p_i in the spin image, β denotes the y coordinate of p_i , and n_i denotes the normal vector of p_i .

We create an 8×8 spin image [27] for each point. To reduce the computation complexity, the first, second, third, fifth, and seventh columns of the spin image are selected to form a 5×8 spin image for the point. As shown in Fig. 2, three spheres of which radii are 0.2, 0.8, and 1.4 m are constructed, respectively. Afterward, the spin features are extracted from the three spheres for generating a $5 \times 8 \times 3$ spin-image feature F_{spin} .

Inspired by [28], we introduce the local feature descriptors and color texture features to distinguish the different categories from the point clouds. The local feature of N_p is expressed by the normal vector n_i of each point p_i . Specifically, we obtain n_i by calculating the eigenvalue and eigenvector of the covariance matrix of p_i . The covariance matrix C_i of p_i is derived from N_p

$$C_i = \frac{1}{g} \sum_{t=1}^g (p_t - \bar{p})(p_t - \bar{p})^T$$

$$C_i \cdot \vec{v}_j = \eta_j \cdot \vec{v}_j, \quad j \in \{0, 1, 2\} \quad (3)$$

where T is the transposition operator, g is the number of points in the support region N_p , \bar{p} is the centroid of the point cluster in N_p , η_j is the j th eigenvalue of C_i , and (\vec{v}_j) is the j th eigenvector of C_i . The eigenvalues η_0 , η_1 , and η_2 ($\eta_0 > \eta_1 > \eta_2$) are obtained by performing the eigendecomposition of the covariance matrix C_i . The eigenvector (\vec{v}_2) of the smallest eigenvalue η_2 in C_i is taken as the normal vector n_i of p_i .

In addition to the spin-image descriptor computed on three scales and the normal vector of each point, we also compute a set of descriptors for each 3-D point: the mean RGB colors of the point as seen in the camera images, the LAB values of that mean RGB [29], the point's height (height) above the estimated ground plane, and its inverse height (height^{-1}), defined as the distance from the uppermost point of the object in the direction of the gravity vector. In this paper, we call all of the multipoint-based features MPFs. Therefore, the MPFs F_i of p_i are expressed as (4). F_i is a 131-dimensional vector

$$F_i = [\text{RGB}_i^T \text{LAB}_i^T n_i^T F_{\text{spin}}^T \text{height}_i^T \text{height}_i^{-1T}] \quad (4)$$

B. Learning Features Using the Linear SVM Classifier

After the MPFs are extracted, the linear SVM classifier is utilized to recognize the point cloud. We label the four categories: cars, trees, pedestrians, and buildings in the training data set. Each sample in the training data set has a label and several attributes. They can be described mathematically as follows.

Given a training data set $(F_i, fy_i), i = 1, \dots, n_{\text{train}}$, $F_i \in R^n$ denotes the feature vectors of the samples, and $fy \in \{-1, 1\}^{n_{\text{train}}}$ denotes the category vector of the samples. δ is the penalty factor which indicates how much penalties to the misclassified points. The SVM classifier is used to compute the optimal solution of the following positive definite programming

$$\min_{w, b, \xi} \frac{1}{2} w^T w + \delta \sum_{i=1}^{n_{\text{train}}} \xi_i$$

$$\text{s.t. } fy_i (w^T \varphi(F_i) + b) \geq 1 - \xi$$

$$\xi \geq 0. \quad (5)$$

Once the linear SVM classifier is trained using the training data set and the MPFs are extracted from the unknown data, we can compute the probability per point belonging to a certain category by using the trained SVM classifier.

III. CLASSIFICATION OPTIMIZATION

A. Optimization Based on Multilabel Graph-Cut

To reduce the misclassification caused by the aforementioned procedure, we adopt the multilabel graph-cut [30], [31] to globally optimize the classification results. The optimization based on multilabel graph-cut is generally energy modeled by the first-order Markov random fields. If each pixel in the image pixel set P is given a label l_p , we find the posterior probability to maximize P . Here, we introduce this idea to segment the point cloud. The classification results obtained by employing the linear SVM classifier are abstracted as a global optimization problem to minimize the Gibbs energy function, which is computed by

$$E(L|\Phi, V) = \sum_{i \in V} D_i(l_i, \Phi) + \lambda \cdot \sum_{(i,j) \in M} V_{ij}(l_i, l_j) \quad (6)$$

where D_i is the data term and l_i is the class label of the i th point. $V_{ij}(i, j)$ is the smooth term of points i and j , and it is an

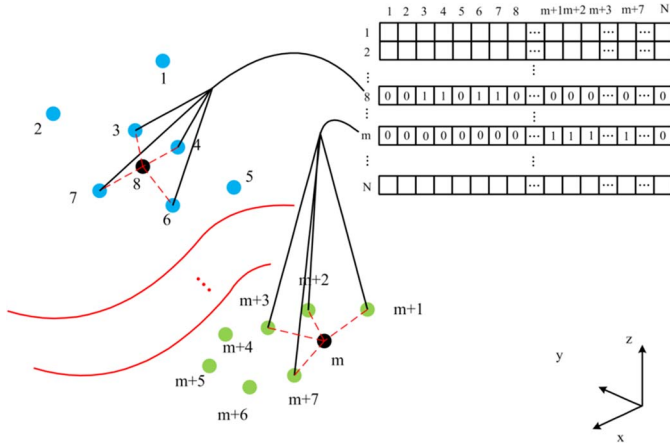


Fig. 3. Adjacency sparse matrix of the point cloud with $k = 4$ in KNN.

$N \times N$ sparse matrix, in which N is the number of points in the point cloud. It enforces spatially smooth labeling solutions over the point by penalizing occurrences of adjacent points i and j , obtaining different labels ($l_i \neq l_j$). M is a neighborhood system. The set V is the point cluster of the point cloud, the label set S is the single connected components within the classified point cloud $\Phi = \{g_i\}$, and the pixel label is analogous to the class label l_i of the current point. λ is the coefficient of the smooth term, i.e.,

$$\lambda = \exp \left(- \left(\frac{d_{ij}}{\sigma} \right)^2 \right) \quad (7)$$

where d_{ij} is the Euclidean distance between points i and j and σ is the expectation of all neighboring distances. All neighboring distances are approximately equal when $k = 4$ in KNN, so $d_{ij}/\sigma \approx 1$. We find the four closest neighboring points of each point by using the KNN algorithm and then generate an adjacency sparse matrix. In the matrix, if the cost value of two points is equal to 1, they are neighbors.

In Fig. 3, the blue and green points belong to different two classes. Nodes 3, 4, 6, and 7 are the neighbors of node 8. Similarly, nodes $m+1$, $m+2$, $m+3$, and $m+7$ are neighborhoods of point m . By using the aforementioned method, we generate an $N \times N$ adjacency sparse matrix as the smooth term of the energy function $E(L|\Phi, V)$. Then, optimize it to minimize the distances in the same classes, and let the separation distance between different classes be as short as possible. The min-cut is utilized to make the optimization convergence. Through this way, the labeling of the point cloud obtained by the optimization is local continuous and global optimal.

B. Local Optimization Based on the Object-Oriented Decision Tree

After the classification results are globally optimized by the multilabel graph-cut, some points are still misclassified. Next, we present the object-oriented decision tree integrated with the prior knowledge to distinguish the confusion categories.

1) *Segmentation-Based DoN*: The difference of normals (DoN) [33] as a multiscale operator is utilized to segment the

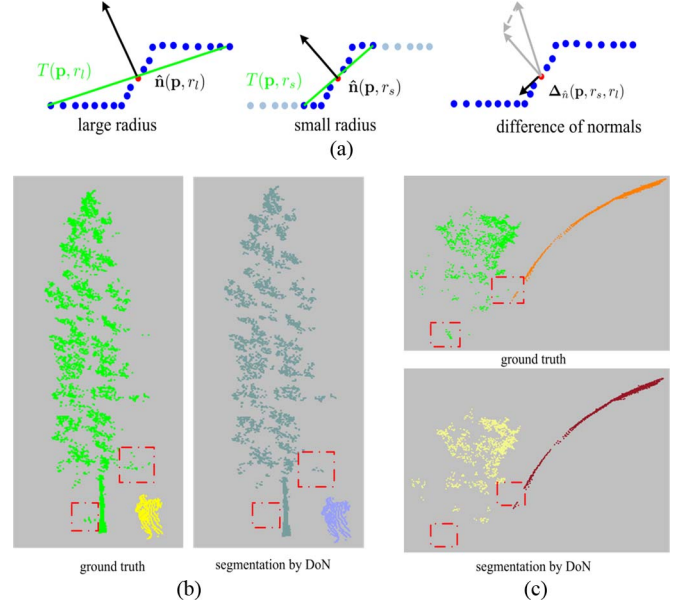


Fig. 4. Effect of the support radius on the estimated surface normals for a point cloud and the segmentation result. (a) comes from [33].

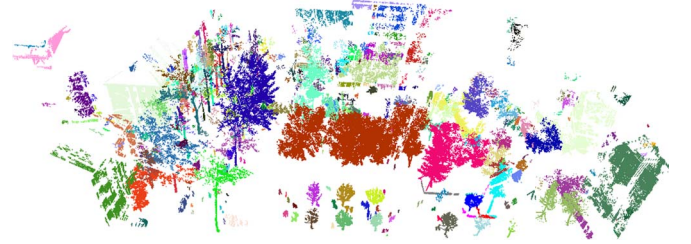


Fig. 5. Segmentation result obtained by employing the normal vectors with two scales. The maximum segmented class has 325 693 points, and the minimum segmented class has 40 points.

point clouds. The primary motivation behind the DoN is that the estimated surface normals at any given radius reflect the underlying geometry of the surface at the scale of the support radius. The normals are always estimated with a support radius which determines the scale in the surface structure which the normal represents. For each point p , two unit point normals $\hat{n}(p, r_l)$ and $\hat{n}(p, r_s)$ are estimated with different radii $r_l > r_s$. The normalized vector difference of these point normals defines the DoN operator

$$\Delta n(p, r_s, r_l) = \frac{n(p, r_s) - n(p, r_l)}{2}. \quad (8)$$

In (8), given the support radius r , $\hat{n}(p, r)$ is the surface normal at point p .

Then, a Euclidean distance threshold-based clustering algorithm [34] is applied to the recognized point cloud by the DoN for clustering the interested objects. In our experiment, the distance tolerance is equal to 0.8 m, and the number of points in each cluster ranges from 40 and 1 000 000.

Fig. 4(a) illustrates the segmentation effect under the perspective of a 1-D. Normals \hat{n} and tangents T estimated with a small support radius r_s are affected by small-scale surface structures. The normals and tangent planes estimated with a

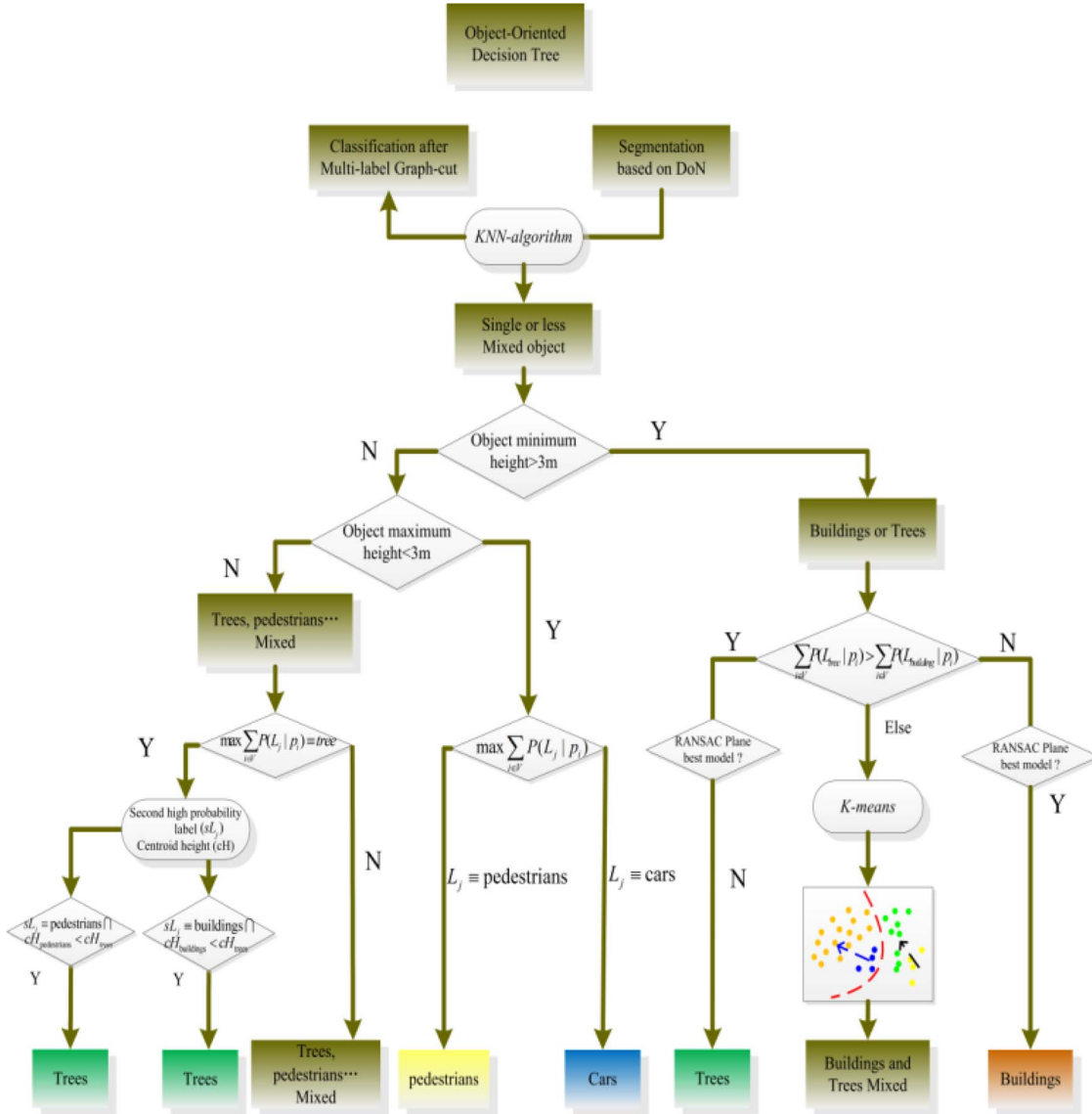


Fig. 6. Procedure for optimizing the classification results based on the object-oriented decision tree.

large support radius r_l are less affected by small-scale structures and represent the geometry of larger scale surface structures. In Fig. 4(b) and (c), the left figures illustrate the ground truth, and the right ones describe the segmentation results through the DoN. It is noted that most of the isolated points (in the red dotted line box) are removed, and different object classes are correctly recognized.

We use the normal vectors with two scales ($r_s = 0.15$ m and $r_l = 1.6$ m) to find the points that fall within the scales and perform a scale-based segmentation. Fig. 5 shows the segmentation result by using the normal vectors with two scales. The points with different colors indicate different object classes. Different from the superpixel segmentation, the aforementioned method guarantees the integrity of the object classes, and each of the segmented classes is independent.

2) *Optimization Through Object-Oriented Decision Tree:* In the TLS point cloud, the shapes of tree trunks are often similar with pedestrians; thus, they are often misclassified into the same class. This type of misclassification is not rectified by using

the multilabel graph-cut. Some weak priors can be utilized to distinguish instances of different categories in the urban environment. For example, the average height of a building above the ground should be greater than a certain value. Inspired by this, we can further improve the classification accuracy with the help of such information.

We have applied the DoN to derive the points of each segmented object class. Then, the KNN algorithm is used to obtain the labeled point set of the segmented object classes in the previous classification results. Afterward, the object segmented object class integrated with the priors is utilized to optimize the classification results. Based on the prior knowledge, we can make initial decisions by using the minimum height H_{\min} of each segmented class, which can be defined as $H_{\min} = \min\{z | (x_i, y_i, z_i) \in P_{\text{object}}\} - \min Z_{\text{ground}}$, where P_{object} is the segmented point set and Z_{ground} is the height of the terrain surface. Generally, the height of a pedestrian or a car is less than 3 m. If $H_{\min} > 3$ m, the class is probably a building or a tree. This means that it is a pure class. Otherwise, this class may

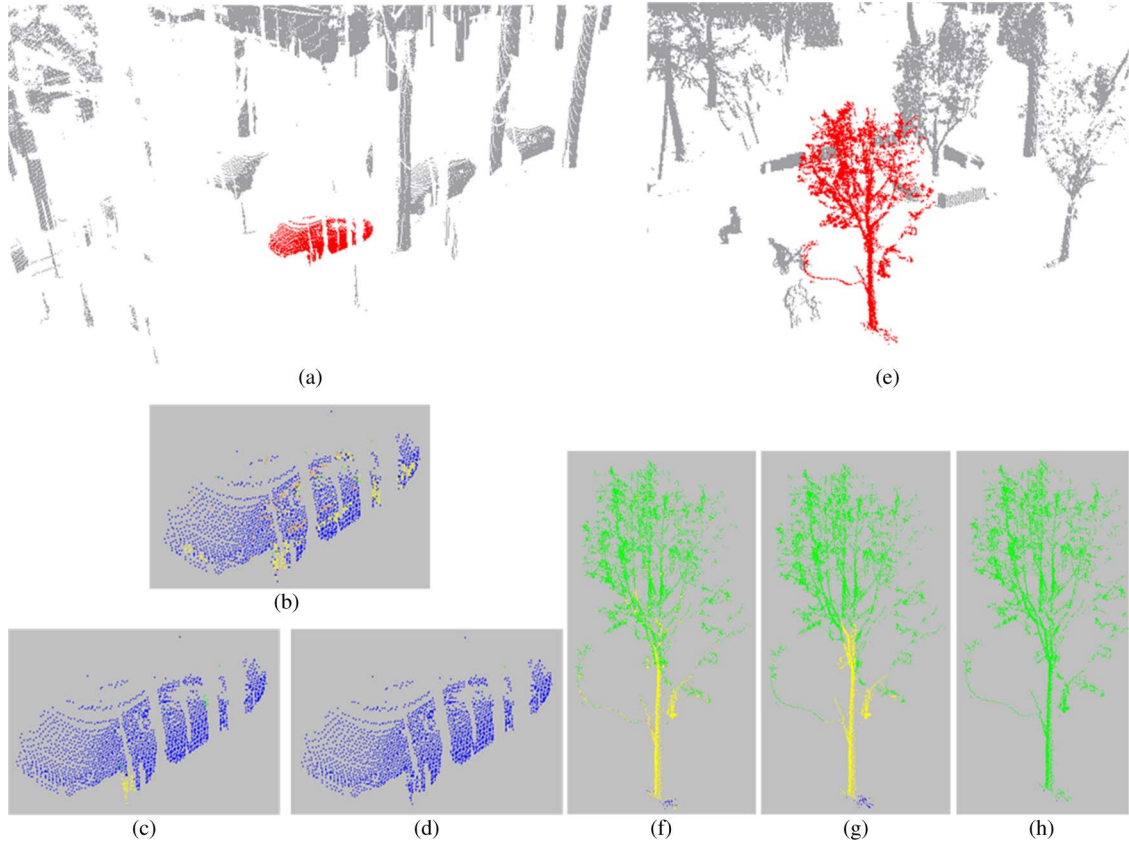


Fig. 7. Comparison of the classification results obtained by using the linear SVM classifier with the results optimized by performing the global and local optimizations.

contain buildings and trees. We need to determine the planarity of the shape of the class through the conditional probability and the RANSAC plane model [32]. Fig. 6 shows the local optimization flow through the object-oriented decision tree.

Specifically, a planar model for all points in the class is estimated by the RANSAC model, and then, this planar model is used to test the remaining points in the class. The points which are suitable for the model are regarded as the correct points. If there are sufficient points that are classified as the correct points, we regard the estimated model as valid. When the ratio between the correct points and the points in the class is larger than 80%, we believe that the object in the class has a good planarity; thus, the probability of the object belonging to the building class is the largest. If the object class is hybrid, it contains at least two different object categories. We first use the k -means algorithm to obtain the cluster centers of pedestrians and cars. Then, we calculate the distances between the centers and the trees or buildings and merge the points into the nearest class (trees or buildings).

If $H_{\min} < 3$ m, the classification results have several probabilities. As shown in the left tree of Fig. 6, we make decisions by using the maximum height H_{\max} of each class to be segmented. $H_{\max} = \max\{z | (x_i, y_i, z_i) \in P_{\text{object}}\} - \min Z_{\text{ground}}$. If $H_{\max} < 3$ m, this class probably belongs to pedestrians or cars. Maybe, it also belongs to the tree or building class. We use $\max \sum_{i \in V} P(l_i | p_i)$, which is the probability of a point p_i that belongs to the label l_i , to determine the type of the object. In the aforementioned optimization based on the multilabel graph-cut,

tree trunks tend to be wrongly classified as buildings or pedestrians. We combine the contextual information to analyze the centroid height (cH), the first high probability label, and the second high probability label (sLj). We first judge that the first high probability label of this object is a tree, and the second probability label is a pedestrian or a building. Second, we decide the centroid heights of these two objects by $cH_{\text{tree}} > cH_{\text{pedestrians}}(cH_{\text{building}})$. By using such information, the accuracy of the final classification labels is much improved.

We compare the classification result obtained by using the linear SVM classifier on the MPFs with the result optimized by performing the global and local optimizations. The red points in Fig. 7(a) and (e) denote cars and trees, respectively. Fig. 7(b) shows the car classification result by using the linear SVM classifier. It is observed that there are many points on the head of the car, which are wrongly classified as the pedestrian or building class. Fig. 7(c) illustrates the classification result which is optimized by the multilabel graph-cut algorithm. Some wrong points on the car have been recognized correctly. However, there are still some wrong points on the wheels. Fig. 7(d) shows the classification result which is optimized by the multilabel graph-cut and object-oriented decision tree. The points belonging to other classes have been removed from the car. Fig. 7(f) shows the tree classification result by using the linear SVM classifier. Because the tree trunk much resembles a pedestrian in shape, it is difficult to distinguish them. Fig. 7(g) shows the similar problem as Fig. 7(f). The final results are satisfactorily refined through the two optimization procedures.

TABLE I
SUMMARY OF THE THREE DATA SETS

| | Scene I | Scene II | Scene III |
|--|-------------------|------------------|------------------|
| Scanners | RIEGL LMS-Z620 | RIEGL vz-1000 | RIEGL vz-1000 |
| Measurement principle | Time of flight | | |
| Measurement range | 2-2000meters | 2.5-1400meters | |
| Accuracy | 10mm | 8mm | |
| Precision | 5mm | 5mm | |
| Beam divergence | 0.15mrad | 0.3mrad | |
| Laser wavelength | Near infrared | 1,550nm | |
| Horizontal and vertical-angle spacing | 0.57° | 0.60° | 0.60° |
| Car points | 31,026 | 171,220 | 72,835 |
| Tree points | 1,427,674 | 506,719 | 726,200 |
| Pedestrian points | 102,326 | 65,372 | 32,686 |
| Building points | 213,504 | 496,307 | 182,483 |
| Total points | 1,774,530 | 1,239,618 | 1,015,204 |

IV. RESULTS AND DISCUSSIONS

To validate the performance of our method, we perform both qualitative and quantitative evaluations on three scanned urban scenes.

A. Data Sets

The point clouds of three scenes were obtained by TLS scanners in a single scan. The majority of the objects that appear in the urban scenes are buildings, trees, cars, and pedestrians. The point density of the point clouds in the three scenes varies a lot with respect to the distance of the scanned objects to the scanners. Moreover, the point clouds acquired by the single scan are often incomplete due to occlusions. For example, the point clouds of the faraway objects are often sparse and fragmented because the laser is blocked by the closer objects. To train the classifiers and evaluate the presented method quantitatively, we further assigned the ground truth labels on the three data sets manually. The statistics of the data sets are shown in Table I.

The selected training data set is the same as that in [6], which is illustrated in Fig. 8. The whole training data set contains 647 635 points including 67 604 pedestrian points, 303 994 tree points, 141 820 building points, and 134 217 car points. The points of trees and buildings in the training data set are mainly obtained from scene I. The points of pedestrians in the training data set are mainly obtained from scene II. The points of cars in the training data set are mainly obtained from scene III. Some particular objects are also selected, e.g., the ladders classified as building only appear in scene II, and the baby carriages classified as cars only appear in scene I.

The following evaluations consist of two parts: 1) the comparison of the precision and recall of our method with those of other methods and 2) the scale selection of the spin images with three scales in the feature extraction and the sensitivity of k in KNN are tested.

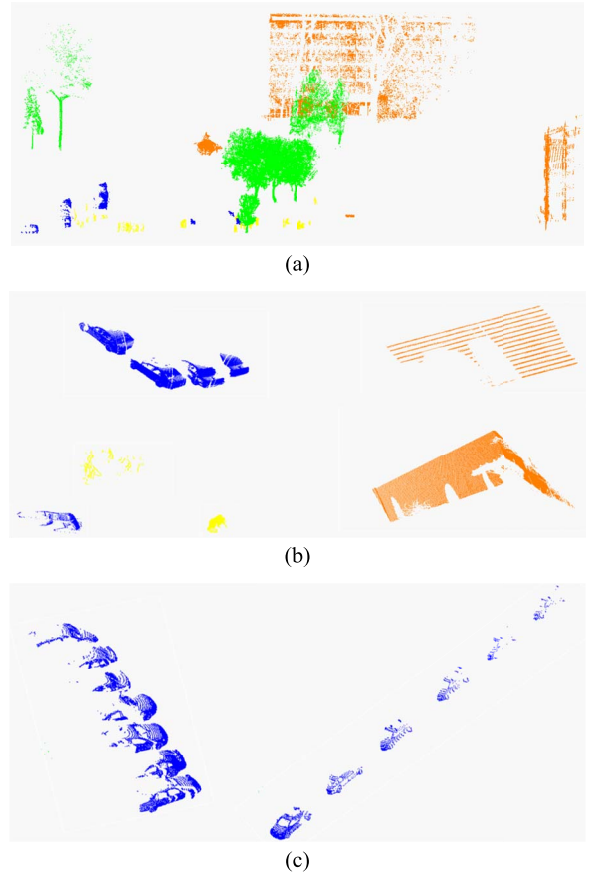


Fig. 8. Training data. Blue points are on cars, green points represent trees, yellow points are on pedestrians, and orange points are buildings. (a) Part of the training data set obtained from scene I. (b) Part of the training data set obtained from scene II. (c) Part of the training data set obtained from scene III.

B. Classification Accuracy Evaluation

We compare our method with the one in [6] in terms of the precision/recall, F_1 measure (9), and accuracy. In [6], multi-scale and hierarchical point clusters (MHPCs) are presented to capture multiscale properties and hierarchical structures of the TLS point clouds. Based on the MHPCs, the features of point clusters are constructed by employing the latent Dirichlet allocation. An AdaBoost classifier is trained in each point-cluster set, and the corresponding classifiers are obtained to separate the TLS point clouds into semantic regions

$$F_1 = \frac{2(\text{recall} \times \text{precision})}{\text{recall} + \text{precision}}. \quad (9)$$

Precision/recall can be used to represent the classification quality [6]. Precision is the fraction of retrieved instances that are relevant, while recall is the fraction of relevant instances that are retrieved. High precision means that an algorithm returned substantially more relevant results than irrelevant, while high recall means that an algorithm returned most of the relevant results. Table II compares the classification performance of our method with that of the method [6] in terms of precision/recall, F_1 measure, and accuracy. In our experiment, the parameter spin radii ($r_0=0.2$ m, $r_1=0.8$ m, $r_2=1.4$ m, and $k=4$ in KNN) are the same in the three scenes. From Table II, it is noted that the precision and recall of buildings and trees are quite stable, while

TABLE II
PRECISION/RECALL, F_1 MEASURE, AND ACCURACY
ON THE THREE SCENES

| Scene I | Cars (%) | Trees (%) | Pedestrian s (%) | Buildings (%) | Accu racy |
|---|-------------|-------------|---------------------|------------------|--------------|
| Classified by MPFs | 59.52/46.16 | 97.54/81.12 | 29.42/92.43 | 72.60/82.19 | 81.2 |
| | 52.00 | 88.58 | 44.63 | 77.10 | 9 |
| Multi-label Graph-cut | 81.95/51.24 | 98.76/84.68 | 32.11/97.90 | 85.81/88.01 | 85.2 |
| | 63.06 | 91.18 | 48.36 | 86.90 | 6 |
| LDA model [6] | 52.9/45.4 | 95.4/98.3 | 82.9/62.7 | 89.9/86.7 | 93.4 |
| | 48.8 | 96.8 | 71.4 | 88.3 | |
| Object-ori ented decision tree | 91.38/93.76 | 96.86/97.47 | 80.37/67.82 | 90.46/93.68 | 95.1 |
| | 92.56 | 97.17 | 73.56 | 91.75 | 7 |
| Scene II | Cars (%) | Trees (%) | Pedestrians (%) | Building (%) | Accu racy |
| Classified by MPFs | 83.82/92.82 | 91.91/84.43 | 66.76/89.01 | 89.85/90.04 | 88.0 |
| | 88.09 | 88.01 | 76.30 | 89.95 | 8 |
| Multi-label Graph-cut | 85.69/96.91 | 95.87/89.31 | 74.1/91.64 | 94.43/93.81 | 92.2 |
| | 90.95 | 92.48 | 81.94 | 94.12 | 9 |
| LDA model [6] | 83.7/86.4 | 95.9/90.1 | 78.8/77.5 | 89/93.3 | 90.2 |
| | 85.03 | 92.91 | 78.14 | 91.1 | 0 |
| Object-ori ented decision tree | 88.68/97.81 | 94.72/94.33 | 86.32/96.47 | 97.52/92.96 | 94.3 |
| | 93.02 | 94.52 | 91.11 | 95.18 | 7 |
| Scene III | Cars (%) | Trees (%) | Pedestrians (%) | Buildings (%) | Accu racy |
| Classified by MPFs | 88.4/89.99 | 95.76/90.02 | 75.14/91.59 | 72.70/86.59 | 88.0 |
| | 89.19 | 92.81 | 82.55 | 79.04 | 8 |
| Multi-label Graph-cut | 92.66/95.67 | 97.16/95.11 | 80.15/92.81 | 86.59/90.29 | 92.2 |
| | 94.14 | 96.12 | 86.01 | 88.40 | 9 |
| LDA model [6] | 77.4/85.6 | 98.2/95.6 | 84.9/69.4 | 83.7/92.3 | 93.3 |
| | 81.29 | 96.88 | 76.37 | 87.8 | |
| Object-ori ented decision tree | 97.96/99.92 | 99.55/99.48 | 97.54/92.66 | 98.04/98.39 | 99.1 |
| | 98.93 | 99.52 | 95.04 | 98.22 | |

those of cars and pedestrians fluctuate greatly. Simultaneously, we find that the numbers of points on cars and baby carriages in this scene are less than 2.0% of tree points, of which the number of points on cars is less than 0.6% of tree points. It leads to a low classification precision due to the insufficient learning during the training stage. Through the global optimization, F_1

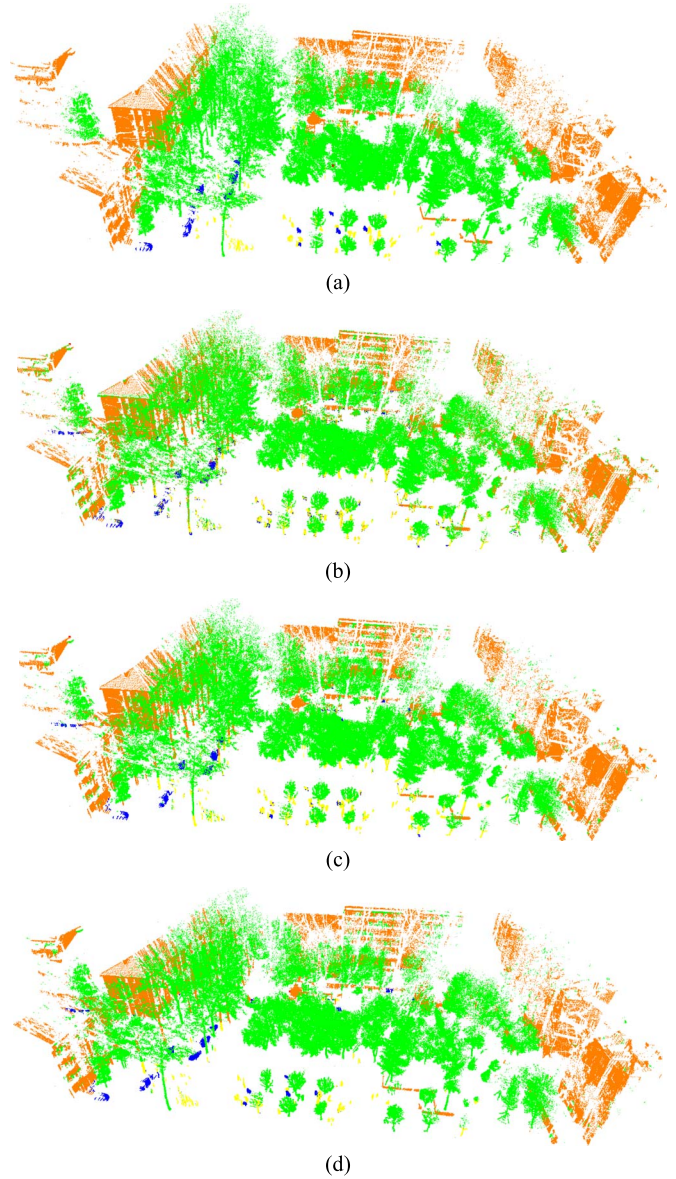


Fig. 9. Classification results obtained by the three-step approach in scene I. Points on cars, trees, pedestrians, and buildings are colored in blue, green, yellow, and orange, respectively. (a) Ground truth. (b) Classification result obtained by using the linear SVM. (c) Classification result optimized by the multilabel graph-cut. (d) Classification result in (c) optimized by the object-oriented decision tree.

reaches 63.06% and then 92.6% after the local optimization is performed. Pedestrians keep on moving throughout the scan, which stretches the pedestrian shape and makes the texture of pedestrians blurry. Moreover, there are only a few points on them. This makes the classification precisions of the pedestrian class in the three scenes not high. In scene I, the recall of pedestrians exceeds 92%, whereas the precision is quite low, i.e., between 30% and 40%. Because the shapes and heights of tree trunks are similar to those of pedestrians, it is easy to wrongly classify the tree trunks as pedestrians, which leads to a low precision. After the optimization, the precisions of all of the classes are improved significantly.

In terms of the F_1 measure, the accuracy of the car class obtained by our method is about 23.1% higher than that obtained by

the method [6] in the three scenes, the accuracy of the pedestrian class is 11.3% higher, the accuracy of the tree class is 1.5% higher, and the accuracy of the building class is 6.1% higher.

As shown at the lower left corner of Fig. 9(b), it is noted that the cars are desirably separated from scene I. However, we observe from Table II that the precision and recall of cars are not very high; the reason is that there are six baby carriages in scene I which are also grouped into the car class. Although two baby carriages were put into the training data, they are only a small part of the training data, and it would not be very helpful for recognizing the baby carriages in the first step of our method. The classification accuracy and recall are up to 90% after the second and third optimization steps of our method are performed. If the points on the baby carriages are removed, most of the real car points are recognized in scene I by using our method, and the classification accuracy is up to 97%.

The point-based features may cause misclassification in cluttered scenes. For example, Figs. 9(b) and 11(b) show that a small number of green points appeared on buildings. That means that some points on buildings are misclassified as tree points. After the global optimization, the recognition quality of the points on the car and building edges is obviously improved. However, in Fig. 10(b), some parts of the staircases in front of the buildings are wrongly classified as trees because of data missing. The global optimization based on the multilabel graph-cut still fails to recognize the points on the staircases correctly. After the local optimization based on the object-oriented decision tree integrated with the prior knowledge is applied to refine the classification, these points are recognized correctly. The classification results are shown in Fig. 10(d).

Small-sized objects, such as cars and pedestrians, not only contain fewer points which are often incomplete due to occlusions, but also, they keep on moving during the scan, and thus, the profiles of the point clouds are stretched. Obviously, it is very challenging to label such relatively small-sized objects. Compared to other methods, our method has a better ability to classify small-sized object classes.

In summary, as shown in Table II and Figs. 9–11, the classification performances of the three scenes achieved by using our method are much higher than the method in [6].

C. Sensitivity of the Parameters

In this section, we analyze the sensitivity of the parameters in our method to the classification results. The parameters include the spin images with three scales and the number of neighborhoods k in the KNN.

Fig. 12 displays the sensitivity of the radii that control the sizes of the three scales in the spin-image extraction to the classification results. The vertical axis denotes the F_1 measure, and the horizontal axis denotes the radii sizes of the three scales. Due to the small sizes and different distances from the scanners throughout the scanning, which make their point density inhomogeneous, the spin images have a difference in characterizing the shape information of the pedestrians and cars in a large scene. From Fig. 12, we observe that, when the size of the radius is the range [0.06, 0.10], the classification results of the car and pedestrian classes fluctuate greatly.

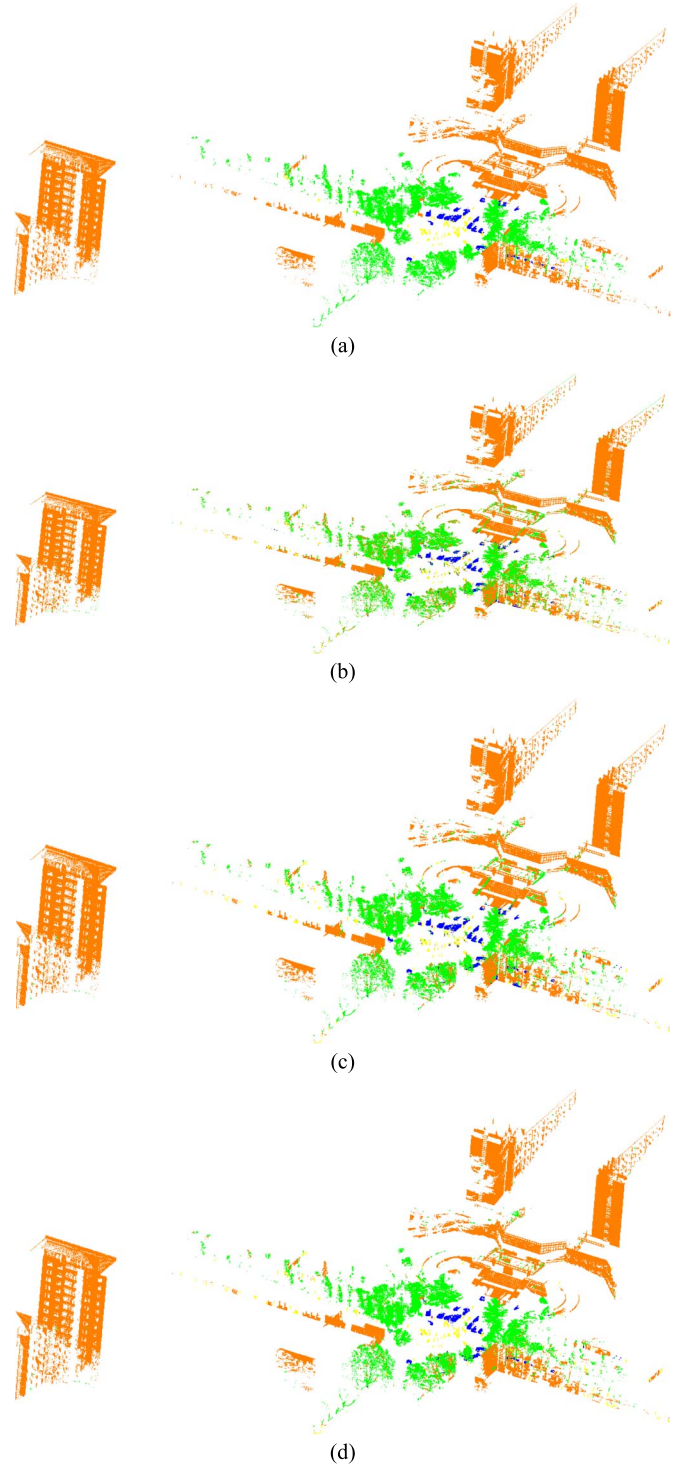


Fig. 10. Classification results obtained by the three-step approach in scene II. Points on cars, trees, pedestrians, and buildings are colored in blue, green, yellow, and orange, respectively. (a) Ground truth. (b) Classification result obtained by using the linear SVM. (c) Classification result optimized by the multilabel graph-cut. (d) Classification result in (c) optimized by the object-oriented decision tree.

Under the three different radii ($r_0 = 0.2$ m, $r_1 = 0.8$ m, and $r_2 = 1.4$ m) of the spheres that determine the scale size of the spin images, we test the sensitivity of different k 's in KNN to the classification result. As shown in Fig. 13, when $k = 14$, F_1 measure keeps a high level.

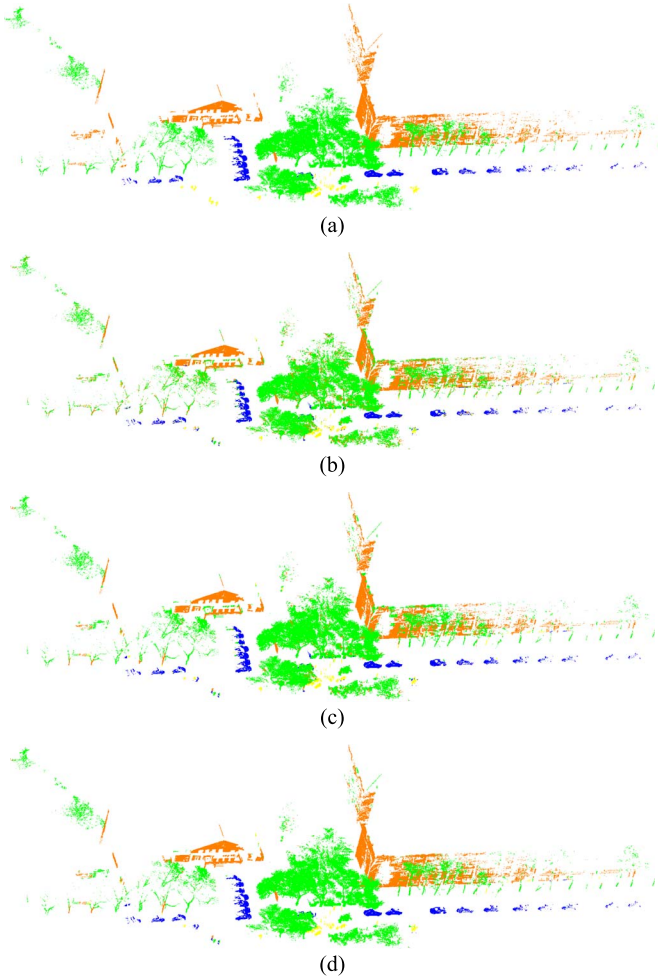


Fig. 11. Classification results obtained by the three-step approach in scene III. Points on cars, trees, pedestrians, and buildings are colored in blue, green, yellow, and orange, respectively. (a) Ground truth. (b) Classification result obtained by using the linear SVM. (c) Classification result optimized by the multilabel graph-cut. (d) Classification result in (c) optimized by the object-oriented decision tree.

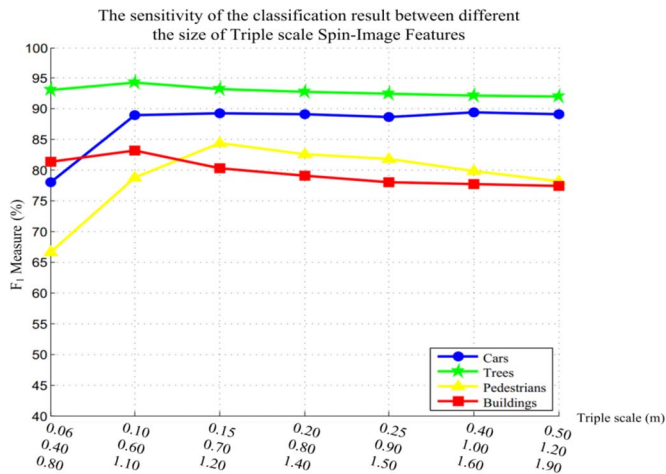


Fig. 12. Sensitivity of the radii that control the sizes of the three scales in spin-image extraction to the classification results in scene III.

Our method is implemented on the machine with a processor of Intel Core i7-4790 K 3.60 GHz and 8-GB RAM. During the whole classification procedure, the feature extraction occupies

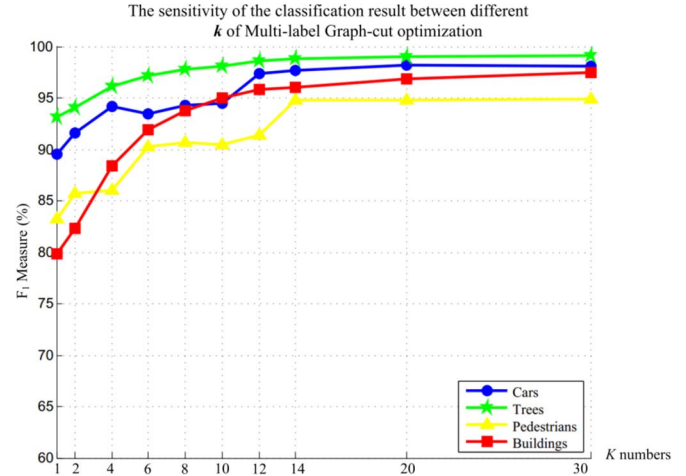


Fig. 13. Sensitivity of the different k 's in KNN for the multilabel graph-cut optimization to the classification results (the radii of the three scales in the spin-image extraction are the following: $r_0 = 0.2$ m, $r_1 = 0.8$ m, and $r_2 = 1.4$ m).

TABLE III
TIME COSTS ON FEATURE EXTRACTION, LEARNING ALGORITHM,
AND OPTIMIZATION IN THE THREE SCENES

| Scenes | MPFs extraction (min) | Linear SVM learning time (s) | Linear SVM testing time (s) | Multi-label time (s) | Object-orient ed decision tree time (min) |
|--------|-----------------------------|---------------------------------------|---|----------------------------|--|
| I | 87.5 | 28.5 | 85.4 | 62.3 | 39 |
| II | 61.3 | 22.2 | 63 | 41.5 | 28 |
| III | 48.6 | 18.2 | 54.1 | 42 | 25 |

much time and nearly occupies 60% of the whole procedure. The time costs on feature extraction, learning algorithm, and optimization in the three scenes are shown in Table III. The feature extraction is able to be implemented in parallel, which can reduce the entire time cost.

V. CONCLUSION

In this paper, we have proposed a three-step method for classifying point clouds captured from large-scale urban environments. As the first step, we have investigated a set of descriptors: the spin-image descriptor computed on three scales, the mean RGB colors of the point as seen in the camera images, the LAB values of that mean RGB, the normal at the 3-D point, the point's height above the estimated ground plane, and its inverse height, to characterize the shape, color, and geometry information of each 3-D point in the point cloud. Then, we have classified each point in the point cloud into four semantic classes (buildings, trees, pedestrians, and cars) by using a linear SVM classifier on the aforementioned descriptors. Next, to reduce misclassification of each class, the global and local optimization algorithms are employed to further enhance the classification performance. The obtained features in our method can capture the characteristics of the object in each class. Therefore, the initial classification of the point clouds generated by utilizing the linear SVM classifier achieves a

good performance. Through the global and local optimizations integrated with the contextual information, the classification accuracy is further enhanced. Experiments have been performed on three different urban environments, and the experimental results show that the proposed method is well capable of dealing with the point clouds of large-scale complex scenes. In particular, compared with the method [6], the classification for small objects achieved by the presented method has a better performance. The data are available at <http://geog.bnu.edu.cn/rydw/zrjs/js/6003.html>.

To significantly reduce the intensive labeling cost, we are planning to extend our method to automatically generate training data from the input data with weak priors like [25] and learning discriminative features [35]. Meanwhile, inspired by the idea that image segmentation and dense 3-D reconstruction can contribute valuable information to each other's task [36], we will solve a joint urban point cloud segmentation and reconstruction problem.

REFERENCES

- [1] J. Xiao, M. Gerke, and G. Vosselman, "Building extraction from oblique airborne imagery based on robust facade detection," *ISPRS J. Photogramm. Remote Sens.*, vol. 68, pp. 56–68, Mar. 2012.
- [2] X. Zhu and M. Shahzad, "Facade reconstruction using multi-view spaceborne TomoSAR point clouds," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3541–3552, Jun. 2014.
- [3] F. Lafarge and C. Mallet, "Creating large-scale city models from 3D-point clouds: A robust approach with hybrid representation," *Int. J. Comput. Vis.*, vol. 99, no. 1, pp. 69–85, Feb. 2012.
- [4] J.-F. Côté, J.-L. Widlowski, R. A. Fournier, and M. M. Verstraete, "The structural and radiative consistency of three-dimensional tree reconstructions from terrestrial LiDAR," *Remote Sens. Environ.*, vol. 113, no. 5, pp. 1067–1081, May 2009.
- [5] Y. Livny *et al.*, "Automatic reconstruction of tree skeleton from point clouds," *ACM Trans. Graph.*, vol. 29, no. 6, pp. 1–8, Dec. 2010.
- [6] Z. Wang *et al.*, "A multiscale and hierarchical feature extraction method for terrestrial laser scanning point cloud classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2409–2425, May 2015.
- [7] D. Angelov *et al.*, "Discriminative learning of Markov random fields for segmentation of 3D scan data," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2005, pp. 169–176.
- [8] A. Golovinskiy, V. G. Kim, and T. Funkhouser, "Shape-based recognition of 3D point clouds in urban environments," in *Proc. 12th Int. Conf. Comput. Vis.*, Sep./Oct. 2009, pp. 2154–2161.
- [9] B. Douillard, D. Fox, and F. Ramos, "Laser and vision based outdoor object mapping," in *Proc. Robot. Sci. Syst.*, 2008, pp. 9–16.
- [10] I. Posner, M. Cummins, and P. Newman, "Fast probabilistic labeling of city maps," in *Proc. Robot. Sci. Syst.*, 2008, pp. 17–24.
- [11] D. Munoz, N. Vandapel, and M. Hebert, "Onboard contextual classification of 3D point clouds with learned high-order Markov random fields," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2009, pp. 2009–2016.
- [12] A. Patterson, IV, P. Mordohai, and K. Daniilidis, "Object detection from large-scale 3D datasets using bottom-up and top-down descriptors," in *Proc. Eur. Conf. Comput. Vis.*, Marseille, France, 2008, pp. 553–566.
- [13] B. K. P. Horn, "Extended Gaussian images," *Proc. IEEE*, vol. 72, no. 12, pp. 1671–1686, Dec. 1984.
- [14] S. H. Sun and C. Salvaggio, "Aerial 3D building detection and modeling from airborne LiDAR point clouds," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 3, pp. 1440–1449, Jun. 2013.
- [15] C. Mallet, F. Bretar, M. Roux, U. Soergel, and C. Heipke, "Relevance assessment of full-waveform LiDAR data for urban area classification," *ISPRS J. Photogramm. Remote Sens.*, vol. 66, no. 6, pp. 71–84, Dec. 2011.
- [16] M. Franceschi *et al.*, "Discrimination between marls and limestones using intensity data from terrestrial laser scanner," *ISPRS J. Photogramm. Remote Sens.*, vol. 64, no. 6, pp. 522–528, Nov. 2009.
- [17] A. Jaakkola, M. Kaasalainen, A. Krooks, and A. Kukko, "Analysis of incidence angle and distance effects on terrestrial laser scanner intensity: Search for correction methods," *Remote Sens.*, vol. 3, no. 10, pp. 2207–2221, Oct. 2011.
- [18] N. Brodu and D. Lague, "3D terrestrial LiDAR data classification of complex natural scenes using a multi-scale dimensionality criterion: Applications in geomorphology," *ISPRS J. Photogramm. Remote Sens.*, vol. 68, pp. 121–134, Mar. 2012.
- [19] A. S. Antonarakis, K. S. Richards, and J. Brasington, "Object-based land cover classification using airborne LiDAR," *Remote Sens. Environ.*, vol. 112, no. 6, pp. 2988–2998, Jun. 2008.
- [20] S. Pu, M. Rutzinger, G. Vosselman, and S. O. Elberink, "Recognizing basic structures from mobile laser scanning data for road inventory studies," *ISPRS J. Photogramm. Remote Sens.*, vol. 66, no. 6, pp. 28–39, Dec. 2011.
- [21] J. Rau, J. Jhan, and Y. Hsu, "Analysis of oblique aerial images for land cover and point cloud classification in an urban environment," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1304–1319, Mar. 2015.
- [22] E. Lim and D. Suter, "3D terrestrial LiDAR classifications with super-voxels and multi-scale conditional random fields," *Comput. Aided Design*, vol. 41, no. 10, pp. 701–710, Oct. 2009.
- [23] E. Lim and D. Suter, "Multi-scale conditional random fields for over-segmented irregular 3D point clouds classification," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Anchorage, AK, USA, 2008, pp. 1–7.
- [24] G. Mills and G. Fotopoulos, "Rock surface classification in a mine drift using multiscale geometric features," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 6, pp. 1322–1326, Jun. 2015.
- [25] H. Zhang, J. Wang, T. Fang, and L. Quan, "Joint segmentation of images and scanned point cloud in large-scale street scenes with low-annotation cost," *IEEE Trans. Image Process.*, vol. 23, no. 11, pp. 4763–4772, Nov. 2014.
- [26] D. Chen, L. Zhang, Z. Wang, and H. Deng, "A mathematical morphology-based multi-level filter of LiDAR data for generating DTMs," *Sci. China Inf. Sci.*, vol. 56, no. 10, pp. 1–14, Oct. 2013.
- [27] Z. Zhang *et al.*, "A multi-level point cluster-based discriminative feature for ALS point cloud classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3309–3321, Jun. 2016.
- [28] A. Martinovic, J. Knopp, H. Riemenschneider, and L. Van Gool, "3D all the way: Semantic segmentation of urban scenes from start to end in 3D," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 4456–4465.
- [29] H. Riemenschneider, A. Bodis-Szomoru, J. Weissenberg, and L. Van Gool, "Learning where to classify in multi-view semantic segmentation," in *Proc. Eur. Conf. Comput. Vis.*, Zürich, Switzerland, Sep. 2014, pp. 516–532.
- [30] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 9, pp. 124–137, Sep. 2004.
- [31] B. Fulkerson, A. Vedaldi, and S. Soatto, "Class segmentation and object localization with superpixel neighborhoods," in *Proc. 12th IEEE Int. Conf. Comput. Vis.*, 2009, pp. 670–677.
- [32] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *ACM Commun.*, vol. 24, no. 6, pp. 381–395, Jun. 1981.
- [33] Y. Ioannou, B. Taati, R. Harrap, and M. Greenspan, "Difference of normals as a multi-scale operator in unorganized point clouds," in *Proc. 2nd Int. Conf. 3D Imaging, Model., Process., Vis. Transmiss.*, 2012, pp. 501–508.
- [34] R. B. Rusu, "Semantic 3D Object Maps for Everyday Manipulation in Human Living Environments," M.S. thesis, Dept. Comput. Sci., Technische Universitaet Muenchen, Munich, Germany, Oct. 2009.
- [35] F. Zhang, B. Du, and L. Zhang, "Saliency-guided unsupervised feature learning for scene classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 2175–2184, Apr. 2015.
- [36] C. Hane, C. Zach, A. Cohen, R. Angst, and M. Pollefeys, "Joint 3D scene reconstruction and class segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Portland, OR, USA, Jun. 25–27, 2013, pp. 97–104.



Zhuqiang Li received the B.S. degree from the China University of Geosciences, Beijing, China. He is currently working toward the Master's degree in the School of Geography, Beijing Normal University, Beijing.

His research interests are LiDAR point cloud processing and 3-D urban modeling.



Liqiang Zhang received the Ph.D. degree in geoinformatics from the Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing, China, in 2004.

He is currently a Professor with the School of Geography, Beijing Normal University, Beijing. His research interests include remote sensing image processing, 3-D urban reconstruction, and spatial object recognition.



Liang Zhang is currently working toward the Ph.D. degree in the School of Geography, Beijing Normal University, Beijing, China.

His research interests are remote sensing imagery processing and 3-D urban modeling.



Xiaohua Tong received the Ph.D. degree from Tongji University, Shanghai, China, in 1999.

He was a Postdoctoral Researcher with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan, China, between 2001 and 2003. He was a Research Fellow with The Hong Kong Polytechnic University, Hung Hom, Hong Kong, in 2006, and a Visiting Scholar with the University of California, Santa Barbara, CA, USA, between 2008 and 2009. He is a Chang-Jiang Scholar Chair

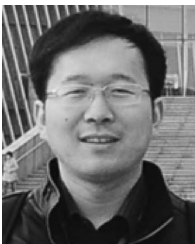
Professor appointed by the Ministry of Education, China. He also received the National Natural Science Funds for Distinguished Young Scholars in 2013. He is the author of more than 40 publications in international journals. His current research interests include remote sensing, GIS, trust in spatial data, and image processing for high-resolution and hyperspectral images.

Prof. Tong serves as the Vice-Chair of the Commission on Spatial Data Quality of the International Cartographical Association and the Cochair of the ISPRS working group (WG II/4) on Spatial Statistics and Uncertainty Modeling. Moreover, he received the State Natural Science Award (second place) from the State Council of the Peoples' Republic of China in 2007.



Zhenxin Zhang is currently working toward the Ph.D. degree in geoinformatics in the School of Geography, Beijing Normal University, Beijing, China.

His research interests include light detection and ranging data processing, quality analysis of geographic information systems, and algorithm development.

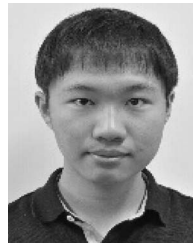


Bo Du (M'10–SM'15) received the B.S. degree and the Ph.D. degree in photogrammetry and remote sensing from the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China, in 2005 and 2010, respectively.

He is currently a Professor with the School of Computer, Wuhan University. He has more than 40 research papers published in the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (TGRS), IEEE TRANSACTIONS ON IMAGE

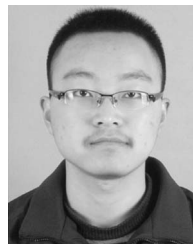
PROCESSING (TIP), IEEE JOURNAL OF SELECTED TOPICS IN EARTH OBSERVATIONS AND APPLIED REMOTE SENSING (JSTARS), IEEE GEOSCIENCE AND REMOTE SENSING LETTERS (GRSL), etc. His major research interests include pattern recognition, hyperspectral image processing, and signal processing.

Dr. Du was the Session Chair of the 4th IEEE Geoscience and Remote Sensing Society (GRSS) Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS). He also serves as a Reviewer of 20 Science Citation Index (SCI) magazines, including IEEE TGRS, TIP, JSTARS, and GRSL. He received the Best Reviewer Award from the IEEE GRSS for his service to the JSTARS in 2011 and ACM Rising Star Award for his academic progress in 2015.



Hao Liu received the Bachelor's degree from Beihang University (BUAA), Beijing, China, in 2012 and the Master's degree from Kyushu University, Fukuoka, Japan, in 2015. He is currently working toward the Ph.D. degree in the School of Geography, Beijing Normal University, Beijing.

His present research interests include machine learning and computer vision.



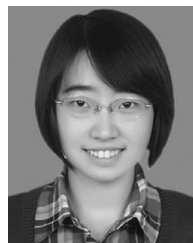
Jie Mei is currently working toward the Bachelor's degree in the School of Geography, Beijing Normal University, Beijing, China.

His research interest is TLS point cloud recognition.



Yuebin Wang is currently working toward the Ph.D. degree in the School of Geography, Beijing Normal University, Beijing, China.

His research interests are remote sensing imagery processing and 3-D urban modeling.



Xiaoyue Xing is currently working toward the Bachelor's degree in the School of Geography, Beijing Normal University, Beijing, China.

Her research interest is TLS point cloud processing.



P. Takis Mathiopoulos (SM'94) received the Ph.D. degree in digital communications from the University of Ottawa, Ottawa, ON, Canada, in 1989.

From 1982 to 1986, he was with Raytheon Canada Ltd., working in the areas of air navigational and satellite communications. In 1988, he joined the Department of Electrical and Computer Engineering (ECE), University of British Columbia (UBC), Vancouver, BC, Canada, where he was a faculty member until 2003, holding the rank of Professor from 2000 to 2003. From 2000 to 2014, he was with

the Institute for Space Applications and Remote Sensing (ISARS), National Observatory of Athens (NOA), first as the Director and then as the Director of Research, where he established the Wireless Communications Research Group. As ISARS' Director (2000–2004), he has led the institute to significant expansion R&D growth and international scientific recognition. For these achievements, ISARS has been selected as a National Centre of Excellence for 2005 to 2008. Since 2014, he has been an Adjunct Researcher with the Institute of Astronomy, Astrophysics, Space Applications and Remote Sensing (IAASARS), NOA. Since 2003, he has also been teaching part-time at the Department of Informatics and Telecommunications, University of Athens, Athens, Greece, where he has been a Professor of digital communications since 2014. From 2008 to 2013, he was appointed as a Guest Professor with Southwest Jiaotong University, Leshan, China. He has also been appointed by the Government of China as a Senior Foreign Expert with the School of Information Engineering, Yangzhou University, Yangzhou, China, in 2014–2015 and as a Visiting Professor with the Department of Information and Computer Science, Keio University, Tokyo, Japan, in 2015–2016 under the Top Global University Project of the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Government of Japan. For the last 25 years, he has been conducting research mainly on the physical layer of digital communication systems for terrestrial and satellite applications, including digital communications over fading and interference environments. He is the coauthor of a paper in GLOBECOM'89, establishing for the first time in the open technical literature the link between MLSE and multiple (or multisymbol) differential detection for the AWGN and fading channels. He is also interested in channel characterization and measurements, modulation and coding techniques, synchronization, SIMO/MIMO, UWB, OFDM, software/cognitive radios, and green communications. In addition, since 2010, he has been actively involved with research activities in the fields of remote sensing, LiDAR systems, and photogrammetry. In these areas, he is the coauthor of more than 100 journal papers, mainly published in various IEEE and IET journals, 4 book chapters, and more than 120 conference papers. He has been the PI of more than 40 research grants and has supervised the dissertation of 11 Ph.D. students and the thesis of 23 Master's students.

Dr. Mathiopoulos has been or currently serves on the editorial board of several archival journals, including the *IET Communications* and the *IEEE TRANSACTIONS ON COMMUNICATIONS* (1993–2005). He has regularly acted as a consultant for various governmental and private organizations. Since 1993, he has served on a regular basis as a Scientific Advisor and a Technical Expert for the European Commission (EC). In addition, from 2001 to 2014, he has served as a Greek representative to high-level committees in the EC and the European Space Agency. He has been a member of the Technical Program Committee (TPC) of more than 70 international IEEE conferences, as well as the TPC Vice Chair of the 2006-S IEEE VTC and 2008-F IEEE VTC as well as Cochair of FITCE2011. He has delivered numerous invited presentations, including plenary and keynote lectures, and has taught many short courses all over the world. As a faculty member at the ECE of UBC, he was elected as an ASI Fellow and a Killam Research Fellow. He was a corecipient of two best paper awards for papers published in the 2nd International Symposium on Communication, Control, and Signal Processing (2008) and 3rd International Conference on Advances in Satellite and Space Communications (2011).