Automatic Detection and Parameter Estimation of Trees for Forest Inventory Applications Using 3D Terrestrial LiDAR

Ahmad Aijazi, Paul Checchin, Laurent Malaterre, Laurent Trassoudaine

To cite this version:

Ahmad Aijazi, Paul Checchin, Laurent Malaterre, Laurent Trassoudaine. Automatic Detection and Parameter Estimation of Trees for Forest Inventory Applications Using 3D Terrestrial LiDAR. Remote Sensing, MDPI, 2017, 9 (9), pp.946. 10.3390/rs9090946. hal-01655563

HAL Id: hal-01655563
https://hal.archives-ouvertes.fr/hal-01655563
Submitted on 28 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License
Automatic Detection and Parameter Estimation of Trees for Forest Inventory Applications Using 3D Terrestrial LiDAR

Ahmad K. Aijazi, Paul Checchin, Laurent Malaterre and Laurent Trassoudaine *
Institut Pascal, Université Clermont Auvergne, CNRS, SIGMA Clermont, F-63000 Clermont-Ferrand, France; kamalaijazi@gmail.com (A.K.A.); paul.checchin@uca.fr (P.C.); laurent.malaterre@uca.fr (L.M.)
* Correspondence: laurent.trassoudaine@uca.fr; Tel.: +33-473-40-7810
Received: 17 July 2017; Accepted: 8 September 2017; Published: 12 September 2017

Abstract: Forest inventory plays an important role in the management and planning of forests. In this study, we present a method for automatic detection and estimation of trees, especially in forest environments using 3D terrestrial LiDAR data. The proposed method does not rely on any predefined tree shape or model. It uses the vertical distribution of the 3D points partitioned in a gridded Digital Elevation Model (DEM) to extract out ground points. The cells of the DEM are then clustered together to form super-clusters representing potential tree objects. The 3D points contained in each of these super-clusters are then classified into trunk and vegetation classes using a super-voxel based segmentation method. Different attributes (such as diameter at breast height, basal area, height and volume) are then estimated at individual tree levels which are then aggregated to generate metrics for forest inventory applications. The method is validated and evaluated on three different data sets obtained from three different types of terrestrial sensors (vehicle-borne, handheld and static) to demonstrate its applicability and feasibility for a wide range of applications. The results are evaluated by comparing the estimated parameters with real field observations/measurements to demonstrate the efficacy of the proposed method. Overall segmentation and classification accuracies greater than 84% while average parameter estimation error ranging from 1.6 to 9% were observed.


Data Set License: ODC Attribute Licence

Keywords: tree segmentation; 3D LiDAR; forest inventory; parameter estimation

1. Introduction

Forests are an important natural resource that require careful attention for sustainability. Proper management of these resources plays an important role in ecological and economic development. For this purpose, forest inventory is essential. Forest inventories present the main source of information for describing the structure and quantifying of forest resources. They help to provide comprehensive information about the state and dynamics of forests for strategic and management planning. In forest inventory, different attributes are studied and measured for both ecological benefits (such as habitat analysis, studying different environmental influences on growth, etc.) and economic reasons (i.e., timber volume estimation for wood production). Other purposes of forest inventory include assessing standing volume to determine potential fire hazards and the risk of fire. The results of such types of inventory can be used for preventive actions against such potential hazards and also awareness.

Even today, in many countries, tree measurement methods are still based on human estimation and experience (e.g., for crown diameters) or field measurements performed with very simple tools which
can be very labour intensive and time-consuming. Lack of automation makes them expensive and subjective (i.e., dependent on the expert). In recent years, with the advent of newer technologies such as Light Detection and Ranging (LiDAR) systems, there has been a move towards full or semi automation.

2. Related Work and Outline

2.1. Employing Airborne Laser Scanning Technology

Airborne Laser Scanning (ALS) has been widely used to measure forest height, individual tree height, crown diameter (depending on the 3D point cloud density) and mapping forested areas [1]. In the literature, we find both area-based [2] and single-tree based approaches [3]. Area-based methods provide statistically-based maps of forest stand parameters such as height of stand and stem density, which are useful for large-area forest inventory and long term management planning. There are various methods proposed to delineate individual trees using ALS data. Li et al. [4] segmented individual trees from point cloud data by taking advantage of the relative spacing between trees using a top-to-bottom region growing approach that segments trees individually and sequentially from the tallest to the shortest. Using small-footprint airborne LiDAR data, Hamraz et al. [5] proposed a method that collects and uses crown information such as steepness and height, on-the-fly, to delineate crown boundaries without a priori assumptions of crown shape and size. In [6], Vega et al. extract trees from ALS data employing a normalization process and using the highest unclassified point instead of the traditionally used local maxima. A multi-scale and multi-criteria framework is introduced to optimize the efficiency of tree detection. In a study by Korpela et al. [7], a multi-scale template matching approach for tree detection and measurement is proposed, in which elliptical and other shaped templates are used to represent tree models. Duncanson et al. [8] used a watershed-based delineation of canopy height model, which is subsequently refined using LiDAR point cloud data to extract individual trees. In [9], Mund et al. investigate the use of full-waveform ALS data for the detection and classification of understory structures by decomposing and analyzing 3D data in multiple layers. Density and time delay of the return pulses are used for classification purposes. However, the information provided by these methods using ALS data is usually insufficient or less accurate on a single tree level. Greater sensor height above ground and lack of penetration of the laser beam to ground level due to dense canopy cover result in insufficient point density within the understory layer as discussed by Amiri et al. [10]. Also, estimating certain important forest parameters such as tree species, tree measurements and the soil’s carrying capacity, relies heavily on understory information as shown by Korpela et al. [11].

2.2. Employing Terrestrial Laser Scanning Technology

Different terrestrial or ground-based laser scanners provide a more effective solution for obtaining detailed understory information important when estimating different tree parameters. Both static and mobile systems provide millions of 3D points from inside the forest at close range. The use of static terrestrial laser scanner (or Terrestrial Laser Scanning, TLS) dominates the related works. In [12], Simonse et al. extracted out layers at heights corresponding to DBH (Diameter at Breast Height) from the point cloud converted into a Digital Terrain Model (DTM). In this layer, circular structures were detected, i.e., stem sections, using a Hough transformation. Thies et al. [13] used registered TLS data obtained from three different scan positions per tree to model the trunk of single trees, by fitting cylinders into the point cloud using a non-linear least squares method. Pfeifer et al. [14] developed a method to fit cylinders into multiple scan mode point clouds. Single trees were extracted and cylinders fitted by the determination of five parameters using the non-linear least squares method. Stem and branches were partially detected. These shape fitting methods may be more useful for modeling purposes but are prone to errors when determining accurate tree parameters. Buksch et al. [15] proposed a technique to determine a tree skeleton using point cloud data by generating an octree containing the TLS points. Using the neighborhood information of the octree cells, a graph was extracted, while the cycles contained in the graph were deleted. The resulting graph was
found to represent the tree skeleton. Xu et al. [16] used a similar technique for tree skeletonization in which each 3D point was linked with points in a neighborhood of 0.2 m, resulting in a connected graph. In the graph, the shortest path from a pre-selected root point was calculated for every node using the Dijkstra algorithm [17]. The lengths of these paths were quantized and clustered into bins and a skeleton was formed by connecting centroids of the adjacent bins. This approach was further applied by Livny et al. [18]. Gorte et al. [19] reconstructed forest trees from TLS data using technique based on mathematical morphology. Identification of tree structures in terms of stem and branches was done in 3D voxel space employing a selection of basic and advanced 2D raster (image) processing algorithms transferred into the 3D domain. Both papers focused on the modeling and visualization of trees rather than on the estimation of tree parameters.

In [20], Watt and Donoghue compared the TLS-based measurements with the field measurements of **DBH** and tree height. Their results indicated that occlusion was an important factor affecting the information obtained from the scanned 3D point clouds. Maas et al. [21] presented an approach to estimate the **DBH** and height of trees using DTM reduction and single tree detection algorithm. They also estimated stem profiles, including shape and straightness based on the **DBH** determination techniques. In a study by Belton et al. [22], volume of trees is calculated after extracting the tree from the raw 3D point cloud. Principal component analysis (PCA) is applied to each point’s neighborhood. The features described by the eigenvalues obtained by PCA were used as input for clustering through a Gaussian mixture model. The points pertaining to the leaves were deleted from the input point cloud data, leaving only the stem and branch points. Tree skeleton was extracted by fitting ellipses and connecting the centers of overlapping ellipses. A cylinder fitting routine was applied to the different sections for parameter estimation. An estimation of the stem diameter based on the intensity profile using the TLS system from a fixed view point is presented by Similarly, Huang et al. [23] also presented an automated method for measuring **DBH** and tree height using TLS data. In addition, using tree height and basal area the stand value and timber production were determined by [24]. Brolly and Kiraly [25] used a least squares adjustment to fit circles at different heights above DTM and matching them. This approach worked better than using a single fitted circle for tree detection but slightly underperformed compared to cylinder fitting for the estimation of **DBH** and tree height. A large number of these studies has been conducted supposing a typical cylindrical stem shape as is often encountered in forest stands of pine and spruce and also forests managed for timber production. However, this is not always true for all types of forests and complex tree structures. A shape independent analysis or even free-form curves provide better results. For example, Pfeifer and Winterhalder [26] presented a method using B-splines to represent tree stem cross sections and to account for ovality and other deformations in the tree stem. Several others studies have tried to extract **DBH** information from TLS data, but are limited to thinned stands [24], limited species [27] and limited samples [28,29], etc. Compared to ALS data, tree height has been retrieved from TLS data with limited success. Often, tree height is taken as the maximum LiDAR return height within a radial distance from detected tree locations or is based on detected stem taper and taper functions [30]. Hopkinson et al. [31] demonstrated in their study that LiDAR data underestimated mean plot-level tree height by about 1.5 m compared to the ground truth (manual field measurements). Another recent study conducted by Yao et al. [32], demonstrated the use of a phase-shift TLS to determine forest inventory parameters, however, phase-shift scanners were found to suffer from many distortion effects, specifically with tree leaves.

Apart from static TLS systems, recently, moving terrestrial scanning systems have also been used for forest inventory. The use of such mobile laser scanners helps to reduce the problem of occlusion and also allows for data acquisition on a larger scale and in a time-efficient manner. Forsman et al. [33] employed a laser scanner mounted on a vehicle to scan the forest in a push-broom configuration. The trajectory and the orientation were used to generate a 3D point cloud. Clusters representing trees were extracted line-wise to cater for the uncertainty in the positioning system. Due to the displacement, the measurements were obtained from large portions of the stem, and the multiple lines
from different views were then used to fit circles. Owing largely to the quality of the point cloud generated, an error of 14% in the circle fitting was reported. In their study [34], Bauwens et al. used a handheld mobile laser scanner to scan forest environment and estimate several forest parameters. Even though, the tree sections at breast height were totally scanned, the canopy was poorly described because of the limited range. Due to the low accuracy and inherent noise of the system, the estimation of different parameters such as DBH still needs to be improved. A comprehensive overview concerning LiDAR-based computational forestry is given by Leeuwen and Nieuwenhuis [35] and Dassot et al. [36] including supplementary information about various TLS devices and functions.

2.3. Outline of this Work

This work presents a method for automatic detection and estimation of trees, especially in forest environments using terrestrial 3D LiDAR data acquired by both moving and static systems (Section 3.1). The proposed method does not rely on any predefined tree shape or model, unlike many previous studies. It uses the vertical distribution of the 3D points partitioned in a gridded Digital Elevation Model (DEM) to extract out ground points (Section 3.2). The cells of the DEM are then clustered together to form super-clusters representing potential tree objects. The 3D points contained in each of these super-clusters are then classified into trunk and vegetation classes using a super-voxel based segmentation method exploiting surface normals. Different tree attributes (such as DBH, BA, height and volume) are then estimated at individual tree level which are then aggregated to generate metrics for forest inventory applications (Section 4). The method was validated and evaluated on real data obtained from three different types of terrestrial sensors (vehicle borne, handheld and static) to demonstrate its applicability and feasibility for a wide range of applications (Section 5). The results were evaluated by comparing the estimated parameters with field observations/measurements. After discussion (Section 6), we finally conclude in Section 7.

3. Materials and Method

3.1. Materials

In this study, we use data sets obtained from three different terrestrial LiDAR systems. Each source uses a different sensor and employs a different acquisition technique. The use of these different data sources not only helps to better evaluate our method but also to demonstrate the viability of our method developed for different types of applications. These different data sets along with their sources are briefly described in the following subsections.

3.1.1. Data_Set-1 (from Mobile Terrestrial Mapping System)

This data set used to evaluate our work belongs to the dynamic data set of the 3D Urban Data Challenge 2011, obtained from a terrestrial mobile mapping system, i.e., the Vis Center’s LiDAR Truck (University of Kentucky) [37] containing two Optech LiDAR sensor heads (high scan frequency up to 200 Hz), a GPS, an inertial measurement unit and a spherical digital camera as shown in Figure 1. The method uses integrated GPS/IMU data to directly orient LiDAR data from its local reference frame to the mapping reference frame (WGS84) to obtain a colored 3D point cloud. The laser is used in a pushbroom configuration sweeping the scene with profiles while fixing the scan angles as the vehicle moves.
Figure 1. (a) The Vis Center’s LiDAR Truck. (b) Two Optech LiDAR sensor heads, a GPS, an inertial measurement unit and a spherical digital camera mounted on a rigid frame.

This data set contains a number of trees scanned in the urban environment around the Kentucky area, USA. Although the environment is less cluttered, the results simply affirm the applicability and utility of the proposed method using such type of scanning technology in the forest environment as also presented by Forsman et al. [33]. In this study, the authors successfully demonstrated that the use of such vehicle-borne mobile terrestrial mapping systems is practically applicable in the forest environment and can be used for collecting data and the status of forest stands after thinning. This makes it more pertinent to also evaluate our method using this type of data. In addition, more complex forest data sets are also used to evaluate our method (presented as follows).

3.1.2. Data_Set-2 (from Handheld Mobile Mapping System)

In this data set, Zebedee, ZEB-1, handheld mapping system [38] is used to obtain a 3D point cloud. The system, as shown in Figure 2, consists of a lightweight laser scanner with 1530 m maximum range (dependent on surface reflectivity and environmental conditions) and an industrial-grade MEMS inertial measurement unit (IMU) mounted on a simple spring mechanism. As an operator holding the device moves through the environment, the scanner loosely oscillates on the spring, thereby producing rotational motion that converts the laser’s inherent 2D scanning plane into a local 3D field of view. With the use of proprietary software, the six degree of freedom sensor trajectory is accurately and continuously calculated from the laser and inertial measurements, and the range measurements are projected into a common coordinate frame to generate a 3D point cloud.

This data set is composed of 3D scans of the parts of the Forêt de Tronçais in the Auvergne-Rhone-Alpes region (département de l’Allier) of Central France as indicated in Figure 3. It consists of sessile oak trees (Quercus petraea (Matt.) Liebl.) and the acquisition was made at the end of the winter season (i.e., end of February 2014).

Figure 2. (a) The Zebedee ZEB-1 handheld mapping system; (b) An operator holding the device moving through the Forêt de Tronçais, France.
Figure 3. Study site Forêt de Tronçais in the Auvergne-Rhone-Alpes region (département de l’Allier) of Central France. Source: Google Maps, 19 August 2017.

3.1.3. Data_Set-3 (from Static Ground-based LiDAR System)

Data are obtained using a ground-based Leica’s P20 laser scanning station as shown in Figure 4. To cover the whole scanning area, multiple scans are obtained from different locations and then registered together to obtain a colored 3D point cloud. The scans are taken so as to ensure some overlapping to facilitate the registration process. In order to further aid the process, additional targets are also placed all around. The scans are registered, one by one, using a standard ICP (Iterative Closest Point) algorithm [39].

Figure 4. Leica P20 terrestrial scanning system in the forest environment.

This data set contains scans of parts of a silver fir (*Abies alba* Mill.) tree forest in the Auvergne-Rhone-Alpes region (département du Puy-de-Dôme) of Central France (Figure 5). The data were obtained during the end of the summer season (i.e., beginning of September 2016).
3.2. Method

The 3D point clouds in the forest environment are obtained from different terrestrial LiDAR systems as explained in Section 3.1. The 3D points are registered together in a global frame of reference to form the point cloud and if available each 3D point in the point cloud is also associated with the corresponding color and laser intensity values.

The 3D point clouds obtained from the different sources are first binned on to a 2D grid (similar to a Digital Elevation Model) as shown in Figure 6b. Each cell $C$ of the grid (constant size of $0.5 \times 0.5$ m) is defined by its center and the maximum and minimum elevation values, based on the 3D points contained, as expressed in (1).

\[
\text{Cell}_i \in \{c(x, y), \text{elev}_{\text{min}}, \text{elev}_{\text{max}}, \text{elev}_{2.5}\} \tag{1}
\]

Here $\text{Cell}_i$ is the $i$th cell of the grid while $c(x, y)$ is the geometrical center (in 2D) of the points contained in the cell while the $\text{elev}_{\text{min}}, \text{elev}_{\text{max}}$ and $\text{elev}_{2.5}$ are the minimum, maximum and maximum elevation value of the 3D points contained in the cell below 2.5 m, respectively.

The 3D points falling into each cell of this elevation map are stored and will be merged to form objects after the clustering phase.

3.2.1. Ground Segmentation

The ground points in each cell were segmented and filtered out. Each 3D point in the cell was analyzed. The profile in the elevation direction was analyzed. It is based on the observation that a continuous vertical distribution of 3D points in the vertical axis represents a tree-like structure whereas an interrupted distribution—such as high density below or high density below and above—indicates the presence of ground points. These two cases along with the vertical distribution are presented in Figure 7.
Figure 6. Segmentation and classification of Tree structure. In (a) the raw 3D point cloud. In (b), the 3D point cloud is represented in DEM while (c) is after ground segmentation/extraction. (d,e) show the tree trunk segmentation using a super-voxel based approach. In (d), it can be clearly seen that for a tree trunk the surface normals are predominantly horizontal. (f) shows the final classification results.
Figure 7. 3D point distribution along the vertical axis to segment out ground points.

The 3D points in the cell are considered as ground points if there is discontinuity in the vertical distribution, i.e., if \( \text{elev}_{\text{min}} \) and \( \text{elev}_{2.5} < 1 \) m (exploiting the vertical distribution). These 3D points are segmented out as ground points while the other 3D points contained in the cell are conserved as non-ground points. This method helps us to segment out the ground even under dense canopy which is not easily segmented using methods based on standard DEM analysis using the elevation height [40] or plane fitting [41] as the criteria for ground segmentation. As these methods also remove important tree points close to the ground, they often result in inaccurate calculation of tree metrics.

On the contrary, the proposed method segments and filters out not only the real ground points but it retains the ground points under the tree trunk as part of the tree. This allows for the calculation of the local ground slope, which is important in accurate calculation of tree height. These ground points are ultimately segmented out during the tree trunk segmentation process explained in Section 3.2.3.

3.2.2. Clustering

Once the ground points are removed from the cells, we cluster the cells together to form larger super-cells. Any two cells are clustered together to form a new larger cell and their 3D points are merged if the following condition is satisfied:

\[
\|c(x, y)_i - c(x, y)_j\| \leq \sqrt{2} \times \text{Cell}_{\text{dist}}
\]  

(2)

Here \( c(x, y)_i \) and \( c(x, y)_j \) are the geometrical centers of the 3D points contained in the \( i \)th and \( j \)th cell respectively and \( \| \| \) is the Euclidean distance. \( \text{Cell}_{\text{dist}} \) is given as: 
\[
\left( \frac{\text{Cell}_{\text{X}} + \text{Cell}_{\text{Y}}}{2} \right)
\]

and \( s_{X,Y} \) is the cell size along the X and Y directions. These cell sizes are initialized at the initial grid-cell size of 0.5 m, however, the new cell size then varies along the X and Y directions depending upon which axis the cells merge.

The 3D points of these two cells are merged together and values of \( c(x, y), \text{elev}_{\text{min}}, \text{elev}_{\text{max}} \) and \( \text{elev}_{2.5} \) are updated. An overview of the algorithm is presented in Algorithm 1. This modified grid
(with different sized cells) now represents a collection of potential tree objects. These objects are then segmented/classified as tree trunks and vegetation in each cluster as explained in the next section.

**Algorithm 1 Clustering.**

**input** Grid of 2D cells  
**repeat**  
1. Select a 2D cell for clustering  
2. Find all neighboring cells satisfying (2) to include in the cluster  
3. Merge these cells to form a cluster  
4. Re-calculate $Cell_{size}$  
5. Update values of $c(x, y)$, $elev_{min}$, $elev_{max}$ and $elev_{2.5}$  
**until** All 2D cells in the grid are used  
**return** New updated grid

3.2.3. Tree Segmentation

The 3D points contained in each of these super-cells are classified into two classes: \{Tree trunk, Vegetation\}. The Tree trunk consists of the actual trunk visible in the 3D point cloud while the vegetation consists of the leafy portion including canopy. In order to classify these 3D points, they are first over-segmented into 3D voxels and then converted into super-voxels. These are then clustered together using a Link-Chain method to form objects [42].

The method uses agglomerative clustering to group 3D points based on $r$-NN (radius Nearest Neighbor). Although the maximum voxel size is predefined, the actual voxel sizes vary according to the maximum and minimum values of the neighboring points found along each axis to ensure the profile of the structure is maintained as shown in Figure 8.

![Figure 8](image-url)  
*Figure 8.* A number of points is grouped together to form cubical voxels of a maximum size $2 \times r$ where $r$ is the radius and the actual voxel sizes vary according to the maximum and minimum values of the neighboring points found along each axis.
A voxel is then transformed into a super-voxel when properties based on its constituting points are assigned to it. These properties mainly include geometrical center, mean hue (H) and saturation (S) values, maximum of the variance of H & S values, mean laser intensity value, variance of laser intensity values, voxel size along each axis X, Y & Z and surface normals of the constituting 3D points. Here H and S are the values corresponding to the R, G & B (Red, Green and Blue) color values transformed in the HSV (Hue, Saturation and Value) space. As the R, G, B color values are prone to lighting variation (especially in dense forest environments), they are converted into HSV color space for each 3D point. This conversion separates the color component from the intensity component. Also, the intuitiveness of the HSV color space is very useful because we can quantize each axis independently. Wan and Kuo [43] reported that a color quantization scheme based on HSV color space performed much better than one based on RGB color space. The component, invariant to the lighting conditions, is then analyzed. It is referred to in this paper as the color component as it provides more stable color information. Based on the description presented by Hughes et al. [44], the following equations were used for the conversion.

\[
\begin{align*}
    h' &= \begin{cases} 
    \frac{(G - B)}{2 + |B - R|} & \text{if } R = \text{MAX} \\
    \frac{2 + |B - R|}{4 + (R - G)} & \text{if } B = \text{MAX} \\
    \delta & \text{if } G = \text{MAX}
    \end{cases} \\
    H &= h' \times 60^\circ, \quad S = \frac{\text{MAX} - \text{MIN}}{\text{MAX}}, \quad \text{and } V = \text{MAX}
\end{align*}
\]

Here MAX = max(R, G, B), MIN = min(R, G, B), δ = MAX − MIN and H, S, V are the corresponding point of R, G, B in the HSV space. Also, to be noted that the normalized values of R, G, B are used, i.e., R, G, B ∈ [0 . . . 1], and so as a result H ∈ [0 . . . 360°] and S, V ∈ [0 . . . 1]. In case of R = G = B = 0, H is undefined, hence it is assumed to be −1. After the conversion, the color component (c = {H, S}) is then used in our analysis.

With the assignment of all these properties, a voxel is transformed into a super-voxel. All these properties are then used to cluster these super-voxels into objects using a link chain method. In this method each super-voxel is considered to be a link of a chain. All secondary links attached to each of these principal links are found. In the final step all the principal links are linked together to form a continuous chain removing redundant secondary links in the process (see Algorithm 2). If \( V_p \) is a principal link and \( V_n \) is the \( n^{th} \) secondary link, each \( V_n \) is linked to \( V_p \) if and only if the following three conditions are fulfilled:

\[
\begin{align*}
    ||V_{p_{X,Y,Z}} - V_{n_{X,Y,Z}}|| &\leq (w_d + c_d) \\
    ||V_{p_{H,S}} - V_{n_{H,S}}|| &\leq 3\sqrt{w_c} \\
    ||V_{p_{I}} - V_{n_{I}}|| &\leq 3\sqrt{w_l}
\end{align*}
\]

where, for the principal and secondary link super-voxels, respectively:

- \( V_{p_{X,Y,Z}}, V_{n_{X,Y,Z}} \) are the geometrical centers;
- \( V_{p_{H,S}}, V_{n_{H,S}} \) are the mean H & S values;
- \( V_{p_{I}}, V_{n_{I}} \) are the mean laser intensity values;
- \( w_c \) is the color weight equal to the maximum value of the two variances \( Var(H, S) \);
- \( w_l \) is the laser intensity weight equal to the maximum value of the two variances \( Var(I) \).

\( w_d \) is the distance weight given as \( \frac{(V_{p_{X,Y,Z}} + V_{n_{X,Y,Z}})}{2} \). Here \( s_{X,Y,Z} \) is the voxel size along X, Y & Z axis, respectively. \( c_d \) is the inter-distance constant (along the three dimensions) added depending upon the density of points and also to overcome measurement errors, holes and occlusions, etc. If the color or the laser intensity values are not available in any particular data set, (4) or (5) could be dropped respectively. The more information there is, the better the results, even though the method continues to work.
Algorithm 2 Segmentation.

input 3D points
repeat
1. Select a 3D point for voxelisation
2. Find all the neighboring points to be included in the voxel using r-NN within the specified maximum voxel length
3. Transform the voxel into super-voxel by first finding and then assigning to it all the properties found using PCA, including surface normal
until All 3D points are used in a voxel
repeat
1. Specify a super-voxel as a principal link
2. Find all the secondary links attached to the principal link
until All the super-voxels are used
Link all the principal links to form a chain removing the redundant links in the process
return Segmented objects

These clustered objects are then classified using local descriptors and geometrical features into 2 main classes: {Tree trunk, Vegetation}. These mainly include:

- **Surface normals**: The orientation of the surface normals is found essential for distinction between Tree trunk and Vegetation as for the first, the surface normals are predominantly (threshold values greater than 80%) parallel to the X-Y axis (ground plane as seen in Figure 6d) whereas for the vegetation the surface normals are scattered in all directions.

- **Color and intensity**: Intensity and color are also an important discriminating factor for the two object classes.

- **Geometrical center and barycenter**: The height difference between the geometrical center and the barycenter along with other properties is very useful in distinguishing objects like tree trunk and vegetation. For the tree trunks both are closely aligned (being a symmetric pole-like structure) whereas for the vegetation they can be different depending on the shape of the vegetation.

- **Geometrical shape**: Along with the above mentioned descriptors, geometrical shape plays an important role in classifying objects. In 3D space, tree trunks are always represented as long and thin while vegetation is usually more spread-out, broad and large with height depending upon the type of vegetation (i.e., tree canopy, ground bushes, etc.).

As the object classes are so distinctly different a simple threshold-based method is used as presented by Aijazi et al. [42] where the values of the comparison thresholds for these features/descriptors are set accordingly. However, they can also be used to train an SVM classifier as described by Aijazi [45].

Some results of this method are shown in Figure 6. The salient features of this method are data reduction, efficiency and simplicity of approach. During this process, the ground under the tree trunk is also segmented out and we use it to determine the local ground slope on which the tree is present as explained in the next section.

4. Generating Tree Metrics

For different forest inventory applications, generating different tree metrics (such as DBH, BA, height, volume, etc.) at individual tree level is essentially required. These can then be used to estimate complete forest information depending upon the application.
4.1. Diameter at Breast Height

In order to calculate the DBH, we first calculate the local ground slope on which the tree is present. The ground points below the trunk (A in Figure 9) are considered to calculate the slope determining the best fit Plane using planar regression, as presented in [46]. A best fit plane is defined with the equation:

\[(x_i - \bar{x}) = B(y_i - \bar{y}) + C(z_i - \bar{z})\]  \hspace{1cm} (6)

where \(\bar{x}, \bar{y},\) and \(\bar{z}\) are the respective mean values of the \(X, Y,\) and \(Z\) coordinates of all the points considered. To find the equation of the best fit plane for a given set of points, Press et al. [47] present the following equations that are solved for \(B\) and \(C:\)

\[\sum ((x_i - \bar{x})(y_i - \bar{y})) = B \sum (y_i - \bar{y})^2 + C \sum ((y_i - \bar{y})(z_i - \bar{z}))\]  \hspace{1cm} (7)

\[\sum ((x_i - \bar{x})(z_i - \bar{z})) = B \sum ((y_i - \bar{y})(z_i - \bar{z})) + C \sum (z_i - \bar{z})^2\]  \hspace{1cm} (8)

The result of the regression is a plane that passes through a point with coordinates \((x, y, z)\) and is returned in the form of a vector normal to the best fit plane. The equations in [47] are corrected to deal with traces/residue, by replacing (6) with the following definition:

\[(y_i - \bar{y}) = A(x_i - \bar{x}) + C(z_i - \bar{z})\]  \hspace{1cm} (9)

and modifying (4) and (5) accordingly.

Once the slope of this plane is estimated, the vertical height perpendicular to the slope is calculated. Although measurements can be calculated at any height, for the DBH calculation a standard height of 1.3 m is considered. At this height the diameter \(D\) of the trunk is calculated.

Apart from the possibility of a sloping ground, tree trunks are not always vertical and so in order to effectively calculate the height one should also know the slope of the tree. Otherwise, this may result in erroneous calculations of the height and diameter. In order to cater for this problem, we estimate the slope of the tree slice at a given height \(L \pm 0.02\) m (B in Figure 9a) using (7)–(9). Examples of the two possibilities are shown in Figure 9. Let \(\alpha_0\) be the slope of the ground and \(\alpha_L\) be the slope of the tree slice at height \(L\) (about 1 m) respectively then corrected height \(h_c\) is given as:

\[h_c = 1.3 \times \cos \phi\]  \hspace{1cm} (10)

where \(\phi = \alpha_L - \alpha_0\). If \(\phi < 0\) then \(\phi < 90^\circ - \phi\).

The diameter is then calculated parallel to the slope of the slice (using (7)–(9)) at \(h_c \pm 0.02\) m height (C in Figure 9a). This allows for accurate calculation of the DBH value taking into account the local deviations as shown in Figure 9c. The diameter DBH is given as:

\[DBH = \frac{d_1 + d_2}{2}\]  \hspace{1cm} (11)

where \(d_1\) and \(d_2\) are the length of the sides of the bounding rectangle as shown in Figure 9c. Usually the value of DBH is calculated in cm.
Figure 9. (a,b) show the case of a sloping tree on a horizontal ground and a vertical tree on a sloping ground, respectively. In (c) we can see the local deformation/deviation which can lead to erroneous diameter estimations.

4.2. Height, Basal Area and Volume

The estimation of the true tree trunk height $H$ is very difficult as normally top part of the tree trunk is covered with soft canopy. As a result, the higher parts of the tree trunk are not fully visible in the laser scans. The tree trunk height $H$ is therefore taken (as per the definition for a standing tree [48]) as:
where \(T\) and \(C\) are the tree trunk and canopy, respectively, while \(h_i\) is simply the height function. \(T \cap C\) represents the visible overlapping portion of the tree trunk and canopy. This height \(H\) is corrected with respect to the slope of the ground and the tree as calculated in the previous Section 4.1.

Using the value of DBH calculated by (11), the Basal Area \(BA\) is given as:

\[
BA = \frac{\pi \times DBH^2}{4000} \tag{12}
\]

where \(DBH\) is in cm and \(BA\) is calculated in \(m^2\).

Calculating the volume \(V\) of the tree trunk is complicated by the fact that different sections of a typical tree trunk can be modeled by different shapes such as a cylinder, cone, parabola and neloid depending upon the species. Hence, approximating a generalized form for these tree trunks for volume calculation purposes is very difficult. Several methods are provided for fallen trees, however for standing trees this is even more complicated due to occlusion because of canopy, etc. For simplification purposes, and also currently used as standard practice in the forestry industry, we use the Huber method as presented by Akossou et al. [49]. In this method, the tree trunk is considered to be a paraboloid and the diameter \(d\) at the center height of the trunk is used for volume calculation as expressed in (13):

\[
V = \frac{\pi \times d^2 \times H}{40,000} \tag{13}
\]

where \(V\) is in \(m^3\), \(d\) in cm and \(H\) in m, respectively.

### 4.3. Map Building

A 2D map of the scanned area is generated containing the position of the tree (taken as the geometrical center of the lowest slice of the tree trunk) associated with the tree metrics (DBH, \(H\) and \(V\)). Using the estimated parameters of individual trees, the complete metrics for the whole of the forest patch is also calculated. This is the aggregation of the individual tree parameters as expressed in Table 1. An example of the generated Map of a forest patch (about 15 m \(\times\) 20 m) from the Forêt de Tronçais (Data_set-2 presented in Section 3.1) is shown in Figure 10.

![Figure 10. Generated map of a forest patch with different estimated parameters.](image-url)
Table 1. Estimation of parameters for a complex forest patch with \( n \) number of trees based on individual tree parameters.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean DBH</td>
<td>( \frac{1}{n} \sum_{i=1}^{n} DBH_i )</td>
</tr>
<tr>
<td>Mean height</td>
<td>( \frac{1}{n} \sum_{i=1}^{n} H_i )</td>
</tr>
<tr>
<td>Average BA</td>
<td>( \frac{1}{n} \sum_{i=1}^{n} BA_i )</td>
</tr>
<tr>
<td>Total Volume</td>
<td>( \sum_{i=1}^{n} V_i )</td>
</tr>
</tbody>
</table>

5. Results

The method was evaluated using 3 different types of data sets as presented in Section 3.1. For all three data sets, ground truth relating to segmentation and classification was obtained by manual annotations of the 3D points by visual inspection, whereas, for Data_set-3, actual measurements of tree metrics were conducted in the field for ground truth as explained in Section 5.2. Some qualitative results for the three data sets are presented in Figures 11–13, respectively.

In order to further evaluate the proposed method, more quantitative analysis was conducted. First, the segmentation and classification results were evaluated and then the parameter estimation results.

Figure 11. (a–c) present some segmentation and classification results for trees in the urban environment around the Kentucky area, USA using a vehicle borne mobile mapping scanning system (Data_set-1).
Figure 12. (a–c) present some segmentation and classification results for trees (*Quercus petraea*) from the Forêt de Tronçais in the Auvergne-Rhone-Alpes region (département de l’Allier) of Central France using a portable handheld scanning system (Data_set-2).

Figure 13. (a–c) present some segmentation and classification results for trees from a silver fir (*Abies alba* Mill.) tree forest in the Auvergne-Rhone-Alpes region (département du Puy-de-Dôme) of Central France using a static terrestrial scanning system (Data_set-3).

5.1. Tree Trunk Segmentation and Classification

The segmentation and classification results were evaluated for the 3 classes: [Tree trunk, Ground, Vegetation] the three different data sets.
The segmentation and classification quality was evaluated point-wise, i.e., the number of 3D points correctly classified as members of a particular class. The results presented in Table 2 are in the form of a confusion matrix in which rows and columns are the class labels from the ground truth and the evaluated method, respectively. The matrix values are the percentage of points with the corresponding labels using the metrics defined in [42]. If $T_i, i \in \{1, \ldots, N\}$ is the total number of super-voxels distributed into objects belonging to $N$ different classes in the ground truth and if $t_{jk}, i \in \{1, \ldots, N\}$ is the total number of super-voxels classified as a particular class of type-$j$ and distributed into objects belonging to $N$ different classes in the algorithm results (for example a super-voxel classified as part of the ground class may actually belong to a tree) then the ratio $S_{jk}$ ($j$ is the class type as well as the row number of the matrix and $k \in \{1, \ldots, N\}$) is given as $S_{jk} = \frac{t_{jk}}{T_k}$.

These values of $S_{jk}$ are calculated for each type of class and used to fill up each element of the confusion matrix, row by row. The segmentation accuracy (SACC) is represented by the diagonal of the matrix while the values of classification accuracy (CACC), overall segmentation accuracy (OSACC) and overall classification accuracy (OCACC) are calculated as explained in [42].

Compared to contemporary evaluation methods such as used by Koch et al. [50], employing a standard confusion matrix, this method is more suitable for this type of work as it provides more insight into segmentation results along with the classification results and directly gives the segmentation accuracy.

Tables 2–4 present the results for the three data sets, respectively. The overall scores of OSACC and OCACC, greater than 0.90 and 0.84, respectively, demonstrate the efficacy of the proposed method.

### Table 2. Segmentation and classification results for Urban Data Challenge data set (Data_set-1).

<table>
<thead>
<tr>
<th>Tree Trunk</th>
<th>Ground</th>
<th>Vegetation</th>
<th>CACC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tree trunk</td>
<td>0.924</td>
<td>0.013</td>
<td>0.005</td>
</tr>
<tr>
<td>Ground</td>
<td>0.018</td>
<td>0.900</td>
<td>0.013</td>
</tr>
<tr>
<td>Vegetation</td>
<td>0.004</td>
<td>0.001</td>
<td>0.910</td>
</tr>
</tbody>
</table>

Overall segmentation accuracy: OSACC 0.911
Overall classification accuracy: OCACC 0.940

### Table 3. Segmentation and classification results for the Forêt de Tronçais data set (Data_set-2), sessile oak trees (Quercus petraea (Matt.) Liebl.).

<table>
<thead>
<tr>
<th>Tree Trunk</th>
<th>Ground</th>
<th>Vegetation</th>
<th>CACC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tree trunk</td>
<td>0.900</td>
<td>0.037</td>
<td>0.270</td>
</tr>
<tr>
<td>Ground</td>
<td>0.081</td>
<td>0.901</td>
<td>0.028</td>
</tr>
<tr>
<td>Vegetation</td>
<td>0.022</td>
<td>0.180</td>
<td>0.900</td>
</tr>
</tbody>
</table>

Overall segmentation accuracy: OSACC 0.900
Overall classification accuracy: OCACC 0.843

### 5.2. Estimated Tree Parameters

In order to evaluate the tree metrics we used the static LiDAR data set (Data_set-3) as the ground truth for this data set was available. In order to obtain the ground truth, a patch of the forest (approx. $25 \times 25$ m containing 28 trees) was scanned and each tree marked/identified in the 3D scans. DBH values were physically measured for individual trees and then these trees were cut and the corresponding $H$ was measured for each one. Due to several limitations and the extensive work involved in obtaining the ground truth, i.e., physically marking the trees, diameter measurements at 1.3 m above ground, and specially cutting down these trees in order to measure the height of the fallen timber, only a limited area was considered similar to several other studies [51,52] which also
used similar plot sizes (and tree number) for evaluation purposes. The estimated parameter values obtained from the proposed method were then evaluated using the ground truth, as shown in Table 5 and regression plots in Figure 14.

**Figure 14.** (a,c,e) show the regression plots while (b,d,f) show the residuals for the estimated DBH, height $H$ and volume $V$, respectively.
Table 4. Segmentation and classification results for the silver fir (Abies alba Mill.) tree forest data set (Data_set-3).

<table>
<thead>
<tr>
<th></th>
<th>Tree Trunk</th>
<th>Ground</th>
<th>Vegetation</th>
<th>CACC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tree trunk</td>
<td>0.914</td>
<td>0.031</td>
<td>0.120</td>
<td>0.880</td>
</tr>
<tr>
<td>Ground</td>
<td>0.110</td>
<td>0.912</td>
<td>0.048</td>
<td>0.877</td>
</tr>
<tr>
<td>Vegetation</td>
<td>0.018</td>
<td>0.220</td>
<td>0.890</td>
<td>0.826</td>
</tr>
<tr>
<td>Overall</td>
<td>0.905</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Overall classification accuracy: OCACC 0.861

The results in Table 5 were evaluated using the following relations:

\[ \text{Average Error (AE)} = \frac{1}{n_d} \sum_{i=1}^{n_d} ||P_i^m - P_i^e|| \]

\[ \text{Root Mean Square Error (RMSE)} = \sqrt{\frac{1}{n_d} \sum_{i=1}^{n_d} (||P_i^m - P_i^e||)^2} \]

\[ \text{Average Relative Error (ARE)} = \frac{1}{n_d} \sum_{i=1}^{n_d} \left( \frac{||P_i^m - P_i^e||}{P_i^m} \right) \times 100 \]

\[ \text{Relative RMSE (RRMSE)} = \sqrt{\frac{1}{n_d} \sum_{i=1}^{n_d} \left( \frac{||P_i^m - P_i^e||}{P_i^m} \right)^2} \times 100 \]

where \( n_d \) is the total number of trees detected, \( P_i^m \) and \( P_i^e \) are the parameter values measured (in ground truth) and estimated for the \( i \)th tree, respectively.

The results show that the tree metrics are well estimated with an ARE less than 9%. DBH value was the best estimated (ARE < 2%).

Table 5. Error evaluation for different estimated parameters.

<table>
<thead>
<tr>
<th>Metric</th>
<th>AE</th>
<th>RMSE</th>
<th>ARE (%)</th>
<th>RRMSE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBH value</td>
<td>0.27 cm</td>
<td>0.31 cm</td>
<td>1.6</td>
<td>2.1</td>
</tr>
<tr>
<td>Height ( H )</td>
<td>1.40 m</td>
<td>1.64 m</td>
<td>7.8</td>
<td>8.6</td>
</tr>
<tr>
<td>Volume ( V )</td>
<td>0.19 m³</td>
<td>0.24 m³</td>
<td>8.5</td>
<td>9.7</td>
</tr>
</tbody>
</table>

6. Discussion

The segmentation and classification of tree trunks as well as the ground were fairly good even in the cluttered forest environment. The overall scores of OSACC and OCACC, greater than 0.90 and 0.84 respectively, demonstrate the efficacy of the proposed method. For Data_set-1, the OSACC and OCACC scores were the best showing that the trees in the urban environment were easy to segment/classify due to less clutter and a more structured environment. However, due to the highly cluttered forest environment, the OSACC and OCACC scores for the Data_set-2 and Data_set-3 were lower. Tables 3 and 4 show a relatively stronger interaction between the Tree trunk and Vegetation classes. This is due to the fact that sometimes parts of tree trunks were wrongly segmented/classified as Vegetation. This was very much evident in the case of Data_set-2 as the top parts of the tree trunks were misclassified as Vegetation because of the sharp sloping angles of the splitting branches, high noise and low point density (due to ZEB-1), see Figure 12c for example. Also, due to the criterion
of surface normals in the coefficients used in the classification, trees or branches with large sloping angles or steep curvatures were not correctly classified.

Sometimes the 3D points belonging to the ground were finally segmented/classified as Tree trunk, as can be seen from the tables. This problem is less evident in Data_set-1 due to a relatively smoother ground surface as compared to the forest environment where small bushes, long grass and fallen leaves deteriorate the conditions. This also results in some 3D points belonging to the ground being classified as Vegetation and vice versa. Good segmentation and classification enabled a fairly accurate calculation of height above ground and eventually the DBH value. The results show that the tree metrics are well estimated with an ARE less than 9%. DBH value was the best estimated (ARE < 2%). This value was just slightly underestimated as also concluded in the study by Calders et al. [53]. The typical cylindrical shape of the tree trunk for this tree species, silver fir (\textit{Abies alba} Mill.), (little variation in diameter), somewhat compensated for any small errors in the height above-ground estimation for the DBH value calculation. In the case of the estimation of height $H$, this was generally over-estimated (see Figure 14c) due to the fact that the top part of the tree trunk was usually not visible in the scans because of the canopy. As a result, the canopy height above the tree trunk was considered as the tree height. This is usually more than the actual height of the tree trunk. The over-estimation of the height $H$ also resulted in the over-estimation of the volume $V$, similar to presented by Hackenberg et al. [54] (see Figure 14e). An ARE of 8.5% was observed. These results are slightly better than some other similar studies conducted, such as by Dassot et al. [55] and Raumonen et al. [56]. Although these results could slightly vary with the tree species, they could definitely improve with less or no canopy (for example for some tree species leaves fall off during the autumn or winter season).

7. Conclusions

Automatic detection and measurement of tree parameters is an essential task for different forest inventory applications. This study presents a method for automatic detection of trees and estimation of some of their attributes, especially in the forest environment, using 3D terrestrial LiDAR data. The proposed method does not rely on any predefined tree shape or model, unlike many previous studies. The trees were successfully detected and different attributes (such as DBH, BA, height and volume) were then estimated at individual tree level and then aggregated to generate metrics for forest inventory applications. The method was validated and evaluated on different data sets obtained from different types of terrestrial sensors (vehicle borne, handheld and static) to demonstrate its applicability and feasibility for a wide range of applications.

The evaluation was done by comparing the estimated results with real field observations/measurements to demonstrate the efficacy of the proposed method. An overall segmentation and classification accuracy of more than 84% was obtained. Average parameter estimation error ranging from 1.6% to 9% was observed for the different parameters. The segmentation and classification of tree trunks as well as the ground were fairly good (even in the cluttered forest environment) enabling accurate calculation of height above ground and eventually the DBH value. However, occlusion of higher parts of tree trunk because of canopy resulted in a slight over-estimation of the height, and thus the volume. Although these results could slightly vary with the tree species, they could definitely improve with less or no canopy (for example for some tree species leaves fall off during the autumn or winter season). In order to further evaluate the robustness of the method, more tests could be conducted on larger forest patches with diverse tree species.

Acknowledgments: The authors would like to thank Dr. Robert Zlot and CSIRO Australia for the acquisition of 3D data using ZEB-1, Zebedee 3D Mapping System developed by CSIRO Australia.

Author Contributions: Laurent Malaterre, Paul Checchin and Laurent Trassoudaine conceived and designed the experiments; Laurent Malaterre and Laurent Trassoudaine performed the experiments concerning the forest in the Auvergne-Rhone-Alpes region (département du Puy-de-Dôme); Ahmad K. Aijazi and Paul Checchin analyzed the data; Ahmad K. Aijazi and Laurent Malaterre contributed materials/analysis tools; Ahmad K. Aijazi, Paul Checchin and Laurent Trassoudaine wrote the paper.
Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations
The following abbreviations are used in this manuscript:

- **ALS** Airborne Laser Scanning
- **BA** Basal Area
- **DBH** Diameter at Breast Height
- **DEM** Digital Elevation Model
- **DTM** Digital Terrain Model
- **GPS** Global Positioning System
- **ICP** Iterative Closest Point algorithm
- **IMU** Inertial Measurement Unit
- **MEMS** Microelectromechanical systems
- **PCA** Principal Component Analysis
- **TLS** Terrestrial Laser Scanning

References


© 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/).