Article

Individual Tree Crown Segmentation of a Larch Plantation Using Airborne Laser Scanning Data Based on Region Growing and Canopy Morphology Features

Zhenyu Ma ^{1,2}, Yong Pang ^{1,*} , Di Wang ³ , Xiaojun Liang ¹, Bowei Chen ^{1,4}, Hao Lu ^{1,5} , Holger Weinacker ² and Barbara Koch ²

¹ Institute of Forest Resource Information Techniques, Chinese Academy of Forestry, Beijing 100091, China; zhenyu.ma@felis.uni-freiburg.de (Z.M.); Stanforldxj@gmail.com (X.L.); chenbw@aircas.ac.cn (B.C.); luhao@bjfu.edu.cn (H.L.)

² Department of Environment and Natural Resource, University of Freiburg, 79104 Freiburg, Germany; holger.weinacker@felis.uni-freiburg.de (H.W.); barbara.koch@felis.uni-freiburg.de (B.K.)

³ Department of Built Environment, Aalto University, 00076 Aalto, Finland; di.wang@aalto.fi

⁴ Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China

⁵ School of Information Science and Technology, Beijing Forestry University, Beijing 100083, China

* Correspondence: pangy@ifrit.ac.cn

Received: 11 February 2020; Accepted: 19 March 2020; Published: 27 March 2020

Abstract: The detection of individual trees in a larch plantation could improve the management efficiency and production prediction. This study introduced a two-stage individual tree crown (ITC) segmentation method for airborne light detection and ranging (LiDAR) point clouds, focusing on larch plantation forests with different stem densities. The two-stage segmentation method consists of the region growing and morphology segmentation, which combines advantages of the region growing characteristics and the detailed morphology structures of tree crowns. The framework comprises five steps: (1) determination of the initial dominant segments using a region growing algorithm, (2) identification of segments to be redefined based on the 2D hull convex area of each segment, (3) establishment and selection of profiles based on the tree structures, (4) determination of the number of trees using the correlation coefficient of residuals between Gaussian fitting and the tree canopy shape described in each profile, and (5) k-means segmentation to obtain the point cloud of a single tree. The accuracy was evaluated in terms of correct matching, recall, precision, and F-score in eight plots with different stem densities. Results showed that the proposed method significantly increased ITC detections compared with that of using only the region growing algorithm, where the correct matching rate increased from 73.5% to 86.1%, and the recall value increased from 0.78 to 0.89.

Keywords: light detection and ranging (LiDAR), airborne laser scanning (ALS), individual tree crown (ITC) segmentation; larch plantation; region growing; canopy morphology

1. Introduction

Natural forests and plantations are significant ecosystems because they strongly modulate stores and fluxes of water and carbon near the Earth's surface. Plantations are an essential renewable resource for the timber and paper industries [1]. The stand density is one of the most important factors controlling the productivity of managed larch forests [2]. Accurate individual tree detection could improve the management efficiency and forestry production.

Light detection and ranging (LiDAR) has shown high potential in characterizing and monitoring forest structures and functions [3,4], and for supporting forest sustainability and conservation [5]. The algorithms of three-dimensional (3D) individual tree crown (ITC) extraction using airborne laser scanning (ALS) data have been commonly exploited to minimize the traditional time-consuming and manpower-demanding forest inventory practices [6,7]. Once a single tree is accurately segmented, tree structural attributes, such as height [8], species [9], crown size [8,10], stem density [11], wood volume [12], and biomass [13,14], can be derived.

Traditional methods for ITC detection were developed from image processing techniques, including brightness variations exploited from optical data [15], and height variation in LiDAR production (e.g., canopy height models (CHMs)) [16], with local filtering and variable window size [17], or using a watershed algorithm with a local minima [18]. A CHM can depict the forest in a two-dimensional (2D) mathematical morphology of structural components of the canopy surface. Various methods attempt to optimize the identification of treetops, including image filtering with fixed or variable window sizes, multi-scale analysis [19], template matching [20,21], and stochastic geometry based on marked point processes [22]. Crown segments were estimated using the region growing algorithm [12,23,24], watershed analysis, template matching [21,25], fitting functions [26], wavelet analysis [27], or a combination of several methods [28]. Moreover, the level of smoothing used in interpolation methods to build crown segments has a direct impact on the quality of the segmentation [19,24].

The drawbacks of using the CHM are the inherent errors of interpolation production and possible limitations of the understory [29]. Especially in high-density forests, spatial error from overlapping canopies is caused during the interpolation process in generating the digital terrain model (DTM) from the point cloud. This error can reduce the accuracy of ITC segmentation and relevant measurements. Thus, methods to accurately detect individual trees directly from LiDAR point clouds need to be developed and enhanced. For example, Morsdorf [30] carried out a voxel-based algorithm based on k-means clustering to detect a single tree. Gupta [31] compared the iterative partitioning and hierarchical clustering methods using full-waveform ALS data. Ferraz [32] implemented a mean-shift clustering method on pre-defined vegetation layers, which required only one parameter of an adaptive kernel bandwidth that had to be optimized per vegetation layer. The method obtained promising results in well-stratified vegetation layers, but more sophisticated algorithms are still required for more complex forest structures. Wang [33] exploited the voxel structure and a hierarchical morphological algorithm to describe canopy regions at different height intervals. Their algorithm identified both canopy and over-topped trees but was sensitive to both the voxel scale and the size of the morphological elements. Using hypotheses regarding the relative spacing between trees, Li [34] proposed a distance-based method for individual tree segmentation, which exploited the appropriate spacing between the canopies to identify and group points into single trees based on simple rules of proximity and plausible tree shapes. However, this method has not been evaluated in dense and multi-layered forests. The normalized cut method has also been applied to ALS data with full-waveform attributions to derive forest parameters and to detect single trees. Nevertheless, the algorithm was focused on the point or voxel adjacent relationship instead of the crown shapes [35,36]. So far, the efficiency and effectiveness have been the main drawbacks of those methods working directly with 3D points clouds, as they require more computation power compared to image-processing-based methods. Furthermore, the canopy vertical morphology has been extended to separate single trees. Hu [37] developed a two-step segmentation method to combine the CHM watershed segmentation method and prior knowledge of tree crown shape. However, the watershed hardly depicted the actual canopy edge. The vertical morphology formed by points in sequential profiles established in four directions showed a weak descriptions representativeness in a high-density forest. This is because the simple plots have a relatively homogeneous architecture in terms of tree sizes and spatial arrangement, while the point distribution in complex plots are more heterogeneous [38]. In addition, it is challenging to extract useful point features from the heterogeneous point cloud generated using a high-density forest.

To combine the advantage of the original point-cloud positional relationship and crown morphology, we present a two-stage approach that extracts ITC using ALS data over plantation forests with different stem densities. In the first step, we extracted a recognizable potential segment using the region growing method [34] to derive potential tree positions. In the second step, the rotary profiles were established to depict the canopy vertical morphology and the tree structure, and Gaussian fitting was applied to the produced profiles to accurately detect the canopy in the potential segments. The following sections provide detailed descriptions of the algorithm in Section 3, performance evaluation in Section 4, and a discussion of the implications of the method in Section 5.

2. Study Site and Datasets

2.1. Study Area

The study area was located at the Mengjiagang Forest Farm (130°32′–130°52′E, 46°20′–46°30′N) in Heilongjiang Province, China (Figure 1). The research area was a 200–300 m above sea level with a 10°–30° slope. The dominant tree species were *Larix olgensis*, *Pinus sylvestris*, *Pinus koraiensis*, and *Picea asperata*. The dominant tree species were four kinds of coniferous trees with relatively cone crowns. Thus, shape-based segmentation performed well in ITC detection.

Figure 1. The location of field plots, and airborne laser scanning (ALS) flight lines in the Mengjiagang Forest Farm, Heilongjiang Province.

2.2. ALS Data

In this study, ALS data were obtained from May 31 to June 15 in 2017 using a Yun-5 airplane at altitudes of approximately 900 m and 1200 m. The full-waveform ALS data were acquired using the Riegl's LMS-Q6800 (RIEGL Laser Measurement Systems, Horn, Austria) sensor installed at the CAF-LiCHy Airborne Observation System (LiCHy) platform [39]. The scanning parameters were 300 kHz and MTA Zone 2 with a 60%–70% side overlap. The raw waveform data were decomposed into discrete returns according to Gaussian decomposition models [40] and georeferenced with position and orientation data. The registration errors in multiple stripes' data were minimized using an iterative closest point (ICP) algorithm [41,42]. After eight ICP iterations, the mean registration error was reduced to 2.5 cm with a standard deviation of 0.32 m. The derived point clouds of discrete returns had a maximum and minimum average point density of 50 pts/m² and 30 pts/m², respectively. The ground

and non-ground points were separated based on a progressive TIN densification algorithm using TerraScan software [43]. The DTM and digital surface model (DSM) were generated based on filtered ground points and all points, respectively. Then, the normalized height (height above ground) of each point was calculated by subtracting the DTM from the original coordinate Z-value. Figure 2 shows examples of derived point clouds of a complex and a simple plot.

Figure 2. Examples of ALS point clouds for plots with different structural complexities. (a) A complex plot A1 that was $30 \times 20 \text{ m}^2$ in size. (b) A simple plot B4 that was $30 \times 30 \text{ m}^2$ in size.

2.3. Ground Survey Data

Along with the ALS campaign, a ground survey was simultaneously conducted by the Chinese Academy of Forestry and the Northeast Forestry University in June 2017. Four $20 \text{ m} \times 30 \text{ m}$ rectangular plots (A1–A4) and four $30 \text{ m} \times 30 \text{ m}$ square larch plantation plots (B1–B4) were measured. Each tree with a diameter at breast height (DBH) of $\geq 2 \text{ cm}$ was measured, of which height, breast diameter, crown width, and geographical location were recorded. A total of 842 single trees were measured in eight plots. Dead and fallen woods due to natural death and selective cutting were removed from the measurement. Finally, the number of single trees for verification was 770. According to the survey measurement, the variance of tree height in the same plot was relative small. Other measured parameters are shown in Table 1.

Table 1. Parameters of field plots.

Plot ID	Mean DBH (cm)	Mean Height (m)	Min Height (m)	Max Height (m)	Crown Width East–West (m)	Crown Width North–South (m)	Stem Density (stems/ha)
A1	25.76	24.23	14.8	28.8	5.26	4.86	1216
A2	31.48	28.93	23.1	33.1	5.01	5.00	1050
A3	23.85	22.82	19.3	28.1	4.89	5.02	1367
A4	21.67	22.19	17.3	26.2	4.21	3.36	1850
B1	19.48	21.11	14.8	25.9	2.98	3.06	1617
B2	17.65	20.58	15.1	26.4	2.65	2.78	1783
B3	16.69	19.66	14.1	24.3	2.88	3.03	1817
B4	15.56	19.92	13.1	27.2	2.77	2.80	2117

3. Methods

This study proposed an algorithm that combined the region growing method with canopy morphology features. The workflow is shown in Figure 3. The purpose of the region growing

algorithm is to generate the segments used to determine possible single-tree positions and point clusters, which represent the coarse ITC structures with branches around the boundary. Then, the multi-angle vertical profile is established in each selected segment and the points within the profile are horizontally projected onto the profile. The best three profiles are selected based on the canopy point characteristics depicted by the tree canopy. The Gaussian fitting is performed for the 2D point cloud in the selected profile and the autocorrelation coefficient of the fitting residual is used to determine the number of single trees in each segment. Finally, the point clusters of the ITC tree separation are obtained using k-means clustering.

Figure 3. Workflow of the proposed method.

3.1. Region Growing

Region growing is the process of aggregating groups of elements into larger regions. Starting from the search for seed points, adjacent elements with similar properties to each seed point are merged [26,44]. We used the region growing algorithm [34] as the first segmentation step to detect the rough ITC positions and corresponding pre-defined segments. This method uses a top-to-bottom region growing approach on normalized height point cloud to sequentially segment trees from the highest to the lowest points. Specifically, there is a horizontal spacing between the canopies, and the spacing between the tops of canopies is greater than that of the bottoms. Points with a spacing larger than a specified threshold, T-spacing, are excluded from the target tree; points with a spacing smaller than the threshold are classified into one tree based on a minimum spacing rule. Referring to the algorithm's principle, the T-spacing should be approximately equal to the crown radius. If the threshold is set too small, trees with elongated branches can be over-segmented; if the threshold is set too large, nearby trees can be missed. Additionally, the point distribution rules of the tree shape are also appended to improve the accuracy of the segmentation, which can reduce the over-segmentation. In a high-density plantation, the empirical threshold resulted in the under-segmentation of segments containing one tree or multiple trees. The exact number of trees in segments is then determined based on the shape morphology, as described in Section 3.3.

3.2. Segments Filter

Some larch plantation canopies were tightly intersected, and the overlap of the crown edge was not obvious enough for individual tree detection using the previous region growing method only. In this regard, the crown shapes used for the secondary morphology segmentation could determine

the number of trees in each segment. However, the secondary morphology segmentation was not applied to those segments with a 2D area smaller than the threshold of 4π because based on the crown width, the areas smaller than this threshold had a low possibility of containing multiple trees. The 2D area was calculated using convex hull outlines of points, which were vertically projected onto the X-Y plane in each segment. The left segments were further processed with the secondary morphology segmentation to identify the exact number of trees.

3.3. Tree Number Definition Using the Profile Morphology

3.3.1. Profile Establishment

The vertical profiles were established in each segment selected in the previous step. Every profile was a vertical cuboid with a 1 m thickness that went through the highest point in the segment, as the green profile shows in Figure 4a,b. The 16 profiles were generated based on the rotation axis with a 11.25° interval angle from 0 to π in the X-Y plane. The 3D points with the original z-value contained in each profile were projected vertically onto the X-Y plane, which reflected the shape of the true canopy without the influence of height normalization. Figure 5 shows three examples of the generated 2D profiles. If a segment contained more than one tree, there should be one or more profiles separating the canopies because the rotary profiles could describe the multifaceted structure. Additionally, the canopy structures and point numbers in adjacent profiles were similar, thus optimal profile selection was necessary to improve the ITC detection accuracy and efficiency.

Figure 4. The diagram of three established profiles in a region growing segmentation segment. (a) 3D view of the segment, where the vertical profile with a 1 m width was established based on the rotation axis that went through the max height point. (b) Top view of the segmentation segment, where the green profile represented the three profiles selected based on the aforesaid rules.

(a)

(b)

(c)

Figure 5. Blue asterisk represents the upper canopy point cloud of the profile selected using 0.2 m width uniformly-spaced lags. The lags are shown as the green rectangle to obtain the 3 m upper canopy points in each profile. (a), (b), and (c) represent the first, second, and third selected profiles, respectively, using the point number rules.

3.3.2. Profile Selection

First, we deleted the profile with a horizontal distance shorter than 2.5 m, since it was implausible that these profiles contained multiple canopies within such a particularly short horizontal distance (the

magenta line in Figure 4b). Second, based on the characteristics of region growing and the observations of numerous profiles, the following assumption was made: there were more points near the center of a tree crown than around its edges [37] owing to the laser beam multiple reflections caused by the complex canopy structure in the canopy center area. Sub-arboreal vegetation was eliminated by establishing an appropriate height threshold of 4 m above the ground. The points under 4 m had a slight impact on the crown shape because of scrub-cleaning and pruning activities in the plantation. We assumed that the more points contained in the profile, the more likely for it to contain multiple canopies. Therefore, profiles were selected according to the decreasing order of points number in each profile. Finally, the number of points of adjacent profiles were generally similar, and they often demonstrated a similar canopy structure. Thus, we set the rules to eliminate similar adjacent profiles according to the following condition: all profiles were initially sorted in a descending order by the number of points in them. When the first profile with the most points was selected, the next profile was not selected if it was located in the area bounded by a 22.5° angle on both sides of the first profile (the grey area in Figure 4). According to this rule, the third profile was selected if it was outside the area bounded by previously selected profiles. This procedure was repeated until three profiles were selected. In practice, three profiles can effectively represent the canopy structure in this segment.

3.3.3. Tree Number Definition in Each Segment

The canopy uppermost points represent the canopy structure, while the canopy inner points influence the Gaussian fitting of canopy shape; thus, the points in the canopy's uppermost 3 m were filtered according to a 0.2 m width that was uniformly spaced to divide the X-axis (Figure 5). Next, to estimate the exact number of trees in each profile, we used the Gaussian model (Equation (1)) from MATLAB to fit the filtered uppermost canopy points:

$$f(x) = \sum_{i=1}^n a_i e^{\frac{-(x-\mu_i)^2}{\sigma_i^2}} \quad (1)$$

where n represents the number of Gaussian functions. a_i , μ_i , and σ_i represent the magnitude, central position, and standard deviation of the i th Gaussian function, respectively. The fitting process was applied sequentially from one to three Gaussian functions (Figure 6), with an increment of one function each time. As the number of functions increased, the Gaussian curve fitting improved with gradually decreased fitting residuals showed in Figure 7. We limited the maximum number of Gaussian functions to three, as the fitting effect was rarely improved after that. The appropriate number of Gaussian functions used to fit a given profile was considered as an estimation of the number of trees. To do so, we used residual correlation coefficients to determine the appropriate number of functions (Table 2). The residuals were calculated using the points within the 99% confidence interval of the fitting with three Gaussian functions. The correlation coefficient threshold was empirically set to 0.9. For instance, if the residual correlation coefficients of the first function and the second function was less than 0.9, this implied that the second function fitting improved significantly compared with the first function fitting. Thus, the appropriate tree number in this profile was two.

(a) The result of Gaussian function fitting in the first profile

(b) The result of Gaussian function fitting in the second profile

(c) The result of Gaussian function fitting in the third profile

Figure 6. Gaussian fitting of three selected profiles, with the number of functions increasing from one to three. The figures show the upper canopy points as a blue asterisk, the combined model as a red line, and the model components using the single function as green imaginary lines. The magenta imaginary lines are the 99% confidence interval of the Gaussian fitting model.

Figure 7. The residual of one to three Gaussian fitting functions using 0.2 m width lags. The bar chart shows the difference between the Z-value of points in the 99% confidence interval of three Gaussian fitting functions and the fitting model.

Table 2. The correlation coefficient of the residual.

Autocorrelation Coefficient	Profile 1	Profile 2	Profile 3
Gaussian function 1–2	0.6433	0.9396	0.9687
Gaussian function 2–3	0.9986	0.9505	0.9362

The process is further explained in Figure 6, which depicts the three profiles selected from the 16 profiles in a segment according to the method described in 3.3.2. The fitting result of the Gaussian functions is shown in Figure 6. It was concluded that as the number of functions increased, the Gaussian curve gradually approached the actual profile shape. When there was only one canopy shape in the profile, one Gaussian function was enough to resemble the canopy curve. If there were two canopies in the profile, such as in Figure 6a, the fitting accuracy improved significantly with two Gaussian functions compared with one function (the residual declined in Figure 7). When the number of Gaussian functions increased from 2 to 3, the residual slightly declined. The correlation coefficient between the residual of one function and that of two functions was $0.6433 < 0.9$. The correlation coefficient between the residual of two functions and that of three functions was $0.9986 > 0.9$. Therefore, we concluded that there were two trees in this profile. All three profiles were processed using the same rule, and the number of trees in each segment was finally determined.

3.4. k-Means Segmentation

Based on the determination of tree numbers in Section 3.3.3, we then applied k-means clustering to group points of individual trees. The default random seeds were chosen. In the k-means algorithm, the sum of absolute differences in Euclidian distance between each point and its closest center is minimized. The squared error function (S_e) was calculated using Equation (2). Additionally, the z-coordinates of the ALS points were scaled down by a factor of six before the initialization of the k-means process [31] (Figure 8). Scaling the height value of the points helped in minimizing the vertical-cluster variance, or the squared error function (S_e), which is the ultimate objective of the k-means method.

$$S_e = \sum_{i=1}^k \sum_{n=1}^n |x_j^i - c_i|^2 \quad (2)$$

where $|x_j^{(i)} - c_i|^2$ represents a selected distance measure between a data point $x_j^{(i)}$ and the mean vector or cluster center c_i . The value of S_e depended on how the samples were grouped into clusters and the number of defined k clusters. At different iterations, new cluster memberships and its center were computed. For the optimal partitioning, S_e was minimized. The number of trees obtained in Section 3.3.3 had a direct relationship with the k-means clusters to be generated.

Figure 8. The result of the k-means clustering based on the tree number defined in Section 3.3.3. The asterisks represent the original tree points belonging to the segment; the different colored asterisks represent the k-means segmentation results. The circles represent the scaled-down canopy points colored using the k-means segmentation result.

3.5. Accuracy Evaluation

The segmented results were evaluated against field reference data. Four situations of the segmentation results were placed into the following three categories. If a tree was correctly segmented, it was called true positive (TP); if a tree is not segmented but assigned to a nearby tree, it was called false negative (FN) or an omission error; if a tree did not exist but was segmented from the point cloud, it was called false positive (FP) or a commission error. TP, FN, and FP indicate perfect segmentation, under-segmentation, and over-segmentation, respectively. To evaluate the accuracy, we calculated the correct matching (c), recall (r), precision (p), and F-score using the following equations [45,46]:

$$c = \frac{TP}{Num_{measured}}, \quad (3)$$

$$r = \frac{TP}{TP + FN}, \quad (4)$$

$$p = \frac{TP}{TP + FP}, \quad (5)$$

$$F = 2 \times \frac{r \times p}{r + p}. \quad (6)$$

The matching value represents the rate of the correct segmentation of trees. The recall value indicates the tree detection rate. The precision value represents the correctness of the method regarding tree detection. The F-score value is the overall accuracy considering both the commission and omission errors. The values of c, r, p, and F vary from 0 to 1.

4. Results and Discussion

4.1. Accuracy of the ITC Segmentation

Table 3 shows the detection accuracy of the region growing algorithm and our morphology method in eight plots with varied stem densities. The ITC convex edge and reference tree position of the low-density A1 and high-density B8 are shown in Figure 9a,b. Figure 9c,d shows the segmented point clouds for plots A1 and B4. It is visually apparent that most of the trees were correctly segmented. The tree segmentation results in other areas in our study were similar and hence are not shown here.

Table 3. The accuracy of individual tree segmentation.

Plot Information				Region Growing Method							Morphology Segmentation Method							
Plot ID	Measured Tree	Density (num/ha)	Segment Trees	TP	FN	FP	c	r	p	F	Segment trees	TP	FN	FP	c	r	p	F
A1	63	1216	53	52	10	1	82.54%	0.84	0.98	0.90	60	58	3	2	92.06%	0.95	0.97	0.96
A2	73	1050	68	67	5	1	91.78%	0.93	0.99	0.96	72	70	3	2	95.89%	0.96	0.97	0.97
A3	82	1367	67	66	13	1	80.49%	0.84	0.99	0.90	79	74	7	4	90.24%	0.91	0.95	0.93
A4	111	1850	78	75	30	2	67.57%	0.71	0.97	0.82	97	93	8	4	83.78%	0.92	0.96	0.94
B1	97	1617	77	74	20	2	76.29%	0.79	0.97	0.87	92	82	13	8	84.54%	0.86	0.91	0.89
B2	107	1783	86	79	22	3	73.83%	0.78	0.96	0.86	96	89	12	6	83.18%	0.88	0.94	0.91
B3	109	1817	8	73	28	3	65.14%	0.72	0.96	0.82	98	91	13	4	83.49%	0.88	0.96	0.91
B4	128	2117	82	80	32	1	62.50%	0.71	0.99	0.83	114	106	19	6	82.81%	0.85	0.95	0.89
Total	770	/	580	566	160	8	73.50%	0.78	0.99	0.87	707	663	78	36	86.10%	0.89	0.95	0.92

Figure 9. Tree segments of plot A1 (a,c) and B4 (b,d) generated using the region growing algorithm and morphology segmentation. In (a) and (b), a blue cross represents a measured single-wood position; a green boundary corresponds to a convex hull edge of a single tree derived from the region growing algorithm's result; and a red boundary corresponds to tree segments that needed further refinement using morphology segmentation, where morphology segmentation successfully detected more than two trees in the region growing segments. (c) and (d) show the segmented point cloud, where the color of points indicates individual tree segments.

According to the accuracy of our morphology segmentation, the overall correct matching rate was 86.1%, with the lowest accuracy of 82.81% and the highest accuracy of 95.89%. The recall in the eight plots varied from 0.85 to 0.96; the precision varied from 0.91 to 0.97. Among the simple plots A1–A3, the mean values of the correct matching with and without morphology segmentation were 84.94% and 92.73%, respectively. Furthermore, for the complex plots A4 and B1–B4, the accuracy using the morphology segmentation was significantly increased compared with that using the region growing algorithm only. The total recall of morphology segmentation was increased from 0.78 to 0.89, and the F-score was increased from 0.87 to 0.92. In these dense plots, trees were under-segmented using the region growing method, and the recall was relatively larger than that of the low-density plot. However, the precision slightly decreased from 0.98 to 0.95 using morphology segmentation, indicating that some trees were over-segmented. For instance, in the dense plots B1 and B4, the precision declined by 0.06 and 0.04, respectively. In general, the number of trees was under-estimated with our method. A

total of 663 out of 770 trees were successfully segmented in our test plots. Our algorithm missed 78 trees and over-segmented 56 trees.

4.2. Discussion

In this paper, we proposed a morphology segmentation approach for 3D ITC segmentation in plantation forests with different densities. The results were significantly improved for the plots dominated by high-density larch plantations. The results demonstrated that morphology segmentation performed better in a higher-density forest, which was difficult to segment using the region growing algorithm alone. However, the shortcoming of this algorithm was that it could not solve the over-segmentation in first segmentation, which was slightly increased in second segmentation caused by definition of tree number in multiple profiles, especially in high-density plots. In conclusion, the region growing method could under-segment the high-density plantation, but the proposed combination of region growing and morphology segmentation algorithms in our study improved the accuracy of ITC detection in different density plots, even though it caused a slight over-segmentation. The algorithm took advantage of prior knowledge of the plantation canopy shape and region growing characteristics. The morphology segmentation based on the original point cloud increased the accuracy and clarity of individual tree structure, as it was not affected by the interpolation and the elevation normalization. Additionally, it was common that threshold T-spacing in region growing segmentation was difficult to set appropriately in a high-density plantation plot. The morphology segmentation algorithm solved the under-segmentation caused by an inappropriate threshold in the region growing method. The segments including multiple crowns were clearly separated with Gaussian functions on the 2D profile. The designed framework was efficient since a detailed examination of 3D ALS points was not needed for all segments, but only for filtered ones based on prior knowledge of segment area. Moreover, the 3D structure in the segments was further described and simplified using 2D Gaussian curves.

4.2.1. Region Growing Segments

Figure 10 shows three typical segments that resulted from the region growing algorithm. Case I represents a perfect segmentation. The segment contained a single canopy that was described in different profiles. Case II represents under-segmentation, which contained two complete canopies. according to the principle of the region growing algorithm, the under-segmentation was generally divided into two types: (1) II-a, where the crown height difference of two consecutive canopies was relatively large and the relative lower canopy could not be separated because the distance between the lower treetop and its highest point nearby was smaller than the spacing threshold in the region growing segmentation. (2) II-b, which was similar to under-segmentation, but the height difference of consecutive canopies was small. Finally, case III represents the over-segmentation segment. The canopy points on edges caused over-segmentation in its adjacent segments. The segments III-a included a dominated canopy and a disconnected adjacent canopy, which resulted in an over-segmentation of the adjacent canopy III-b. The region growing algorithm produced a fusion at the edges of some canopies, causing the segment to contain the point clouds of another over-segmented canopy. We discuss the morphology segmentation performance for different segments.

Figure 10. The diagram of region growing algorithm segments.

4.2.2. Morphology Segments

For case I, morphology segmentation determined that there was only one tree according to the residual correlation coefficient of Gaussian fitting. For case II, there was a possibility that the two trees could be identified using the residual correlation coefficient of two Gaussian functions as long as the two canopies could be fully described in the selected profiles, as shown in Figure 11a,b. If the crowns were severely overlapped, the morphology segmentation was difficult because the residual was slightly influenced by the adjacent canopies. In particular, the residual correlation coefficient was set to 0.9 to guarantee the detection in a high-density forest area since the distance between the two trees was relatively close and the intersection was not clearly separated. In comparison with II-a, the profile of adjacent canopies in II-a formed the Gaussian functions with a proximate proportionate width such that the structural morphology of the two trees could be better described. In addition, in both II-a and II-b, several points were mixed indiscriminately at the edge of the canopy. For case III, the possibility to detect two trees depended on the proportion of the over-segmentation canopy to the whole segments. For example, the over-segmentation canopies could be separated in Figure 11c. However, if the over-segmentation part only occupied a very small part of the segments (Figure 11d), these canopy edges only had a minor impact on the fitting process of the dominated crown. In this case, the point cloud at both the beginning and end of the profile slightly influenced the residual of the crown edge at the Gaussian fitting edge.

Figure 11. Examples of refinement results using morphology segmentation: (a) the segment contained two canopies with a large tree height difference, (b) the segment contained two canopies with relatively equal heights, (c) the refinement segment contained one over-segmentation crown and one complete crown, and (d) the unchanged segment contained one over-segmentation crown and one complete crown.

Furthermore, having three crowns fall within one segment was a rare occasion (Figure 12a). This was difficult to segment using Gaussian fitting because the clumped canopies formed a smoothly curved surface that resembled a single tree. The overlap ratio of the canopies was too high, meaning that the crowns were unable to be segmented.

Figure 12. A tree segment with more than three crowns contained a multiple-tree centralized distribution. (a) Points in the segment. (b–d) Gaussian fitting of the profile selected from the segment that contained three canopies, where the number of functions increased from one to three. Blue points represent the upper canopy points, with the combined model represented with a red line. The magenta lines are the 99% confidence interval of three Gaussian function fitting models.

5. Conclusions

In conclusion, we developed a new algorithm to segment individual trees from point clouds. Our two-step method used canopy morphological clues to improve the under-segmentation that occurs using the region growing algorithm. First, initial segments containing several trees were retrieved using region growing segmentation in a 3D spatial domain. Then, rotating profiles were established to delineate the canopy morphology in different directions, with the inclusion of multi-angle profile analysis and the 2D canopy structure. The number of trees in each initial segment was determined using Gaussian fitting. Finally, k-means segmentation was used to obtain the point cloud of a single tree. The algorithm was tested in larch plantation plots with different stem densities. The results showed that the proposed method significantly increased ITC detections compared with that of using only the region growing algorithm, especially for the clumped trees whose canopies were hard to segment relying only on region growing features, where the correct matching rate increased from 73.5% to 86.1%, and the recall value increased from 0.78 to 0.89.

Although the refinement could not recover every missed tree due to under-segmentation, it solved the problem that was difficult for dense forests and mitigated the omission errors to some extent. The new algorithm improved single tree detection in a larch plantation and allows for the processes of forest stands that have different stem densities. The accuracies, in terms of recall, precision, and F-score, were improved compared with results from using only the region growing algorithm, indicating that the new algorithm had a good potential for dense forests with cone crown shapes. The method could detect 86.1% of trees over a range of forest stem densities. This method could efficiently improve the under-segmentation caused by the region growing algorithm. However, the over-segmentation needed a rough region growing threshold to be reduced, which should be developed through a combination of the point distance relationship and canopy morphology to minimize over-segmentation detections.

Author Contributions: Conceptualization, Y.P.; Methodology, Z.M.; Data curation, X.L.; Data analysis, B.C. and H.L.; Supervision, Y.P. and B.K.; Writing-original draft, Z.M.; Writing-review and editing, D.W., Y.P. and H.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Key Research and Development Program (2017YFD0600404) and CAF Foundation (CAFYBB2016ZD004).

Acknowledgments: The research is grateful for the scholarship under the China Scholarship Council.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sakurai, S. *Plantation Forestry in the Tropics*; Springer: Berlin, Germany, 1982.
2. Walters, G.A. *Saligna Growth in a 15-Year-Old Spacing Study in Hawaii*; Research Paper PSW-RP-151; U.S. Department of Agriculture: Berkeley, CA, USA, 1980.
3. Næsset, E.; Gobakken, T.; Solberg, S.; Gregoire, T.G.; Nelson, R.; Ståhl, G.; Weydahl, D. Model-assisted regional forest biomass estimation using LiDAR and InSAR as auxiliary data: A case study from a boreal forest area. *Remote Sens. Environ.* **2011**, *115*, 3599–3614.
4. Rosenqvist, Å.; Milne, A.; Lucas, R.; Imhoff, M.; Dobson, C. A review of remote sensing technology in support of the Kyoto Protocol. *Environ. Sci. Policy* **2003**, *6*, 441–455. [\[CrossRef\]](#)
5. Wulder, M.A.; White, J.C.; Nelson, R.F.; Næsset, E.; Ørka, H.O.; Coops, N.C.; Hilker, T.; Bater, C.W.; Gobakken, T. Lidar sampling for large-area forest characterization: A review. *Remote Sens. Environ.* **2012**, *121*, 196–209. [\[CrossRef\]](#)
6. Wang, Y.; Hyypä, J.; Liang, X.; Kaartinen, H.; Yu, X.; Lindberg, E.; Holmgren, J.; Qin, Y.; Mallet, C.; Ferraz, A.; et al. International benchmarking of the individual tree detection methods for modeling 3-D canopy structure for silviculture and forest ecology using airborne laser scanning. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 5011–5027. [\[CrossRef\]](#)
7. Eysn, L.; Hollaus, M.; Lindberg, E.; Berger, F.; Monnet, J.-M.; Dalponte, M.; Kobal, M.; Pellegrini, M.; Lingua, E.; Mongus, D. A benchmark of lidar-based single tree detection methods using heterogeneous forest data from the alpine space. *Forests* **2015**, *6*, 1721–1747. [\[CrossRef\]](#)
8. Unger, D.R.; Hung, I.-K.; Brooks, R.; Williams, H. Estimating number of trees, tree height and crown width using Lidar data. *GISci. Remote Sens.* **2014**, *51*, 227–238. [\[CrossRef\]](#)
9. Harikumar, A.; Bovolo, F.; Bruzzone, L. An internal crown geometric model for conifer species classification with high-density lidar data. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 2924–2940. [\[CrossRef\]](#)
10. Popescu, S.C.; Wynne, R.H.; Nelson, R.F. Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. *Can. J. Remote Sens.* **2003**, *29*, 564–577. [\[CrossRef\]](#)
11. Richardson, J.J.; Moskal, L.M. Strengths and limitations of assessing forest density and spatial configuration with aerial LiDAR. *Remote Sens. Environ.* **2011**, *115*, 2640–2651. [\[CrossRef\]](#)
12. Hyypä, J.; Kelle, O.; Lehikoinen, M.; Inkinen, M. A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 969–975. [\[CrossRef\]](#)
13. Ene, L.T.; Næsset, E.; Gobakken, T.; Bollandsås, O.M.; Mauya, E.W.; Zahabu, E. Large-scale estimation of change in aboveground biomass in miombo woodlands using airborne laser scanning and national forest inventory data. *Remote Sens. Environ.* **2017**, *188*, 106–117. [\[CrossRef\]](#)
14. Vauhkonen, J.; Næsset, E.; Gobakken, T. Deriving airborne laser scanning based computational canopy volume for forest biomass and allometry studies. *ISPRS J. Photogramm. Remote Sens.* **2014**, *96*, 57–66. [\[CrossRef\]](#)
15. Erikson, M. Segmentation of individual tree crowns in colour aerial photographs using region growing supported by fuzzy rules. *Can. J. For. Res.* **2003**, *33*, 1557–1563. [\[CrossRef\]](#)
16. Bongers, F. Methods to assess tropical rain forest canopy structure: An overview. In *Tropical Forest Canopies: Ecology and Management*; Springer: Dordrecht, The Netherlands, 2001; pp. 263–277.
17. Popescu, S.C.; Wynne, R.H.; Nelson, R.F. Estimating plot-level tree heights with lidar: Local filtering with a canopy-height based variable window size. *Comput. Electr. Agric.* **2002**, *37*, 71–95. [\[CrossRef\]](#)
18. Zhao, D.; Pang, Y.; Li, Z.; Liu, L. Isolating individual trees in a closed coniferous forest using small footprint lidar data. *Int. J. Remote Sens.* **2014**, *35*, 7199–7218. [\[CrossRef\]](#)
19. Véga, C.; Durrieu, S.; Morel, J.; Allouis, T. A sequential iterative dual-filter for LiDAR terrain modeling optimized for complex forested environments. *Comput. Geosci.* **2012**, *44*, 31–41.
20. Brandtberg, T.; Warner, T.A.; Landenberger, R.E.; McGraw, J.B. Detection and analysis of individual leaf-off tree crowns in small footprint, high sampling density lidar data from the eastern deciduous forest in North America. *Remote Sens. Environ.* **2003**, *85*, 290–303. [\[CrossRef\]](#)
21. Korpela, I.; Tuomola, T.; Välimäki, E. Mapping forest plots: An efficient method combining photogrammetry and field triangulation. *Silva Fenn.* **2007**, *41*, 457–469. [\[CrossRef\]](#)

22. Zhou, J.; Proisy, C.; Descombes, X.; Hedhli, I.; Barbier, N.; Zerubia, J.; Gastellu-Etchegorry, J.-P.; Couteron, P. Tree crown detection in high resolution optical and LiDAR images of tropical forest. In Proceedings of the Remote Sensing for Agriculture, Ecosystems, and Hydrology XII, Toulouse, France, 20–22 September 2010; p. 8240Q.
23. Persson, A.; Holmgren, J.; Soderman, U. Detecting and measuring individual trees using an airborne laser scanner. *Photogramm. Eng. Remote Sens.* **2002**, *68*, 925–932.
24. Solberg, S.; Næsset, E.; Bollandsas, O.M. Single tree segmentation using airborne laser scanner data in a structurally heterogeneous spruce forest. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 1369–1378. [[CrossRef](#)]
25. Pollock, R. The Automatic Recognition of Individual Trees in Aerial Images of Forests Based on a Synthetic Tree Crown Image Model. Ph.D. Thesis, University of British Columbia, Vancouver, BC, Canada, 1996.
26. Popescu, S.C.; Zhao, K. A voxel-based lidar method for estimating crown base height for deciduous and pine trees. *Remote Sens. Environ.* **2008**, *112*, 767–781. [[CrossRef](#)]
27. Falkowski, M.J.; Smith, A.M.; Hudak, A.T.; Gessler, P.E.; Vierling, L.A.; Crookston, N.L. Automated estimation of individual conifer tree height and crown diameter via two-dimensional spatial wavelet analysis of lidar data. *Can. J. Remote Sens.* **2006**, *32*, 153–161. [[CrossRef](#)]
28. Koch, B.; Heyder, U.; Weinacker, H. Detection of individual tree crowns in airborne lidar data. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 357–363. [[CrossRef](#)]
29. Guo, Q.; Li, W.; Yu, H.; Alvarez, O. Effects of topographic variability and lidar sampling density on several DEM interpolation methods. *Photogramm. Eng. Remote Sens.* **2010**, *76*, 701–712. [[CrossRef](#)]
30. Morsdorf, F.; Meier, E.; Allgöwer, B.; Nüesch, D. Clustering in airborne laser scanning raw data for segmentation of single trees. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2003**, *34*, W13.
31. Sandeep, G.; Holger, W.; Barbara, K. Comparative analysis of clustering-based approaches for 3-D single tree detection using airborne fullwave lidar data. *Remote Sens.* **2010**, *2*, 968–989.
32. Ferraz, A.; Bretar, F.; Jacquemoud, S.; Gonçalves, G.; Pereira, L.; Tomé, M.; Soares, P. 3-D mapping of a multi-layered Mediterranean forest using ALS data. *Remote Sens. Environ.* **2012**, *121*, 210–223. [[CrossRef](#)]
33. Wang, Y.; Weinacker, H.; Koch, B. A lidar point cloud based procedure for vertical canopy structure analysis and 3D single tree modelling in forest. *Sensors* **2008**, *8*, 3938–3951. [[CrossRef](#)] [[PubMed](#)]
34. Li, W.; Guo, Q.; Jakubowski, M.K.; Kelly, M. A new method for segmenting individual trees from the lidar point cloud. *Photogramm. Eng. Remote Sens.* **2012**, *78*, 75–84. [[CrossRef](#)]
35. Polewski, P.; Yao, W.; Heurich, M.; Krzystek, P.; Stilla, U. Detection of fallen trees in ALS point clouds using a Normalized Cut approach trained by simulation. *ISPRS J. Photogramm. Remote Sens.* **2015**, *105*, 252–271. [[CrossRef](#)]
36. Yao, W.; Krzystek, P.; Heurich, M. 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.* **2012**, *123*, 368–380. [[CrossRef](#)]
37. Hu, B.; Li, J.; Jing, L.; Judah, A. Improving the efficiency and accuracy of individual tree crown delineation from high-density LiDAR data. *Int. J. Appl. Earth Observ. Geoinf.* **2014**, *26*, 145–155. [[CrossRef](#)]
38. Dai, W.; Yang, B.; Dong, Z.; Shaker, A. A new method for 3D individual tree extraction using multispectral airborne LiDAR point clouds. *ISPRS J. Photogramm. Remote Sens.* **2018**, *144*, 400–411. [[CrossRef](#)]
39. Pang, Y.; Li, Z.; Ju, H.; Lu, H.; Jia, W.; Si, L.; Guo, Y.; Liu, Q.; Li, S.; Liu, L.; et al. LiCHy: The CAF's LiDAR, CCD and hyperspectral integrated airborne observation system. *Remote Sens.* **2016**, *8*, 398. [[CrossRef](#)]
40. Wagner, W.; Ullrich, A.; Ducic, V.; Melzer, T.; Studnicka, N. Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner. *ISPRS J. Photogramm. Remote Sens.* **2006**, *60*, 100–112. [[CrossRef](#)]
41. Besl, P.J.; McKay, N.D. Method for registration of 3-D shapes. In Proceedings of Sensor Fusion IV. *Control Paradig. Data Struct.* **1992**, *14*, 586–607.
42. Glira, P.; Pfeifer, N.; Briese, C.; Ressel, C. A Correspondence Framework for ALS Strip Adjustments based on Variants of the ICP Algorithm Korrespondenzen für die ALS-Streifenausgleichung auf Basis von ICP. *Photogramm. Fernerkundung Geoinf.* **2015**, *4*, 275–289. [[CrossRef](#)]
43. Axelsson, P. Processing of laser scanner data—algorithms and applications. *ISPRS J. Photogramm. Remote Sens.* **1999**, *54*, 138–147. [[CrossRef](#)]
44. Adams, R.; Bischof, L. Seeded region growing. *IEEE Trans. Pattern Anal. Mach. Intell.* **1994**, *16*, 641–647. [[CrossRef](#)]

45. Sokolova, M.; Japkowicz, N.; Szpakowicz, S. Beyond accuracy, F-score and ROC: A family of discriminant measures for performance evaluation. *Australas. Joint Conf. Artif. Intell.* **2006**, *4304*, 1015–1021.
46. Goutte, C.; Gaussier, E. A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. *Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med.* **2005**, *51*, 345–359.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).