POINT CLOUD CLASSIFICATION BY FUSING SUPERVOXEL SEGMENTATION WITH MULTI-SCALE FEATURES

Wei Ao[1], Lei Wang[2], Jie Shan[3, *]

[1]School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China – wei_ao@whu.edu.cn
[2]State Key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, China - wlei@whu.edu.cn
[3]Lyles School of Civil Engineering, Purdue University, West Lafayette, IN 47907, USA - jshan@purdue.edu

KEY WORDS: Point Cloud Classification, Supervoxel Segmentation, Multi-Scale Feature Extraction, Feature Modification, Global Optimization

ABSTRACT:

Point cloud classification is quite a challenging task due to the existence of noises, occlusion and various object types and sizes. Currently, the commonly used statistics-based features cannot accurately characterize the geometric information of a point cloud. This limitation often leads to feature confusion and classification mistakes (e.g., points of building corners and vegetation always share similar statistical features in a local neighbourhood, such as curvature, sphericity, etc). This study aims at solving this problem by leveraging the advantage of both the supervoxel segmentation and multi-scale features. For each point, its multi-scale features within different radii are extracted. Simultaneously, the point cloud is partitioned into simple supervoxel segments. After that, the class probability of each point is predicted by the proposed SegMSF approach that combines multi-scale features with the supervoxel segmentation results. At the end, the effect of data noises is supressed by using a global optimization that encourages spatial consistency of class labels. The proposed method is tested on both airborne laser scanning (ALS) and mobile laser scanning (MLS) point clouds. The experimental results demonstrate that the proposed method performs well in terms of classifying objects of different scales and is robust to noise.

1. INTRODUCTION

Point cloud classification, which aims to assign each point a proper class label, is a basic problem in 3D scene understanding for intelligent robots, digital cities and unmanned vehicles (Nguyen and Le, 2013). Extracting the discriminative features of ground objects from noisy, sparse, unstructured data is a key step in producing accurate classification results. Currently, the mainstream features for point cloud classification are statistics-based, which can be derived from the local covariance matrix at a point's neighbourhood. However, because of the uneven density of point clouds obtained by different sampling intervals and ranges of laser scanners, these features may not sufficiently characterize the geometric information of point clouds. For instance, the statistical features extracted at the intersection of two planes (e.g., building corner) are likely to be similar to those of a bunch of scattered points (e.g., vegetation points) in a local neighbourhood. Therefore, it leads to confusions in places like corners of buildings, and cannot accurately characterize complex objects in the point cloud.

These problems have been extensively studied for many years. Several researchers tried to improve the classification accuracy by extracting features of point cloud at different scales (Chen and Maggioni, 2011; Wang et al., 2015; Zhang et al., 2016). In the training stage, multi-scale processing of point cloud is carried out, which is similar to the scale-invariant feature transform (SIFT, Lowe, 2004) pyramid. Technically, point clouds are resampled into different scales. Then the neighbourhoods at different scales are applied for neighbouring points searching and feature extraction (Brodu and Lague, 2012). The advantage of this method is that it can be adapted to various sizes of objects. But it

is sensitive to noise and is difficult to accurately characterize the features of points at the boundary of multiple objects or planes.

Another way to address these issues is supervoxel segmentation. Just as superpixel in 2D image processing, it can be leveraged to substantially reduce the number of points to work with. Most importantly, this provides a more natural and compact representation of 3D point clouds, which enables the operations to be performed on regions instead of the scattered points (Lin et al., 2018). Therefore, supervoxel segmentation can well adapt to the situation with curvature changes (e.g., building corner). However, many existing methods of supervoxel segmentation employ fixed resolution for each voxel. Meanwhile, the accuracy of segmentation result is excessively dependent on the initialization of the seed points. Besides, as the principle of supervoxel segmentation is based on the similarity of low-level geometrical features, the segmentation results are always piecemeal and have no meaningful object information.

In this paper, a novel point cloud classification approach, SegMSF, is proposed that combines the supervoxel segmentation and multi-scale features. The local statistical features are fused with the supervoxel segments to leverage both of their advantages to achieve a better performance. At the end, a global energy function which encourages the spatial consistency among neighbouring labels is optimized to supress the impact of the inevitable noises. In order to demonstrate the superiority of this method, we use both airborne laser scanning (ALS) and mobile laser scanning (MLS) point clouds which contain objects of different sizes and scanning densities to validate its good stability and self-adaptability. Finally, the proposed method is proved to be able to distinguish objects of various sizes and handle the effect of occlusion to the ground objects.

---

* Corresponding author

2. RELATED WORK

In this section, we will discuss the related prior works in terms of two aspects: point cloud feature extraction and point cloud classifier.

Point cloud feature extraction. Good feature is important for high-precision point cloud classification. Commonly used point cloud features can be generally divided into point-based (Blomley et al., 2014; Weinmann et al., 2015; Zhang et al., 2016), segment-based (Aijazi et al., 2013; Xiang et al., 2018; Zhou et al., 2012) and learning-based (Graham et al., 2018; Tchapmi et al., 2017) levels.

Point-based feature describes the local properties of a point cloud according to its neighbourhood, such as the surface normals, curvatures, and eigenvalues of the covariance matrix (Guinard and Landrieu , 2017; Wang et al., 2015). However, such local descriptors are sensitive to the scale of neighbourhood as objects often present different properties at different spatial scales. To avoid the influence of scale, the multiscale or multi-resolution spatial feature descriptors (Chen and Maggioni, 2011; Wang et al., 2015) are proposed to describe the objects across different scales. Segment-based feature first partitions the point cloud into small segments and then calculates the properties of each segment, such as length ratio, surface ratio, volume ratio, and so on (Landrieu and Simonovsky, 2018; Xiang et al., 2018). However, as the segments are partitioned according to the local geometrical similarity, the segments are always piecemeal and has no meaningful object information.

Learning-based method aims at extracting features of a point cloud by using machine learning method with enormous training samples. Current learning-based method can be divided into four categories: multiview-based, voxel-based, set-based and graph-based. Multiview-based and voxel-based methods (Le et al., 2017; Qi et al., 2016) represent the 3D shape into a set of images or regular volumetric occupancy grids, so that the feature learning method on regular arrays can be directly used. However, it is difficult to determine the distribution of the views, and the voxel-based method inevitably leads to memory and computation consumptions as they increase cubically with respect to the voxel's resolution. Set-based method (Qi et al., 2017a, 2017b; Zaheer et al., 2017) aims to find a symmetric function which can aggregate the information of the set with any input orders. Graph-based method (Bronstein et al, 2017; Simonovsky and Komodakis, 2017; Yi et al., 2017) first represents the point cloud into a graph according to their neighbourhoods, and then the graph neural networks are applied for feature learning. However, learning-based method needs enormous training samples and yields huge computations.

Point cloud classifier. After the point cloud features are extracted, proper classifier is required to assign each point with certain class label. Generally, classifiers for point cloud can be divided into 1) individual point classifier and 2) contextual classifier (Martin Weinmann et al., 2015). The commonly used individual point classifiers include support vector machine (Ghamisi and Höfle, 2017; Zhang et al., 2013), cascaded AdaBoost (Lodha et al., 2007), random forest (Li and Cheng, 2018; Ni et al., 2017), extreme learning machine (Strom et al., 2010), cascade classifiers (Carlberg et al., 2009), and so on. Although these methods can perform well in point classification, the topological relationship among different objects in the urban environment is usually ignored. Moreover, these methods are sensitive to noise.

In order to account the relationship among neighbouring points, contextual classifier regards the point cloud as a graph. Each vertex corresponds to a point in the data and the edges connect neighbouring points (Kang and Yang, 2018; Yan et al, 2014). After that, respective approaches such as Markov random fields (MRF, Najafi et al., 2014; Shapovalov et al., 2010), conditional random field (CRF, Kalogerakis et al., 2010; Niemeyer et al., 2014, 2012) can be applied. MRF/CRF converts the classification as a multi-labeling optimization problem. By minimizing their corresponding energy functions, the feature differences are minimized in the same class and are maximized among different classes (Lozes et al., 2013). Contextual classifier is insensitive to noise, but its segmented results depend on the initial values and are usually piecemeal.
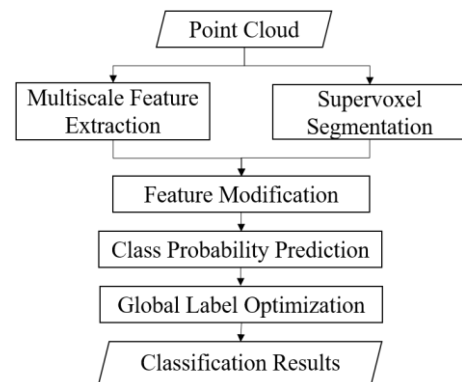

Figure 1. Flowchart of the proposed SegMSF method.

3. METHODOLOGY

The flowchart of proposed method is presented in Figure 1. Multi-scale feature extraction and supervoxel segmentation are conducted simultaneously on the original point cloud. According to the principle of segmentation, the features of points belonging to the same segment tend to be consistent. Therefore, for each point, its multi-scale features are then modified by concatenating with the features of the segment it belongs to. After that, the random forest classifier is leveraged to produce preliminary classification probabilities. At last, considering the effect of noise, a global energy function which encourages the spatial consistency between neighbouring labels is used for further label optimization.

3.1 Multi-Scale Feature Extraction

Feature design is of great importance to the process of point cloud classification. Considering the various size of objects in urban scenes, statistics-based geometrical features of point cloud at multi-scales are extracted in this work.

Geometrical Features	Formula
Linearity	$(\lambda_1 - \lambda_2)/\lambda_1$
Planarity	$(\lambda_2 - \lambda_3)/\lambda_1$
Sphericity	λ_3/λ_1
Omnivariance	$(\lambda_1\lambda_2\lambda_3)^{1/3}$
Anisotropy	$(\lambda_1 - \lambda_3)/\lambda_1$
Eigenentropy	$-\Sigma_{i=1}^{3}\lambda_i \ln \lambda_i$
Curvature	$\lambda_3/(\lambda_1 + \lambda_2 + \lambda_3)$
Sum of eigenvalues	$\lambda_1 + \lambda_2 + \lambda_3$

Table 1. Local geometrical features.

Specifically, for the given point cloud $P = \{p_1, p_2, ..., p_N\}$, we first down-sample the point cloud into multiple resolutions

$\{P_1, ..., P_L\}$ and then a series of search radii $\{r_1, ..., r_L\}$ are used for neighbouring points searching. For each point $p_i(x, y, z) \in P$, denote $\mathcal{N}_l(p_i)$ as the set of neighbouring points of p_i within radii r_l. The geometrical features at each scale are derived from the normalized eigenvalues $\lambda_1, \lambda_2, \lambda_3$ of the covariance matrix of the neighbour set $\mathcal{N}_l(p)$. Assume that these eigenvalues are sorted as $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$, the local geometrical feature $f_g(p_i)$ applied in this paper contains the linearity, planarity, sphericity, omnivariance, anisotropy, eigenentropy, curvature and sum of eigenvalues, whose definitions are provided in Table 1.

3.2 Supervoxel Segmentation

The local features described in Section 3.1 characterize the geometrical attributes of point cloud within a sphere neighbourhood according to their Euclidean distances. However, they neglect the semantic differences of points and therefore cannot precisely represent the feature of points. For instance, the points at a building corner should actually belong to one side of the building, but its local covariance features are similar to the vegetation points as its spatial neighbourhood contains the points of both sides of the building (Figure 2). Therefore, for complex classification involving multiple types of objects, it is required to over-segment the original point cloud into small regions or objects simultaneously. This will not only enhance the capacity of distinguishing the points that are likely to be confused, but also greatly reduce the amount of computation.
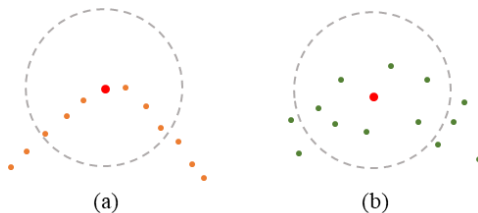

Figure 2. Illustration of the neighbourhood searching for local feature extraction. (a) Points at a building corner. (b) Points of vegetation. The local geometrical features (e.g., linearity, planarity, curvature, etc.) of these two situations are similar. Thus, they are difficult to distinguish with only the local geometrical features.

In order to well preserve the boundaries of ground objects during segmentation, voxel cloud connectivity segmentation (VCCS, Papon et al., 2013) is adopted in our study. This approach is achieved by a seeding method based on 3D space and local iterative clustering with flow-constraint using colour and geometrical features. Firstly, the point clouds are voxelized by octree, and then the initial supervoxels are extracted by dividing the 3D space evenly. The local k-means clustering method is then used to grow these initial supervoxels. Especially, VCCS is reported to be highly efficient in avoiding the problem of under-segmentation and can conform to real geometric relationships.

3.3 Feature Modification

Statistics-based geometrical features represent the local feature of point cloud within a sphere neighbourhood, while they cannot accurately characterize the points at the boundary of objects. The supervoxel segmentation method aggregate the points lying on one local plane into a segment, but the segmentation results are always piecemeal and have no meaningful object information. In this section we describe how to combine the supervoxel segmentation information and the local geometrical features, while maintaining the advantages of both.

Denote the segments as $\{S_1, S_2, ..., S_K\}$, where S_k is set points belonging to the k-th segment and $\bigcup_{k=1}^K S_k = P$. We first calculate the feature attribute of each segment as f_{S_k} according to the covariance matrix of set S_k as described in Section 3.1. The modified feature of the i-th point p_i is defined as follows:

$$f_i = \left[f_{S_k}, \frac{1}{|S_k|} \sum_{p \in S_k} f_g(p), f_g(p_i) \right], p_i \in S_k \qquad (1)$$

where the second term represents the mean feature of the points belonging to the same segment. The modified feature f_i leverages both the advantages of the supervoxel segments and the local geometrical features.

3.4 Class Probability Prediction with Random Forest

In this section, the random forest (RF) algorithm is used as the classifier for preliminary class probability prediction of each point. RF is an integrated algorithm that creates a set of decision trees from a randomly selected subset of the training set, then aggregates votes from different decision trees to determine the final class of the test point. It has the advantages of no data preprocessing, convenient and fast processing of multi-class problems and stable classification results. Therefore, RF has already been widely used and shown good classification results for LiDAR point clouds (Sun and Lai, 2014; Hackel et al., 2017).

In this study, to analyse the performance on different datasets, we input the modified features of each training set into the RF separately. During the prediction phase, the class probability $\hat{p}_i$ of the corresponding point p_i in the testing set is predicted and is used as the initial value for global label optimization with CRF (Section 3.5).

3.5 Global Label Optimization

To deal with the spatial inconsistency of the predicted label caused by noise, a CRF model with spatial consistency is constructed for global label optimization. For this purpose, we construct a graph $G(V, E)$ with vertex $v \in V$ and edge $e \in E$. Each vertex is associated to a point, and the edges are added between the point and its K-nearest points of the point cloud.
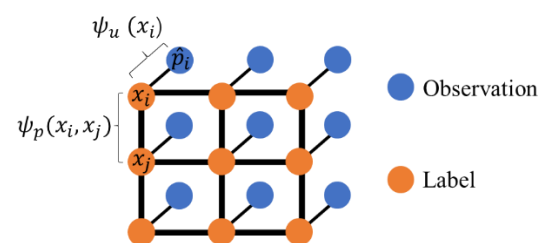

Figure 3. The unary and pairwise term of the CRF model.

Let random variable X_i be the label of vertex i, whose value domain is a set of labels $\mathcal{L}^C = \{l_1, l_2, ..., l_C\}$, where C is the number of class. Random variable X consists of $X_1, X_2, ..., X_N$, where N is the total number of points. Regard vertex V of the graph $G(V, E)$ as the random variable of label, i.e. $V = \{X_1, X_2, ..., X_N\}$. Given a global observation (point cloud) P, the pair (P, X) can be model as a conditional random field characterized by a Gibbs distribution $p(X = x|P) = \frac{1}{Z(P)} \exp(-E(x|P))$, which can also be seen as the posterior probability of the point cloud assigned to label l under the global observation P. The Gibbs energy of labelling $x \in \mathcal{L}^N$ is $E(x|P)$ and $Z(P)$ indicates the normalized index (Krähenbühl and

Koltun, 2012). The rest of the paper drops the conditioning, i.e., denote $E(x|P)$ as $E(x)$ for notational convenience.

In the CRF model about graph $G(V, E)$, the corresponding Gibbs energy is formed as:

$$E(x) = \sum_{i \in V} \psi_u(x_i) + \sum_{(i,j) \in E} \psi_p(x_i, x_j) \qquad (2)$$

where the unary term $\psi_u(x_i)$ measures the cost of assigning label x_i to the i-th point, and the pairwise term $\psi_p(x_i, x_j)$ measures the cost of assigning labels x_i, x_j to neighbouring points p_i, p_j (Figure 3). In this study, the unary term is defined as $\psi_u(x_i) = -ln(\hat{p}_i)$, where $\hat{p}_i$ is the assignment label probability for i-th point predicted by the RF classifier (Section 3.4).

The pairwise term integrates the spatial smooth relationship, which encourages assigning similar labels to the neighbouring points. It can be formulated as

$$\psi_p(x_i, x_j) = \mu(x_i, x_j) \sum_{m=1}^{K} w^{(m)} \kappa^{(m)}(f_i, f_j)$$
$$= \mu(x_i, x_j) \kappa(f_i, f_j) \qquad (3)$$

where the first term learns the penalty of global co-occurrence

between any pair of labels, e.g., the output value of $\mu(x_i, x_j)$ is large if x_i and x_j should not coexist as neighbouring. Each $k^{(m)}$ is a kernel function that measures the distance between the feature pair (f_i, f_j) corresponding to the feature of i-th and j-th point respectively. Commonly, $\kappa^{(m)}$ is usually defined as Gaussian kernel. Inference of $E(x)$ can be achieved with mean field algorithm (Zhang and Chen, 2012).

4. EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Experimental Data

The test data includes both ALS and MLS point clouds which contain objects of different sizes and scanning densities. They are over the same region of Wuhan University, China. The MLS data with two blocks were manually labelled into the following seven classes: vegetation, building, car, pedestrian, lamp, fence and others. The ALS data, which is relatively sparser and more fragmented, was only divided into three classes (i.e., vegetation, building and car). Besides, all the ground points have been removed in advance. Each dataset is divided into two parts as shown in Figure 4. One part is for training while the other for testing.
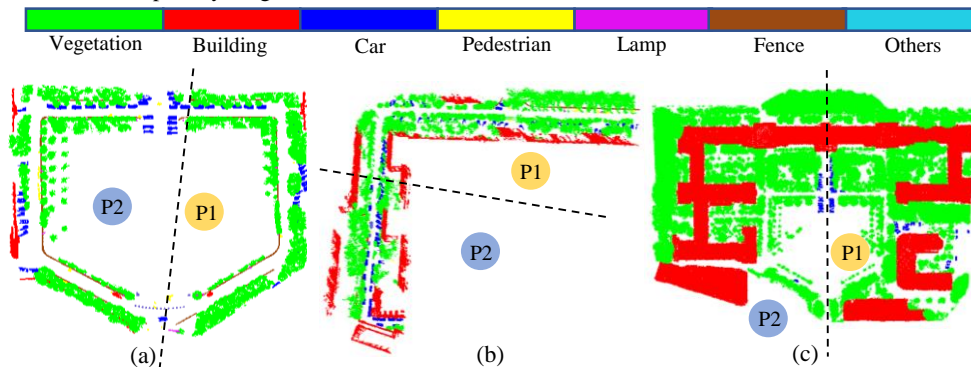

Figure 4. Experimental datasets. (a) and (b) are the MLS point clouds and (c) is the ALS point cloud. Each block is divided into two parts, P1 and P2, by a vertical plane shown with a dash line. P1 is for training and P2 for testing.
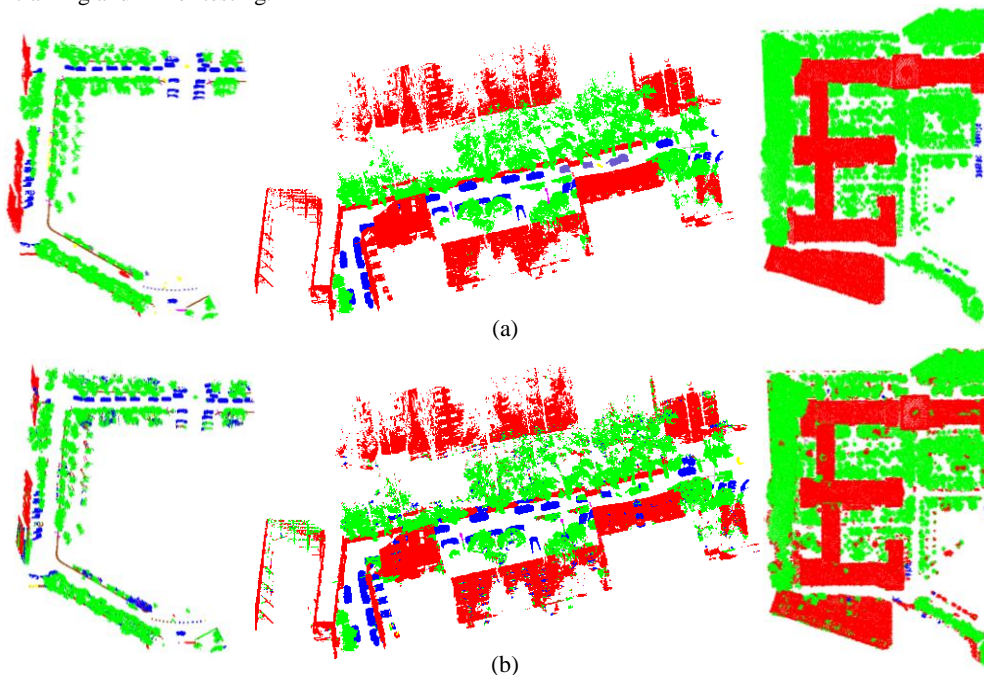

Figure 5. Classification results of the testing data. The first row shows the ground truth of the three testing sets; the second row represents the corresponding classification results with the proposed SegMSF.

Method	Testing data	Vegetation (%)	Building (%)	Car (%)	Pedestrian (%)	Lamp (%)	Fence (%)	Others (%)	Overall accuracy (%)
SegMSF	MLS-I	91.73/90.06	61.67/94.13	94.31/61.96	0.14/0.14	31.92/46.33	76.36/77.18	0/0	83.48
	MLS- II	95.62/88.47	88.97/98.68	95.98/76.36	3.19/1.38	27.31/43.54	-	0/0	90.94
	ALS	91.52/94.93	93.72/89.93	67.22/43.73	-	-	-	-	91.48
MSNet	MLS- I	88.88/96.21	93.11/52.13	56.5/93.57	-	85.02/38.21	99.8/58.77	0/0	83.18
	MLS- II	76.84/93.87	99.64/73.59	65.77/98.52	-	11.93/31.44	-	0/0	82.98
	ALS	92.38/97.81	96.55/89.88	0/0	-	-	-	-	94.06

- represents that there is no such kind of object in the data.

Table 2. The precision/recall and overall accuracy of the test data with the proposed method and MSNet.

4.2 Experimental Results

To adapt to the various sizes of objects, point cloud features with four scales are applied in this experiment. In addition, considering the difference of point densities and object sizes between the MLS and the ALS point clouds, different sizes of neighbourhoods are used for local geometrical feature extraction. Specifically, we set the radii for the MLS point cloud to be 0.25m, 0.5m, 1m, and 2m. The radii for ALS point cloud are set to be 0.5m, 1m, 2m, and 4m due to its relatively larger objects and sparser point density.

For quantitative evaluation, three metrics including overall accuracy (OA), precision and recall over each class are applied. The three metrics are defined as follows:

$$OA = \frac{TP + TN}{TP + FP + TN + FN}$$
$$Percision = \frac{TP}{TP + FP} \qquad (4)$$
$$Recall = \frac{TP}{TP + FN}$$

where TP, FP, TN and FN is the number of true positives, false positives, true negatives and false negatives respectively. OA represents the overall performance of the classification results, recall indicates the completeness of each class, and precision is a measure of exactness or quality.

Table 2 lists the quantitative results on three different testing sets. Overall accuracies greater than 90% and 83% have been achieved on ALS and MLS dataset respectively. The performance on ALS data is relatively higher than MLS data due to its sparser point density and simpler categories. Large-size objects (i.e., vegetation and buildings) always get higher classification accuracies than small-size objects (i.e., pedestrian and lamps). The poor accuracy on small-size objects is mainly caused by incomplete shapes and deficient training samples. The visualization of the classification results on the three datasets are shown in Figure 5.

To demonstrate the advantages of this method, we compare it with MSNet (Wang et al, 2018) which used the same datasets as we did. As it indicated in Table 2, the proposed method can achieve higher overall accuracy on the MLS datasets. Obviously, this method has more robust classification ability for small

objects (i.e., lamps and cars). Additionally, because the building walls are easily obscured by surrounding plants, the recalls of buildings are not high in the classification results of MSNet. However, in our method, this problem does not influence the classification results as the recalls of buildings are higher than 88%. Thus it is concluded that the proposed method is able to adaptively characterize and classify all the objects in spite of various sizes and mutual occlusions among ground objects.

4.3 Analysis and Discussion

To further verify the effectiveness of the proposed method, we remove the supervoxel segmentation and only use the multi-scale geometrical features for classification. The experimental results are provided in Table 3. By comparison, the overall accuracies drop by 7.83%, 5.21% and 5.60% for the ALS, MLS (block I and block II) point clouds respectively. The mean recall and mean precision overall three datasets drop by 11.92% and 7.75%. Results show that the supervoxel segments serve as the clustering information which encourages the points lying on one plane to share the same class label. It complements the drawbacks of local geometrical features that are difficult to characterize the feature of points at the boundary of multiple objects or planes.

Method	Testing data	Mean recall	Mean precision	Kappa	Overall accuracy
SegMSF	MLS-I	61.85	59.36	0.719	83.48
	MLS-II	51.41	51.84	0.852	90.93
	ALS	76.20	84.15	0.848	92.41
w/o segmentation	MLS-I	49.20	46.35	0.627	78.27
	MLS-II	40.62	46.07	0.764	85.33
	ALS	63.89	78.68	0.695	84.58

Table 3. Classification results of test data with the proposed method.

5. CONCLUSION

In this paper, we propose a novel point cloud classification approach, SegMSF, which combines the supervoxel segmentation results and the multi-scale features of each point for precise point cloud classification. It leverages both the advantages of the segmentation information and multi-scale features of point clouds. For one thing, it can accurately characterize the features of points at object boundaries. For another, it performs well on classifying objects of different scales

and is robust to noise in the data. The method is tested on both ALS and MLS point cloud datasets. The experimental results show that the SegMSF method can get higher precision and recall rate in comparison to the traditional methods that without fusion, and achieve remarkable performance on three testing sets.

REFERENCES

Aijazi, A.K., Checchin, P., Trassoudaine, L., 2013. Segmentation based classification of 3D urban point clouds: A super-voxel based approach with evaluation. *Remote Sensing*, 5(4), 1624-1650.

Blomley, R., Weinmann, M., Leitloff, J., Jutzi, B., 2014. Shape distribution features for point cloud analysis – A geometric histogram approach on multiple scales. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2(3), 9-16.

Brodu, N., Lague, D., 2012. 3D terrestrial lidar data classification of complex natural scenes using a multi-scale dimensionality criterion: Applications in geomorphology. *ISPRS Journal of Photogrammetry and Remote Sensing*, 68, 121–134.

Bronstein, M.M., Bruna, J., Lecun, Y., Szlam, A., Vandergheynst, P., 2017. Geometric deep learning: going beyond euclidean data. *IEEE Signal Processing Magazine*, 34(4), 18-42.

Carlberg, M., Gao, P., Chen, G., Zakhor, A., 2009. Classifying urban landscape in aerial lidar using 3D shape analysis. *International Conference on Image Processing*.

Chen, G., Maggioni, M., 2011. Multiscale geometric dictionaries for point-cloud data. *Int. Conf. Sampl. Theory Appl*.

Ghamisi, P., Höfle, B., 2017. LiDAR data classification using extinction profiles and a composite kernel support vector machine. *IEEE Geoscience and Remote Sensing Letters* ,14(5), 659–663.

Graham, B., Engelcke, M., van der Maaten, L., 2018. 3D semantic segmentation with submanifold sparse convolutional networks. *Proc. IEEE Comput. Vis. Pattern Recognit*, 9224-9232.

Guinard, S., Landrieu, L., 2017. Weakly supervised segmentation-aided classification of urban scenes from 3D lidar point clouds. *ISPRS. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLII-1/W1,* 151-157.

Hackel, T., Wegner, J.D., Schindler, K., 2017. Joint classification and contour extraction of large 3D point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 130, 231–245.

Kalogerakis, E., Hertzmann, A., Singh, K., 2010. Learning 3D mesh segmentation and labeling. *ACM Trans. Graph*.

Kang, Z., Yang, J., 2018. A probabilistic graphical model for the classification of mobile LiDAR point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 143, 108-123.

Krähenbühl, P., Koltun, V., 2012. Efficient inference in fully connected CRFs with Gaussian edge potentials. *International Conference on Neural Information Processing Systems Curran Associates Inc*, 109-117.

Landrieu, L., Simonovsky, M., 2018. Large-scale point cloud semantic segmentation with superpoint graphs. *Proc. IEEE Comput. Vis. Pattern Recognit*, 4558-4567.

Le, T., Bui, G., Duan, Y., 2017. A multi-view recurrent neural network for 3D mesh segmentation. *Computers & Graphics*, 66, 103-112.

Li, Q., Cheng, X., 2018. Comparison of different feature sets for TLS point cloud classification. *Sensors*, 18(12), 4026.

Lin, Y. B., Wang, C., Zhai, D. W., Li, W., Li, J., 2018. Toward better boundary preserved supervoxel segmentation for 3D point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 143, 39-47.

Lodha, S.K., Fitzpatrick, D.M., Helmbold, D.P., 2007. Aerial Lidar data classification using AdaBoost. *Sixth Int. Conf. 3-D Digit. Imaging Model*, 435–442.

Lozes, F., Hidane, M., Elmoataz, A., Lezoray, O., 2013. Nonlocal segmentation of point clouds with graphs. *IEEE Global Conference on Signal and Information Processing*, 459–462.

Najafi, M., Taghavi Namin, S., Salzmann, M., Petersson, L., 2014. Non-associative higher-order markov networks for point cloud classification. *Lecture Notes in Computer Science,* 500–515.

Nguyen, A., Le, B., 2013. 3D point cloud segmentation: A survey. *IEEE Conference on Robotics, Automation and Mechatronics, RAM – Proceeding*, 225–230.

Ni, H., Lin, X., Zhang, J., 2017. Classification of ALS point cloud with improved point cloud segmentation and random forests. *Remote Sensing*, 9(3), 288.

Niemeyer, J., Rottensteiner, F., Soergel, U., 2012. Conditional random fields for LiDAR point cloud classification in complex urban areas. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci*. I-3, 263–268.

Niemeyer, J., Rottensteiner, F., Soergel, U., 2014. Contextual classification of lidar data and building object detection in urban areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, 87, 152–165.

Papon, J., Abramov, A., Schoeler, M., Worgotter, F., 2013. Voxel cloud connectivity segmentation - supervoxels for point clouds. *Computer Vision & Pattern Recognition*, 2027-2034.

Qi, C.R., Su, H., Mo, K., Guibas, L.J., 2017a. PointNet: Deep learning on point sets for 3D classification and segmentation. *Proc. IEEE Comput. Vis. Pattern Recognit*, 77–85.

Qi, C.R., Su, H., Niessner, M., Dai, A., Yan, M., Guibas, L.J., 2016. Volumetric and multi-view CNNs for object classification on 3D data. *Proc. IEEE Comput. Vis. Pattern Recognit*, 5648–5656.

Simonovsky, M., Komodakis, N., 2017. Dynamic edge-conditioned filters in convolutional neural networks on graphs. *Proc. IEEE Comput. Vis. Pattern Recognit*, 29–38.

Strom, J., Richardson, A., Olson, E., 2010. Graph-based segmentation for colored 3D laser point clouds. *International Conference on Intelligent Robots & Systems*, 2131–2136.

Sun, J., Lai, Z. L., 2014. Airborne LiDAR feature selection for urban classification using random forest. *Geomatics and Information Science of Wuhan University*, 39(11), 1310-1313.

Tchapmi, L.P., Choy, C.B., Armeni, I., Gwak, J., Savarese, S., 2017. SEGCloud: Semantic segmentation of 3D point clouds. *3D Vis*. 537–547.

Qi, C.R., Yi, L., Su, H., Guibas, L.J., 2017b. PointNet++: Deep hierarchical feature learning on point sets in a metric space. *Adv. Neural Inf. Process. Syst.*

Shapovalov, R., Velizhev, A., Barinova, O., 2010. Non-associative markov networks for 3d point cloud classification. *ISPRS Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*

Wang, L., Shan, J., Huang, Y. C., He, L., 2018. MSNet: Multi-scale convolutional network for point cloud classification. *Remote Sensing*, 10(4), 612.

Wang, Z., Zhang, L. Q., Fang, T., 2015. A multiscale and hierarchical feature extraction method for terrestrial laser scanning point cloud classification. *IEEE Trans. Geosci. Remote Sens*, 53(5), 2409-2425.

Weinmann, M., Urban, S., Hinz, S., Jutzi, B., Mallet, C., 2015. Distinctive 2D and 3D features for automated large-scale scene analysis in urban areas. *Computers & Graphics*, 49, 47-57.

Xiang, B., Yao, J., Lu, X., Li, L., Xie, R., Li, J., 2018. Segmentation-based classification for 3D point clouds in the road environment. *International Journal of Remote Sensing*, 1-31.

Yan, J., Shan, J., Jiang, W., 2014. A global optimization approach to roof segmentation from airborne lidar point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing,* 94, 183–193.

Yi, L., Su, H., Guo, X., Guibas, L., 2017. SyncSpecCNN: Synchronized spectral CNN for 3D shape segmentation. *Proc. IEEE Comput. Vis. Pattern Recognit*, 6584–6592.

Zaheer, M., Kottur, S., Ravanbakhsh, S., Poczos, B., Salakhutdinov, R., Smola, A., 2017. Deep sets. *Adv. Neural Inf. Process. Syst.*

Zhang, J., Lin, X., Ning, X., 2013. SVM-Based classification of segmented airborne LiDAR point clouds in urban areas. *Remote Sensing*, 5(8), 3749–3775.

Zhang, Y., Chen, T., 2012. Efficient inference for fully-connected CRFs with stationarity. *Proc. IEEE Comput. Vis. Pattern Recognit*, 582-589.

Zhang, Z., Zhang, L., Tong, X., Mathiopoulos, P.T., Guo, B., Huang, X., Wang, Z., Wang, Y., 2016. A multilevel point-cluster-based discriminative feature for ALS point cloud classification. *IEEE Trans. Geosci. Remote Sens*, 54, 3309–3321.

Zhou, Y., Yu, Y., Lu, G., Du, S., 2012. Super-segments based classification of 3D urban street scenes. *International Journal of Advanced Robotic Systems*, 9(6), 248-256.