Int J Appl Earth Obs Geoinformation 60 (2017) 99–110

Contents lists available at ScienceDirect

Int J Appl Earth Obs Geoinformation

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

Selection of LiDAR geometric features with adaptive neighborhood size for urban land cover classification

Weihua Dong[a,*], Jianhang Lan[a], Shunlin Liang[a,b], Wei Yao[c], Zhicheng Zhan[a]

[a] State Key Laboratory of Remote Sensing Science, Beijing Key Laboratory for Remote Sensing of Environment and Digital Cities & Faculty of Geography, Beijing Normal University, China

[b] Department of Geographical Sciences, University of Maryland, College Park, MD, USA

[c] Photogrammetry and Remote sensing, Technische Universitaet Muenchen, Munich, Germany

ARTICLE INFO

Keywords:
LiDAR data
Geometric feature
Neighborhood
Land cover classification

ABSTRACT

LiDAR has been an effective technology for acquiring urban land cover data in recent decades. Previous studies indicate that geometric features have a strong impact on land cover classification. Here, we analyzed an urban LiDAR dataset to explore the optimal feature subset from 25 geometric features incorporating 25 scales under 6 definitions for urban land cover classification. We performed a feature selection strategy to remove irrelevant or redundant features based on the correlation coefficient between features and classification accuracy of each features. The neighborhood scales were divided into small (0.5–1.5 m), medium (1.5–6 m) and large (> 6 m) scale. Combining features with lower correlation coefficient and better classification performance would improve classification accuracy. The feature depicting homogeneity or heterogeneity of points would be calculated at a small scale, and the features to smooth points at a medium scale and the features of height different at large scale. As to the neighborhood definition, cuboid and cylinder were recommended. This study can guide the selection of optimal geometric features with adaptive neighborhood scale for urban land cover classification.

1. Introduction

Light Detection and Ranging (LiDAR) is an important data source for generating DTM, topographic maps, 3D city models, land cover classifications(Yan et al., 2015; Rottensteiner, 2012), ecosystem studies (Yao et al., 2012) and natural hazard assessments (Jaboyedoff et al., 2012). However, due to the diversity of object classes, the complexity of object structures and the variability of point features, classification is still an active field of research. Early studies focused on classifying the points into ground points and non-ground points, which is also called filtering (Axelsson, 1999; Meng et al., 2010). With the development of LiDAR technology, denser points with high precision are obtained in urban areas. LiDAR data are also used for the classification of non-ground objects, such as buildings or vegetation (Axelsson, 1999). In different scenes, many methods are developed to achieve this goal. Classification can be applied directly to the 3D points or pixels of a Digital Surface Model derived from the LiDAR points. Each point or pixel has its own potential features which are applied to different methods of classification. Supervised classification is the most common method, including maximum likelihood classification, artificial neural networks, adaptive boosting, support vector machines and Random Forest (Yan et al., 2015; Rottensteiner, 2012). Relevant features are often uncertain for supervised classifiers and it is a common strategy to introduce more candidate features to get a better representative of the domain (Dash and Liu, 1997). As a new candidate of feature space, LiDAR point cloud data describes the geometrical information of objects in 3D space directly (Demantké et al., 2011). The geometry is computed for each point with its neighbor as a set of geometric features. Therefore, geometric features can be used together with intensity (Zhou, 2013), full-wave features (Alexander et al., 2010; Alexander et al., 2011) and high-resolution satellite imagery (Guan et al., 2013; Yanfeng et al., 2015) Feature extraction and feature selection are the two broad categories for reducing not important features. Since feature selection preserve the physical information of the original features, previous work explored the importance of features to select a subset of highly discriminant features. Chehata et al. (2010) compared fifteen geometric features and other six non-geometric features with Random Forest classifier for urban scene classification. The results demonstrate that height difference and height variance are the top two most important features and that height differences are more important than the others. Guo et al. (2011) explored the relevance of airborne LiDAR and multispectral image data for urban scene classification. The feature

* Corresponding author.
E-mail address: dongweihua@bnu.edu.cn (W. Dong).

http://dx.doi.org/10.1016/j.jag.2017.04.003
Received 26 October 2015; Received in revised form 31 March 2017; Accepted 11 April 2017
Available online 03 May 2017

vector was composed of optical features, multi-echo LiDAR features, full waveform LiDAR features and three geometric features. Unlike height difference, the plane angle and residuals are of low importance in the comparison. Niemeyer et al. (2012) presented a context-based Conditional Random Field classifier with LiDAR features to classify urban scenes and obtained reliable classification results. Although the geometric features are often used in the classification, the importance of each feature nor the effect of combining features is clear.

Moreover, the selection of neighbor of points is a critical factor for calculating geometric features and influences the classification accuracy of land cover. The neighbors of points are typically chosen as the k nearest or all points in a cylinder or sphere (Weinmann et al., 2014). As to grid based features, the neighbor of pixel refers to window size. The neighborhood scale are traditionally fixed with an empirical value (Alexander et al., 2010; Tang et al., 2014). The optimal neighborhood scale may be different for each feature. Some studies explored the optimal neighborhood scale for LiDAR features. Niemeyer et al. (2011) found that overall accuracy had a local maximum at seven neighbors. The selected features included residuals of an estimated plane and as well as the features dimensionality based on the eigenvalues. Demantk & et al. (2011) retrieved the optimal neighborhood scale of eigenvalue-based features for labeling point dimensionality by minimizing entropy feature and maximizing similarity index. Lack of research on neighborhood scale makes it necessary to explore the optimal neighborhood scale for each geometric feature.

This study aimed to explore the relationship between classification accuracy and geometric features with different neighborhood scale under different neighborhood definition and to propose criterion for selecting optimal geometric feature with adaptive window size for urban land cover classification. The paper was structured as follows. The geometric features were described in Section 2. A feature selection strategy was developed in Section 3. The selected features and classification results were presented in Section 4. The results were discussed in Section 5, and conclusions were drawn in Section 6.

2. Geometric features of LiDAR points

2.1. Geometric features of LiDAR points

Height difference (Δ_{min}) between the current point and the lowest point is the most commonly used feature, as it roughly measures local variation. However, Δ_{min} takes only the current height and the lowest point into account and loses useful information of the other points. Axelsson (1999) uses the second derivatives of interpolated raster images to enhance variations. Maas (1999) introduces the Laplace operator, maximum slope measures and original height data to classify the data. Unlike previous features constructed by variation, Filin (2002) presented surface parameters to describe the planarity. Additionally, Gross and Thoennessen (2006) proposed eigenvalue-based values to depict 3D characters. Chehata et al. (2009) proposed grouped point cloud features, including height-based LiDAR features (height difference compared to its neighbors, height difference between first and last pulses, height variance and local curvature), eigenvalue-based LiDAR features computed with the variance-covariance matrix of the local neighborhood (anisotropy, planarity, sphericity and linearity), and local-plane-based LiDAR features (deviation angle, variance of deviation angles and residual of the local plane estimated in a cylinder). Geometric features used for LiDAR classification can be shown in Table 1.

In addition to the geometric features listed in Table 1, some geometric features are calculated after rasterizing LiDAR points, such as the Laplace filter, Sobel operator (Maas, 1999) and texture measures (height homogeneity, height contrast, height entropy, height correlation) (Im et al., 2008). Previous studies demonstrate that the most importance metrics are mean height, height standard variance (Gross and Thoennessen, 2006), height difference (Alexander et al., 2010), obvious height and minimum value of height in a neighborhood (Charaniya et al., 2004).

2.2. Feature selection method

To preserve the physical information of the original features, previous work reduced not important features with the effectiveness of features for classification, which belongs to embedded models of feature selection. However, to evaluate the importance of selected features, it's necessary to compare different feature selection and feature extraction method. In the section, we introduce the major method for feature selection.

Assuming features are independent, feature selection is further divided into three groups − filter models, wrapper models, and embedded models. Filter models depends on the characteristics of features. It first ranks features with criteria, such as fisher score (Duda et al., 2012), mutual information (Koller and Sahami, 1996) and feature relief (Robnik-Šikonja and Kononenko, 2003), and then select the highest ranked features. However, the optimal features subset would relate to classifier, which is ignored in filter method. Wrapper models compare the effectiveness of classifier with all the combination of features and then select the subset with highest quality. Forward selection and backward elimination the two most frequently used method (Mao, 2004). However, the combination count for m features is 2^m, which make it impractical for a large m. Embedded models embed feature selection with classifier construction, including pruning methods, build-in mechanism and regularization models (Dash and Liu, 1997). Random Forest is a decision tree-based ensemble classifier and provides feature importance after classification. Chehata et al. (2009) used the Random Forest to classify full-waved LiDAR points, achieving an overall accuracy of approximately 95% for the classes of building, vegetation, artificial ground and natural ground. Wei et al. (2012) evaluated the feature relevance of point cloud provided with an AdaBoost classifier and verified with the importance of Random Forest. In this paper, we proposed a feature selection method to select features. To evaluate the method, we compared the selected features with other feature selection method.

3. Methodology

3.1. Geometric features extraction

Geometric feature was influenced with the types of features, neighborhood definition and neighborhood scale. Apart from using most of most of features extracted from point data listed in Table 1, we added additional statistical features, such as the min, max, mean and the medium of height of points within neighborhood. These features consisted of four series: (1) Ten height-statistics-based features: min (S_{min}), max(S_{max}), mean(S_{mean}), mod(S_{mod}), median(S_{med}), data range (S_{dr}), standard variance(S_{std}), Coefficient of Variation(S_{coev}), skewness (S_{skw}), and kurtosis(S_{krt}); (2) Four height-texture-based features: z-min (Δ_{min}), z-max(Δ_{max}), z-mean(Δ_{mean}), and max slope divided by π to normalization(Δ_{slope}); (3) Five fitting-plane-based features: the parameter to coordinate x of the fitting plane (Π_a), the parameter to coordinate y of the fitting plane (Π_b), the correlation coefficient of the fitting plane (Π_{R^2}), the root-mean-square error of the fitting plane (Π_{RMSE}), and the normal vector angle of the fitting plane divided by π (Π_{ang}); and (4) Six eigenvalue-based features: λ_1–λ_3, linearity(λ_L), planarity(λ_P), and sphericity (λ_S). The neighborhood was defined as cuboid (D_{cubd}), cube (D_{cube}), cylinder (D_{cyln}), sphere (D_{sphr}), k nearest points in 3D (D_{k3d}) and k nearest points in 2D projection (D_{k2d}).

With the increase of neighborhood scale, the proportion of additional neighborhood points to previous neighborhood points would decrease. With a large scale, expansion of scale would almost not affect geometric features. We recommended selecting scales corresponding to geometric sequences by Eq. (1).

Table 1
List of geometric features.

Geometric feature	symbol	Category	Definition	Description	Type	Reference
Height data	z	Original	The original height data	The original height data will allow for a segmentation between high objects, such as buildings and trees on one hand, and objects such as streets or plain ground on the other.	Point	Maas (1999)
Point coordinates	(x, y, z)	Original	The original point coordinates	The z is the same as Height data. The points within a short distance of (x, y) may be the same object.	Point	Filin (2002)
Data range	S_{dr}	Statistical	The data range of Z_s^i	The data range will be zero on flat roofs or streets, small on tilted roofs and large within trees and at the edges of objects.	Point, Raster	Maas, (1999), Charaniya et al. (2004)
Variance	S_{var}	Statistical	The variance of Z_s^i	Variance will show characteristics similar to the data range.	Point, Raster	Bartels and Wei,(2006)
Skewness	S_{skw}	Statistical	The skewness of Z_s^i	Skewness measures the asymmetry compared to a normal distribution.	Point, Raster	Antonarakis et al. (2008)
Kurtosis	S_{krt}	Statistical	The kurtosis of Z_w^i	Kurtosis measures the peakedness of Z_w^i.	Point, Raster	Antonarakis et al. (2008)
Second derivative	δ	Height-texture	$\delta = \frac{\partial^2 z}{\partial x^2}$, where x is the direction along the scan line and z is the elevation.	The Second derivative will be zero for a straight line and have random behavior at breakpoints or vegetation.	Raster	Axelsson (1999)
Mean height difference	Δ_{mean}	Height-texture	The difference of z^i and the mean of Z_s^i	The mean height difference decreases from high vegetation to low vegetation, smooth surface and planar surface.	Point, Raster	Filin (2002)
Height difference	Δ_{min}	Height-texture	The difference of z^i and the minimum of Z_s^i	Height difference helps in discriminating ground and off-ground objects and is used most frequently.	Point	Chehata et al. (2009), Guo et al. (2011)
Max slope	Δ_{slope}	Height-texture	The maximum slope from p^i to each point in P_s^i	The max slope is valuable for distinguishing tilted roofs from flat roofs or street from trees, where the slope will reach very large values	Point	Maas (1999)
Curvature	Δ_{curv}	Height-texture	The difference between two adjacent gradients	Curvature is similar to the second derivative. It will detect the discontinuity of points.	Raster	Chehata et al. (2009), Vogtle (2001)
Echo height difference	Δ_{fl}	Height-texture	The height difference between first echo and last echo	Echo height difference would help discriminating building roofs and ground.	Point	Antonarakis et al. (2008)
Surface parameters	Π_a, Π_b, Π_c	Surface-plane	A fitting plane Π_w^i ($z = ax + by + c$) estimated with P_s^i	The two parameters can also be represented by the angle and the direction of the plane. The angle of ground and building will be fixed and the angle of vegetation will be random.	Point	Filin (2002)
Normal vector angle	Π_{ang}	Surface-plane	The angle of the fitting plane normal vector in the vertical direction	The normal vector angle will highlight the ground. Angle is the complementary angle of the plane angle, which is the angle between the plane and the horizontal plane.	Point	Chehata et al. (2009), Guo et al. (2011)
Residual	$\Pi_{residual}$	Surface-plane	The sum of the distance between points and fitting plane.	Residuals should be high for vegetation.	Point, Raster	Chehata et al. (2009), Guo et al. (2011)
Angle Variance	Π_{angvar}	Surface-plane	The variance of deviation angles within the fitting plane.	Angle variance discriminates planar surfaces such as roads and building roofs from vegetation.	Point, Raster	Chehata et al. (2009)
Eigenvalues	λ_1, λ_2, λ_3 ($\lambda_1 > \lambda_2 > \lambda_3$)	Eigenvalue-based	The eigenvalues of the variance-covariance matrix computed within P_s^i	The eigenvalues provide the magnitude of the three principal directions of P_w^i. λ_3 increases as the degree of dispersion grows. The eigenvalues are used to compute spatial distribution.	Point	Demantké et al. (2011), Chehata et al. (2009), Gross and Thoennessen (2006)
Linearity	λ_L	Eigenvalue-based	$\lambda_L = \frac{\lambda_1 - \lambda_2}{\lambda_1}$	The Linearity is large if the points of P_w^i are linear.	Point	
Planarity	λ_P	Eigenvalue-based	$\lambda_P = \frac{\lambda_2 - \lambda_3}{\lambda_1}$	The planarity is large if the points of P_w^i are coplanar.	Point	
Sphericity	λ_S	Eigenvalue-based	$\lambda_S = \frac{\lambda_3}{\lambda_1}$	The sphericity is large if the points of P_w^i are discrete.	Point	
Anisotropy	λ_A	Eigenvalue-based	$\lambda_A = \frac{\lambda_1 - \lambda_3}{\lambda_1}$	The anisotropy is equal to 1 minus sphericity.	Point	

p^i is the current point, z^i is the height of p^i, P_s^i represents the points in a scale (s) around the point p^i, and Z_s^i is the height of P_s^i.

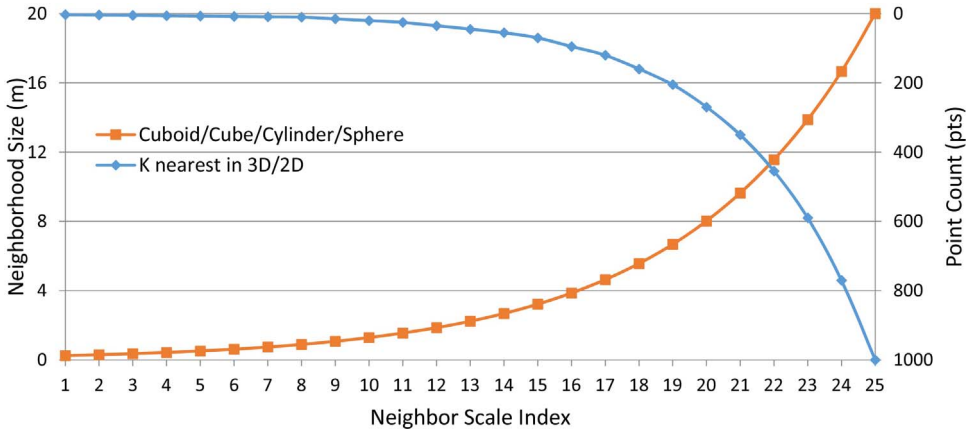

Fig. 1. Neighborhood scale for each type of neighborhood definition.

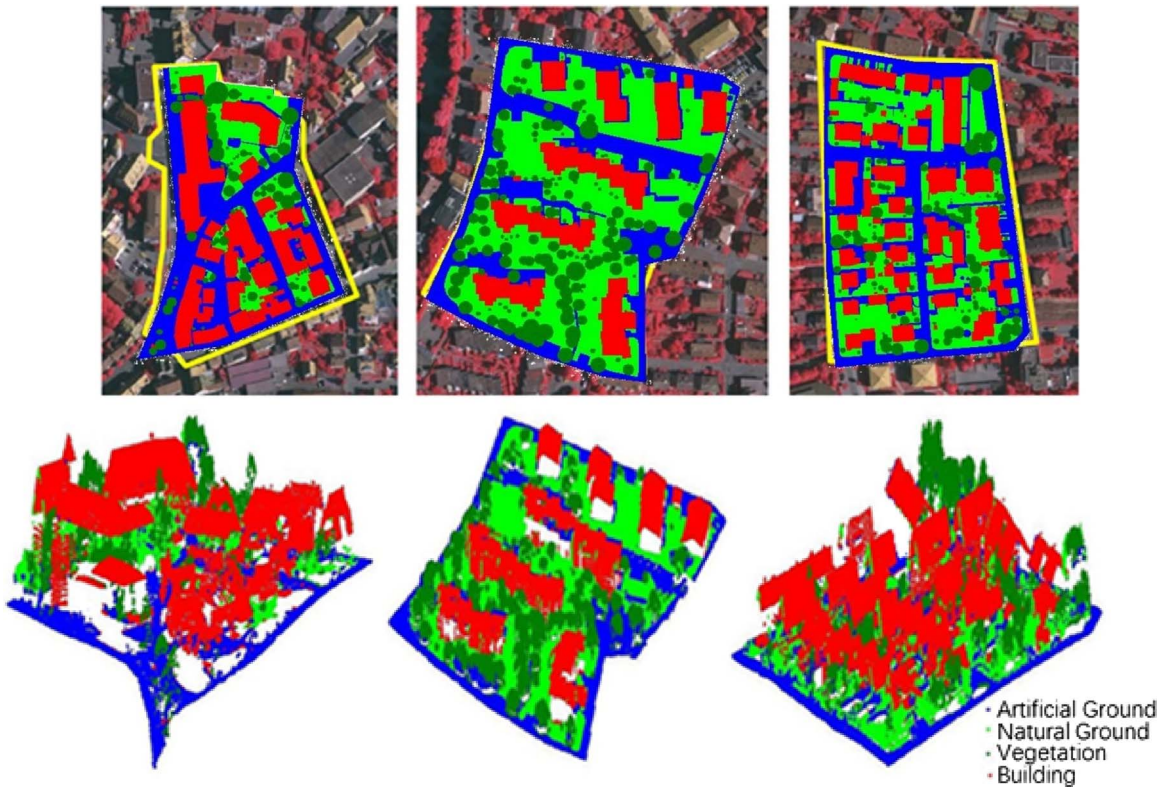

Fig. 2. 2D reference data and labeled point cloud.

Table 2
Point count and distribution rate of four object classes in three areas.

Object Classes	Area 1		Area 2		Area 3		Total	
	point	Rate	point count	Rate	point	Rate	point	Rate
Artificial Ground	29935	27.25%	45909	23.26%	67481	31.27%	143325	27.40%
Natural Ground	24292	22.11%	65677	33.28%	56642	26.25%	146611	28.03%
Vegetation	17151	15.61%	43189	21.88%	24531	11.37%	84871	16.23%
Building	38493	35.03%	42581	21.58%	67147	31.12%	148221	28.34%
Total	109871	100.00%	197356	100.00%	215801	100.00%	523028	100.00%

$$s_i = s_{min} \cdot r^{i-1},\ r = \sqrt[n-1]{s_{max}/s_{min}} \quad (1)$$

Where s_i was the i-th neighborhood scale, $s_{\min}$ and $s_{\max}$ were the minimum and maximum scale, r was the common ration of geometric sequences and n was the number of scales. The selected scales distributed dense in small scale and sparse in large scale. For each neighborhood definition, we selected 25 scales. As to D_{cubd}, D_{cube}, D_{cyln} and D_{sphr}, the minimum and maximum scales were 0.25 m and 20 m. As to D_{k3d} and D_{k2d}, the minimum and maximum numbers of points were 3 and 1000. The neighborhood scales were shown in Fig. 1. The resulting feature vector $\mathcal{F}$ for each point is given by Eq. (2).

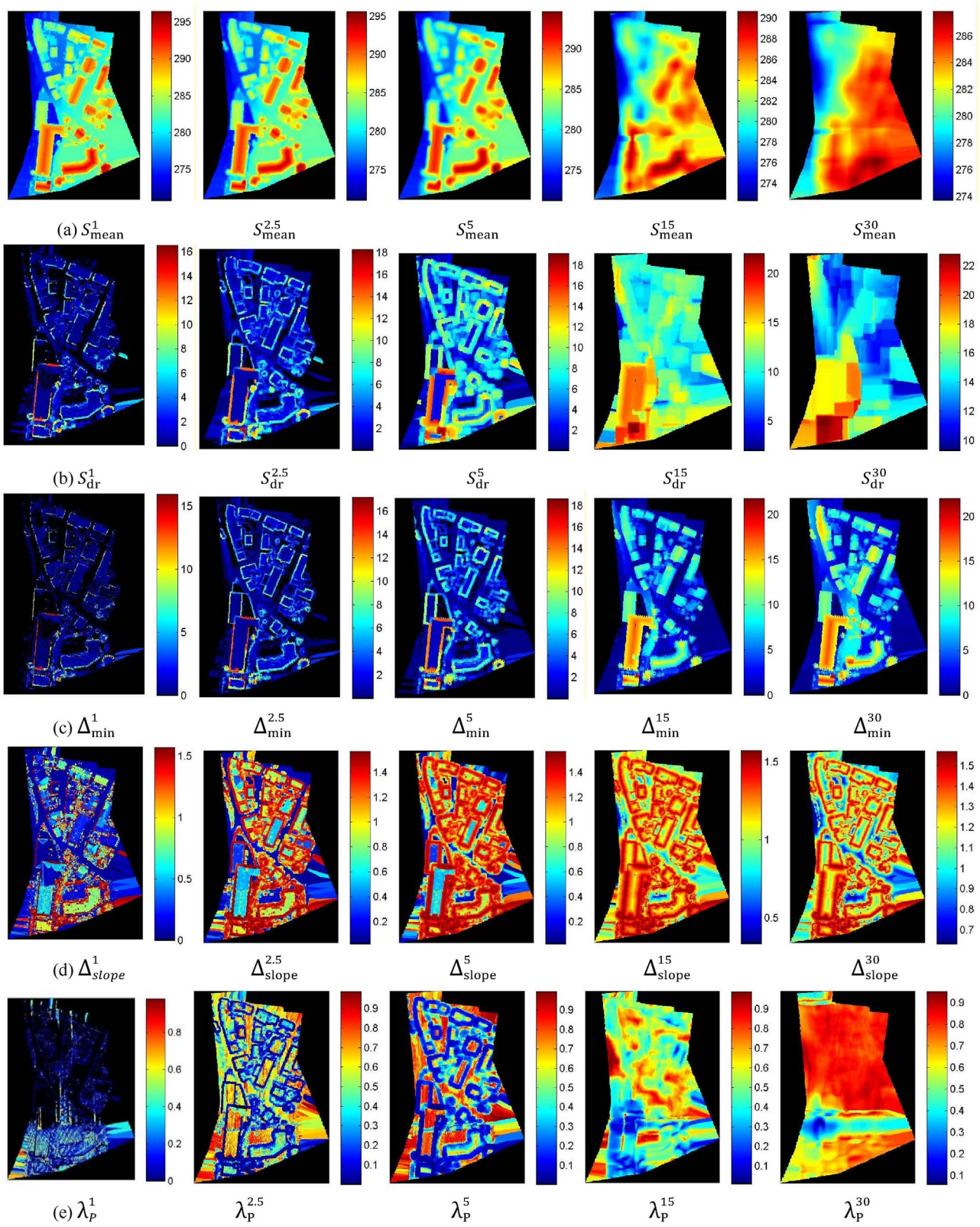

Fig. 3. Representative geometric features with different neighborhood scale.

$$\mathcal{F} = \{T, D, S\} \quad (2)$$

Where T represented the type of geometric features, D represented the definition of neighborhood and S represented the neighborhood scale. There were up to 3750 features by combining the 25 types of geometric features, 6 types of neighborhood and 25 neighborhood scales.

3.2. Feature selection strategy

Too many features would bring a lot of irrelevant or redundant information for classification. Adding irrelevant or redundant information would not improve the classification accuracy. However, it was not

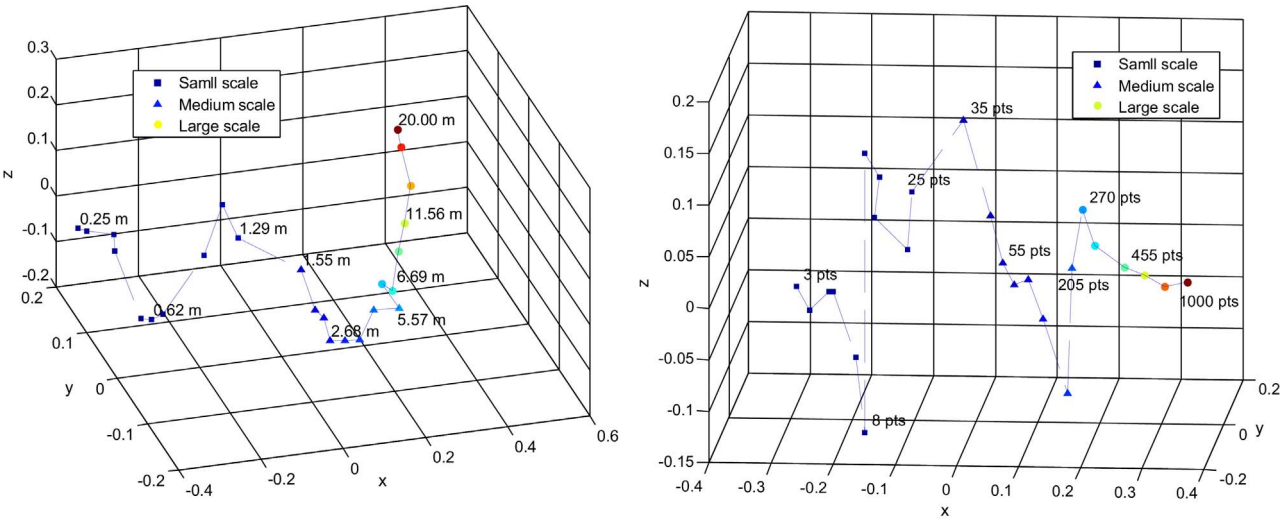

Fig. 4. Cluster of neighborhood scale. Left, based on meter; Right, based on points.

Table 3
Table Clustering of feature types under different neighborhood definition.

	Group 1	Group 2	Group 3	Group 4
D_{cubd}	S_{min}, S_{max}, S_{mean}, S_{mod}, S_{med}	Π_{ang}, λ_L, λ_P	Δ_{mean}, Π_b, S_{krt}	S_{dr}, S_{std}, S_{coev}, S_{skw}, Δ_{min}, Δ_{max}, Δ_{slope}, Π_a, Π_{R^2}, Π_{RMSE}, λ_1, λ_2, λ_3, λ_S
D_{cube}		Π_a, Π_b, Π_{R^2}, λ_2, λ_L, λ_P	S_{skw}, S_{krt}, Δ_{mean}	S_{dr}, S_{std}, S_{coev}, Δ_{min}, Δ_{max}, Δ_{slope}, Π_{RMSE}, Π_{ang}, λ_1, λ_3, λ_S
D_{cyln}		Π_{ang}, λ_L, λ_P,	S_{krt}, Δ_{mean}, Π_a, Π_b	S_{dr}, S_{std}, S_{coev}, S_{skw}, Δ_{min}, Δ_{max}, Δ_{slope}, Π_{R^2}, Π_{RMSE}, λ_1, λ_2, λ_3, λ_S
D_{sphr}		Π_a, Π_{R^2}, λ_1, λ_2, λ_L, λ_P,	S_{skw}, S_{krt}, Δ_{mean}, Π_b	S_{dr}, S_{std}, S_{coev}, Δ_{min}, Δ_{max}, Δ_{slope}, Π_{RMSE}, Π_{ang}, λ_3, λ_S
D_{k3d}		S_{krt}, Π_{R^2}, λ_2, λ_3, λ_S	S_{skw}, Δ_{min}, Δ_{max}, Δ_{mean}, Π_a, Π_b,	S_{dr}, S_{std}, S_{coev}, Δ_{slope}, Π_{RMSE}, Π_{ang}, λ_1, λ_L, λ_P
D_{k2d}		Π_a, Π_b	S_{skw}, S_{krt}, Δ_{min}, Δ_{max}, Δ_{mean}, λ_L, λ_P	S_{dr}, S_{std}, S_{coev}, Δ_{slope}, Π_{R^2}, Π_{RMSE}, Π_{ang}, λ_1, λ_2, λ_3, λ_S

a good strategy to remove all features with low relevant or high redundant information: 1) a low relevant features with little redundant information would raise classification accuracy, 2) a high redundant but relevant feature might also help improve classification. To select geometric features for classification, we should take into account the type of features, the definition of neighborhood and neighborhood scale.

To search for well-performed geometric feature subset, we considered three cases: 1) The optimal neighborhood scale of each geometric features. We calculated kappa coefficient of classification with each geometric feature incorporating 25 neighborhoods separately. The neighborhood scale achieving highest classification accuracy would be the optimal scale for the feature. 2) Optimal feature subset under each neighborhood definition. Adding more feature types, we needed to select a subset of features to achieve preferable classification. 3) Optimal feature subset of the full feature space. For each neighborhood definition, we extracted 625 features combining different feature type and neighbor scale, it was hard to compare classification accuracy with all possible subset and select the optimal subset.

Here, we proposed a methodology to select features based on correlation coefficient and classification accuracy of each feature. The correlation coefficient could reflect the redundant information between features, and classification accuracy of each feature would indicate how much the feature was relevant to classification. Therefore, selecting features with high classification and low correlation coefficient would improve classification.

After extracting all features, we first calculated classification accuracy with each feature. We then calculated the correlation coefficient between each pair of neighborhood definition COF_{ij}^{D}, neighborhood scale COF_{ij}^{S} and feature types COF_{ij}^{T}. The distance was estimate as Eq. (3).

$$d = 1 - abs(COF) \tag{3}$$

With distance matrix, we performed multidimensional scaling (MDS) to compute the coordinates of neighborhood scale X^s and feature types X^t and obtained the coordinate of each feature in different neighborhood definition $X^d = [X^s, X^t]$. To select feature subset with K feature, we used K-means clustering to partition the features into K clusters. Finally, we selected the feature with highest classification accuracy from each cluster and combined them as optimal subset.

3.3. Classification accuracy assessment

Accuracy assessment is an important task of classification. Producer accuracy PA_F^C, user accuracy UA_F^C, overall accuracy OA_F, kappa coefficient Kp_F (Liu et al., 2007) and the variable importance of each feature f by mean decrease in permutation accuracy VI_f^C (Breiman, 2001) are computed to evaluate the urban land cover classification results. F refers to the combination of features. C refers to object class, including artificial ground, natural ground, vegetation and building. PA_F^C and UA_F^C could be combined into their harmonic mean, known as F-measure FA_F^C (Hripcsak and Rothschild, 2005). F-measure is a balance of producer accuracy and user accuracy. F-measure can be formulated by Eq. (4).

$$FA_F^C = \frac{2 \cdot PA_F^C \cdot UA_F^C}{PA_F^C + UA_F^C} \tag{4}$$

Kappa coefficient is an effective metric for evaluating the classification results. Kappa coefficient range from 0 to 1. The classification can be interpreted to poor(less than 0.2), fair (0.2–0.4), moderate (0.4–0.6), good (0.6–0,8) and very good (0.8–1.0) (Alexander et al., 2010). To prove the capacity of features for classification, we compared our selected subset with other feature selection method, such as features with K highest fisher score or importance index computed by random forest. For each feature selection model, a 10-fold cross validation was performed to evaluate the classification result.

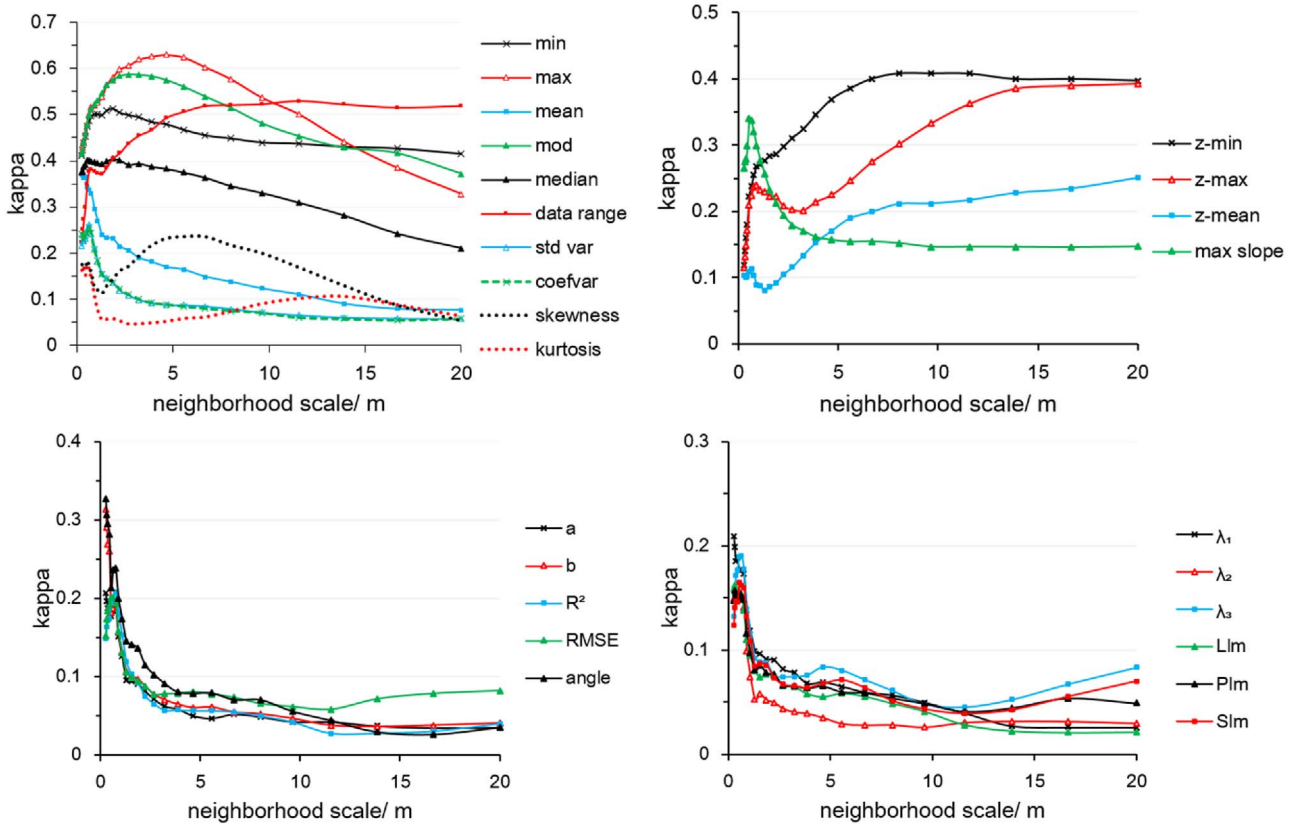

Fig. 5. Kappa coefficient of a single feature for Area 1.

4. Experiment and results

4.1. Study area

The airborne LiDAR data were collected using a Leica ALS50 system over Vaihingen, Germany (mean flying height above ground: 500 m, field of view: 45°) (Cramer, 2010). The average strip overlap is 30% and the point density ranges from 4 to 8 points/m^2. Multiple echoes and intensities were recorded. Three test sites with different scenes were considered. Area 1 consisted of roads, trees and buildings with complex structure. In Area 2, the high-rise residential buildings were surrounded by trees. Area 3 was composed of detached houses and many surrounding trees.

With the 2D reference data for these three areas (Duda et al., 2012), we labeled the point cloud by matching its coordinates with reference data and fixed error labels by manual correlation. For each area, we discerned the following four object classes: artificial ground, natural ground, vegetation and building (Guo et al., 2011). The point count and distribution of the object classes in each area was shown in Fig. 2 and Table 2.

4.2. Geometric features

We extracted totally 3750 features with 25 types of geometric features incorporating 6 neighborhood definitions and 25 neighborhood scales. Fig. 3 showed how geometric features varied as neighborhood scale increase. S_{mean} would smooth the surface of objects, such as the root of buildings. However, a large neighborhood scale would involve several objects and make S_{mean} hard to distinguish objects. S_{dr}, $\Delta_{\min}$ and Δ_{slope} had similar performance with small neighborhood scale. They were all large at the edge of buildings and within trees and were small on the roots of building and streets. With a large neighborhood scale, S_{dr} became fuzzy while $\Delta_{\min}$ made the building more significant and Δ_{slope} remained the information of edges. In contrast with S_{dr}, Π_{R^2} depicted the homogeneity of height within a neighborhood. Π_{R^2} was large within the roof of buildings. λ_L, λ_P and λ_S could reveal the dimensionality of points.

It was obvious that the same types of features with large neighborhood scale were highly relevant. When a large neighborhood scale expanded, the additional points would play a minor role in extracting features. In contrast when expanding a small neighborhood scale, the additional points would account for a large proportion and influence the features greatly. We calculated the correlation coefficient of each couple of neighborhood scale for each of the 25 types of features, and then computed the average of the correlation coefficient. Features with adjacent scale had high correlation coefficient. As the same times, features with large size difference had low correlation coefficient, but it was inconspicuous when the neighborhood scale are both large. To categorize the neighborhood scale, we took "1 − absolute correlation coefficient" as the distance between neighborhood scales, and then performed multidimensional scaling (MDS) to map every neighborhood scales as points in multi-dimensional space. The results were shown in Fig. 4. Finally, we categorize the neighborhood scale as small scale (0.25–1.5 m or 3–25 pts), medium scale (1.5–6 m or 25–250 pts) and large scale (larger than 6 m or more than 250 pts) by k-means clustering.

We also tried to cluster the feature types into four groups based on correlation under different neighborhood definition. Table 3 showed the result of clustering. The clustering of different neighborhood definition was inconsistent. However, there were still two stable feature type group. $S_{\min}$, $S_{\max}$, S_{mean}, S_{mod} and S_{med} were close to each other, which smooth the height value of points. S_{dr}, S_{std}, S_{coev}, Δ_{slope}, and Π_{RMSE} with large value would indicate edges of objects or trees. The rest of feature types would reflect whether the object was smooth or not.

4.3. Classification with geometric feature

Fig. 5 showed the kappa coefficient of classification by Random

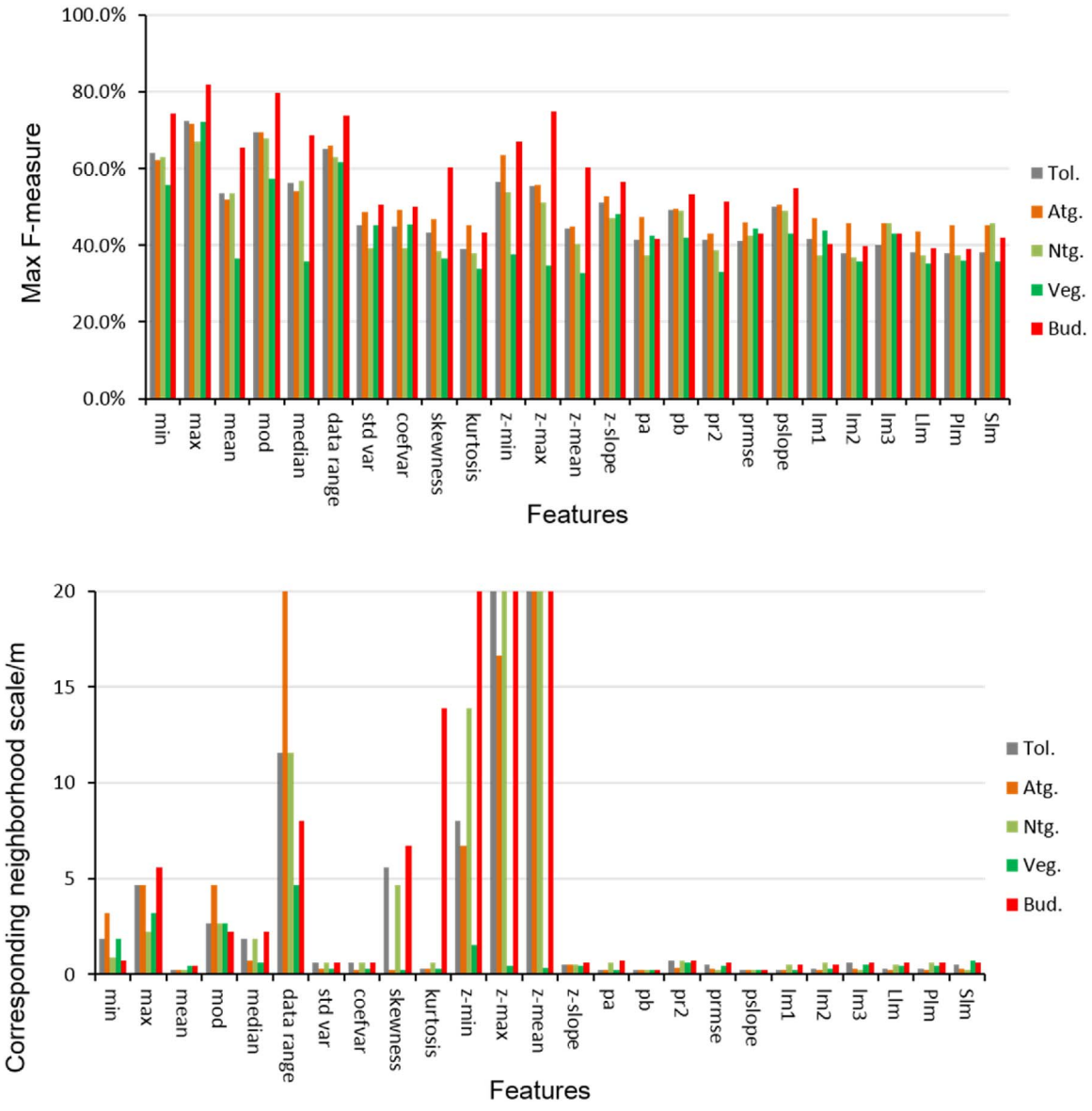

Fig. 6. Maximum of F-measure and corresponding neighborhood scale.

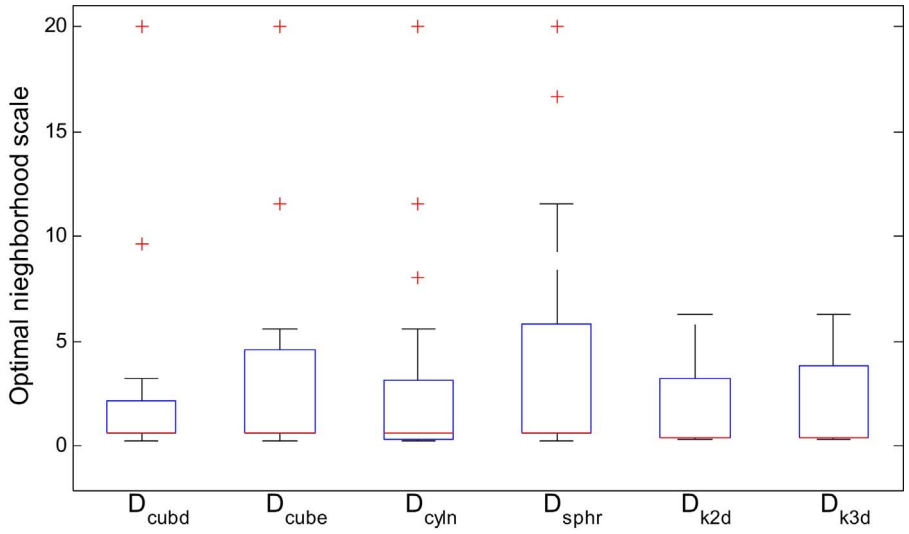

Fig. 7. Optimal neighborhood scale of different definition.

Table 4
Classification Result using all features with different neighborhood types.

index	D_{cubd}	D_{cube}	D_{cyln}	D_{sphr}	D_{k2d}	D_{k3d}
kappa	0.9671	0.9581	0.9659	0.9574	0.9595	0.9506
FA_{art}	0.9788	0.9744	0.9785	0.9740	0.9734	0.9690
FA_{ntr}	0.9690	0.9614	0.9676	0.9605	0.9614	0.9543
FA_{veg}	0.9602	0.9457	0.9591	0.9451	0.9511	0.9370
FA_{bul}	0.9882	0.9848	0.9872	0.9846	0.9862	0.9824
Accuracy	0.9757	0.9690	0.9748	0.9685	0.9700	0.9634

Forest using only one geometric feature under cylinder. The results showed that $K_{S_{min}}$ first increased and then decreased, reaching the local maximum of 0.51 at 1.86 m. The minimum value was higher than 0.41. $K_{S_{max}}$ was an inverted u-curve and ranged from 0.33 to 0.63. $K_{S_{mean}}$ only decreased with the maximum value of 0.37 at 0.25 m $K_{S_{mod}}$ decreased constantly and more gradually than $K_{S_{mean}}$. Similar to the $K_{S_{max}}$, $K_{S_{med}}$ was also an inverted u-curve and ranges from 0.21 to 0.40. When the neighborhood scale was larger than 1.86 m, $K_{S_{med}}$ decreased less than $K_{S_{max}}$. $K_{S_{dr}}$ increased to a stable value of 0.53. $K_{S_{std}}$ reached a local maximum value of 0.26 at 0.62 m. K_{cof} coincided with K_{std}. $K_{S_{skw}}$ was also an inverted u-curve, with a vertex value of 0.24. $K_{S_{krt}}$ was less than 0.2 for all neighborhood scale. $K_{\Delta_{min}}$ increased slowly and reached a stable value of 0.41. $K_{\Delta_{max}}$ could be divided into three sections. From 0.5 m to 1 m, $K_{\Delta_{max}}$ increased rapidly. From 1 m to 3 m, $K_{\Delta_{max}}$ fell by 0.03. In the final section, $K_{\Delta_{max}}$ increased and reached a stable value similar to $K_{\Delta_{min}}$. $K_{\Delta_{mean}}$ increased slowly from 0.08 to 0.25. $K_{\Delta_{slope}}$ showed a steep peak at 0.52 m. When the neighborhood scale was larger than 3 m, $K_{\Delta_{slope}}$ did not change. For fitting-plane-based features and eigenvalue-based features, the classification result would decrease or have a peak at about 0.5 m.

For each feature type, we found the maximum of F-measure over the entire neighborhood scale and got the corresponding optimal neighborhood scale. Fig. 6 showed the maximum F-measure for each type of features and the corresponding neighborhood scale. The F-measures of building and artificial ground were the two highest and range from 39.1% to 81.9%, while most of the F-measures of natural ground and vegetation are lower than 40%. The height-statistics-based measures, such as S_{mean}, S_{max}, S_{med} and S_{dr}, performed well in classification results for the four classes. The maximum F-measures were concentrated at small scale. The optimal neighborhood scale for Δ_{min}, Δ_{max} and Δ_{mean} were large scale

We compared the optimal neighborhood scale of different neighborhood definition. The neighborhood of k nearest in 3D and 2D had different metrics with the number of points, we estimated the radius of the circle or spherical with the density of 8 pts/m^2. Fig. 7 showed the box plot of optimal neighborhood scale under different definition. The optimal scale under different definition had the similar performance. Most of optimal scale ranged from 0.25 m to 1.5 m and a part of feature achieved highest accuracy with largest scale. The correlation coefficient of optimal scale between cuboid, cube, cylinder and sphere neighborhood were all larger than 0.92. K nearest points in 2D could be considered as a cylinder as characteristic case. The average of kappa coefficient of classification with different neighborhood were 0.2104(D_{cubd}), 0.1994(D_{cube}), 0.2036(D_{cyln}), 0.1937(D_{sphr}), 0.2277(D_{k2d}) and 0.2187(D_{k3d}).

We also compared the classification result using all features with different neighborhood types. The result was shown as Table 4. With all features, the classification was improved significantly. All the kappa coefficient was higher than 0.95. The most of indexes from high to low were D_{cubd}, D_{cyln}, D_{k2d}, D_{cube}, D_{sphr} and D_{k3d}. Only the FA_{art}^{cube} was higher than FA_{art}^{k2d}. In the four objects, the F-measure from high to low were building, artificial ground, natural ground and vegetable.

4.4. Feature subset with feature selection strategy

Based on D_{cyln}, we compared our feature selection method with fisher score and importance of Random Forest. The top ten features were selected sequentially. The classification results were shown in Fig. 8 and Table 5. Features based on fisher score were almost the types of Δ_{min} and Δ_{mean} with a series of successive scale. The high correlation coefficient between the features resulted in little improvement on classification. Features based on importance of random forest were similar to fisher score. It chose Δ_{min}, Δ_{max} and Δ_{mean} as optimal subset and achieved worse classification than fisher score. Our new methods clustered the features to N partition and selected the best performance feature from each partition to compose of optimal subset with N features. The result showed that our new methods performed the best. Kappa coefficient improved from 0.6299 to 0.9285. In the different neighborhood definition, the kappa coefficient would be stable after selecting five or seven features.

We also selected the best 10 features from the other definition of features and compared the kappa coefficient and F-measure (Table 6). In the six neighborhood definitions, D_{cubd} performed the best with highest kappa coefficient and F-measure, followed with and cube. 2D and 3D nearest neighborhood had similar kappa coefficient, but the FA_{nstc}^{art} and FA_{nstc}^{ntr} are higher than FA_{nsts}^{art} and FA_{nsts}^{ntr} while FA_{nstc}^{veg} and FA_{nstc}^{bul} are lower than FA_{nsts}^{veg} and FA_{nsts}^{bul}.

Each times we selected different combination of features, we counted the times of the type of feature had been selected (Fig. 9). S_{dr}, S_{max} (S_{mod} in D_{k3d}), Δ_{min} and Π_b were the most frequently selected feature types. S_{skw}, Δ_{max}, Δ_{slope} and Π_{R^2} were also selected many times. Some features with low kappa coefficient of single feature would play an important role in classification, such as λ_S under D_{cubd}, Π_{R^2} under D_{cyln}, Π_{R^2} and λ_L under D_{k3d}. In the most of these selected feature types, the corresponded neighborhood scales were consistent with the optimal scale under single features. However, some features would range from the whole scale. For example, S_{dr} was selected with scale of 0.75 m, 3.22 m and 11.56 m.

5. Discussion

5.1. Influence of geometric feature for classification

A geometric feature was determined by neighborhood definition,

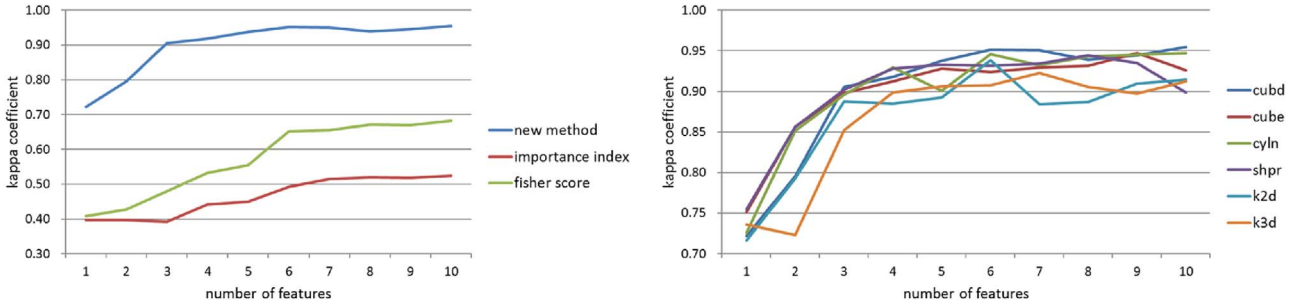

Fig. 8. Classification kappa coefficient with different number of features. Left was selected by different methods and right was selected with different neighborhood definition.

Table 5
The selected features by different methods.

		1	2	3	4	5	6	7	8	9	10
Fisher score		$\Delta_{min}^{8.03}$	$\Delta_{min}^{9.63}$	$\Delta_{min}^{11.56}$	$\Delta_{min}^{13.89}$	$\Delta_{min}^{6.68}$	$\Delta_{min}^{16.66}$	Δ_{min}^{20}	Δ_{mean}^{20}	$\Delta_{min}^{5.57}$	$\Delta_{mean}^{16.66}$
importance		Δ_{max}^{20}	$\Delta_{max}^{16.66}$	Δ_{mean}^{20}	$\Delta_{mean}^{16.66}$	$\Delta_{min}^{9.63}$	$\Delta_{min}^{8.03}$	$\Delta_{mean}^{13.88}$	$\Delta_{min}^{6.68}$	$\Delta_{min}^{13.89}$	$\Delta_{min}^{11.56}$
New method	1	$S_{max}^{4.64}$									
	2	$S_{max}^{4.64}$	$S_{dr}^{11.56}$								
	3	$S_{max}^{4.64}$	$S_{dr}^{11.56}$	$S_{dr}^{0.75}$							
	4	$S_{max}^{4.64}$	$S_{dr}^{11.56}$	Δ_{mean}^{20}	$\Pi_{b}^{0.25}$						
	5	$S_{max}^{4.64}$	$S_{dr}^{11.56}$	Δ_{mean}^{20}	$\Pi_{b}^{0.25}$	$S_{dr}^{3.22}$					
	6	$S_{max}^{4.64}$	$S_{dr}^{11.56}$	$\Pi_{b}^{0.25}$	$\Pi_{R^2}^{0.75}$	Δ_{max}^{20}	$\Delta_{min}^{8.03}$				
	7	$S_{max}^{4.64}$	$S_{dr}^{11.56}$	Δ_{mean}^{20}	$S_{dr}^{3.22}$	$\Pi_{ang}^{0.25}$	$\Delta_{max}^{0.90}$	$S_{krt}^{0.52}$			
	8	$S_{max}^{4.64}$	$S_{dr}^{11.56}$	$S_{max}^{3.22}$	$\Delta_{slope}^{0.52}$	$\Pi_{ang}^{0.25}$	$\lambda_{L}^{0.30}$	$\Pi_{b}^{0.25}$	Δ_{max}^{20}		
	9	$S_{max}^{4.64}$	$S_{dr}^{11.56}$	$S_{max}^{3.22}$	$\Delta_{slope}^{0.52}$	$\Pi_{R^2}^{0.75}$	Δ_{mean}^{20}	$\Pi_{b}^{0.25}$	Δ_{max}^{20}	$S_{skw}^{5.58}$	
	10	$S_{max}^{4.64}$	$S_{dr}^{11.56}$	$S_{max}^{3.22}$	$S_{krt}^{13.88}$	$S_{krt}^{0.52}$	$\Delta_{min}^{8.03}$	$\Pi_{b}^{0.25}$	$\lambda_{L}^{0.30}$	$S_{dr}^{3.22}$	Δ_{mean}^{20}

Table 6
Classification result of top 10 features selected from different neighborhood definition.

index	D_{cubd}	D_{cube}	D_{cyln}	D_{sphr}	D_{k2d}	D_{k3d}
kappa	0.9391	0.8995	0.9285	0.8630	0.8839	0.8812
FA_{art}	0.9589	0.9363	0.9503	0.9029	0.9147	0.9237
FA_{ntr}	0.9424	0.9101	0.9320	0.8730	0.8895	0.8930
FA_{veg}	0.9298	0.8650	0.9188	0.8433	0.8672	0.8526
FA_{bul}	0.9780	0.9656	0.9754	0.9521	0.9651	0.9546
Accuracy	0.9550	0.9257	0.9472	0.8987	0.9142	0.9122

neighborhood scale and the type of geometric features. In our experiments, we tested the influence of these three factors to classification.

In urban scene, the building, roads or natural ground were planning as regular shapes. Compared with D_{cyln} and D_{sphr}, D_{cubd} and D_{cube} could reflect the directional properties and got better classification than cylinder. On the other hand, the difference in the vertical direction would help to distinguish ground point or non-ground point. D_{cubd} and D_{cyln} performed better than D_{cube} and D_{sphr}. Since the LiDAR points were not evenly distributed, D_{cubd} and D_{k3d} look like a mixture of D_{cyln} and D_{sphr} with different scale and got worse classification result. The results showed that whether the neighborhood definition could reflect the height difference was the most importance factor to select neighborhood definition. The kappa coefficients of classification with all features extracted by different neighborhood definition were D_{cubd} (0.9671), D_{cyln} (0.9659), D_{k2d} (0.9595), D_{cube} (0.9581), D_{cube} (0.9574) and D_{k3d} (0.9505).

The type of features would affect classification a lot. S_{mean}, S_{mod}, $S_{\max}$, $S_{\min}$ and S_{med} were similar as smooth operator and obtained better classification than the other type geometric features. The height-texture based features would reflect the difference of current point from ground points. To ensure the neighborhood of current points contain at least a ground point, a larger neighborhood would be selected. S_{dr}, S_{std}, S_{coev} and Π_{RMSE} would reveal the edge feature of object. In small scale, the types of features got higher value at the edges of objects. With increasing scale, the features become smooth and finally hard to distinguish the edge of object. The remaining geometric features depicted the homogeneous and heterogeneous of the neighborhood points. Small scale was usually selected to calculate these features.

The neighborhood scale could be separated into small scale from 0.5 m to 1.5 m, medium scale from 1.5 m to 6 m and large scale greater than 6 m. Any two features in the same scale category would have higher correlation. With a larger scale, the neighbors of a point would include more distinct structures or objects. On the other hand, with small scale, the features of objects performed similar. Furthermore, the mixtures also occurred at the borders of each scale. These characteristics made it more difficult to describe objects. It is necessary to use multiple scale to improve classification accuracy.

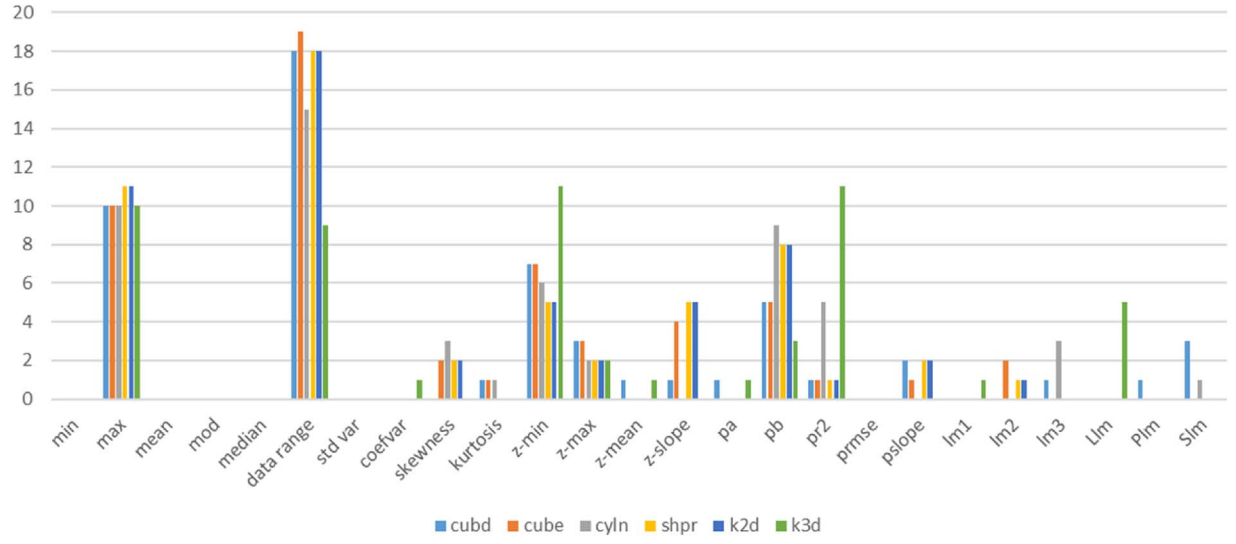

Fig. 9. The times of the feature had been selected.

5.2. The effect of feature selection

Feature selection based on random forest importance or fisher score could get a subset of features with high correlation. In this paper, we first calculated the correlation coefficient of each pair of neighbor definition and then used multidimensional scaling and k-means clustering to select optimal features.

D_{cubd} and D_{cyln} were the two best neighborhood definition for classification performance, followed by D_{cube}, D_{k2d}, $D_{sphr}$and D_{k3d}. The order was consistent with classification with full feature space. The new method would select a subset of features with low correlation coefficient. $S_{\min}$, $S_{\max}$, S_{mean}, S_{mod} and S_{med} had play a role in smoothing height. After selecting features, only $S_{\max}$ or S_{mod} with a medium scale were chosen with other features with low correlation. $\Delta_{\min}$ with a large scale was used frequently and was selected as expected. What's more, $\Delta_{\max}$ with a large scale was also selected many times. S_{dr} was selected most times with different scale. S_{dr} with small scale would help to judge whether the object was smooth or not. S_{dr} with large scale was often combined with $S_{\max}$. A large S_{dr} indicated that there were different objects in the neighbor of points and $S_{\max}$ discriminated ground and off-ground objects. Π_a and Π_b were the parameter to coordinate x and coordinate y of the fitting plane. However, $\Pi_b$with a small scale was selected more times than Π_a. Different performance in Π_a and Π_b would result from the orientation of objects, such as buildings. The other selected features were concentrated in small scale, such as Δ_{slope}, Π_{R^2}, Π_{ang}, λ_L and λ_S. These features would reflect the homogeneity of points in a neighborhood. When using D_{k2d} or D_{k3d}, the nearest 5 five points were selected the most times, while Niemeyer et al. (2011) recommended seven neighbors. Generally speaking, features types owned specific neighborhood scale. It was useful to select features with low correlation and the features performed well when using single features.

6. Conclusions

This paper explored the performance of a series of geometric features with a range of neighborhood scale under different neighborhood definition for urban land cover classification. To select optimal geometric feature subset, we proposed a feature selection method based on correlation coefficient, multidimensional scaling and k-means clustering. According to the clustering results, we separated neighborhood scales into small scales, medium scale and large scales, and selected two stable feature type groups under different neighborhood definitions. It was recommended to select features with low correlation coefficient between each other and with high kappa coefficient of classification with single feature. We suggested to select Δ_{slope}, Π_{R^2}, Π_{ang}, λ_L and λ_S at small scales, $S_{\max}$ or S_{mod} at medium scale and $\Delta_{\min}$ or $\Delta_{\max}$ at large scale. We also compared different neighborhood definition and D_{cubd} and D_{cyln} would be the best choice for classification under urban scenes. This study could be used to select appropriate geometric features with adaptive neighborhood scales of LiDAR data for more accurate land cover classification in urban areas. The feature selection strategy could be transferred to classification with other platform data.

Conflicts of interest

The authors declare no conflict of interest.

Author contributions

This survey has been conducted by the stated authors. The idea was conceived by Weihua Dong, performed and written by Jianhang Lan and Weihua Dong, revised by Shunlin Liang and Wei Yao, and polished by Zhicheng Zhan.

Acknowledgment

The Vaihingen data set was provided by the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF) (Nourzad and Pradhan, 2014).

References

Alexander, C., Tansey, K., Kaduk, J., Holland, D., Tate, N.J., 2010. Backscatter coefficient as an attribute for the classification of full-waveform airborne laser scanning data in urban areas. ISPRS J. Photogramm. Remote Sens. 65, 423–432.
Alexander, C., Tansey, K., Kaduk, J., Holland, D., Tate, N.J., 2011. An approach to classification of airborne laser scanning point cloud data in an urban environment. Int. J. Remote Sens. 32, 9151–9169.
Antonarakis, A., Richards, K.S., Brasington, J., 2008. Object-based land cover classification using airborne lidar. Remote Sens. Environ. 112, 2988–2998.
Axelsson, P., 1999. Processing of laser scanner data—algorithms and applications. ISPRS J. Photogramm. 54, 138–147.
Bartels, M., Wei, H., 2006. Rule-based improvement of maximum likelihood classified lidar data fused with co-registered bands. In: Annu. Conf. Remote Sens. Photogramm. Soc., CD Proceedings. 05–08 Sep.. pp. 1–9.
Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.
Charaniya, A.P., Manduchi, R., Lodha, S.K., 2004. Supervised parametric classification of aerial lidar data. In: IEEE Workshop on Real Time 3D Sensor and Their Use. IEEE : Washington DC. pp. 25–32.
Chehata, N., Guo, L., Mallet, C., 2009. Airborne lidar feature selection for urban classification using random forests. In: Paris, France. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., vol. 39. pp. 207–212.
Cramer, M., 2010. The dgpf-test on digital airborne camera evaluation–overview and test design. Photogramm. Fernerkund. Geoinform. 2010, 73–82.
Dash, M., Liu, H., 1997. Feature selection for classification. Intelligent data analysis 1, 131–156.
Demantké, J., Mallet, C., David, N., Vallet, B., 2011. Dimensionality based scale selection in 3d lidar point clouds. In: ISPRS Workshop on Laser Scanning 2011. Calgary, Canada, 29–31 Aug.. pp. 29–31.
Duda, R.O., Hart, P.E., Stork, D.G., 2012. Pattern Classification. John Wiley & Sons.
Filin, S., 2002. Surface clustering from airborne laser scanning data. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. 34, 119–124.
Gross, H., Thoennessen, U., 2006. Extraction of lines from laser point clouds. Symposium of ISPRS Commission III: Photogrammetric Computer Vision PCV06. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 36, 86–91.
Guan, H., Ji, Z., Zhong, L., Li, J., Ren, Q., 2013. Partially supervised hierarchical classification for urban features from lidar data with aerial imagery. Int. J Remote Sens. 34, 190–210.
Guo, L., Chehata, N., Mallet, C., Boukir, S., 2011. Relevance of airborne lidar and multispectral image data for urban scene classification using random forests. ISPRS J. Photogramm. 66, 56–66.
Hripcsak, G., Rothschild, A.S., 2005. Agreement, the f-measure, and reliability in information retrieval. J Am. Med. Inform. Assn. 12, 296–298.
Im, J., Jensen, J.R., Hodgson, M.E., 2008. Object-based land cover classification using high-posting-density lidar data. GISci. Remote Sens. 45, 209–228.
Jaboyedoff, M., Oppikofer, T., Abellán, A., Derron, M.-H., Loye, A., Metzger, R., Pedrazzini, A., 2012. Use of lidar in landslide investigations: a review. Nat. Hazards 61, 5–28.
Koller, D., Sahami, M., 1996. Toward optimal feature selection.
Liu, C., Frazier, P., Kumar, L., 2007. Comparative assessment of the measures of thematic classification accuracy. Remote Sens. Environ. 107, 606–616.
Maas, H.-G., 1999. The potential of height texture measures for the segmentation of airborne laserscanner data. In: Fourth Int. Airborne Remote Sensing Conference and Exhibition/21st Canadian Symposium on Remote Sensing. Ottawa, Ontario, Canada. pp. 154–161.
Mao, K., 2004. Orthogonal forward selection and backward elimination algorithms for feature subset selection. Syst. Man Cybernet. B: Cybernet. IEEE Trans. 34, 629–634.
Meng, X., Currit, N., Zhao, K., 2010. Ground filtering algorithms for airborne lidar data: a review of critical issues. Remote Sens. 2, 833–860.
Niemeyer, J., Wegner, J., Mallet, C., Rottensteiner, F., Soergel, U., 2011. Conditional Random Fields for Urban Scene Classification with Full Waveform Lidar Data. Springerpp. 233–244.
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. Inform. Sci. 1, 263–268.
Nourzad, S.H.H., Pradhan, A., 2014. Ensemble methods for binary classifications of airborne lidar data. J. Comput. Civil Eng. 28.
Robnik-Šikonja, M., Kononenko, I., 2003. Theoretical and empirical analysis of relieff and rrelieff. Mach. Learn. 53, 23–69.
Rottensteiner, F., 2012. Advanced methods for automated object extraction from lidar in urban areas. Geosci. Remote Sens. Symp. (IGARSS), 2012 IEEE Int. IEEE 5402–5405.
Tang, J., Alelyani, S., Liu, H., 2014. Feature selection for classification: a review. Data Class. Algor. Appl. 37.
Vogtle, E.S., 2001. Automated extraction and reconstruction of buildings in laser scanning data for disaster management. In: Automatic Extr. Manmade Objects from Aerial and Satellite Images III, Balkema (Swets & Zeitlinger). Lisse, The Netherlands, 10–15 Jun..

pp. 309–318.
Wei, Y., Yao, W., Wu, J., Schmitt, M., Stilla, U., 2012. Adaboost-based feature relevance assessment in fusing lidar and image data for classification of trees and vehicles in urban scenes. In: Melbourne, Australia. ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., vol. I-2. pp. 323–328.
Weinmann, M., Jutzi, B., Mallet, C., 2014. Semantic 3d scene interpretation: a framework combining optimal neighborhood size selection with relevant features. In: Zurich, Switzerland. ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., vol. II-3. pp. II–I3.
Yan, W.Y., Shaker, A., El-Ashmawy, N., 2015. Urban land cover classification using airborne lidar data: a review. Remote Sens. Environ. 158, 295–310.
Yanfeng, G., Qingwang, W., Xiuping, J., Benediktsson, J.A., 2015. A novel mkl model of integrating lidar data and msi for urban area classification. IEEE T Geosci. Remote Sens. 53, 5312–5326.
Yao, W., Krzystek, P., Heurich, M., 2012. Tree species classification and estimation of stem volume and dbh based on single tree extraction by exploiting airborne full-waveform lidar data. Remote Sens. Environ. 123, 368–380.
Zhou, W., 2013. An object-based approach for urban land cover classification: integrating lidar height and intensity data. IEEE Geosci. Remote Sens. Lett. 10, 928–931.