Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Plot-level wood-leaf separation of trees using terrestrial LiDAR data based on a segmentwise geometric feature classification method

Peng Wan (wanpeng@mail.crsri.cn)
Beijing Normal University https://orcid.org/0000-0002-5366-835X
Wuming Zhang (zhangwm25@mail.sysu.edu.cn)
School of Geospatial Engineering and Science, Sun Yat-sen University
Shuangna Jin
Faculty of Geographical Science, Beijing Normal University
Tiejun Wang
Faculty of Geo-Information Scienence and Earth Observation, University of Twente
Shengmei Yang
Changjiang River Scientific Research Institute
Guangjian Yan
Faculty of Geographical Science, Beijing Normal University

Research

Keywords: wood-leaf separation, segmentwise method, terrestrial laser scanning (TLS), connected component segmentation

Posted Date: July 24th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-42032/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Plot-level wood-leaf separation of trees using terrestrial LiDAR data based on a segmentwise geometric feature classification method

Peng Wan [a, b], Wuming Zhang [c*], Shuangna Jin [b], Tiejun Wang [d], Shengmei Yang [a], Guangjian Yan [b]

[a] Changjiang River Scientific Research Institute (CRSRI), Wuhan 430010, China

[b] Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China

[c] School of Geospatial Engineering and Science, Sun Yat-sen University, Zhuhai 519082, China

[d] Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

[*] Corresponding author.

Email address: zhangwm25@mail.sysu.edu.cn (W. Zhang)

Abstract

Background: Detailed and accurate information about tree crown structures is crucial for researching tree physiology. Terrestrial laser scanning (TLS) is a promising technique for retrieving crown structure parameters but still faces bottlenecks in terms of data processing such as wood-leaf separation. Currently, most wood-leaf separation methods use pointwise and supervised classification strategies at the forest plot level, which is very time consuming for large-volume TLS datasets.

Methods: In this study, we proposed a novel classification method to separate wood and leaf points based on connected component segmentation and segmentwise classification by geometric features. We tested the proposed method on both needleleaf and broadleaf forest plots and compared the results to those from a widely used pointwise classification method, CANUPO.

Results: The results showed that the proposed segmentwise classification method is superior to CANUPO in terms of both classification accuracy and time efficiency, which were improved by 1.5 times and over 10 times, respectively.

Conclusions: This study indicated that the segmentwise classification strategy is applicable to dense TLS data and can enhance the accuracy and efficiency of wood-leaf separation, which can in turn facilitate the retrieval of crown structure parameters.

Keywords: wood-leaf separation, segmentwise method, terrestrial laser scanning (TLS), connected component segmentation

1. Introduction

Tree architecture, e.g., the stem form, branching pattern, and spatial distribution of leaves, directly influences tree photosynthesis and evapotranspiration and ultimately influences forest carbon and water storage (Lau et al., 2018). Quantifying the variation in tree architecture across species is important in understanding how tree architecture is related to the physiological function of trees (Disney et al., 2018). Additionally, tree architectural characteristics can be used to explore the allometric relationship among, for example, tree height, diameter at breast height (DBH), and distribution of biomass within trees (Abd Rahman et al., 2017). However, it is difficult and time consuming to accurately quantify tree architecture with manual measurement (Quammen, 2012).

Terrestrial laser scanning (TLS) is an efficient technology for collecting dense, highly detailed three-dimensional (3D) point clouds of trees. The point clouds can be used for the reconstruction of quantitative 3D tree models and the quantitative analysis of tree architecture features, such as branching structures (Pyörälä et al., 2018) and stem curves (Liang et al., 2013). In the data processing procedures, an important prerequisite step in reconstructing the quantitative tree models from TLS data is wood-leaf separation, since the wood structural reconstruction is based on the wood points (Raumonen et al., 2013). Moreover, the leaf points can be used for the estimation of the canopy gap fraction and the leaf area index, which are very important indicators in vegetation science (Chen et al., 2018). However, separating the wood and leaf points of trees in TLS data is

challenging due to the complicated radiometric and geometric characteristics of the LiDAR points from trees, the occlusions within canopy, the roughness of the wood surface, and the influence of site conditions, such as the terrain slope, stem density and the presence of understory vegetation. Over the past decade, various kinds of methods have been proposed to solve this problem, and it has become a research focus in the recent years. In general, the types of features used in wood-leaf separation include radiometric features, waveform features, and geometric features.

Radiometric features refer to the intensity and reflectivity of laser returns at specific wavelengths. Côté et al. (2009) differentiated wood and leaf points before the reconstruction of branch structure based on the intensity of laser returns. The threshold values of the point intensity to discriminate wood and leaves were chosen manually. Then, manual classification was performed to further remove unwanted parts from each component. However, the accuracy of the point classification was not reported in Côté's study. It is known that the intensity of laser returns is related not only to the spectral properties of the target but also to the incidence angle, the traveling distance of the laser beam (Kukko et al., 2008; Pfeifer et al., 2007), and the roughness of the reflecting surface (Pesci and Teza, 2008). Therefore, the radiometric calibration of point intensity is also a challenging problem. In recent years, some advanced TLS equipment has been developed and used. Yao et al. (2011) and Yang et al. (2013) used the relative widths of returned waveforms to separate wood and leaf points. Danson et al. (2014) tested a dual-wavelength full-waveform TLS, and Li et al. (2013) explored the feasibility of multiwavelength LiDAR for wood-leaf separation. However, there were no quantitative assessments or comparisons of the classification accuracy in these studies. In addition, the use of dual-wavelength and full-waveform information is limited by the availability of equipment.

Therefore, most recently developed methods depend on geometric features, such as eigenvalue-based features and density-based features. Some methods have been proposed for individual trees, and others have been developed for forest plots. For example, Tao et al. (2015) developed a geometric

method based on circle detection in point slices and the shortest-path algorithm and tested it using two broad-leaved trees and a virtual tree produced by an L-system. Since the shortest-path algorithm is used to extract the skeletal structures from the point cloud of a single tree, Tao's method is applicable only to individual trees. Similarly, the shortest-path algorithm has also been used for wood-leaf separation in the methods proposed by Xu et al. (2007) and Livny et al. (2010). By contrast, Ferrara et al. (2018) proposed an approach using the density-based clustering algorithm DBSCAN. Ferrara's method was also tested on individual trees; however, it is potentially applicable at the plot level and thus requires plot-level tests. Raumonen et al. (2013) proposed a segment-based approach for wood-leaf separation in individual trees. The point cloud was first segmented into small nonoverlapping cover sets, and then the cover sets were classified as belonging to leaves and wood components based on their geometric features, such as the tangent and normal vectors. However, the size of the cover sets was obtained by the pointwise computation. Wang et al. (2017) compared four machine learning classifiers for the wood-leaf separation of individual trees based on both geometric and radiometric features. The results indicated that the machine learning classifiers can obtain high accuracies for wood-leaf separation on individual trees.

For plot-level classification, methods based on machine learning algorithms are more frequently used. Lalonde et al. (2006) introduced a method that characterizes point clouds in terms of three salient features, namely, "scatter", "linear" and "surface". The features of the target objects are learned by fitting a Gaussian mixed model to manually labeled training data. Bayesian classification is then used to label the entire point cloud. Ma et al. (2016) improved Lalonde's method by adding two additional filters based on geometric information, and the overall classification accuracy reportedly improved from 84.28% to 97.8%. Brodu and Lague (2012) extended the local geometric features to multiple scales. The classifiers, which include the least absolute deviation and support vector machine (SVM), were used for sample and classification training. Zhu et al. (2018) combined various local geometric features and radiometric features to separate foliar and woody materials by

using a random forest classifier. The reported overall classification accuracies in their study ranged from 80% to 90%. In addition, the deep learning method has also been introduced for the wood-leaf separation problem (Xi et al., 2018). Table 1 shows a summary of the abovementioned methods.

Table 1. List of published methods for wood-leaf separation using TLS data

Authors (Publish years)	Input points	Method categories	Reported mean accuracy
Xi et al. (2018)	Forest plots	# Deep 3D FCN* # Point-wise classification	94%
Ferrara et al. (2018)	Individual trees	# Unsupervised # Geometric feature # Point-wise classification	96%
Zhu et al. (2018)	Forest plots	# Random Forest # Radiometric & Geometric features # Point-wise classification	84.40%
Wang et al. (2017)	Individual trees	# Machine learning methods # Radiometric & Geometric features # Point-wise classification	Support vector machine: 93.5% Naïve Bayes: 89% Random forest: 96% Gauss mixture model: 91%
Ma et al., (2016)	Forest plots	# Gauss Mixture Model plus six additional filters # Geometric features # Point-wise classification	94%
Tao et al. (2015)	Individual trees	# Short-path algorithm # Geometric features # Point-wise classification	84%
Li et al., (2013); Danson et al. (2014)	Forest plots	# Multi-wavelength intensity # Point-wise classification	No reports
Côté et al. (2009); Wu et al., 2013	Individual trees	# Intensity # Point-wise classification	No reports
Yao et al. (2011); Zhao et al., 2011; Yang et al. (2013)	Forest plots	# Return waveforms # Point-wise classification	No reports
Brodu and Lague, (2012)	Non-forest natural scenes	# Support Vector Machine # Multi-scales geometric features # Point-wise classification	98% Non-forest natural scenes
Raumonen et al. (2013)	Individual trees	# Unsupervised # Geometric features # Segment-wise classification	No reports

Xu et al. (2007); Livny et al., (2010)	Individual trees	# Short-path algorithm # Geometric features # Point-wise classification	No reports
Lalonde et al. (2006)	Non-forest natural scenes	# Gauss Mixture Model # Geometric features # Point-wise classification	88% for Non-forest natural scenes

Note：FCN stands for **F**ully **C**onvolutional **N**etworks.

As seen from the existing algorithms listed above, most algorithms have adopted local geometric features for the wood-leaf separation of TLS data. This occurs mainly for two reasons. 1) The spatial distribution characteristics of wood points and leaf points are very different. The distribution of wood points is regular and continuous, and the density is relatively large. The leaf points are scattered due to canopy gaps, and the overall point cloud density is relatively smaller. 2) The 3D coordinates are the fundamental information provided in LiDAR point clouds. Algorithms based on geometric features have broader application scenarios and work well with the data provided by LiDAR equipment. Although some methods based on geometric features and machine learning have achieved good wood-leaf separation accuracy, the main disadvantages of these methods are their high computing demand and slow computing speed (Liang et al., 2012) in plot-level classification tasks. There are three main reasons for these disadvantages: 1) the geometric features are calculated point by point, which is time consuming due to the large volume of TLS datasets; 2) the geometric features vary with the radius of the neighborhood, so the latest algorithms usually extract two-scale or even multiscale geometric features by using different neighborhood sizes to achieve better classification results (Brodu and Lague, 2012); (Xia et al., 2015), which greatly increases the computational burden; and 3) a supervised classification method contributes to the improvement of classification accuracy but requires high-quality training samples (Wang et al., 2017).

Here, we propose a segmentwise classification method for the rapid wood-leaf separation of standing trees from TLS data at the plot level. Compared with the pointwise method, the segment-based method first divides a point cloud into segments and then classifies the segments based on the

features of each segment. This strategy can reduce the computational burden and the uncertainties of point cloud classification.

2. Study areas and data

2.1 Study areas

We tested the proposed method on the TLS datasets from three forest plots of different stem density, terrain and tree species, i.e., a white birch (*Betula papyrifera*) plot, a Dahurian larch (*Larix gmelinii*) plot and a Chinese scholar tree (*Styphnolobium japonicum*) plot. The white birch (WB) plot and Dahurian larch (DL) plot are located in the Saihanba National Forest Park, Hebei Province, China (42.40°N, 117.31°E), and the Chinese scholar tree (CST) plot is located in a pure forest southwest of Beijing, China (39.76°N, 116.22°E). The WB plot is a 15 m × 30 m that includes 21 white birches on a slope of approximately 22 degrees. The DL plot is 15 m × 15 m and includes 15 trees on flat terrain. The CST plot is 30 m × 30 m and includes 37 trees on flat terrain. Figure 1 shows the topography and the forest condition of each test plot. The mean DBH and tree height, stem density and the density of ground vegetation of each sample plot are summarized in Table 2.

Table 2. The statistics of the sample forest plots.

Plot ID	Stem density (stems/ha)	Mean DBH (cm)	Mean tree height (m)	Ground vegetation
WB	467	19.34	13.38	Dense
DL	667	28.33	14.23	Moderate
CST	411	24.2	10.66	Sparse



A B C

Figure 1. The topography and the forest condition of each test plot. A. The white birch plot. B. The Dahurian larch plot. C. The Chinese scholar tree plot.

2.2 TLS data

The TLS datasets were collected in July 2018 by using a Riegl VZ-1000 (Riegl GmbH, Horn, Austria) terrestrial laser scanner in multiscan mode. The scan angle resolution is 0.03°, and the vertical and horizontal scanning ranges are 30°-130° and 0°-360°, respectively. In each sample plot, five to seven scans were implemented at the center and the periphery of the plots. The scans were registered using the RiSCAN Pro software package (Riegl GmbH, Horn, Austria). The registered point clouds were clipped to exclude trees outside each rectangular plot (Figure 2).

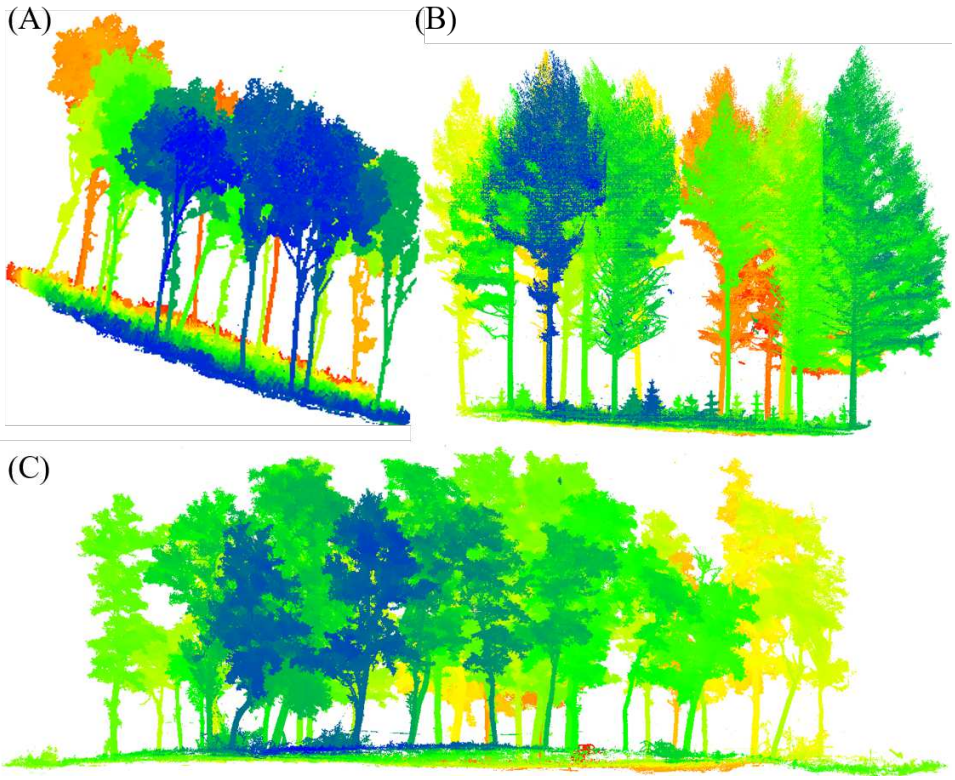


Figure 2. Three field collected TLS datasets. (A) White birch plot. (B) Dahurian larch plot. (C) Chinese scholar tree plot.

3. Methods

3.1. Wood-leaf separation

3.1.1 Overview of the proposed method

The proposed method adopted a segmentwise classification strategy by using a connected component segmentation algorithm. The wood and leaf points were separated based on the geometric features of each segment formed by connected points. It is not possible to segment the point cloud into small segments that consist of a single classification target, i.e., the wood and leaf point. This is mainly because of the very high density of the TLS point clouds and the tight connections between wood and leaf points. Therefore, we introduced a point cloud decomposing step to divide the point cloud into three parts in order to reduce the point density and break the connections between wood and leaf points. Figure 3 shows an overall workflow of the proposed method.

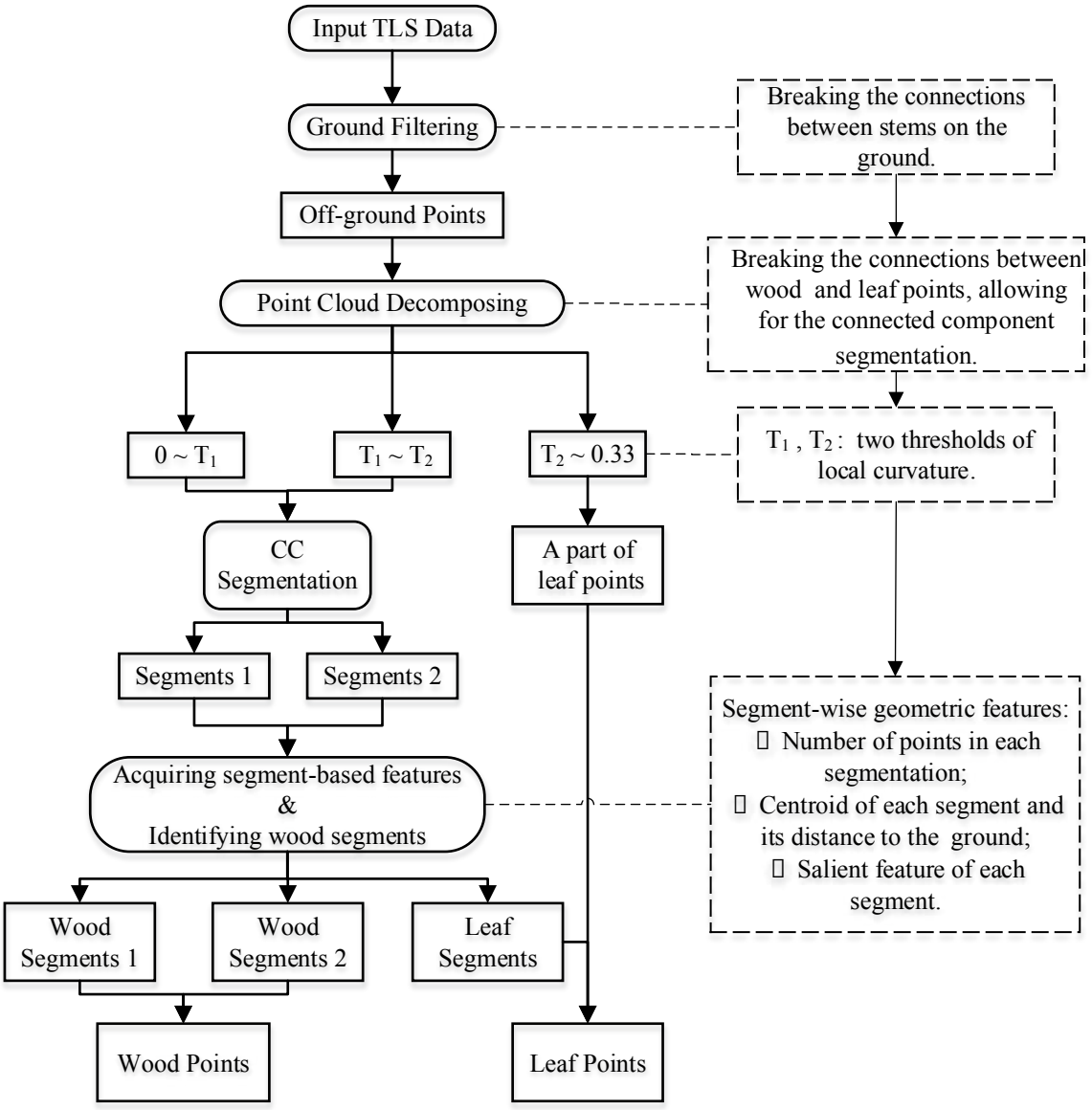


Figure 3. The overall workflow of the proposed method. The left column shows the main steps of the method, and the right column shows the annotation of each main step.

The following subsections detailed methods used in the main steps shown in the flowchart.

3.1.2 Ground filtering

Ground filtering was a prerequisite step in the proposed method for reducing the data volume and breaking the ground connection between stems. In this study, we adopted an open source ground filtering algorithm, Cloth Simulation Filtering (CSF), for its fast processing speed and reliable results.

CSF is based on the principle of cloth simulation (Zhang et al., 2016). By iteratively dropping a simulated cloth over an inverted point cloud, the simulated cloth sticks to the ground points and drapes over the object points due to a certain degree of cloth hardness. The simulated cloth is a set of regular particles with mass. The interconnection between the particles is made of "virtual springs". In each iteration, the particles drop down under the force of gravity and then spring back under the spring force. Finally, the simulated cloth settles to form a digital terrain model of the point cloud.

We implemented the CSF with the following parameters. The cloth resolution and distance threshold (classification threshold) of the CSF were both set to 0.1 m, and the maximum iterations were set to 50 for all sample datasets. The off-ground point clouds from the CSF were used as the inputs in the next step.

3.1.3 Point cloud decomposing

The method of point cloud decomposing was based on the local curvature feature of each point. Local curvature features can be used to roughly separate tree components (Zhang et al., 2019). The points on stems and boughs usually have the smallest local curvatures due to their large diameters and regular point distribution. The branches and twigs, which have smaller diameters and are surrounded by leaves, usually have moderate local curvatures. The leaf points usually have the largest local curvatures because of their scattered distribution pattern and because of canopy gaps. In this study, the surface variation (*SV*) was adopted as a curvature feature to divide the point cloud into three parts. For a point $p_1 = (x_1, y_1, z_1)$ and the points around it within a neighborhood $p_i = (x_i, y_i, z_i)_{i=1}^{n}$, the calculation of the *SV* was based on the eigenvalues of principal component analysis, which was defined as

$$SV(p_1) = \frac{\lambda_2}{\lambda_0 + \lambda_1 + \lambda_2} \quad (1)$$

where $\lambda_k (k = 0,1,2)$ are the eigenvalues of a covariance matrix C sorted in descending order as $\lambda_0 \geq \lambda_1 \geq \lambda_2$. The covariance matrix C is defined as

$$C_{3\times3} = \frac{1}{n}\sum_{i=1}^{n}(p_i - \bar{p})(p_i - \bar{p})^T \quad (2)$$

where $\bar{p} = (1/n)\sum_{i=1}^{n}(p_{xi}, p_{yi}, p_{zi})$ is the centroid of P.

The *SV* has a limited range from 0 to 1/3 for any point. To divide the point cloud into three parts, two thresholds (T_1, T_2) of the *SV* were chosen according to experiments in previous studies (Zhang et al., 2019), i.e., T_1 was set to 0.1, and T_2 was set to 0.2. The same set of thresholds was adopted for all sample plots. For example, Figure 4 shows the three parts of the divided point cloud of an individual tree. The first part, in which the *SV* ranged from 0 to 0.1, comprised almost all the stem points, most of the branch points and nearly half of the leaf points. The second part, in which the *SV* ranged from 0.1 to 0.2, was mainly composed of leaf points and a fraction of branch points. The third part, in which the *SV* ranged from 0.2 to 1/3, was composed of leaf points. The first and second parts that contained wood points were used in the next step for segmentation.

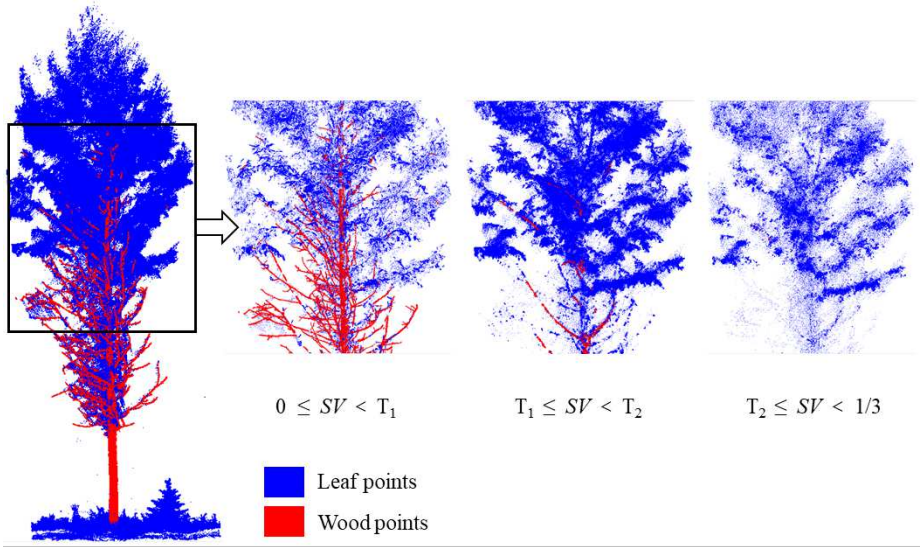


Figure 4. An example of the decomposed point cloud. T_1 and T_2 refer to the thresholds of the surface variation (*SV*). In this example, T_1 equals to 0.1, T_2 equals to 0.2.

3.1.4 Connected component segmentation

We applied the connected component segmentation algorithm to segment the point cloud into a large number of segments in this study. Connected component segmentation is a simple and fast algorithm based on the proximity of points. As shown in Figure 5, the point clouds were first voxelized using 3D grids by building an octree. Then, the 3D grids that contained at least one point were assigned a value of 1. The vacant 3D grids were assigned a value of 0. Finally, the points in adjacent grids with a value of 1 were merged into the same segment. The vacant grids with a value of 0 became the gaps between the segments. The size of the 3D grids determined how many segments were produced and how small the segments were. Smaller 3D grids tend to produce more and smaller segments, which are more likely to contain a single class of tree components; however, these segments will increase the computational burden. Therefore, we set the grid size to a compromise value (0.01 m) in this study. In addition, we set a lower limit on the number of points within the segments at 100 to reduce the number of segments and filter out noise points.

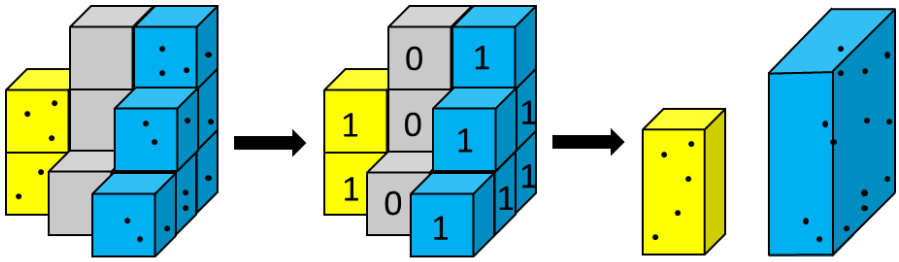


Figure 5. Schematic diagram of the connected component segmentation. Here, a point cloud was segmented into two segments.

3.1.5 Geometric features

We introduced two geometric features to identify the wood segments, i.e., the salient features and the distance between the centroid of each segment and the ground. The salient features were the main features used for the wood-leaf separation, and the distance feature was used to remove the ground vegetation.

3.1.5.1 Salient features

The salient features describe the dominant geometry and distribution pattern of a set of points as linear, planar and scattered. The salient features have been used in most wood-leaf separation algorithms for TLS data. However, their calculation in existing algorithms is pointwise and based on neighboring points. This has caused problems with the large amount of computation required and inconsistency in the salient features of a point when using different neighborhood sizes. Therefore, we applied a segmentwise and neighborhood-based computing method to the point cloud segments from the previous step. This method can greatly reduce the computation load, and the salient features of the points within a segment will be consistent and stable. For the points $p_i = (x_i, y_i, z_i)_{i=1}^{n}$ in each segment, we first calculated their covariance matrix C by using Formula 2. Then, the eigenvalues $\lambda_k (k = 0,1,2)$ of C were obtained by singular value decomposition and sorted in descending order as $\lambda_0 \geq \lambda_1 \geq \lambda_2$. Based on the eigenvalues, the salient features of a segment can be defined as (Demantke et al., 2011)

$$L = \frac{\sqrt{\lambda_0} - \sqrt{\lambda_1}}{\sqrt{\lambda_0}};\ P = \frac{\sqrt{\lambda_1} - \sqrt{\lambda_2}}{\sqrt{\lambda_0}};\ S = \frac{\sqrt{\lambda_2}}{\sqrt{\lambda_0}} \tag{3}$$

where L, P, and S refer to the linear feature, planar feature and scattered feature, respectively, and $L + P + S = 1$.

Figure 6 shows an example of connected component segmentation and the salient features of the segments. It can be found that the stem segment and branch segment were all dominated by the linear feature, which has significantly higher values than the other two features. In contrast, the leaf segment has no dominant feature. Therefore, the segments dominated by the linear feature can be labeled as wood segments. In this study, we introduced an index, i.e., the **S**ignificance **o**f **D**ifference (*SoD*), to evaluate the significance of the linear/planar/scattered features among the three salient features. Specifically, the *SoD* of the linear feature was defined as

$$SoD(L) = L + (1 - L) \times [L - \max(P, S)] \quad (4)$$

where L, P, and S refer to the **L**inear feature, **P**lanar feature and **S**cattered feature, respectively.

The $SoD(L)$ is an index ranging from -1 to 1. The larger the value of the $SoD(L)$, the more likely it is that a segment is dominated by a linear feature. In this study, segments with a $SoD(L)$ greater than 0.7 were considered wood segments. The setting of 0.7 was a conservative threshold for retaining the wood segments as much as possible.

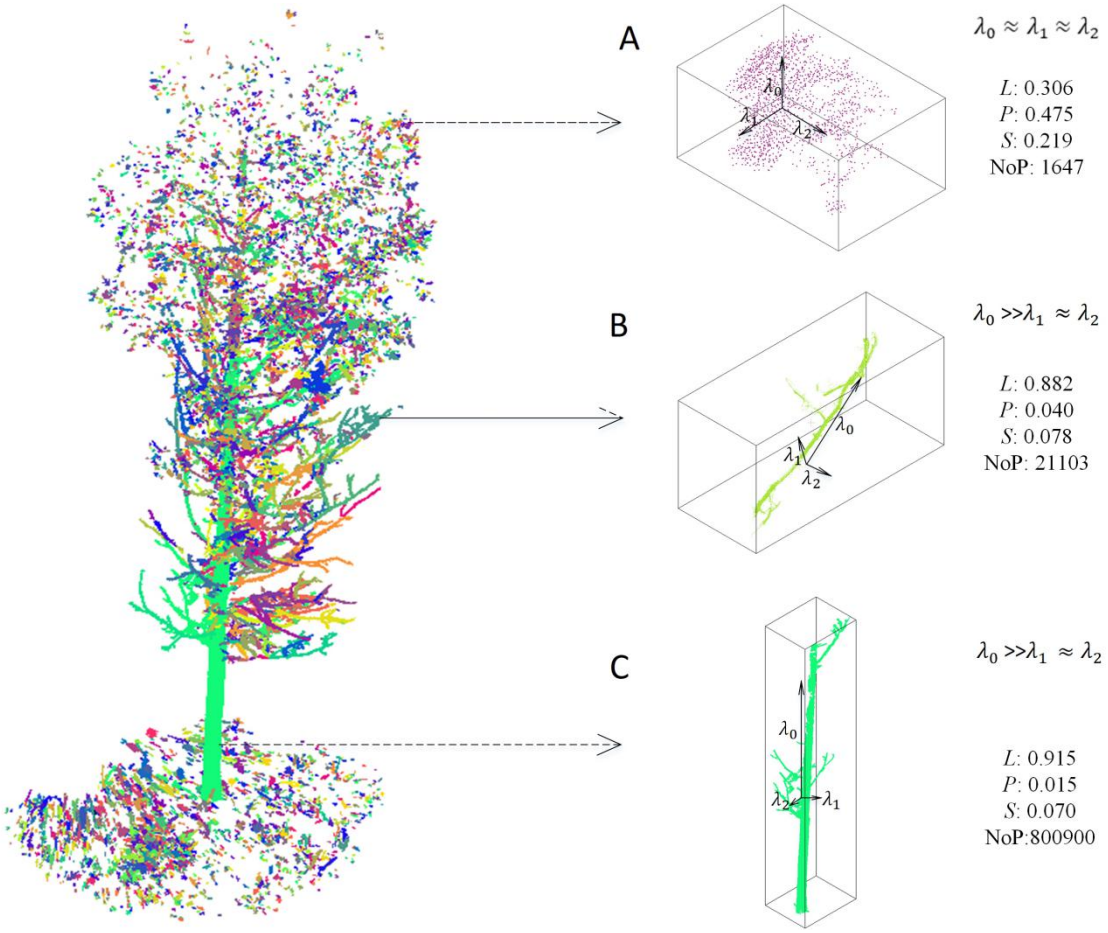


Figure 6. An example of the connected component segmentation and the salient features of the segments. The different colors were drawn for the different segments. L, P, S refer to the attribute value of the linear feature, planar feature, scattered feature and the number of points of each segment. $\lambda_k (k = 0,1,2)$ are the eigenvalues of each segment.

3.1.5.2 The distance between the centroid of each segment and the ground

Ground vegetation, such as grass and small shrubs, is hard to remove from a data set based on salient features. This is because some ground vegetation also has stems that may be wrongly recognized as tree branches. Although ground vegetation can be entirely removed by setting a threshold of distance to the DTM, the true stem points may also be removed in this procedure. For pointwise wood-leaf separation algorithms, it is difficult to solve this problem. However, for the segmentwise method, we can easily recognize segments as ground vegetation based on the centroid of the segments. The centroid of the segments of trees would have a larger distance to the ground surface. For example, points near the ground surface would be classified as wood points if they belong to a segment that has a higher centroid. The distance to the ground surface is an input threshold of the proposed method and was set to 1 m in this study. The centroid of the segments can be calculated by using the formulas shown in section 2.3.3.

3.1.6 Merging the wood segments

Following the procedures shown in Figure 2, we identified the wood segments by their geometric features. The input parameters have been introduced in the sections above. As mentioned in section 3.1.3, the point cloud was decomposed into three parts, and the third part was composed of leaf points, which were directly removed. Therefore, we applied the proposed wood-leaf separation algorithm separately to the first and second parts of the point cloud. After connected component segmentation, the segments identified in the first and second parts were in separate segment groups, i.e., segments 1 and segments 2. Then, we merged the extracted wood segments from the two groups of segments.

3.2. Performance evaluation

3.2.1 Classification accuracy and processing time

The accuracy of the wood-leaf separation method was assessed by the percentage of error-classifying points against the referenced classification results. The references were acquired by visual inspection

for each tree in each plot. Three types of error (type I error, type II error and total error) and the kappa coefficient were quantified. The type I error (T. I), which is also known as the omission error, was defined as the number of wood points that were wrongly classified as leaf points divided by the number of referenced wood points. The type II error (T. II), which is also known as commission error, was the number of leaf points that were wrongly classified as wood points divided by the number of referenced leaf points. The total error (T.E) was the number of wrongly classified points divided by the total number of points. The kappa coefficient was calculated from the statistics of the wrongly classified points. The calculation of the accuracy indexes is shown in Table 3.

Table 3. The calculation of the accuracy evaluation indexes.

Reference	**Result**			**Error（%）**
	Wood	**Leaf**		
Wood	A	b	$f=(a+b)/n$	**Type I error**：$b/(a+b)$
Leaf	C	d	$g=(c+d)/n$	**Type II error**：$c/(c+d)$
	$h=(a+c)/n$	$a_i=(b+d)/n$		
	$n=(a+b+c+d)$			**Total error**：$(b+c)/n$
	$pra=(a+d)/n;\ pre=f*h+g*a_i$			**Kappa**：$(pra-pre)/(1-pre)$

Note: a to d refers to the number of points.

In addition, the processing times were recorded to assess the time efficiency of the wood-leaf separation algorithms. All tests were conducted on a typical desktop computer with an Intel Core i7-6700 CPU (3.40 GHz) and 16 GB of RAM.

3.2.2 Method comparison

To compare the difference in performance between the proposed segment-based method and the point-based method, we processed the sample datasets using a point-based method proposed by Brodu and Lague (Brodu and Lague, 2012), CANUPO, for the classification of natural scene point clouds. The reason for adopting Brodu's method for comparison is that 1) it is a point-based method

using only geometric features that has reported good classification accuracy (98%) in the literature; 2) it adopts multiscale geometric features to improve the results, which were more robust in theory than those from algorithms using single-scale and double-scale geometric features; 3) it is based on machine learning classification algorithms and thus is representative of many supervised classification methods; and 4) it has free and open source codes that are easy to access. Therefore, CANUPO was a representative comparison algorithm for the performance evaluation of the method proposed in this study.

4. Results and Discussion

4.1 Classification accuracy

Table 4 shows the comparison of the accuracy of the two methods in the three sample plots. The classification accuracy of the proposed segmentwise method was higher than that of the pointwise method in all sample plots. The total errors of the proposed method in the three sample plots were 5.39%, 4.74%, and 5.93%, respectively, while they were 8.89%, 9.32%, and 10.05% in the pointwise method. On average, the errors of the pointwise method were 1.76 times those of the proposed method. The kappa coefficients showed a similar trend to that in the total errors. In other words, the proposed segmentwise method is superior to the CANUPO pointwise method based on our analysis.

Table 4. The comparison of time spent between the two methods.

		T. I (%)	T. II (%)	T. E (%)	Kappa (%)
WB	point-wise	14.77	8.08	8.89	64.93
	segment-wise	10.02	4.75	5.39	77.11
DL	point-wise	16.07	7.54	9.32	73.02
	segment-wise	8.85	3.66	4.74	85.90
CST	point-wise	18.41	8.99	10.50	65.02
	segment-wise	21.46	2.98	5.93	77.38

Note: WB, DL and CST represent the plots white birch, Dahurian larch and the Chinese scholar tree. T.I, T.II and T.E stand for the type I error, type II error and total error, respectively.

Figure 7 shows the results of the two methods in the three different plots. It is easy to see that there were many wrongly classified leaves points in the results of the pointwise method. Fewer leaf points were wrongly classified as wood points in the results of the proposed method compared with the CANUPO results. There were some differences between the two methods in the spatial distribution of the classification errors. The residual leaf points were distributed mainly around the branches in the results of the proposed method; however, the residual leaf points were uniformly distributed in CANUPO results. The spatial distribution of the omission and commission is of utmost importance from the standpoint of using the results of this study to extrapolate models of plant physiology. If you were to model the light transmission of these canopies using the segmentwise method, and extrapolating that to plant photosynthetic resource allocation, it would provide a much different result than by using the pointwise method. However, the main topic of this study is how to separate the wood and leaf components of trees, so we did not take the subsequent application into consideration. Actually, the automatic methods for wood-leaf separation cannot be perfect so far, so the further processing such as visual inspection is needed if the leaf distribution should be modeled accurately.

What should be note is that ground points have been removed in the ground filtering step, so the bule points under canopies showed in Figure 7 stand for the points on ground vegetation, such as glass and shrub. It is clear that the CANUPO method cannot separate the ground vegetation points from wood points. In contrast, the proposed method can filter out the ground vegetation points according to the distance between the centroid of the segment and the ground, so there were seldom residual ground vegetation points in the wood class. In conclusion, the proposed method achieved better classification results in this study.

CANUPO method showed around 2% error in literature (Brodu and Lague, 2012). However, it reached 10% error and even 20% error in our tests. It is mainly because the tested datasets in Brodu and Lague's study were collect in a small non-forest natural scene which includes rocks and low vegetation. The separability between rocks and vegetation points is better than wood and leaf points.

In this study, we compared our method against CANUPO method rather than other methods which also reported good classification accuracies, such as the methods proposed by Lalonde et al. (2006) and Ma et al. (2016). It is because the classification accuracies of wood-leaf separation are highly dependent on the characteristics of test data, and we cannot directly evaluate the methods only by literature reports. However, we believe that the CANUPO method is more representative in principle as it used multiscale geometric features and machine learning classifiers. Also, the CANUPO method is easier to access because of its open source codes. That is why a comprehensive comparison between different wood-leaf separation methods based on benchmark datasets is imperative, which should be done by cooperation.

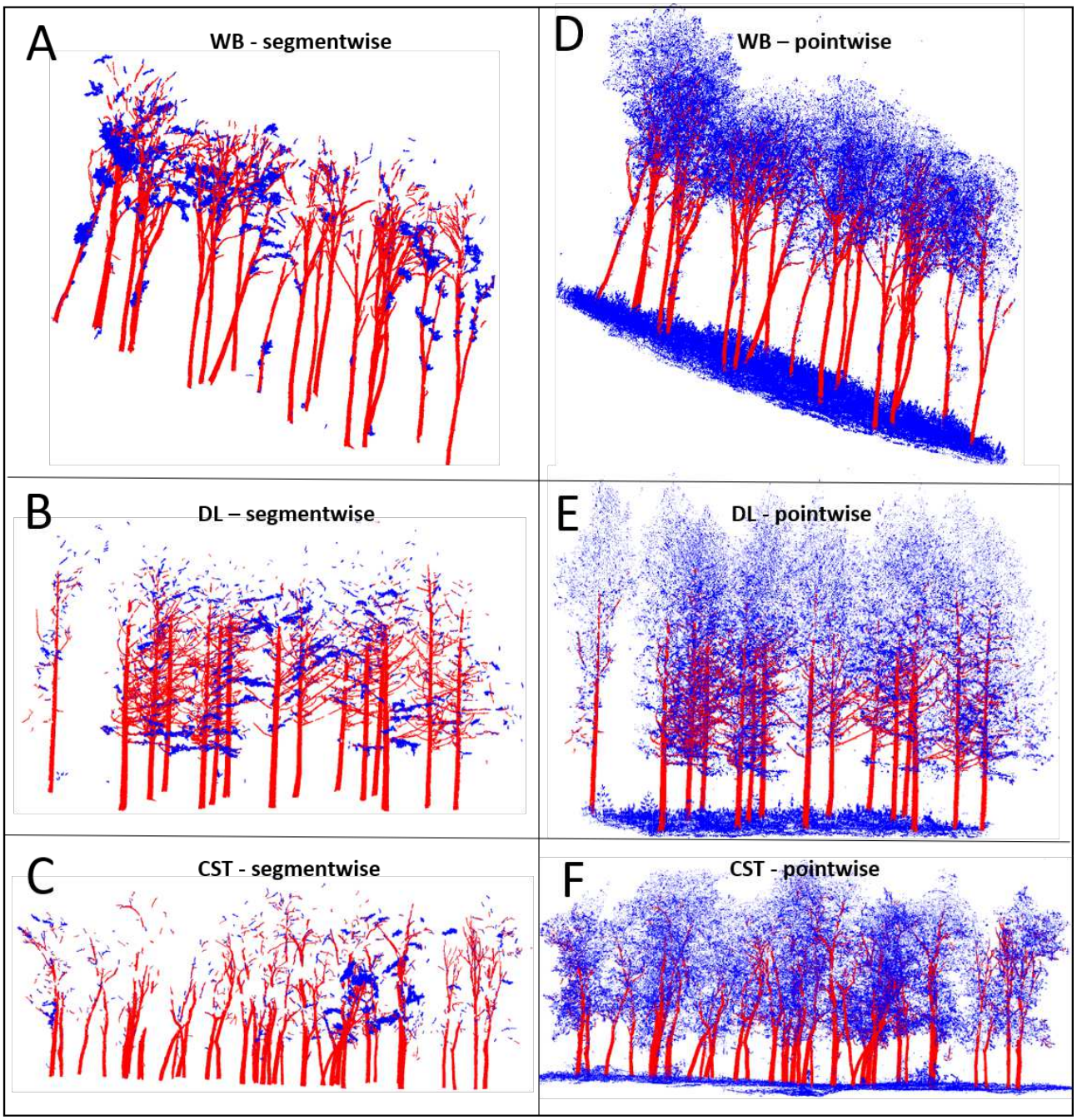


Fig 7. Visualization results of the classification from the two methods. WB, DL and CST represent the plots white birch, Dahurian larch and the Chinese scholar tree.

4.2 Time efficiency

In this study, both methods were run on the same desktop computer. The processing time of the two methods is shown in Table 5. This time does not include the time spent on data reading and writing or parameter tuning or the time for sample selection and training in the pointwise method. Table 5

shows a substantial difference in processing time between the two methods. The proposed method required only 2.57 to 14.41 minutes, while the pointwise method required 28.67 to 235.67 minutes. The time efficiency of the proposed method was approximately 11 times faster than that of the pointwise method in the CST plot, 16 times faster in the DL plot, and 41 times faster in the WB plot. These results fully demonstrated the significant advantage of the proposed method in terms of time efficiency.

Table 5. The comparison of time spent between the two methods.

Species	Methods	Time cost (min)
WB	point-wise	146.22
	segment-wise	3.5
DL	point-wise	235.67
	segment-wise	14.41
CST	point-wise	28.67
	segment-wise	2.57

5 Conclusion

Wood-leaf separation is an important data processing step in the application of TLS to forest scenes. It is required for extracting tree crown parameter data and is also a challenging hot topic that has attracted the attention of many researchers. Because many existing methods require a large amount of calculation, we proposed a novel segmentwise classification method for wood-leaf separation that is fast, accurate and simple. The proposed method requires less than 15 minutes to achieve good classification accuracy. Compared with the CANUPO pointwise classification method, in this study, the classification accuracy was improved by more than 1.5 times, and the processing time was decreased by 11 to 41 times. In addition, the proposed method was robust on sample plots of different species and forest conditions. This study verified the feasibility of the segmentwise classification strategy for the classification of dense terrestrial laser scanning point clouds.

Acknowledgments

Not applicable.

Funding

This work was supported by the National Natural Science Foundation of China (grant Nos. 41671414, 41971380) and the National Key R&D Program of China (2017YFC1502500). This work was also supported by the Guangxi Natural Science Fund for Innovation Research Team (2019GXNSFGA245001).

Availability of data and materials

Not applicable.

Authors' contributions

Peng Wan designed the experiments, processed the data, and analysed the results; Peng Wan, Wuming Zhang, Shuangna Jin and Tiejun Wang coordinated the manuscript preparation; All authors contributed to the manuscript writing and editing.

Ethics approval and consent to participate

The subject has no ethic risk.

Consent for publication

Not applicable.

Consent for publication

No competing interests.

6 References

Abd Rahman, M., Abu Bakar, M., Razak, K., Rasib, A., Kanniah, K., Wan Kadir, W., Omar, H., Faidi, A., Kassim, A., Abd Latif, Z., 2017. Non-destructive, laser-based individual tree aboveground biomass estimation in a tropical rainforest. Forests 8, 86.

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

Chen, Y., Zhang, W., Hu, R., Qi, J., Shao, J., Li, D., Wan, P., Qiao, C., Shen, A., Yan, G., 2018. Estimation of forest leaf area index using terrestrial laser scanning data and path length distribution model in open-canopy forests. Agric. For. Meteorol. 263, 323–333. https://doi.org/https://doi.org/10.1016/j.agrformet.2018.09.006

Côté, J.-F., Widlowski, J.-L., Fournier, R.A., Verstraete, M.M., 2009. The structural and radiative consistency of three-dimensional tree reconstructions from terrestrial lidar. Remote Sens. Environ. 113, 1067–1081.

Danson, F.M., Gaulton, R., Armitage, R.P., Disney, M., Gunawan, O., Lewis, P., Pearson, G., Ramirez, A.F., 2014. Developing a dual-wavelength full-waveform terrestrial laser scanner to characterize forest canopy structure. Agric. For. Meteorol. 198, 7–14.

Demantke, J., Mallet, C., David, N., Vallet, B., 2011. Dimensionality based scale selection in 3D lidar point clouds. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci 38, W12.

Disney, M.I., Boni Vicari, M., Burt, A., Calders, K., Lewis, S.L., Raumonen, P., Wilkes, P., 2018. Weighing trees with lasers: advances, challenges and opportunities. Interface Focus 8, 20170048.

Ferrara, R., Virdis, S.G.P., Ventura, A., Ghisu, T., Duce, P., Pellizzaro, G., 2018. An automated approach for wood-leaf separation from terrestrial LIDAR point clouds using the density based clustering algorithm DBSCAN. Agric. For. Meteorol. 262, 434–444.

Kukko, A., Kaasalainen, S., Litkey, P., 2008. Effect of incidence angle on laser scanner intensity and surface data. Appl. Opt. 47, 986–992.

Lalonde, J., Vandapel, N., Huber, D.F., Hebert, M., 2006. Natural terrain classification using three-dimensional ladar data for ground robot mobility. J. F. Robot. 23, 839–861.

Lau, A., Bentley, L.P., Martius, C., Shenkin, A., Bartholomeus, H., Raumonen, P., Malhi, Y., Jackson, T., Herold, M., 2018. Quantifying branch architecture of tropical trees using terrestrial LiDAR and 3D modelling. Trees 32, 1219–1231.

Li, Z., Douglas, E., Strahler, A., Schaaf, C., Yang, X., Wang, Z., Yao, T., Zhao, F., Saenz, E.J., Paynter, I., 2013. Separating leaves from trunks and branches with dual-wavelength terrestrial LiDAR scanning, in: 2013 IEEE International Geoscience and Remote Sensing Symposium-IGARSS. IEEE, pp. 3383–3386.

Liang, X., Hyyppä, J., Kankare, V., Holopainen, M., 2013. Stem curve measurement using terrestrial laser scanning. IEEE Trans. Geosci. Remote Sens. 52, 1739–1748.

Liang, X., Litkey, P., Hyyppa, J., Kaartinen, H., Vastaranta, M., Holopainen, M., 2012. Automatic Stem Mapping Using Single-Scan Terrestrial Laser Scanning. IEEE Trans. Geosci. Remote Sens. 50, 661–670. https://doi.org/10.1109/TGRS.2011.2161613

Livny, Y., Yan, F., Olson, M., Chen, B., Zhang, H., El-Sana, J., 2010. Automatic reconstruction of tree skeletal structures from point clouds. ACM Trans. Graph. 29, 151.

Ma, L., Zheng, G., Eitel, J.U.H., Moskal, L.M., He, W., Huang, H., 2016. Improved salient feature-based approach for automatically separating photosynthetic and nonphotosynthetic components within terrestrial lidar point cloud data of forest canopies. IEEE Trans. Geosci. Remote Sens. 54, 679–696.

Pesci, A., Teza, G., 2008. Effects of surface irregularities on intensity data from laser scanning: an experimental approach. Ann. Geophys.

Pfeifer, N., Dorninger, P., Haring, A., Fan, H., 2007. Investigating terrestrial laser scanning intensity data: quality and functional relations. na.

Pyörälä, J., Liang, X., Saarinen, N., Kankare, V., Wang, Y., Holopainen, M., Hyyppä, J., Vastaranta, M., 2018. Assessing branching structure for biomass and wood quality estimation using terrestrial laser scanning point clouds. Can. J. Remote Sens. 1–14.

Quammen, D., 2012. Scaling a Forest Giant. Natl. Geogr. Mag. 222, 28–41.

Raumonen, P., Kaasalainen, M., Åkerblom, M., Kaasalainen, S., Kaartinen, H., Vastaranta, M., Holopainen, M., Disney, M., Lewis, P., 2013. Fast Automatic Precision Tree Models from Terrestrial Laser Scanner Data. Remote Sens. . https://doi.org/10.3390/rs5020491

Tao, S., Guo, Q., Xu, S., Su, Y., Li, Y., Wu, F., 2015. A geometric method for wood-leaf separation using terrestrial and simulated lidar data. Photogramm. Eng. Remote Sens. 81, 767–776.

Wang, D., Hollaus, M., Pfeifer, N., 2017. FEASIBILITY OF MACHINE LEARNING METHODS FOR SEPARATING WOOD AND LEAF POINTS FROM TERRESTRIAL LASER SCANNING DATA. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. 4.

Xi, Z., Hopkinson, C., Chasmer, L., 2018. Filtering Stems and Branches from Terrestrial Laser Scanning Point Clouds Using Deep 3-D Fully Convolutional Networks. Remote Sens. 10, 1215.

Xia, S., Wang, C., Pan, F., Xi, X., Zeng, H., Liu, H., 2015. Detecting stems in dense and homogeneous forest using single-scan TLS. Forests 6, 3923–3945.

Xu, H., Gossett, N., Chen, B., 2007. Knowledge and heuristic-based modeling of laser-scanned trees. ACM Trans. Graph. 26, 19.

Yang, X., Strahler, A.H., Schaaf, C.B., Jupp, D.L.B., Yao, T., Zhao, F., Wang, Z., Culvenor, D.S., Newnham, G.J., Lovell, J.L., 2013. Three-dimensional forest reconstruction and structural parameter retrievals using a terrestrial full-waveform lidar instrument (Echidna®). Remote Sens. Environ. 135, 36–51.

Yao, T., Yang, X., Zhao, F., Wang, Z., Zhang, Q., Jupp, D., Lovell, J., Culvenor, D., Newnham, G., Ni-Meister, W., 2011. Measuring forest structure and biomass in New England forest stands using Echidna ground-based lidar. Remote Sens. Environ. 115, 2965–2974.

Zhang, W., Qi, J., Wan, P., Wang, H., Xie, D., Wang, X., Yan, G., 2016. An easy-to-use airborne LiDAR data filtering method based on cloth simulation. Remote Sens. 8. https://doi.org/10.3390/rs8060501

Zhang, W., Wan, P., Wang, T., Cai, S., Chen, Y., Jin, X., Yan, G., 2019. A Novel Approach for the Detection of Standing Tree Stems from Plot-Level Terrestrial Laser Scanning Data. Remote Sens. https://doi.org/10.3390/rs11020211

Zhu, X., Skidmore, A.K., Darvishzadeh, R., Niemann, K.O., Liu, J., Shi, Y., Wang, T., 2018. Foliar and woody materials discriminated using terrestrial LiDAR in a mixed natural forest. Int. J. Appl. earth Obs. Geoinf. 64, 43–50.

Figures



Figure 1

The topography and the forest condition of each test plot. A. The white birch plot. B. The Dahurian larch plot. C. The Chinese scholar tree plot.

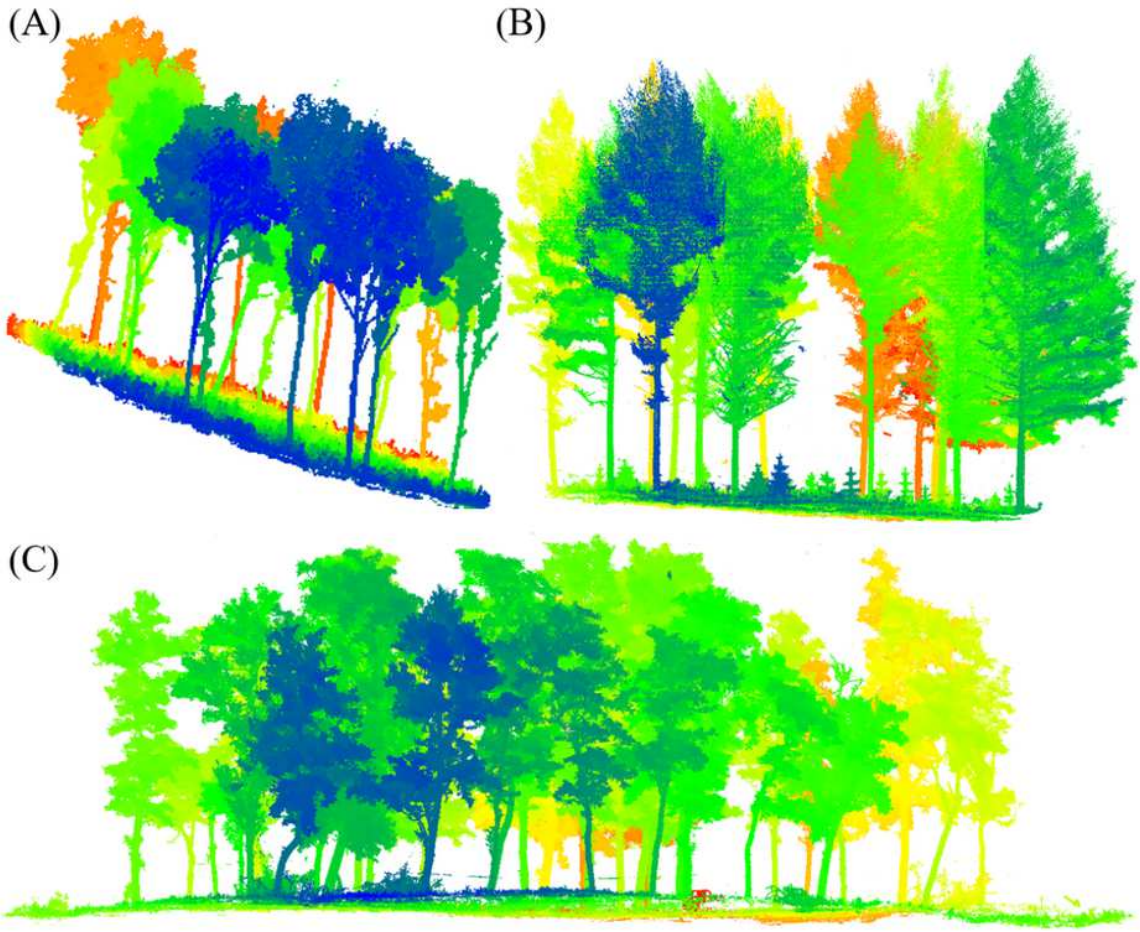


Figure 2

Three field collected TLS datasets. (A) White birch plot. (B) Dahurian larch plot. (C) Chinese scholar tree plot.

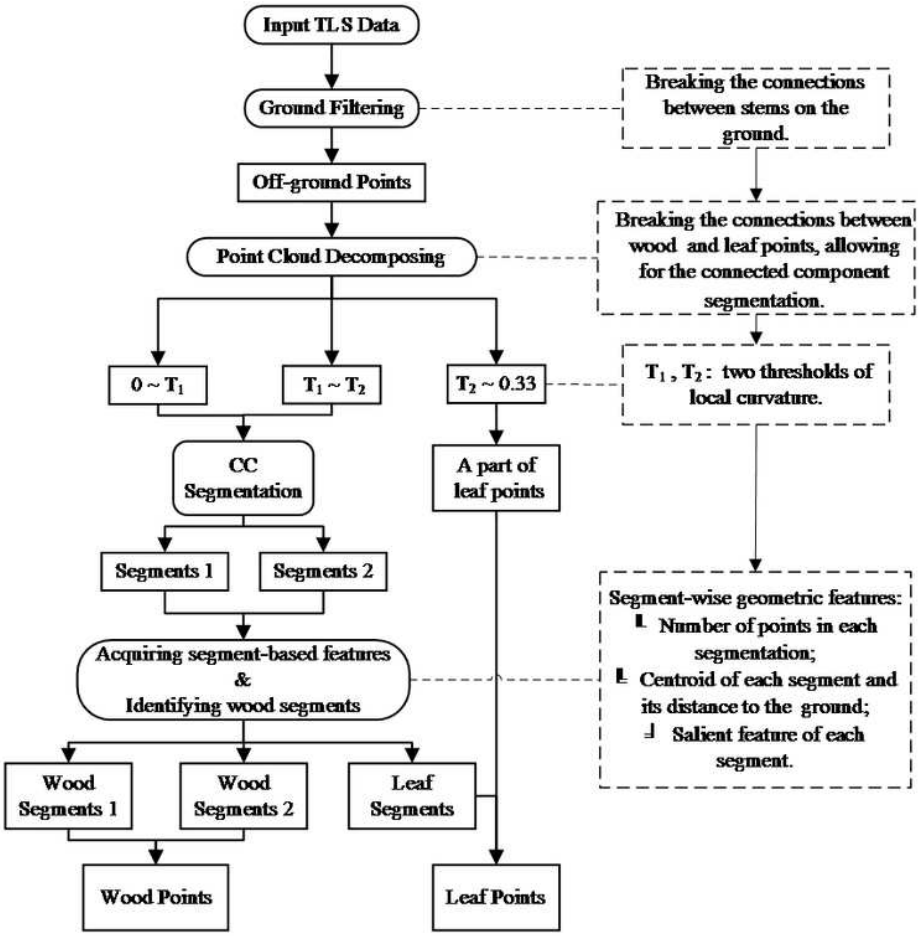


Figure 3

The overall workflow of the proposed method. The left column shows the main steps of the method, and the right column shows the annotation of each main step.

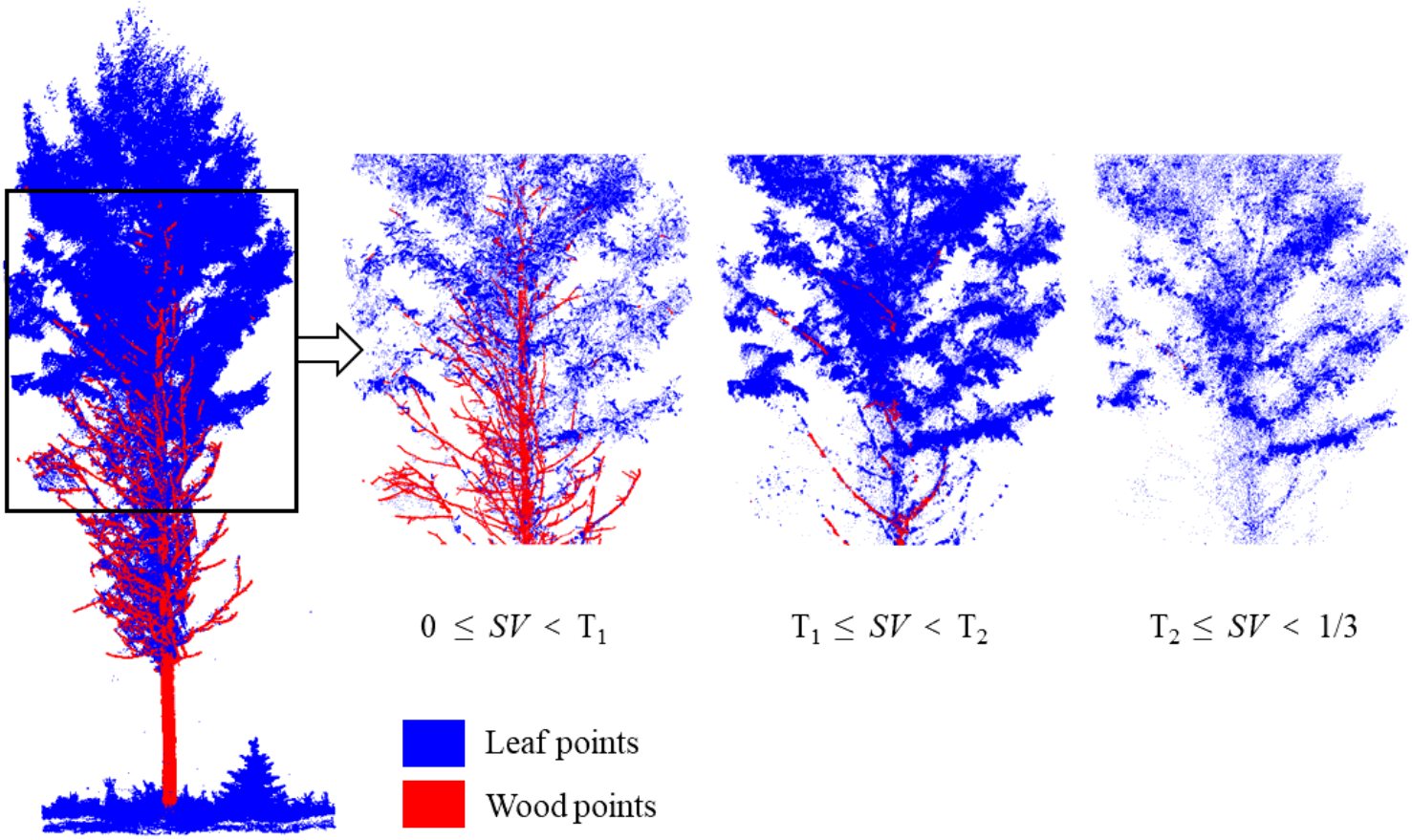


Figure 4

An example of the decomposed point cloud. T1 and T2 refer to the thresholds of the surface variation (SV). In this example, T1 equals to 0.1, T2 equals to 0.2.

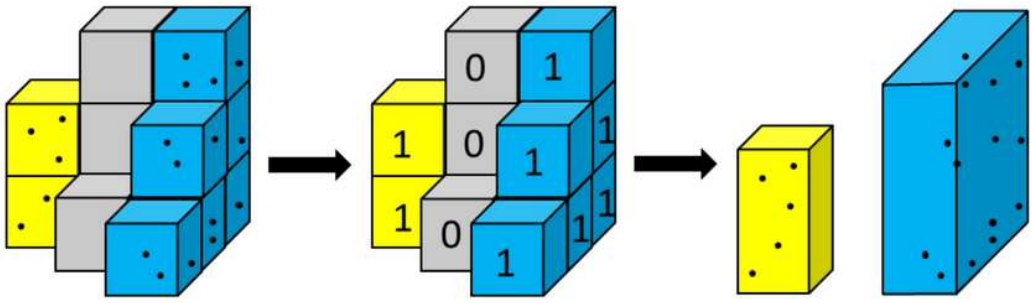


Figure 5

Schematic diagram of the connected component segmentation. Here, a point cloud was segmented into two segments.

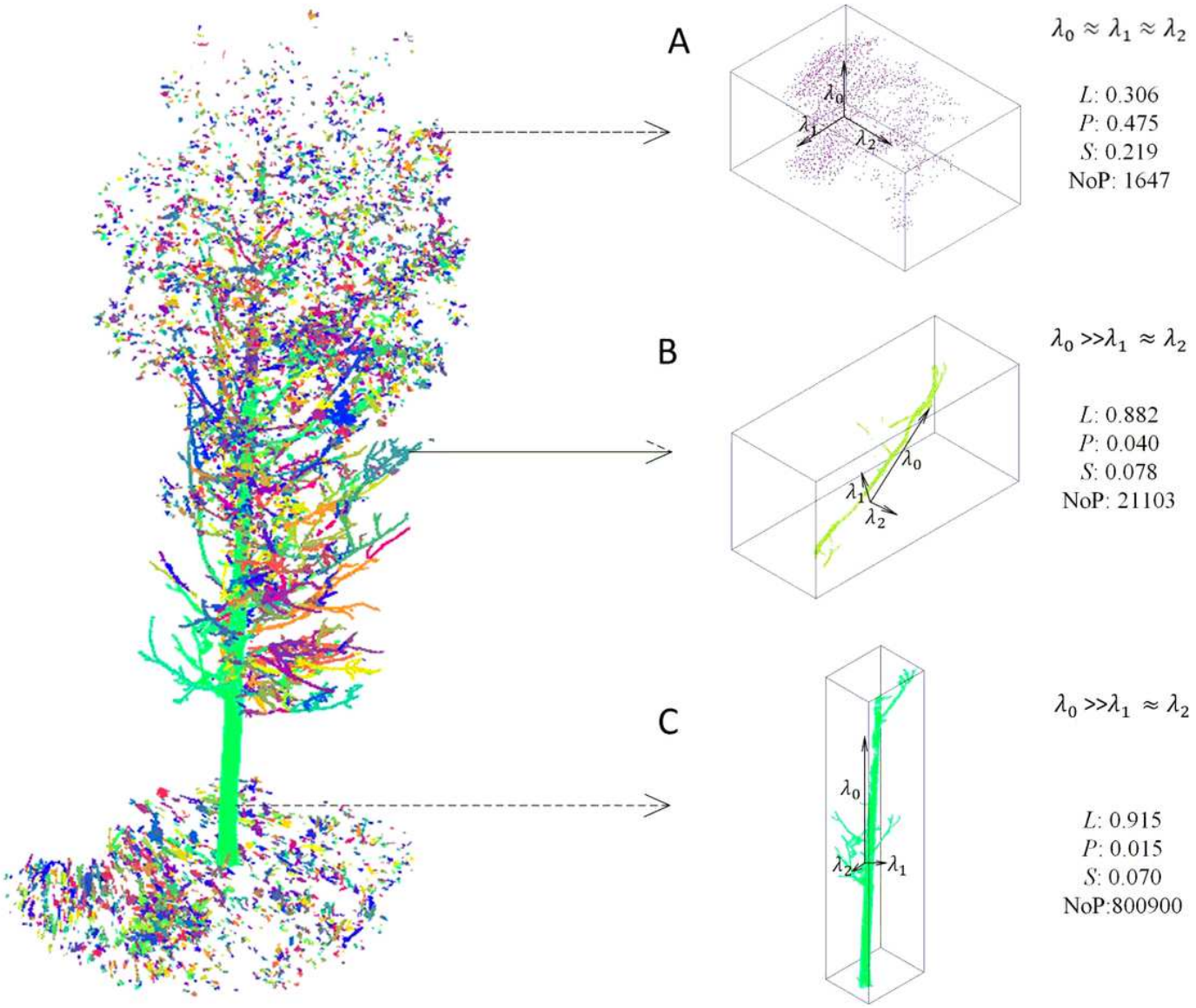


Figure 6

An example of the connected component segmentation and the salient features of the segments. The different colors were drawn for the different segments. L, P, S refer to the attribute value of the linear feature, planar feature, scattered feature and the number of points of each segment. λk(k=0,1,2) are the eigenvalues of each segment.

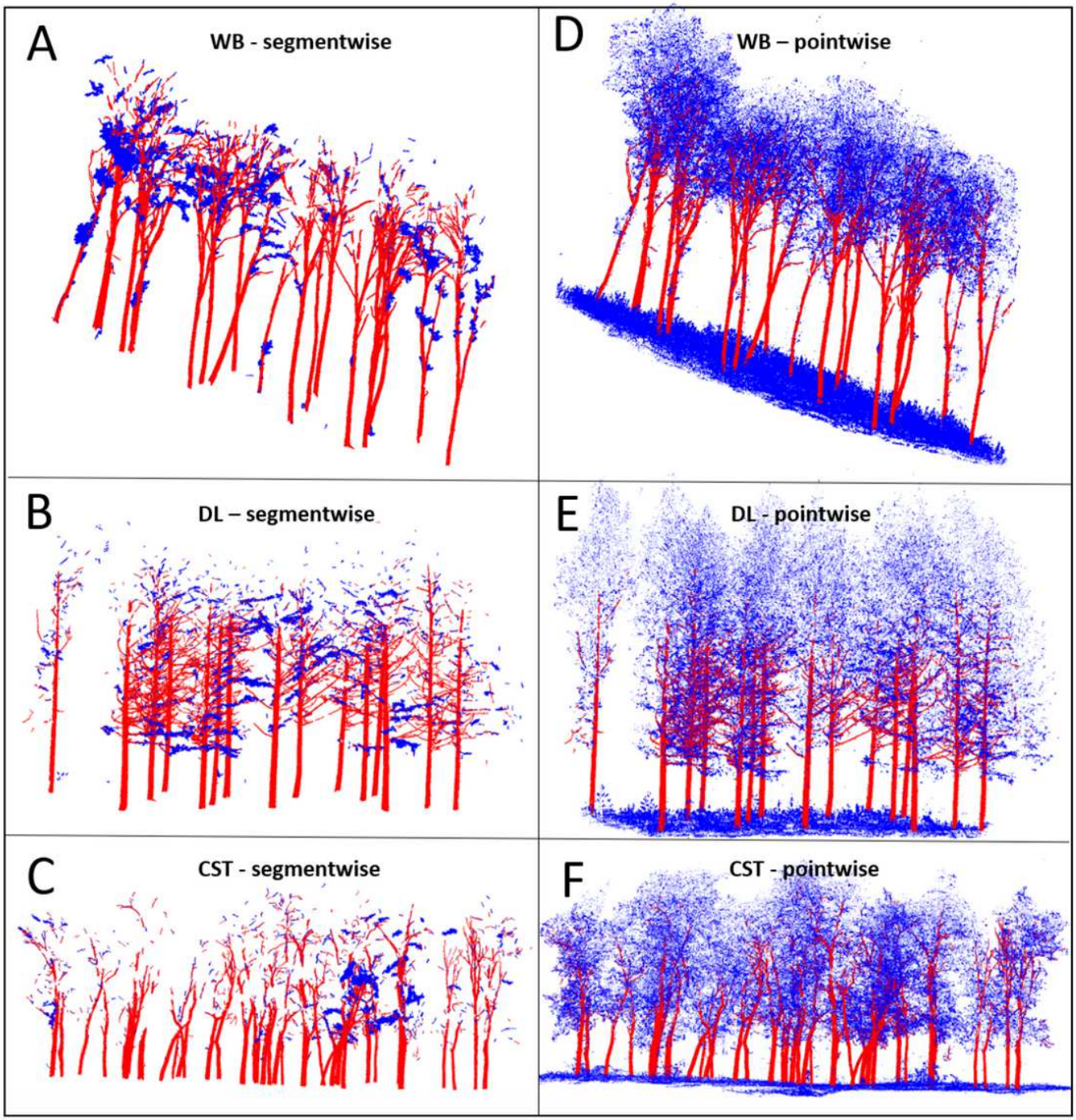


Figure 7

Visualization results of the classification from the two methods. WB, DL and CST represent the plots white birch, Dahurian larch and the Chinese scholar tree.