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

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 ($\Delta_{\min}$) between the current point and the lowest point is the most commonly used feature, as it roughly measures local variation. However, $\Delta_{\min}$ 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 Thoenessen (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 ($\Delta_{\min}$), z-max($\Delta_{\max}$), 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).