

Article

Adaptive Mean Shift-Based Identification of Individual Trees Using Airborne LiDAR Data

Xingbo Hu ¹, Wei Chen ^{2,3} and Weiyang Xu ^{4,*}

¹ School of Information Science and Technology, East China Normal University, Shanghai 200241, China; xbhu@ee.ecnu.edu.cn

² School of Geosciences and Info-Physics, Central South University, Changsha, Hunan 410083, China; lajiao.1225@163.com

³ East China Forest Inventory and Planning Institute, State Forestry Administration, Hangzhou 310019, China

⁴ College of Communication Engineering, Chongqing University, Chongqing 400044, China

* Correspondence: weiyangxu@cqu.edu.cn; Tel./Fax: +86-23-6510-3544

Academic Editors: Guangxing Wang, Erkki Tomppo, Dengsheng Lu, Huaqing Zhang, Qi Chen, Lars T. Waser, Randolph H. Wynne and Prasad S. Thenkabail

Received: 30 August 2016; Accepted: 3 February 2017; Published: 10 February 2017

Abstract: Identifying individual trees and delineating their canopy structures from the forest point cloud data acquired by an airborne LiDAR (Light Detection And Ranging) has significant implications in forestry inventory. Once accurately identified, tree structural attributes such as tree height, crown diameter, canopy based height and diameter at breast height can be derived. This paper focuses on a novel computationally efficient method to adaptively calibrate the kernel bandwidth of a computational scheme based on mean shift—a non-parametric probability density-based clustering technique—to segment the 3D (three-dimensional) forest point clouds and identify individual tree crowns. The basic concept of this method is to partition the 3D space over each test plot into small vertical units (irregular columns containing 3D spatial features from one or more trees) first, by using a fixed bandwidth mean shift procedure and a small square grouping technique, and then rough estimation of crown sizes for distinct trees within a unit, based on an original 2D (two-dimensional) incremental grid projection technique, is applied to provide a basis for dynamical calibration of the kernel bandwidth for an adaptive mean shift procedure performed in each partition. The adaptive mean shift-based scheme, which incorporates our proposed bandwidth calibration method, is validated on 10 test plots of a dense, multi-layered evergreen broad-leaved forest located in South China. Experimental results reveal that this approach can work effectively and when compared to the conventional point-based approaches (e.g., region growing, *k*-means clustering, fixed bandwidth or multi-scale mean shift), its accuracies are relatively high: it detects 86 percent of the trees (“recall”) and 92 percent of the identified trees are correct (“precision”), showing good potential for use in the area of forest inventory.

Keywords: mean shift; adaptive; bandwidth calibration; LiDAR; airborne laser scanning; point cloud; individual trees; identification; segmentation; forest

1. Introduction

In the past two decades, efforts have been made to decrease forest inventory costs by using data obtained from remote sensing technology. Airborne LiDAR (Light Detection And Ranging) is an active remote sensing technology that provides distance measurements between an aircraft (or any other platform) and the surface illuminated by the laser beam [1,2]. Due to its ability to generate 3D (three-dimensional) data with high spatial resolution and accuracy, the airborne small-footprint LiDAR is becoming a promising technique for modeling the forest canopies and thus for completing

several inventory tasks [3,4]. Numerous methods or algorithms have been proposed to identify or detect individual trees or tree crowns using airborne LiDAR data. Raster-based approaches utilizing the canopy height model (CHM) derived from raw LiDAR data are widely adopted for individual tree segmentation [5]. In a typical CHM-based scenario, a local maximum filter is used to detect treetops and then the individual tree crowns are delineated with a marker-controlled watershed segmentation scheme or a pouring algorithm [6–10]. Region growing, as a category of simple region-based segmentation methods, can be adopted for segmentation of 3D point cloud generated by the LiDAR [11] and it is the most common approach to segmenting trees in original LiDAR data [1,3,12]. Extracting individual trees directly from the original LiDAR data can also resort to clustering-based methods. One of the most popular clustering methods is the k -means algorithm, which aims to partition n observations into k clusters where each observation belongs to the cluster with the nearest mean by trying to minimize the overall sum of Euclidean distances of the points in feature space to their cluster centroids [13–15]. Most of these methods fall into one of the two categories: raster-based approaches or point-based approaches. Usually, raster-based approaches are more automatic and require fewer parameters than point-based approaches and thus have a leading edge in processing speed, but they tend to merge crowns in dense stands of deciduous trees or fail to detect understory trees in a multi-layered forest as the process of producing a CHM could lead to loss of information from different levels in the canopy [9]. Alternatively, point-based approaches work directly on the original airborne laser scanning (ALS) data (namely the “point cloud”, a discrete set of point locations) rather than rasterized images. This category of methods is preferable because it provides a promising way of producing relatively high accuracy by detecting smaller trees in the understory than CHM-based methods.

In recent years, mean shift-based clustering techniques have been applied in individual tree delineation using LiDAR data [16–19]. Mean shift is a powerful and versatile feature-space analysis technique and has found much interest in the image processing and computer vision community [20,21]. As it does not depend on any geometric model assumptions, its applications can be extended to the processing of unstructured point cloud data [22,23]. For example, Melzer explored the mean shift’s uses in LiDAR data processing and firstly utilized the mean shift procedure as a tool to automatically extract and reconstruct power lines from ALS point clouds [22]. Different from the region growing method and the k -means algorithm, mean shift does not require seed points and thus it is not sensitive to initializations. Moreover, the mean shift algorithm does not assume anything about the number of clusters and does not assume any predefined shape on data clusters, and the only parameter that has to be specified is the so-called kernel bandwidth for each attribute used. Therefore, it can handle arbitrarily shaped clusters. It is confirmed that the mean shift clusters can better indicate smaller or suppressed trees in the understory, which is not possible to do using other kinds of methods [19]. For simultaneous segmentation of vertical and horizontal structures of forest canopies, Ferraz et al. proposed a multi-scale mean shift method where modes are computed with a stratum-dependent kernel bandwidth whose values are increased as a function of the vegetation layer’s height [16,17]. In a multi-layered Mediterranean forest mainly covered with eucalyptus and maritime pine trees, they announced an overall detection rate of 69.3%. The specific detection rates were observed to decrease with the dominance position: from 98.6% for the dominant trees to only 12.8% for the suppressed trees [17]. A mean shift-based two-staged recursive computing scheme on the data sets in a defined complex multimodal feature space was employed to cluster points with similar vegetation modes together and an overall detection accuracy of more than 80% was reported in [18]. Mean shift clustering was also utilized to construct small parts of trees in a two-tiered 3D segmentation of forest point clouds into segments associated with regeneration trees and up to 70% of the regeneration coverage in a multilayered temperate forest could be correctly estimated [19].

For the mean shift clustering, calibration of the kernel bandwidth is very critical, since this parameter has a strong impact on the quality and speed of the decomposition and clustering. The mean shift procedure with a small kernel bandwidth is time-consuming but will lead to more distinct modes

(small basins of attraction, more and smaller objects), while a large bandwidth has a fast convergent rate but will aggregate small structures into larger ones (small number of modes with large basins of attraction) [16]. The calibration of an optimal value for the kernel bandwidth is actually a major challenge for the clustering techniques based on mean shift. Obviously, an adaptive selection scheme for the kernel bandwidth considering local structure information outperforms that using a single bandwidth for the whole feature space. Methods proposed in [16–19] are all fixed bandwidth mean shift procedures, where a single kernel bandwidth in a forest stand or a specific vegetation stratum is not adapted to heterogeneous crown sizes and thus could result in obvious under-segmentations or over-segmentations. Ferraz et al. designed a technique called 3D adaptive mean shift (AMS3D), in which three bandwidths of different sizes were applied to three forest layers: overstory, understory and ground vegetation, to decompose the entire point cloud into 3D clusters that correspond to individual tree crowns [17]. Yet, such an approach is essentially based on fixed bandwidth mean shift procedures and not suitable for complex forest environments because the boundaries of the vegetation layers are fuzzy and difficult to identify. A self-adaptive approach calibrating the kernel bandwidth as a function of a local tree allometric model that can adapt to the local forest structure was developed also by Ferraz et al. in [24]. It consists of two steps: in the first step, a limited number of easily recognizable individual trees are extracted from the LiDAR data to derive local tree allometry relating tree height to crown size and the second step uses the derived tree allometry model to provide the a priori information required by the AMS3D algorithm. However, detection of isolated individual tree samples would rely on a CHM-based and object-oriented approach, which requires complicated image analysis and processing. Furthermore, local allometric models established by linear regressions using estimated crown size data of individual tree samples might not be suitable for other types of forest.

In this paper, we propose an adaptive mean shift-based clustering scheme dedicated to the segmentation of 3D forest point clouds and identification of individual tree crowns. Our major contribution is in the development of a novel method to calibrate the adaptive mean shift procedure's kernel bandwidth by applying a "divide and conquer" processing strategy directly on the raw point cloud to improve computational efficiency. The kernel bandwidth of the adaptive mean shift procedure can be changed according to estimated crown sizes of distinct trees without relying on any local allometric model. To acquire approximate information about local canopy structure, some processing is a must: each test area is firstly divided into small spatial units, called "partitions" here, by performing a single-scale mean shift-based coarse segmentation, where an original small square grouping technique is used to accomplish rough partitioning of the region of interest; and then rough estimation of crown sizes based on another original 2D (two-dimensional) incremental grid projection technique is performed in each partition, providing a basis for dynamical calibration of the kernel bandwidth. Our proposed method is experimentally applied to small-footprint LiDAR data acquired in a field survey performed in an evergreen broad-leaved forest located in South China, and we validate its performance using 10 sample plots.

2. Study Area and Data

2.1. Study Area

The study zone is in the center of Dinghushan National Nature Reserve, which is situated in Zhaoqing, Guangdong Province, China. It is located in $112^{\circ}30'39''$ – $112^{\circ}33'41''$ E, $23^{\circ}09'21''$ – $23^{\circ}11'30''$ N and covers a total area of 1155 ha with the altitude (height above mean sea level) of 14.1–1000.3 m. This area is mainly covered by a low subtropical evergreen broad-leaved forest more than 400 years old and a coniferous and broadleaf mixed forest as well as montane evergreen shrubs and bushes. The trees under investigation in the site belong to the following species: *pinus massoniana*, *schima superba*, *castanopsis chinensis*, *gironniera subaequalis*, and *eucalyptus robusta*, etc.

A field survey was performed on several representative plots selected from the study zone to collect the ground truth data. The design pattern for sample plots and subplots is assumed to reflect

influences exerted on the forest ecosystem by terrain, soil, man-made environment and other factors, so each plot should contain at least two identical subplots. Figure 1 shows the study area and the forest plot of the Dinghushan National Nature Reserve. In this study, every sample plot was set up as a square area with the size of 100 m \times 100 m, which includes two subplots with the same size of 30 m \times 30 m. A total of 10 typical plots (i.e., 20 subplots) were selected to represent different regions of the forest under investigation. These plots can be divided into two types: single-tree species dominated (e.g., DHS05, DHS06, DHS10) and multi-tree species dominated (e.g., DHS01, DHS03, DHS04). Statistics data about tree species distribution for all sample plots with tree names, ages, percentages and counts are listed in Table 1.

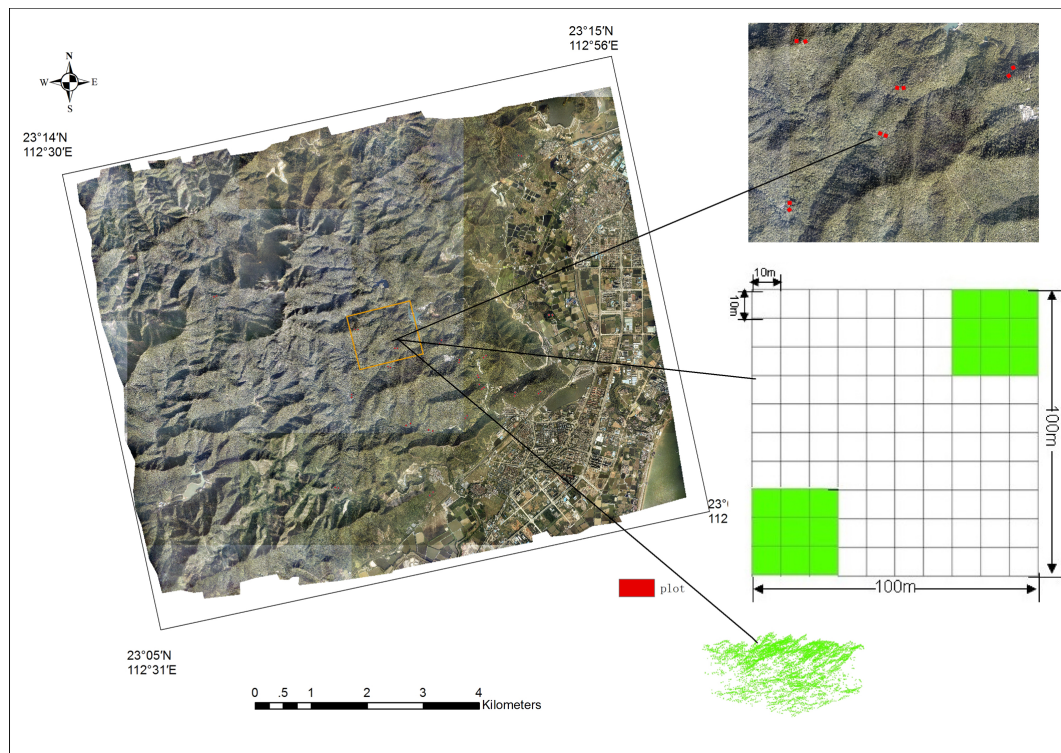


Figure 1. The study zone of Dinghushan National Nature Reserve, Guangdong Province, China.

Table 1. Statistics of tree species distribution for 10 forest sample plots.

Plot ID	Dominant Tree Species	Average Age	Count	Percentage (%)
DHS01	<i>Schima superba</i>	14	73	37
	<i>Castanopsis chinensis</i>	26	57	29
	<i>Pinus massoniana</i>	25	24	12
DHS02	<i>Schima superba</i>	18	133	50
	<i>Pinus massoniana</i>	18	24	9
	<i>Machilus velutina</i>	30	12	5
	<i>Eucalyptus robusta</i>	25	5	2
DHS03	<i>Rhododendron</i>	34	36	16
	<i>Schima superba</i>	25	33	15
	<i>Cryptocarya chinensis</i>	17	26	12
	<i>Craibiodendron kwangtungense</i>	16	25	11
	<i>Castanopsis chinensis</i>	31	19	8

Table 1. Cont.

Plot ID	Dominant Tree Species	Average Age	Count	Percentage (%)
DHS04	Castanopsis chinensis	27	27	17
	Schima superba	15	19	12
	Rhododendron	32	15	10
	Cryptocarya chinensis	20	14	9
	Artocarpus styracifolius	21	12	8
	Syzygium hancei	14	12	8
DHS05	Pinus massoniana	28	85	60
	Mallotus paniculatus	19	29	20
	Evodia lepta	12	16	11
DHS06	Schima Superba	15	172	78
	Castanopsis chinensis	20	17	8
	Pinus massoniana	32	16	7
DHS07	Gironniera subaequalis	24	40	29
	Castanopsis chinensis	28	26	19
	Schima superba	29	10	7
DHS08	Schima superba	21	111	59
	Pinus massoniana	37	42	22
	Eucalyptus robusta	20	7	4
DHS09	Castanopsis chinensis	28	91	48
	Machilus kwangtungensis	20	54	29
	Chinese aporosa	17	26	14
DHS10	Pinus massoniana	28	40	70
	Pinus elliotii	30	8	14

2.2. LiDAR Data

The full-waveform LiDAR data along with high-resolution CCD (charge-coupled device) image data were acquired through a flight experiment conducted on the Dinghushan study site in November 2012. The LiDAR system used in this experiment is the LiteMapper 5600, which integrates a Riegl LMS-Q560 laser scanner operating in a full-waveform mode with a wavelength of 1550 nm, a pulse repetition of 150 kHz, a pulse width of 3.5 ns and a beam divergence of no more than 0.5 mrad. The small-footprint LiDAR system can yield high-density data in both forms of digitally sampled full waveforms and discrete point clouds. The point cloud data studied in this paper were extracted from the waveform data by decomposing the full waveform rather than using the raw point cloud data delivered by the LiDAR system, since the full waveform data can yield more detailed and additional information about the structure of the illuminated surfaces and give an end user more control in the interpretation process of the physical measurement than the classical multi-echo point cloud data. The derived point cloud assumed the WGS84 coordinate system with UTM zone 49N and permitted an average footprint density of 15 pt/m². As the concern of this study is not the land but the forest, the vegetation points were firstly separated with the ground points using a commercial LiDAR processing software—TerraScan—and then further processed by normalizing all coordinate values in the z-dimension that denote the height of the vegetation objects by subtracting the ground points from the LiDAR point cloud [1].

2.3. Ground Truth Data

The ground truth data (including sample data and validation data) acquired in a field survey constitute reference data against which the LiDAR derived tree parameters can be compared and the accuracies of the tree detection approach can be assessed. To collect the coordinates of individual tree locations (namely the trunk positions), a total station assisted by satellite positioning with the real-time kinematic (RTK) technology was used to create a network of ground control points over the

study zone. The plot positions in the WGS84 coordinate system can be determined by a GNSS (Global Navigation Satellite System) RTK system (South S86) set up in open areas with an absolute accuracy of 10 cm. However, as poor visibility to the satellites due to the tree canopy will lead to serious loss of GNSS signals under a tree [25], satellite positioning-based surveys in forest conditions become unworkable. Thus, measurements of individual tree locations in a dense forest must resort to the total station that dispenses with the satellite signal. A South NTS-355R total station equipment, which can measure distances with an accuracy of $\pm(5 + 5 \text{ ppm} \times D)$ mm and angles with a rated accuracy of $5''$, was utilized to record the coordinate locations for trees in four of the 10 plots, which are denser than the other 6 plots. All trees of 2 m or taller were measured. To deal with the issue of poor visibility caused by trees and shrubs, this total station survey adopted a free-station positioning scheme that employed a two-sided linear intersection method. Such a technical scenario resulted in an average positioning error of approximately 0.06 m. The coordinate data of trees in the remaining six plots can be measured using only a South S86 RTK system. The accuracy of such a pure GNSS measurement was estimated to be 0.4–0.5 m.

3. Methodology

3.1. Mean Shift Theory

The mean shift is a recursive computing procedure using a non-parametric probability density estimator based on the Parzen window kernel function [21]. When used for clustering, mean shift treats the feature space as an empirical probability density function and considers the input set of points as sampled from the underlying probability distribution. Accordingly, clusters are defined as areas of higher density than the remainder of the data set and dense regions in feature space correspond to the modes (or local maxima) of the probability density function. The aim of the mean shift procedure is to move each data point to the densest area in its vicinity by iteratively performing the shift operation based on kernel density estimation. Eventually, the points will converge to local maxima of density, or that is to say, mean shift will associate each point with the nearest peak of the surface reflecting the data set's probability density distribution. Mean shift is essentially a hill climbing algorithm which involves shifting a kernel iteratively to a higher density region until convergence. The detailed mathematical derivation of the mean shift procedure is given in the Appendix, or refer to [17].

Mean shift works by placing a kernel on each point in the data set. The kernel, a function that determines the weight of nearby points for re-estimation of the mean, is iteratively shifted with a bandwidth (namely the diameter of the hyper-sphere) to a denser region until it converges to a stationary point. Choice of the kernel bandwidth— h is very critical. Many adaptive or variable bandwidth selection techniques have been proposed for mean shift with regard to 2D imagery analysis. There is a variable bandwidth technique that imposes a local structure on the data to extract reliable scale information by maximizing the magnitude of the normalized mean shift vector [26]. Separability in feature space or local homogeneity can be exploited to adaptively select bandwidth parameters for remote-sensing image classification and object recognition [27]. The local scale and structure information of the neighborhood around individual samples can also be utilized to calibrate the kernel bandwidth in an adaptive mean shift procedure to find arbitrary density, size and shape clusters in remote sensing imagery [28].

3.2. Methodology Workflow

In this paper, we develop an adaptive mean shift scheme to segment the forest point clouds into 3D clusters and properly identify and extract individual tree crowns over a multi-layered forest in China. The mean shift procedure builds upon the concept of kernel density estimation, where the kernel is a weighting function to estimate the probability density function of a random variable. Extraction of individual trees from LiDAR data utilizing the mean shift procedures implies finding modes in a huge data set by performing analysis of a feature space, which is constituted by feature

vectors extracted from the the original 3D point cloud data. Definition of the feature space is an issue in the mean-shift based clustering. As the tree in a forest is usually modelled as a porous medium and a linear structure, very much different from an object with piece-wise continuous surfaces [29], surface features—such as texture, local curvature and normal vector—will fail to function and the only feature that can be exploited is each point’s spatial location, whose mutual distance denotes spatial coherence and connectivity. Similar to the work reported in [17,24], pure spatial 3D coordinates are adopted to constitute the feature space in this study.

Using a single or multi-scale kernel bandwidth over the entire space is not optimal for the analysis of forest environments, because one or several global values might result in under-segmentations (a cluster assigned to more than one field measured trees) or over-segmentations (a single tree segmented into several LiDAR clusters). An approach for adaptive calibration of kernel bandwidth should be developed to deal with complex forest environments. We realize that the kernel bandwidth in the spatial domain should reflect the size of each tree crown. However, it is difficult to acquire exact quantitative information about each individual tree prior to the clustering analysis. Since the bandwidth calibration is critical, in this paper, we focus on designing a novel method to calibrate the kernel bandwidth in an adaptive mean shift-based computational scheme of 3D forest point cloud segmentation, which does not rely on a rasterized digital model such as CHM or canopy maxima model (CMM), but works directly on the original point cloud data and requires no expensive computation. Considering the time complexity of the mean shift algorithm is $O(TN^2)$ using big O notation, where T is the number of iterations and N denotes the data set’s size, a “divide and conquer” strategy is adopted in this method to reduce the computational load by partitioning a large area into smaller ones and perform the standard mean shift procedure in the small area.

Taking into account all necessary mathematical processing used in this study, we present a flowchart of methods used in this study, as illustrated in Figure 2, where the ellipses represent data or parameters and the rectangles represent processes or calculations. From the point of view of image segmentation, our proposed mathematical scheme is mainly a two-staged sequential segmentation procedure: coarse segmentation and fine segmentation. The single-scale mean shift-based area partitioning can be deemed a process of coarse segmentation while the adaptive mean shift-based clustering implements a fine segmentation process. The preprocessing step has been mentioned in Section 2.2 and other processing steps will be described in detail in the following part of this section.

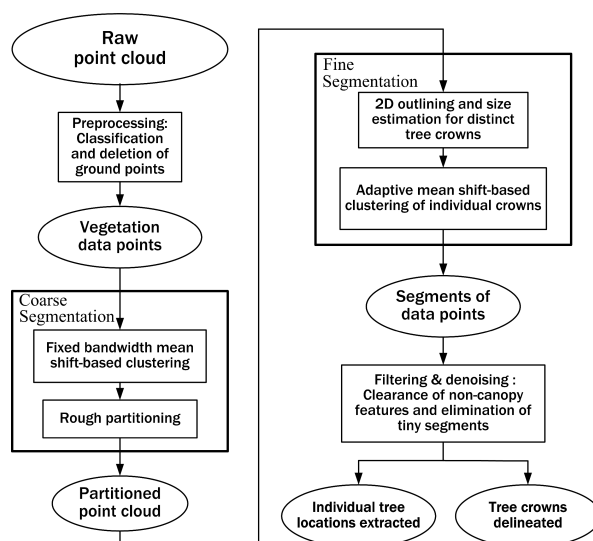


Figure 2. Flowchart of methods used in this study.

3.3. Rough Partitioning of Test Area

Due to the huge amount of point cloud data, it is computationally intensive to analyze all the points for the whole test area in one step. Therefore, the first stage of the two-staged sequential segmentation is to partition the 3D space over the forest area under test into many heterogeneous vertical spatial units with irregular column-like shapes, each of which includes at least one actual tree. The sizes of the spatial units are much smaller than that of the original test area and thus when we perform a more thorough analysis on these small spatial units, which implies more complex mathematics, the total computational load could be reduced to a range that can be affordable. The aim of the rough partitioning is to obtain rough information about tree distributions by performing coarse segmentation and draw a sketch of the forest for each test subplot, where local structure information can be extracted to provide the basis for the subsequent variable bandwidth mean shift procedures. This processing stage consists of two steps: single-scale mean shift clustering and cluster-based spatial partition. The basic idea behind the methods is to divide a portion of the data set into several clusters, which can reflect the basic structure of one or several trees, with a small computational cost and then expand these clusters to the whole data set by exploring the spatial neighborhood relation between data points. Not the entire LiDAR data set is used in the two steps. The whole data set is divided into three parts according to the nature of returned signals: the first pulse data, the last pulse data and the intermediate pulse data. As the intermediate pulse data depict the internal spatial structures of tree crowns or lower vegetation, it can be used alone in the first step. The data contained in LiDAR return's first echoes mainly reflect the information about the vegetation surface and they can be employed to estimate the maximum tree height and accomplish the rough partitioning in the second step.

In the first step, a fast mean shift procedure is utilized to achieve the goal of rough partition. In order to reduce computing expense and accelerate processing speed, only a part of the data set, the intermediate pulse data, is extracted to constitute the input set of points on which the mean shift segmentation is performed. For the same purpose of speeding up the process, the simplest kernel function—the uniform (or flat) kernel, which implies that all observations in the neighborhood of a data point are given the same weight—is used in this processing step. Thus, mean shift-based clustering analysis is performed in the three-dimensional Euclidean space using a uniform kernel function:

$$g_u\left(\left\|\frac{\mathbf{x}}{h_{fix}}\right\|^2\right) = \begin{cases} 1 & \|\mathbf{x}\| \leq h_{fix} \\ 0 & \text{else,} \end{cases} \quad (1)$$

where, \mathbf{x} denotes the data point and h_{fix} is a single-scale kernel bandwidth or window size. Owing to the lack of any information about forest structures in advance, a fixed kernel bandwidth setting is adopted in this stage. We determine the fixed value for the mean shift's kernel bandwidth according to the maximum tree height for a particular subplot, which can be estimated by finding N_{hp} highest points (considered as tree tops) in the LiDAR data set preserving just the first returns and then taking the average of these points' elevation values. The value of N_{hp} relies on the density of the raw point cloud. According to our experiments, a value of 30 is the ideal value for a data set with an average of 15 points per m^2 . Because a single tree in the woods, with smaller crown size than the one growing alone, usually assumes a high and straight form and its aspect ratio (i.e., the ratio between the crown diameter and height of the tree) is far less than 1, the kernel bandwidth should be set to a value less than the maximum tree height:

$$h_{fix} = Q \times \hat{H}_{\max}, \quad (2)$$

where, \hat{H}_{\max} is the estimate of the maximum height for trees in a specific subplot and Q is a positive constant. To avoid over-segmentation that implies segmenting a single tree into several segments, Q should be a number of less than 1 and its specific value will be discussed in Section 4.2.

The single-scale mean shift procedure mentioned above would yield many segmented clusters of points extracted from the LiDAR's intermediate pulse data. Figure 3a shows an example of these

mean shift clustering results for a small selected area ($10\text{ m} \times 10\text{ m}$). As can be seen from this map, the specific middle pulse data set has been divided into several distinct clusters rendered in different colors. However, it is not enough to perform 3D space division of the forest area with only these clusters because the contour profile of an actual tree crown is larger than the shape from a cluster of the middle pulse data point. The first pulse data from the vegetation surface should be added to complete the rough partitioning in the subsequent step. This step is to divide the horizontal plane into small squares of identical sizes, assign each data point (either the intermediate pulse data or the first pulse data) to a specific square and finally group these squares into different partitions based on the mean shift clustering results. Such a technique, called the small square grouping technique in this paper, is illustrated in Figure 3b. The data sets within the region of interest, formed by both the first and intermediate returns with points lower than 2 m excluded, are all projected onto the horizontal plane, which is divided into identical squares with sizes of $0.5\text{ m} \times 0.5\text{ m}$. Each individual square might contain one or several points from the first returns and/or the middle echoes. For a specific square, if there is no point in it, leave it alone. If the square contains at least a middle pulse data point, it will be associated with the cluster that the point belongs to. If there is only the first pulse data in the square, assign the nearest middle pulse point to it and then associate it with the corresponding cluster.

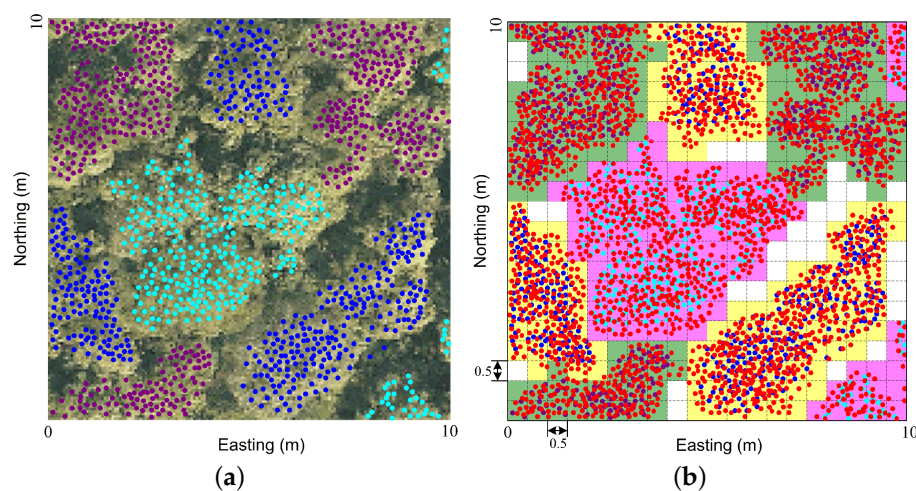


Figure 3. An example of the rough partitioning for a portion of a typical subplot (DHS0101). This is a downward-looking view of the area. (a) Fixed bandwidth mean shift clustering results overlaid onto a digital orthophoto map (DOM). Spatially adjacent clusters are rendered in different colors. (b) Rough partitioning of the forest area based on the small square grouping technique. Connected squares marked in the same color belong to the same partition.

3.4. Estimation of Distinct Tree Crown Sizes

The above described rough partitioning process adopts a fixed bandwidth mean shift algorithm, whose global kernel bandwidth is far greater than the canopy size, and thus obvious under-segmentations would inevitably come into being. Such under-segmentations can be reduced by using a variable bandwidth mean shift algorithm in the subsequent processing stage, which automatically calibrates the kernel bandwidth to the local spatial structures of the forest canopy. Thus, a more sophisticated mean shift clustering scheme with an adaptive kernel bandwidth optimized for local canopy structure should be designed to deal with complex patterns in each coarsely segmented forest partition. The adaptive mean shift procedure necessitates information about the local canopy structure for each partition, which provides a basis for adaptive determination of the kernel bandwidth. There exists a method to derive tree sizes from linear fitting equations to locally calibrate the bandwidth model [24]. However, since this method assumes that there is a linear relationship between tree height and both crown depth and crown width, which is not the case for many forest scenarios, it is essentially

an estimation of tree allometry. Actually, it is impossible to accurately extract data of individual tree crown sizes for a particular forest area prior to complete clustering of individual trees.

In this section, we propose a novel method to roughly estimate the size data of distinct tree crowns adopted in the second segmentation stage, which works directly on the point cloud and does not resort to raster images and sophisticated image processing techniques. The goal of this processing step is to identify distinct tree crowns in a partition, delineate their approximate 2D profiles and determine numerical values as the estimates of these tree crown sizes. Considering that the tree crowns in the study area assume different shapes (e.g., cone, umbrella, spindle, irregular shape, etc.), for the sake of processing simplicity, we model these crowns as spheres or hemispheres. Thus, the estimation problem can be reduced to find the maximum approximate intercepting area for each distinct tree crown and determine its effective diameter. The first returns are selected from the data points, and as they are most likely to be from tree canopy surfaces, these first pulse data are utilized to perform such an estimation to minimize the computational cost. This process is divided into two steps: 2D outlining of distinct tree crowns and determination of the effective crown diameter.

Firstly, rough boundaries of all distinct tree crowns in a partition are outlined by a simple region growing method based on incremental grid projection that we put forward particularly for this study. As shown in Figure 4, suppose there is a plane (or screen) horizontally suspended over a partition of the forest area under study, project all the first pulse data points downward onto the plane and the 3D spatial distributions of the tree crowns will be reduced to 2D horizontal features on the plane. The projected 2D clusters on the horizontal plane represent the distribution of tree crowns at a certain height level and they will increase as the horizontal plane moves from top to bottom. The elevation value (z coordinate) of the horizontal gridded plane decreases in a uniformly discrete mode, which implies shifting the projection plane at equal spacing from top to bottom. All of the points above the height-variable plane in the first return data set are projected downward onto this plane. It is possible to get the 2D horizontal distribution of individual tree crowns at different height levels when more than one 2D horizontal projections are available. Define Z_{\min} and Z_{\max} as the minimal and maximal coordinate values respectively in the z -axis of the axis-aligned minimum bounding box for the data set, and I as the total number of shift layers (each layer corresponds to an iteration of the region growing process), and then the elevation value for the projection plane used in the i -th shift can be determined as

$$Z_i = Z_{\max} - \frac{i}{I}(Z_{\max} - Z_{\min}), \quad i = 1, \dots, I. \quad (3)$$

To speed up the whole estimation process, a small positive integer not exceeding 10—say, six or eight—should be adopted as the value of I in this study.

In order to simplify further analyses, the plane is divided into small grids by a raster net and points within the partition are mapped into grid cells, which can be regarded as pixels of an image. A non-empty cell (containing one or more projected points) is marked as a set pixel while an empty cell corresponds to an unset pixel. Thus, a series of 2D horizontal binary images can be generated. The clusters on the horizontal projection image at each layer represent the distribution of tree crowns in the correspondent height level. Therefore, a distinct tree crown should be visible at the same location of several vertical neighboring layers. The basic concept of the simple region growing method is to trace the presence of crown outlines on the binary images from top to bottom through projection images at different height levels. Outlines of the crown at the upper layers will present a cluster feature on the projection image, thus the task of detecting a distinct tree crown turns to an issue of extracting cluster feature of the treetop and tracing the crown outline from top to bottom. Figure 5 depicts an example of such a region growing process, whose aim is to obtain an approximate intercepting area in terms of grid cell numbers for each distinct tree crown. The axis-aligned 2D horizontal plane is divided into grid cells with identical size of $0.25 \text{ m} \times 0.25 \text{ m}$. In each iteration, points no lower than the elevation of the moving projection plane are vertically mapped to their corresponding grid cells, and each non-empty cell might be determined as a new seed, if the distance from the existing clusters exceeds the threshold value of $(Z_{\max} - Z_{\min})/I$, or else it would be added into an existing group based

on spatial adjacency using a morphological analysis. The last iteration completes the region growing process and the area enclosed by the outmost grid cells in a region will be considered as the estimate of the intercepting area for a particular tree crown. Let N_{GC} be the number of grid cells contained in a grown region, and the approximate intercepting area for the corresponding tree crown will be $0.25 \times 0.25 \times N_{GC} = \frac{1}{16} N_{GC}$, whose unit is m^2 . Thus, the effective crown diameter for a particular tree can be estimated as

$$D_{eff}(r) = \frac{1}{2} \sqrt{\frac{N_{GC}(r)}{\pi}}, \quad (4)$$

where r denotes the index number of a particular distinct crown region.

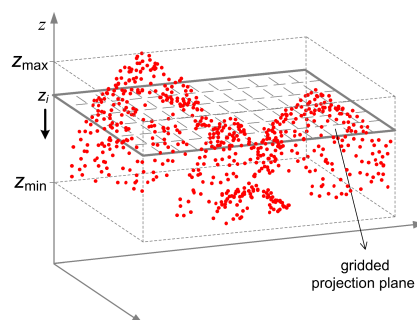


Figure 4. Downward projection of the first pulse data points onto an elevation-variable horizontal gridded plane.

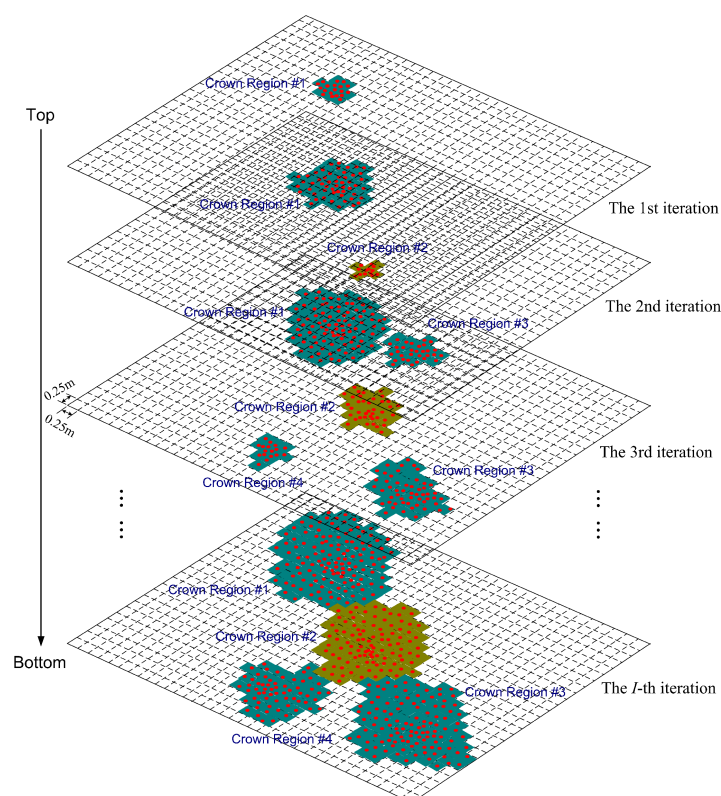


Figure 5. An example of the region growing based on a 2D (two-dimensional) incremental grid projection.

3.5. Adaptive Mean Shift-Based Clustering

With the estimated tree crown diameters, which reflect the local canopy structure in a forest area of interest, we can perform an adaptive mean shift analysis on an entire data set, using a combination of the first, intermediate and last pulse data points whose elevations are above 2 m, in the three-dimensional geometrical space for each partition. The aim of this adaptive mean shift procedure is to cluster points with similar modes (in the forestry scenario, the term “mode” has close relations with each single tree’s location and structure) together and delineate individual tree crowns. Euclidean distance between each point and a stationary point (represent some mode) is used as the measure of a tree’s spatial and structural properties. As mentioned previously, due to different crown shapes assumed by different tree species in the study area, we use a spherical kernel instead of a cylinder-shaped kernel proposed in [17]. Such a unified spherical kernel, although at the expense of a little modeling performance, will facilitate mathematical processing. The uniform kernel is no longer used again because it is reasonable to attribute more weights to points that are near to that mean shift point and less weight to distant points at the stage of fine segmentation. As the Gaussian or normal function has excellent mathematical properties and similar characteristics with the points from a tree crown in statistical distributions and peak shapes, we choose the truncated, multivariate Gaussian function as the kernel (or window) function

$$g_n\left(\left\|\frac{\mathbf{x}}{h_{var}}\right\|^2\right) = \exp\left(-\frac{1}{2}\left\|\frac{\mathbf{x}}{h_{var}}\right\|^2\right), \quad (5)$$

where h_{var} is the variable kernel bandwidth that will adapt to the estimate of each crown diameter.

Let $\mathbf{X} = \{\mathbf{x}_i\} (i = 1, \dots, N)$ be the input data set of the three-dimensional feature space. Run the mean shift computation on the data set \mathbf{X} for each data point \mathbf{x}_i and each iterative process will converge towards their corresponding stationary points or modes. Let $\mathbf{Y} = \{\mathbf{y}_i\} (i = 1, \dots, N)$ denote the data set of mean shift points, i.e., values at point \mathbf{x}_i after its mean shift iteration. The detailed mean shift procedure proposed for the identification and delineation of individual trees is described as follows:

- I. Specify a value for the allowable error— ε (set to 0.0025). Gaussian kernels are also associated with data points in the feature space.
- II. For each point \mathbf{x}_i , execute the following computations:
 - Initialization: associate a mean shift point \mathbf{y}_i with the point \mathbf{x}_i , and initialize it to coincide with that point. Let $t = 0$ and $\mathbf{y}_i(0) = \mathbf{x}_i$, where t denotes the current number of mean shift iterations.
 - For $0 \leq t \leq c$ (c represents the word “convergence”), repeat the following three steps until convergence.
 - Map the point $\mathbf{y}_i(t)$ to the horizontal gridded projection plane defined in Figures 4 and 5 to see which distinct crown region it belongs to. If this point is not in any region, associate it with the crown region nearest to it. Assume the corresponding crown region has the index number of r , and then determine the variable kernel bandwidth as follows:

$$h_{var} = B \times D_{eff}(r), \quad (6)$$

where B is a constant coefficient of greater than 1 and less than 2.

- Calculate the mean shift vector:

$$\mathbf{m}_{h,G}[\mathbf{y}_i(t)] = \frac{\sum_{j=1}^n \mathbf{y}_j(t) g_n\left(\left\|\frac{\mathbf{y}_j(t) - \mathbf{x}_i}{h_{var}}\right\|^2\right)}{\sum_{j=1}^n g_n\left(\left\|\frac{\mathbf{y}_j(t) - \mathbf{x}_i}{h_{var}}\right\|^2\right)} - \mathbf{y}_i(t). \quad (7)$$

- Update the mean shift point:

$$\mathbf{y}_i(t+1) = \mathbf{y}_i(t) + \mathbf{m}_{h,G}[\mathbf{y}_i(t)], \quad (8)$$

stopping when $\mathbf{m}_{h,G}[\mathbf{y}_i(t)]$ is less than the specified value of ε .

- Set $\mathbf{y}_i = \mathbf{y}_i(c)$.

III. Merge all modes whose mutual distance is less than the specific value of the variable kernel bandwidth, i.e., closer than h_{var} , by concatenating the basins of attraction of the corresponding stationary points, to produce the clusters $\{\mathbf{C}_p\} (p = 1, \dots, M)$.

3.6. Removal of Non-Canopy Features

Application of the above described mean shift procedures in the forest point clouds would result in different types of vegetation segments of a sample area: tree canopy, shrubs, bramble, weeds and grasses, etc. Since our study focuses on the identification and delineation for trees, the non-canopy objects are the unwanted surface features whose data points should be recognized, marked and then cleared. Considering that the non-canopy features lie in the undersized vegetation strata, we could identify these objects by deciding whether their heights fall in a predefined height range whose upper limit is a preset threshold value— H_{SV} . The points belonging to these non-canopy features are not taken into account to final decisions. This step would remove the scrubby layers which are usually denser than the higher vegetation layers. Tiny segments containing less than a given number of data points, M_F , should also be eliminated. These filtering operations would improve the ultimate detection results by eliminating possibly detrimental influences of non-canopy data points.

Clustering results are presented to validate the effectiveness of the above described adaptive mean shift scheme. Figure 6 shows two different views (from top and side, respectively) of a segmented point cloud in a typical subplot (DHS0201) after two round, mean, shift-based iterative computations, with spatially adjacent point clusters rendered in different colors.

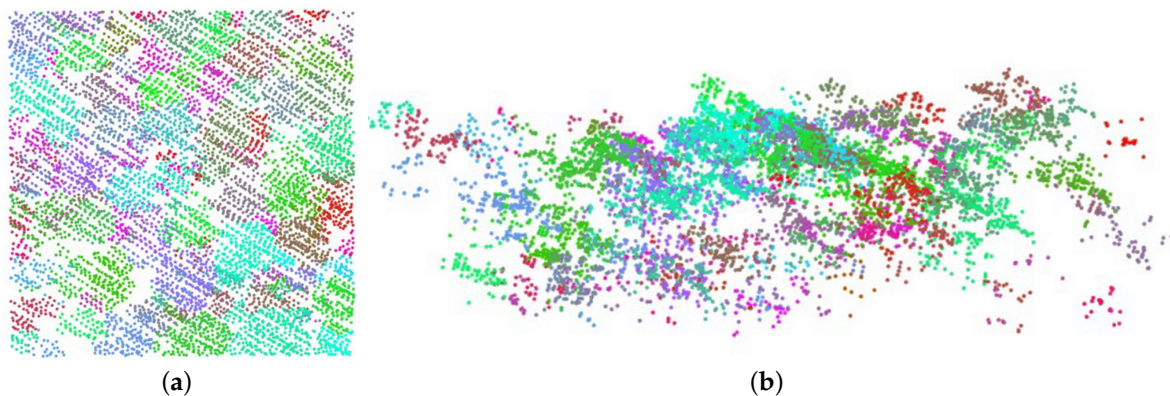


Figure 6. An example of the ultimate clustering results for the point cloud in the test subplot of DHS0201. (a) Top view. (b) Side view.

3.7. Implementation

The overall mathematical processing flow depicted in Figure 2 is performed on a C/C++ software platform for 3D point cloud data processing that we specifically developed for research purposes based on the Point Cloud Library (PCL), which is an open-source library of algorithms for point cloud processing tasks and 3D geometry processing, such as occur in three-dimensional computer vision [30]. Generic techniques for the classification, filtering, segmentation, interpretation, modeling, and visualization of 3D point clouds are implemented with the methods and classes contained in PCL's

modules. The computing procedures for the mean shift-based clustering in two segmentation stages are accomplished by calling a C++ wrapper which implements a standard mean shift algorithm.

We have run the LiDAR data processing software on an HP workstation computer with a 2.5 GHz Intel Core i7-6700 processor, 16 GB of RAM and 512 GB of solid state drive. The time it takes to analyze a test subplot (an area of 30 m \times 30 m with 8000–15,000 laser points) using this software is about 8 s. Considering that the standard mean shift algorithm's running time increases exponentially with the size of the whole data set, if we apply the proposed adaptive mean shift-based scheme directly to a much larger forest area of 1000 m by 1000 m, the computation time will be more than 20 h, which is extremely time consuming and inefficient. Therefore, in order to improve the computational efficiency, it is necessary to divide the area of 1000 m \times 1000 m into 100 plots with the same size of 120 m \times 120 m, each of which includes four 10-meter-long buffer areas on four sides to counteract edge effects. With such a division process, the total computation time taken to run this mean shift algorithm for such a large area of 1000 m \times 1000 m will be drastically reduced to less than 4 h.

4. Results

4.1. Parameters and Sensitivity Analysis

The proposed scheme requires proper values for the kernel radius h of the adaptive mean shift algorithm and the control parameters of the coarse segmentation and rough estimation. Theoretically, the value of Q (a positive number of less than 1) will influence the coarse segmentation, the values of I (a positive integer greater than 1) and B (a real number of greater than 1 and less than 2) will influence the adaptive calibration of the kernel bandwidth, and in turn they will influence the ultimate performance of this method. We demonstrate the performance of tree detection with these control parameters of the method. The results of sensitivity analysis, including both the numbers of truly detected trees, which are related to omission errors, and the falsely detected tree numbers reflecting commission errors, from the variations of the three parameters are shown in Figure 7.

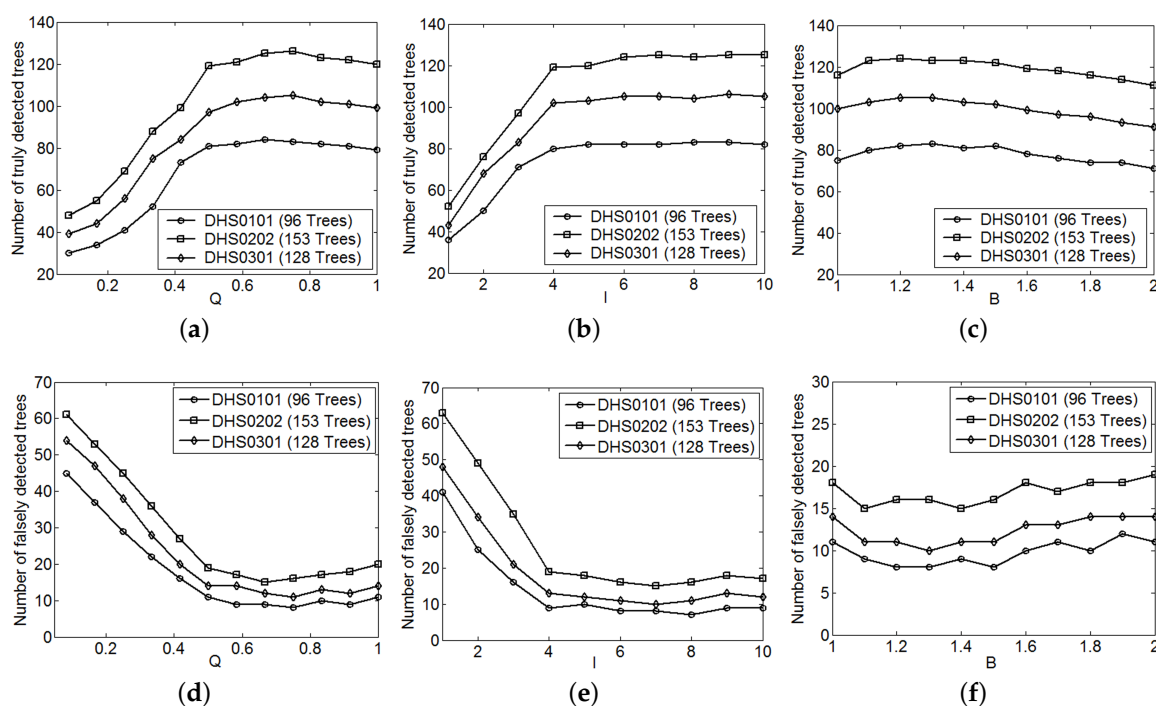


Figure 7. (a–f) Sensitivity analysis for three different data sets. (a,d) With respect to Q ($I = 6$ and $B = 1.2$). (b,e) With respect to I ($Q = 2/3$ and $B = 1.2$). (c,f) With respect to B ($Q = 2/3$ and $I = 6$).

Through experimental trials, we have found that as long as $Q \geq 1/2$, $I \geq 4$ and $1.1 \leq B \leq 1.5$, the method will not be sensitive to the change of parameters. Besides, tree species distribution has little influence on the selection of these parameters. The two parameters required in the filtering step— H_{SV} and M_F —are determined by the forest’s specific conditions and the average laser point density. Method-specific parameters used in these experiments are listed in Table 2. These parameters are uniformly applied to all sample plots under test.

Table 2. Experimental parameters.

Parameter	Value
Coefficient for fixed kernel bandwidth determination— Q	2/3
Iteration number of this region growing process— I	6
Coefficient for adaptive kernel bandwidth determination— B	1.2
Maximum height of scrubby vegetation— H_{SV}	3 m
Maximum number of points of tiny fragment— M_F	50

4.2. Performance Evaluation

Experiments are conducted over all trees in 10 sample plots (or 20 test subplots) to validate and evaluate the performance of our proposed adaptive mean shift-based approach for identification and delineation of individual tree crowns. The performance will be evaluated by means of both qualitative analysis and quantitative calculations. The approximate positions of individual trees acquired through these mathematical procedures are overlaid onto a digital orthophoto map (DOM) with known tree locations of a typical measurement area (Subplot ID: DHS0401) to see if the detected tree locations coincide with the regions of actual trees. The information about actual tree locations comes from the reference data collected through field measurements conducted in the specific area. The approximate tree positions are estimated by finding each cluster’s stationary points via gradient ascent, which represent modes of the underlying density function. The reference trees (or field measured trees) and the estimated trees (or LiDAR determined trees) are paired by assigning each reference tree with the closest cluster, using both distance and tree height as matching criteria. A cluster is linked with a reference tree provided that (i) the distance to the reference tree is smaller than 60% of the mean tree distance within the subplot; and (ii) the difference between the derived tree height and the height of the reference tree is smaller than 15% of the top height of the subplot. A LiDAR tree can be matched with more than one reference tree, compensating the effect of “tree clusters”, with a dominant overstory tree being surrounded by smaller suppressed trees. The results of this matching can be seen in Figure 8. Figure 9 shows matching outcome in other two test subplots. These figures indicate that the spatial fitness between the point clusters and reference trees is fairly good. However, if the mutual distance between two adjacent point clusters is less than the threshold, which is the estimated value of local crown diameter, the two clusters cannot be distinguished and will be identified as a single cluster.

To quantitatively assess the accuracy of the detection results, we compare the identified trees with the reference trees in 20 testing plots. As mentioned in [3], the measures of perfect segmentation, under-segmentation and over-segmentation can be indicated by true positive (TP), false negative (FN) and false positive (FP), respectively. If a cluster is assigned to more than one reference tree, the reference tree with shortest distance to the estimated tree position is considered a true positive and others as false negative trees. Clusters linked with no reference tree are classified as false positive (FP). We can evaluate the detection accuracy in terms of “recall” ($r = TP / (TP + FN)$), which indicates the tree detection rate, and “precision” ($p = TP / (TP + FP)$), which indicates the correctness of the detected trees [31].

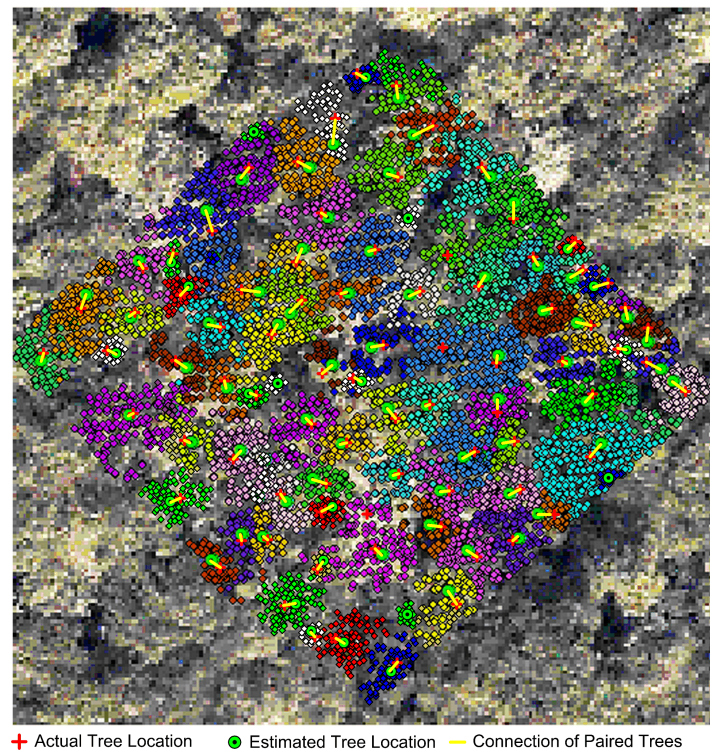


Figure 8. Digital orthophoto map overlaid with detected approximate positions of individual trees for a representative test subplot (DHS0401). A uniform colour is used to indicate the set of LiDAR (Light Detection And Ranging) points that are clustered as a given tree inside the measurement area.

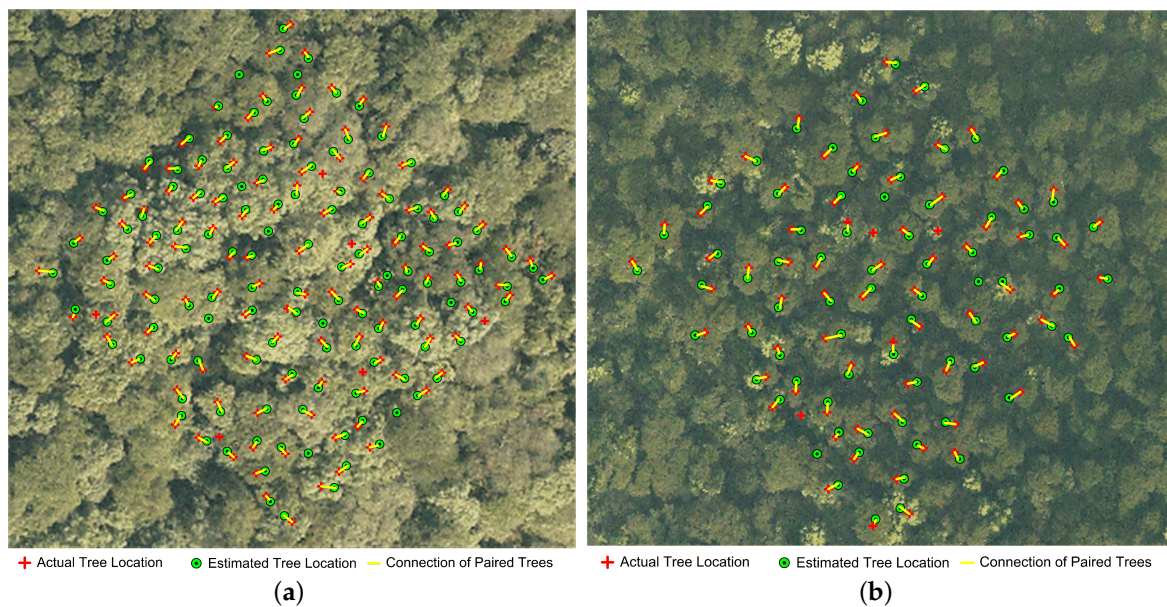


Figure 9. Matching of the reference tree positions and the estimated positions for two different test subplots. (a) A dense single-tree species dominated subplot (DHS0602). (b) A sparse multi-tree species dominated subplot (DHS0702).

Table 3 shows the accuracy assessments for trees located in different forest storeys within six test subplots with ID of DHS0101, DHS0102, DHS0201, DHS0202, DHS0301 and DHS0302. Similar to the results reported in [17], the “recall” and “precision”, which indicate the detection accuracy,

also decrease with the forest storey, but nearly 50 percent of the suppressed trees can be detected and the overall detection rate (“recall”) reaches 86 percent.

Table 3. Accuracy assessments for different tree classes in six test subplots.

Class of Trees	Reference Trees	Identified Trees	TP	FP	FN	Recall (<i>r</i>)	Precision (<i>p</i>)
Dominant trees	234	232	230	2	4	0.983	0.991
Codominant trees	209	205	201	4	8	0.962	0.980
Intermediate trees	151	138	117	21	34	0.775	0.848
Suppressed trees	93	67	45	22	48	0.484	0.672
Total	687	642	593	49	94	0.863	0.924

TP—true positive, FP—false positive, FN—false negative.

We assess the overall detection performance of the proposed method for all trees in 10 test plots (i.e., 20 subplots) by comparing the results, in terms of the indicators of “recall” and “precision”, obtained with several different approaches: one is our proposed mean shift-based algorithm, the others are existing point-based methods including a region growing approach, a *k*-means clustering approach and a mean shift-based approach described in [17]. The region growing approach adopts the same idea as put forward in [1]: the clusters of LiDAR points associated with each seed point are progressively grown by an incremented labelling procedure and the step size is set to 0.25 m, roughly the average LiDAR point spacing. We find the seed points using the raw point data rather than rasterized images by searching for the highest LiDAR point in the data set inside a circular region with the radius *R* [1]. Different sizes of *R* are empirically tested to determine an appropriate value for our particular study area and the interval of $1.25 \leq R \leq 2$ is found to be a reasonable region in which optimal performance can be achieved. For the sake of comparison with other methods, *R* = 1.5 m is adopted as the appropriate size for the search radius. The *k*-means algorithm used in this experiment clusters the data in an iterative process divided into two steps: batch updates and online updates, implemented as described in [14]. Considering that its performance is heavily dependent on the numbers and locations of seeds, we do not use seeds randomly chosen from the entire data set but adopt external seed points, generated in the same way as mentioned above, for initialization of the iterative process. For optimal detection accuracy, *R* = 1.5 m is chosen as the proper value for the search radius. The mean shift algorithm presented in [17] adapts the kernel bandwidth to the thickness of each stratum and uses a single bandwidth for a specific vegetation stratum. Based on our data set, we do not use an asymmetric kernel but an isotropic kernel in this experiment. The bandwidth values we used are given below: 2.7 (understory) and 5.9 (overstory) for plots of DHS01 and DHS02; 3.2 (understory) and 8.3 (overstory) for plots of DHS03, DHS04, DHS05 and DHS06; 2.9 (understory) and 6.6 (overstory) for plots of DHS07 and DHS08; 2.3 (understory) and 4.1 (overstory) for plots of DHS09 and DHS10.

Table 4 lists the data reflecting detection results of our proposed method as well as the existing mean shift-based method proposed in [17] for each test subplot. The overall detection performance (namely the indicators of “recall” and “precision”) of our proposed approach is compared to other point cloud-based approaches, as shown in Table 5. The CHM-based method (multi-resolution segmentation using watersheds) adopted in [32] is also implemented in this comparative experiment to provide a reference. From the data listed in Tables 4 and 5, it is visible that the proposed adaptive mean shift-based method can produce greater values for the overall detection indicators (“recall” and “precision”) than other methods, which indicates that its detection performance outperforms the CHM-based and point-based approaches reported in representative academic literature.

Table 4. Detection result comparison of mean shift-based approaches for all trees in 20 test subplots: our proposed adaptive mean shift algorithm vs. the existing one proposed in [17].

Subplot ID	Reference Trees	Our Proposed Method				Existing Method Presented in [17]			
		Identified Trees	TP	Recall	Precision	Identified Trees	TP	Recall	Precision
DHS0101	96	90	82	0.854	0.911	86	77	0.802	0.895
DHS0102	101	93	85	0.842	0.914	89	80	0.792	0.899
DHS0201	111	104	93	0.838	0.894	98	84	0.757	0.857
DHS0202	153	140	124	0.810	0.886	128	113	0.738	0.883
DHS0301	128	116	105	0.820	0.905	107	94	0.734	0.878
DHS0302	98	93	85	0.867	0.914	89	81	0.826	0.910
DHS0401	87	84	79	0.908	0.940	78	71	0.816	0.910
DHS0402	71	70	64	0.901	0.928	65	60	0.845	0.923
DHS0501	68	66	61	0.897	0.924	62	57	0.838	0.919
DHS0502	74	72	66	0.892	0.917	68	62	0.838	0.912
DHS0601	116	106	98	0.845	0.924	102	91	0.784	0.892
DHS0602	105	100	91	0.867	0.910	94	84	0.800	0.894
DHS0701	71	68	64	0.901	0.941	65	59	0.831	0.908
DHS0702	67	64	61	0.910	0.953	60	56	0.836	0.933
DHS0801	112	103	91	0.812	0.883	101	89	0.795	0.881
DHS0802	75	72	67	0.893	0.931	69	62	0.827	0.898
DHS0901	78	75	68	0.872	0.907	73	66	0.846	0.904
DHS0902	111	101	92	0.829	0.911	98	87	0.784	0.888
DHS1001	28	28	28	1.000	1.000	28	26	0.928	0.928
DHS1002	29	29	28	0.966	0.966	29	26	0.896	0.896
Overall	1779	1674	1532	0.861	0.915	1589	1425	0.801	0.897

Table 5. Comparison of overall detection performance among different approaches.

Method	Identified Trees	TP Count	Recall	Precision
Proposed adaptive mean shift algorithm	1674	1532	0.861	0.915
Mean shift algorithm in [17]	1589	1425	0.801	0.897
Region growing	1531	1343	0.755	0.877
<i>k</i> -means clustering	1548	1367	0.768	0.883
CHM-based method	1408	1191	0.669	0.846

5. Discussion

The results obtained for individual tree identification differ significantly from study to study. It is unclear how much of the variation is caused by the applied methods and how much by the forest conditions. Accordingly, it is difficult to compare our results with those reported in existing literature using different methods since the experimental scenarios (e.g., the LiDAR configuration and the forest morphology and architecture) are not exactly the same. As the effect of variability of forest conditions has a high impact on the achieved accuracy, the performance of one method versus other methods cannot be verified without testing these methods in the same forest conditions [33]. We make a comparative analysis between our proposed method and three typical point cloud-based methods, as well as a conventional CHM-based method, using the same data sets acquired in the study area. Obviously, the CHM-based method produces lower detection accuracy than point-based methods because it works on a raster surface model derived from the raw point cloud that always includes loss of information during transformation, but not directly on the original airborne LiDAR data that is not altered in any way. Furthermore, using the original point cloud would find the smaller trees more accurately than the traditional CHM-based methods. As shown in Table 5, mean shift-based methods could achieve better performance for individual tree detection when compared to region growing or *k*-means clustering approaches. This is mainly due to the fact that mean shift is a non-parametric algorithm, not sensitive to initializations (does not require detecting treetops) and it can handle arbitrarily shaped clusters. In order to evaluate the performance of the proposed mean shift-based method, we compare it to the similar method presented in [17]. Our proposed new method has the similar performance with this existing method in identifying dominant and codominant trees, but can

do better in detecting intermediate and suppressed trees. The mean shift algorithm described in [17] is essentially a fixed bandwidth mean shift procedure, where a single kernel bandwidth in a specific vegetation stratum can result in obvious under-segmentations or over-segmentations. We apply a variable bandwidth mean shift algorithm, which automatically adapts the kernel bandwidth to local crown size, to reduce under- and over-segmentations. The comparison results listed in Tables 4 and 5 reveal that our proposed adaptive mean shift-based method could improve the overall detection accuracy relative to standard mean shift algorithms. Ferraz et al. proposed an approach to adaptively calibrate the kernel bandwidth model required by the so-called AMS3D algorithm [24]. However, accuracies of individual tree crown detection were not reported in [24], and thus we cannot compare our method to it in terms of tree detection performance.

As stated in Section 4, the tree detection accuracy can be assessed in terms of “recall” and “precision”, which are considered to be indices for commission error and omission error respectively. The commission errors usually result from over-segmentation that implies segmenting a single tree into several segments, whereas under-segmentation of tree crowns will lead to omission errors. The two types of errors are due to both the inability of the airborne LiDAR to characterize tree crowns and the detection or extraction algorithm itself. According to [33], the LiDAR point density at which the crowns are sampled, has an impact on the identification of individual vegetation features. Misclassifications may happen where the canopy is unequally sampled by the laser pulses, but this problem can be reduced by increasing the point density [3]. An average laser point density of 15 pt/m² used in this study is enough to capture the 3D spatial structure of individual tree crowns. When the LiDAR point density is not an issue, it is confirmed that the extraction or detection method is the main factor impacting achieved accuracy [33].

The main problem encountered when using mean shift-based clustering algorithms to detect tree crowns is omission and commission errors, that is, such a kind of segmentation methods tends to under-detect or over-detect individual trees. Omission errors are more prone to happen than the commission ones. The main reason is that inadequate tree finding capability inherent in these methods will result in missing of trees, especially small trees. Besides, due to overlapping point cloud patches caused by complex terrain, closed forest canopies or other reasons, groups of trees standing very close to each other, identified as several single trees in the field inventory, might be treated as one cluster composed of all of these trees while performing the mean shift algorithm to segment the LiDAR point cloud. As for our proposed method, the uncertainty in tree segmentations mainly derives from the kernel bandwidth that is adapted to the estimated value of crown diameter, and estimation errors will influence the detection performance. Thus, a more sophisticated adaptive selection scheme for the kernel bandwidth might be developed to further improve the detection accuracy. Also, in dense forests with complex forms, object-oriented and shape-related rules should be utilized to improve the degree of discrimination between trees. We plan to address these issues in future work.

6. Conclusions

In this paper, we propose an adaptive mean shift-based clustering approach to segment the 3D forest point cloud data for the aim of identification and delineation of individual tree crowns. In order to reduce over- and under-segmentations caused by the fixed bandwidth mean shift-based approaches, a novel computationally efficient scheme to adaptively calibrate the kernel bandwidth is devised and validated on test plots of the study zone. The kernel bandwidth is not set to a fixed value or determined as a multi-scale function, but can be varied in a manner that reflects the quantitative information about the local canopy structure. This study demonstrates the ability of our proposed approach to segment the vegetated area and identify individual tree crowns. We show that the proposed approach works well at detecting trees from the point cloud by comparing segmentation results to the ground truth data acquired in the field survey. Its performance is also quantitatively assessed by calculating the overall tree detection rates for all trees in the test plots and making comparisons among different approaches. Test data reveal that the accuracies, in terms of “recall” and “precision”, are relatively high

when compared to the conventional point-based approaches, indicating that the proposed adaptive mean shift-based approach has good potential for use in the area of forestry inventory and digital forest resource monitoring.

Acknowledgments: This work was financially supported by the Special Scientific Research Fund of Public Welfare Profession (Forestry) of China under Grant No. 201504323.

Author Contributions: Xingbo Hu and Weiyang Xu performed the overall mathematical analyses, designed, realized and analysed the experiments. In addition, they wrote and revised the manuscript. Wei Chen contributed to the ideas of this work, design of the experiment and statistical analysis of the data.

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

Appendix A. Mathematical Derivations of the Mean Shift Algorithm

The mean shift algorithm can estimate the stationary points (modes) of the underlying probability density function without explicitly estimating the density. Define

$$K_{\mathbf{H}}(\mathbf{x}) = |\mathbf{H}|^{-1/2} K(\mathbf{H}^{-1/2}\mathbf{x}), \quad (\text{A1})$$

where, $K(\mathbf{x})$ is the d -variate kernel function and \mathbf{H} is a symmetric positive definite $d \times d$ bandwidth matrix. Given n data points $\mathbf{x}_j (j = 1, \dots, n)$ in the d -dimensional space R^d , the kernel estimator for the probability density function $f(\mathbf{x})$ at point \mathbf{x} can be written as

$$\hat{f}_K(\mathbf{x}) = \frac{1}{n} \sum_{j=1}^n K_{\mathbf{H}}(\mathbf{x} - \mathbf{x}_j). \quad (\text{A2})$$

The multivariate kernel can be generated from a radially symmetric univariate kernel, assuming the following form:

$$K(\mathbf{x}) = c_{k,d} k(\|\mathbf{x}\|^2), \quad (\text{A3})$$

where, $c_{k,d}$ is the normalization constant making $K(\mathbf{x})$ integrate to one and $k(\cdot)$ is called the profile of the kernel that describes how strongly the data points are taken into account in the estimation. For the sake of reducing the estimation complexity, the bandwidth matrix \mathbf{H} is chosen as proportional to the identity matrix

$$\mathbf{H} = h^2 \mathbf{I}, \quad (\text{A4})$$

in which the only one parameter $h > 0$. Employing only one bandwidth parameter and the profile notation, the kernel density estimator Equation (A2) becomes the following expression [20]:

$$\hat{f}_{h,K}(\mathbf{x}) = \frac{c_{k,d}}{nh^d} \sum_{j=1}^n k\left(\left\|\frac{\mathbf{x} - \mathbf{x}_j}{h}\right\|^2\right). \quad (\text{A5})$$

The mean shift tries to determine the local modes of the density function $f(\mathbf{x})$, which correspond to the zeros of the gradient $\nabla f(\mathbf{x}) = 0$. The linearity of Equation (A5) is exploited to estimate the density gradient from the gradient of the density estimator:

$$\hat{\nabla} f_{h,K}(\mathbf{x}) \equiv \nabla \hat{f}_{h,K}(\mathbf{x}) = \frac{2c_{k,d}}{nh^{d+2}} \sum_{j=1}^n (\mathbf{x} - \mathbf{x}_j) k'\left(\left\|\frac{\mathbf{x} - \mathbf{x}_j}{h}\right\|^2\right). \quad (\text{A6})$$

Define

$$g(x) = -k'(x), \quad (\text{A7})$$

and use it for profile, then the kernel $G(\mathbf{x})$ can be defined as

$$G(\mathbf{x}) = c_{g,d} g(\|\mathbf{x}\|^2), \quad (\text{A8})$$

where $c_{g,d}$ is the corresponding normalization constant. Introducing $g(x)$ into Equation (A6) yields

$$\begin{aligned}\hat{\nabla} f_{h,K}(\mathbf{x}) &= \frac{2c_{k,d}}{nh^{d+2}} \sum_{j=1}^n (\mathbf{x}_j - \mathbf{x}) g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_j}{h}\right\|^2\right) \\ &= \frac{2c_{k,d}}{nh^{d+2}} \left[\sum_{j=1}^n g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_j}{h}\right\|^2\right) \right] \left[\frac{\sum_{j=1}^n \mathbf{x}_j g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_j}{h}\right\|^2\right)}{\sum_{j=1}^n g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_j}{h}\right\|^2\right)} - \mathbf{x} \right].\end{aligned}\quad (\text{A9})$$

or

$$\hat{\nabla} f_{h,K}(\mathbf{x}) = \frac{2}{h^2 c} \hat{f}_{h,G}(\mathbf{x}) \cdot \mathbf{m}_{h,G}(\mathbf{x}), \quad (\text{A10})$$

where $c = c_{g,d}/c_{k,d}$. In Equation (A10), the first term is the density estimate at point \mathbf{x} computed with kernel G :

$$\hat{f}_{h,G}(\mathbf{x}) = \frac{c_{g,d}}{nh^d} \sum_{j=1}^n g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_j}{h}\right\|^2\right), \quad (\text{A11})$$

and the second term is the so-called mean shift vector:

$$\mathbf{m}_{h,G}(\mathbf{x}) = \frac{\sum_{j=1}^n \mathbf{x}_j g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_j}{h}\right\|^2\right)}{\sum_{j=1}^n g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_j}{h}\right\|^2\right)} - \mathbf{x}, \quad (\text{A12})$$

which is the difference between the weighted mean, using the kernel G for weights, and \mathbf{x} , the center of the kernel.

Rewriting Equation (A10) can yield

$$\mathbf{m}_{h,G}(\mathbf{x}) = \frac{1}{2} h^2 c \frac{\hat{\nabla} f_{h,K}(\mathbf{x})}{\hat{f}_{h,G}(\mathbf{x})}. \quad (\text{A13})$$

This equation shows that the mean shift vector at point \mathbf{x} computed with kernel G is proportional to the normalized density gradient estimate obtained with kernel K . The mean shift vector thus always points towards the direction of the maximum increase in the probability density. By setting

$$\mathbf{x}(t+1) \leftarrow \mathbf{x}(t) + \mathbf{m}_{h,G}[\mathbf{x}(t)], \quad (\text{A14})$$

where t denotes the iteration number, the iterative process converge towards the local maxima. Thus, it can be seen that the mean shift procedure does not need to evaluate the density function $\hat{f}_{h,K}(\mathbf{x})$ itself but only the induced kernel $g(x)$.

References

1. Lee, H.; Slatton, K.C.; Roth, B.E.; Cropper, W.P., Jr. Adaptive clustering of airborne LiDAR data to segment individual tree crowns in managed pine forests. *Int. J. Remote Sens.* **2010**, *1*, 117–139.
2. Lefsky, M.A.; Cohen, W.B.; Parker, G.G.; Harding, D.J. Lidar remote sensing for ecosystem studies. *BioScience* **2002**, *1*, 19–30.
3. 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**, *1*, 75–84.
4. Hyypä, J.; Yu, X.; Hyypä, H.; Vastaranta, M.; Holopainen, M.; Kukko, A.; Kaartinen, H.; Jaakkola, A.; Vaaja, M.; Koskinen, J.; et al. Advances in forest inventory using airborne laser scanning. *Remote Sens.* **2012**, *4*, 1190–1207.
5. Chen, Q.; Baldocchi, D.; Gong, P.; Kelly, M. Isolating individual trees in a savanna woodland using small footprint lidar data. *Photogramm. Eng. Remote Sens.* **2006**, *8*, 923–932.
6. 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**, *5*, 564–577.

7. Tiede, D.; Hochleitner, G.; Blaschke, T. A full GIS-based workflow for tree identification and tree crown delineation using laser scanning. In Proceedings of the ISPRS Workshop CMRT (City Models, Roads and Traffic), Vienna, Austria, 29–30 August 2005; pp. 9–14.
8. Zhen, Z.; Quackenbush, L.; Zhang, L. Impact of tree-oriented growth order in marker-controlled region growing for individual tree crown delineation using airborne laser scanner (ALS) data. *Remote Sens.* **2014**, *6*, 555–579.
9. Koch, B.; Heyder, U.; Weinacker, H. Detection of individual tree crowns in airborne lidar data. *Photogramm. Eng. Remote Sens.* **2006**, *4*, 357–363.
10. Reitberger, J.; Schnorr, C.; Krzystek, P.; Stilla, U. 3D segmentation of single trees exploiting full waveform LIDAR data. *ISPRS J. Photogramm. Remote Sens.* **2009**, *6*, 561–574.
11. Vo, A.V.; Truong-Hong, L.; Laefer, D.F.; Bertolotto, M. Octree-based region growing for point cloud segmentation. *ISPRS J. Photogramm. Remote Sens.* **2015**, *6*, 88–100.
12. Alexander, C. Delineating tree crowns from airborne laser scanning point cloud data using Delaunay triangulation. *Int. J. Remote Sens.* **2009**, *14*, 3843–3848.
13. Lloyd, S.P. Least square quantization in PCM. *IEEE Trans. Inf. Theory* **1982**, *2*, 129–137.
14. Morsdorf, F.; Meier, E.; Allgöwer, B.; Nüesch, D. Clustering in airborne laser scanning raw data for segmentation of single trees. *Int. Arch. Photogram. Remote Sens. Spatial Inf. Sci.* **2003**, *3*, 27–33.
15. Gupta, S.; Weinacker, H.; Koch, B. Comparative analysis of clustering-based approaches for 3-D single tree detection using airborne fullwave lidar data. *Remote Sens.* **2010**, *2*, 968–989.
16. Ferraz, A.; Bretar, F.; Jacquemoud, S.; Gonçalves, G.; Pereira, L. 3D segmentation of forest structure using a mean-shift based algorithm. In Proceedings of the IEEE 17th International Conference on Image Processing, Hong Kong, China, 26–29 September 2010; pp. 1413–1416.
17. 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**, *6*, 210–223.
18. Hu, X.; Xie, Y. Segmentation and clustering of 3D forest point cloud using mean shift algorithms. In Proceedings of the 4th International Conference on Machinery, Materials and Computing Technology (ICMMCT 2016), Hangzhou, China, 23–24 January 2016; Atlantis Press: Amsterdam, The Netherlands, 2016; pp. 1275–1279.
19. Amiri, N.; Wei, Y.; Heurich, M.; Krzystek, P.; Skidmore, A.K. Estimation of regeneration coverage in a temperate forest by 3D segmentation using airborne laser scanning data. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *52*, 252–262.
20. Cheng, Y. Mean shift, mode seeking, and clustering. *IEEE Trans. Pattern Anal. Mach. Intell.* **1995**, *8*, 790–799.
21. Comaniciu, D.; Meer, P. Mean shift: A robust approach toward feature space analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *5*, 603–619.
22. Melzer, T. Non-parametric segmentation of ALS point clouds using mean shift. *J. Appl. Geod.* **2007**, *3*, 159–170.
23. Liu, C.; Wu, H.; Zhang, Y. Extraction of urban 3D features from lidar data fused with aerial images using an improved mean shift algorithm. *Surv. Rev.* **2011**, *9*, 402–414.
24. Ferraz, A.; Saatchi, S.; Mallet, C.; Meyer, V. Lidar detection of individual tree size in tropical forests. *Remote Sens. Environ.* **2016**, *183*, 318–333.
25. Bakula, M.; Oszczak, S.; Pelc-Mieczkowska, R. Performance of RTK positioning in forest conditions: Case study. *J. Surv. Eng.* **2009**, *3*, 125–130.
26. Comaniciu, D. An algorithm for data-driven bandwidth selection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2003**, *2*, 281–288.
27. Huang, X.; Zhang, L. An adaptive mean-shift analyses approach for object extraction and classification from urban hyperspectral imagery. *IEEE Trans. Geosc. Remote Sens.* **2008**, *12*, 4173–4185.
28. Bo, S.; Ding, L.; Li, H.; Di, F.; Zhu, C. Mean shift-based clustering analysis of multispectral remote sensing imagery. *Int. J. Remote Sens.* **2009**, *4*, 817–827.
29. Lalonde, J.F.; Vandapel, N.; Hebert, M. *Automatic Three-Dimensional Point Cloud Processing for Forest Inventory*; Carnegie Mellon University: Pittsburgh, PA, USA, 2006; pp. 1–16.
30. Rusu, R.B.; Cousins, S. 3D is here: Point cloud library (PCL). In Proceedings of the International Conference on Robotics and Automation, Shanghai, China, 9–13 May 2011; pp. 1–4.

31. Goutte, C.; Gaussier, E. A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. In *Advances in Information Retrieval, Proceedings of the 27th European Conference on IR Research (ECIR 2005), Santiago de Compostela, Spain, 21–23 March 2005*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 345–359.
32. Duan, Z.; Zhao, D.; Zeng, Y.; Zhao, Y.; Wu, B.; Zhu, J. Assessing and correcting topographic effects on forest canopy height retrieval using airborne LiDAR data. *Sensors* **2015**, *6*, 12133–12155.
33. Kaartinen, H.; Hyyppä, J.; Yu, X.; Vastaranta, M.; Hyyppä, H.; Kukko, A.; Holopainen, M.; Heipke, C.; Hirschmugl, M.; Monsdorf, F.; et al. An international comparison of individual tree detection and extraction using airborne laser scanning. *Remote Sens.* **2012**, *4*, 950–974.



© 2017 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/>).