Automatic detection of single street trees from airborne LiDAR data based on point segmentation methods

Zehra Cetin 1*, Naci Yastıklı 1

1Yıldız Technical University, Faculty of Civil Engineering, Department of Geomatics, Istanbul, Türkiye

**Abstract**

As a primary element of urban ecosystem, street trees are very essential for environmental quality and aesthetic beauty of urban landscape. Street trees play a crucial role in everyday life of city inhabitants and therefore, comprehensive and accurate inventory information for street trees is required. In this research, an automatic method is proposed to detect single street trees from airborne Light Detection and Ranging (LiDAR) point cloud instead of traditional field work or photo interpretation. Firstly, raw LiDAR point cloud data have been classified to obtain high vegetation class with a hierarchical rule-based classification method. Then, the LiDAR points in high vegetation class were segmented with mean shift and Density Based Spatial Clustering of Applications with Noise (DBSCAN) algorithms to acquire single urban street trees in the Davutpasa Campus of Yıldız Technical University, Istanbul, Turkey. The accuracy assessment of the acquired street trees was also conducted using completeness and correctness analyses. The acquired results from urban study area approved the success of the proposed point-based approach for automatic detection of single street trees using LiDAR point cloud.

1. Introduction

The trees on the streets are an important component of urban vegetation as creating shades, decorating roads, alleviating urban environmental pollution, reducing street noise, decreasing CO₂ emissions and building energy consumption, moderating heat accumulation in urban street canyons [1-4]. However, growth conditions of street trees can be very harsh as they have little space on the roadsides, and they can be affected by spread of diseases besides many natural and abiotic factors in single-species plantations [2, 5]. Also, street trees should be carefully positioned not to block utility lines below or above the ground and street luminaries [6]. Therefore, detection of the single street trees in urban areas is necessary for local governments to plan urban horticulture, manage and maintenance land use and land covers.

The inventory studies of urban street trees have been usually carried out by field investigation or manual visual interpretation of aerial images [4, 6]. The recent advent of the LiDAR systems provides rapid and cost-effective three-dimensional (3D) data acquisition of street trees [1]. Several segmentation approaches have been recommended to detect single trees using airborne laser scanning data [3, 7]. The initial techniques for identification of individual trees from LiDAR point cloud have been based on the methods which were developed to process optical imagery [8-10]. Region growing and watershed segmentation are the most popular and often used raster-based segmentation methods for airborne LiDAR data to obtain single street trees [2, 11-12]. Solberg et al. [13] employed a region-growing algorithm for single tree segmentation and their approach produced results equally good with other studies. Kwak et al. [14] used the watershed segmentation for delineation of individual trees and they concluded that...
LiDAR data can be successfully utilized for detecting single trees and estimating their heights. The template-matching [15], wavelet analysis [16], multi-stage filtering [17] and fitting functions [18] are one of the other methods for segmentation of the individual tree crowns. The loss of information owing to the interpolation of initial 3D point cloud to the grid structure is the major drawback of these segmentation methods [10, 19].

In this paper, we aimed to detect single street trees in the urban study area using raw LiDAR data with the suggested 2D point-based segmentation methods. An automatic hierarchical rule-based classification approach was firstly proposed to acquire high vegetation class. The mean shift and DBSCAN clustering methods had been utilized for automatic point-based segmentation of the high vegetation points to obtain single street trees in the Davutpasa Campus of Yildiz Technical University, Istanbul, Turkey as study area. The accuracy assessment of the acquired single street trees was also conducted using completeness and correctness analyses.

The study consists of five main sections. Study area and dataset is given in section 2. Section 3 contains the methods, including high vegetation classification, point-based segmentation, accuracy assessment subsections. The results and discussion section (section 4) presents the analysis and an experimental evaluation of 2D point-based urban street tree segmentation process. The study concludes with Section 5.

2. Study area and dataset

The Davutpasa Campus of Yildiz Technical University, which is located in Istanbul, Turkey was selected as the urban study area for this research (Fig. 1). There are different types of buildings, a wide variety of plant species and trees, driveways, walking paths, parking lots, recreation areas, as well as many street trees in the urban study area. Two different test areas, test area A and B, were used to automatic detection of single street trees from airborne raw LiDAR point cloud in the urban study area (Fig. 2 and Fig. 3).

The LiDAR data with the density of 16 points/m² was collected by Metropolitan Municipality of Istanbul in September, 2013 in the study area with “Riegli LSM-Q680i” laser scanner mounted on “Eurocopter AS350”. The flying height and speed of the helicopter were approximately 600 m and 148 km/h, respectively during the data acquisition with integrated “IGI DigiCam” camera, “IGI Aerocontrol” georeference system, and the LiDAR system. The ground truth data has been obtained by field investigation for accuracy assessment process of the proposed point segmentation methods.

Figure 1. The Davutpasa Campus of Yildiz Technical University (yellow line) (Google Earth, 2013)

Figure 2. Test area A: City map image (2013) (a), street trees (b), and the raw LiDAR point clouds colored by intensity (c)
3. Methods

In the present study, the point-based workflow to detect single street trees can be split in two parts as classification and segmentation. The flowchart of the proposed approach can be seen in Fig. 4.

### 3.1. Point-based classification of high vegetation

The point-based classification approaches of LiDAR points aim to determine an object class for every single laser point [20-21]. Different point-based classification methods including machine learning-based and rule-based classifications are available for the classification of LiDAR dataset. The 3D LiDAR point clouds are classified with the point-based classification methods using different classification features such as height features, eigenvalues, surface-based features, local plane features, multiple returns features, echo amplitude, echo width, etc. which are calculated for all individual LiDAR points [22-25].

In this paper, classification of LiDAR data of the test areas (test area A and B) was executed with proposed an automatic hierarchical rule-based classification method. The hierarchical rule set was constituted using the selected geometric features for point-based classification and after parameter analyses, the ground, low, medium and high vegetation, building, low point, air point, and default classes were acquired with the determined parameters (Table 1). The all-high vegetation points acquired as a result of the point-based classification were separated from the points of other terrain classes. The classification process of LiDAR point cloud has been achieved using TerraScan module of Terrasolid. The details of the used point-based classification method to acquire high vegetation class can be found in Yastikli and Cetin [12].

### Table 1. The hierarchical rule set and obtained classes

<table>
<thead>
<tr>
<th>Rules</th>
<th>Classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>By class</td>
<td>Default</td>
</tr>
<tr>
<td>Low points</td>
<td>Low point</td>
</tr>
<tr>
<td>Ground</td>
<td>Ground</td>
</tr>
<tr>
<td>Below surface</td>
<td>Low point</td>
</tr>
<tr>
<td>Air points</td>
<td>Air point</td>
</tr>
<tr>
<td>By height from ground 1</td>
<td>Low vegetation</td>
</tr>
<tr>
<td>By height from ground 2</td>
<td>Medium vegetation</td>
</tr>
<tr>
<td>By height from ground 3</td>
<td>High vegetation</td>
</tr>
<tr>
<td>Building</td>
<td>Building</td>
</tr>
</tbody>
</table>

### 3.2. Point-based segmentation of single trees

All individual laser points are grouped into subsets according to their similar characteristics with the point-based segmentation methods. Point-based methods mainly segment the data using geometric features [26]. The commonly adopted three important strategies for point-based segmentation of LiDAR data are geometric fitting, region growing, and clustering [26-27]. The high vegetation points, obtained from the automatic point-based classification of LiDAR data, were segmented using popular mean shift and DBSCAN clustering algorithms after detailed parameter analysis for best segmentation results in this research. 2D tree segmentation process for the detection of single street trees was carried out using Python programming language (Python 3.6.4) in Jupyter Notebook.
3.2.1. **Mean shift clustering**

Mean shift [28] is a nonparametric, recursive, and kernel-based clustering approach which shifts each data to local maximum of density function [29]. Mean shift does not require a pre-determined number of clusters and nor restrict the shape of clusters [30]. Mean shift is a center-based algorithm, and firstly, the algorithm chooses a random point from the data set as cluster center. Typically, it examines the center of mass of its local neighborhood and then shifts the point in the general direction of that center of mass [31].

In the mean shift clustering algorithm, the bandwidth of a kernel function is the most significant parameter [32]. Bandwidth, which determines the size of the region to be searched, can be adjusted manually or estimated using the bandwidth function. A specific bandwidth selection considerably decreases the bias while the variance value remains theoretically unchanged [32].

3.2.2. **DBSCAN clustering**

Density Based Spatial Clustering of Applications with Noise (DBSCAN) [33] is based on the concept of dense regions and performs well in determining arbitrarily-shaped clusters [34]. The three important elements for DBSCAN algorithms are the seeds members of the group, borders of the group, and noise which does not have any effect in the group [35-36]. The main idea of the DBSCAN is based on the requirement of a certain number of neighbor points at a certain radius for each core point [37]. The algorithm starts with an arbitrary core point, and controls if the neighboring points create a dense region [38]. A cluster is started if there are enough points in the neighborhood of the core point; otherwise, the point is labeled as noise [37]. The iterative process ends when no new points are added to any cluster [39]. DBSCAN algorithm requires two input parameters; the radius (the maximum distance of one core point to its neighbors) and minimum samples (density threshold of points in a neighborhood for a point to be regarded as a core point) [34, 40].

3.3. **Accuracy assessment**

Accuracy assessment is a necessary process to determine the performance of the segmentation methods. In this study, accuracy assessment process for the proposed segmentation methods was carried out according to completeness (Eq. (1)) and correctness (Eq. (2)) analyses [41-43].

\[
\text{Completeness} = \frac{(TP)}{(TP + FN)} \tag{1}
\]

\[
\text{Correctness} = \frac{(TP)}{(TP + FP)} \tag{2}
\]

TP, FP, and FN define perfect segmentation, oversegmentation, and under-segmentation, respectively [44]. TP refers to the true positive entities segmented correctly, FP is the false positive entities that were obtained in the segmentation but do not correspond to an entity in the ground truth data, and FN refers to the negative entities available in the ground truth data which were not acquired in the segmentation [45].

4. **Results and Discussion**

The automatic 3D point-based classification results of LiDAR point cloud in the test areas (test area A and test area B) using geometric features were given in Fig. 5, and only high vegetation points separated from the other terrain classes were shown in Fig. 6. According to the point-based classification results (see Fig. 5 and Fig. 6), it can be realized that the points of high vegetation class is obtained precisely for the tree crown segmentation of single street trees in the study area.

The 2D single tree segmentation results in test area A and B with mean shift clustering and DBSCAN clustering are given in Fig. 7 and Fig. 8, respectively, including the raw LiDAR point clouds, the segmented street trees, and the segmented street trees overlaid on the grey coded Digital Surface Model (DSM). When the segmentation results are analyzed, it is seen that the single street trees were successfully determined with the point-based segmentation using both mean shift and DBSCAN clustering algorithms. The segmentation results performed by mean shift and DBSCAN are close to each other in test area A and exactly the same in the test area B.

In Fig. 9, the reference urban single street trees, TP, FP, FN acquired with mean shift and DBSCAN clustering algorithms are given as a result of accuracy assessment of detected trees in the test area A. In the test area A, 17 clusters were determined as single street trees correctly (TP) using both mean shift and DBSCAN clustering algorithms. While there is no incorrectly clustered mean shift and DBSCAN clustering were obtained as 100% in the test area B (Table 2). The completeness and correctness results for both two segmentation methods are very satisfying as expected because of the accurate point-based classification of high vegetation class, and detailed parameter analyses for point-based segmentations. Mean shift clustering method single street tree with mean shift clustering, only 1 cluster were determined as single street tree wrongly (FP) with DBSCAN clustering. The points of 2 street trees couldn’t also be segmented as tree clusters (FN) using both mean shift and DBSCAN clustering. The results of the segmentation were 89.47% completeness and 100% correctness for mean shift clustering method, and 89.47% completeness and 94.44% correctness for DBSCAN clustering method (Table 2) in the test area A. In Fig. 10, reference urban single street trees, TP trees acquired with mean shift and DBSCAN clustering algorithms are given as a result of accuracy assessment of detected trees in the test area B. In the test area B, all 31 clusters were also detected as single street trees correctly (TP) using both mean shift and DBSCAN clustering algorithms. Since there were no single street trees identified wrongly (FP), and there were any single street trees couldn’t be segmented (FN) in the test area B, the completeness and correctness values for both mean shift and DBSCAN clustering were obtained as 100% in the test area B (Table 2). The completeness and correctness results for both two segmentation methods
are very satisfying as expected because of the accurate point-based classification of high vegetation class, and detailed parameter analyses for point-based segmentations. Mean shift clustering method outperformed slightly to DBSCAN clustering according to the correctness value only in the test area A.

Table 2. Completeness and correctness values

<table>
<thead>
<tr>
<th></th>
<th>Mean shift clustering</th>
<th>DBSCAN clustering</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test area A</td>
<td>89.47%</td>
<td>100%</td>
</tr>
<tr>
<td>Test area B</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

**Figure 5.** The automatic point-based classification results of test area A (a) and test area B (b)

**Figure 6.** The high vegetation points in test area A (a) and in test area B (b)
Figure 7. The test area A: raw LiDAR clouds colored by height (a), the segmented street trees using mean shift clustering (b) and the segmented street trees (red circles) overlaid on the grey coded DSM (c), the segmented street trees using DBSCAN clustering (d) and the segmented street trees (red circles) overlaid on the grey coded DSM (e).

Figure 8. The test area B: raw LiDAR clouds colored by height (a), the segmented street trees using mean shift clustering (b) and the segmented street trees (red circles) overlaid on the grey coded DSM (c), the segmented street trees using DBSCAN clustering (d) and the segmented street trees (red circles) overlaid on the grey coded DSM (e).
Figure 9. The test area A: the reference trees (a), TP, FP, FN trees acquired with mean shift clustering (b), and TP, FP, FN trees acquired with DBSCAN clustering (c)

Figure 10. The test area B: the reference trees (a), TP trees acquired with mean shift clustering (b), and TP trees acquired with DBSCAN clustering (c)

5. Conclusion

In this study, an automatic point-based segmentation approach is proposed to detect single street trees using raw airborne LiDAR point cloud. The LiDAR data were classified with hierarchical rule-based classification method, and acquired high vegetation class' points were segmented with mean shift and DBSCAN algorithms to detect single street trees automatically in the test area A and test area B which were located in Davutpasa Campus of Yildiz Technical University, Istanbul, Turkey. The accuracy assessment had been performed with respect to the detection rate of the single street trees in the test areas. The results of completeness and correctness were acquired 89.47% and 100%, respectively for mean shift clustering algorithm and 89.47% and 94.44%, respectively for DBSCAN clustering algorithm in the test area A. In the test area B, all completeness and correctness results are obtained 100% for both mean shift and DBSCAN clustering algorithms. Obtained results verified the success of proposed methods for automatic detection of single street trees using airborne raw LiDAR data without any information loss. The automatic segmentation results are quite satisfactory for both mean shift and DBSCAN clustering algorithms. The proposed approach for automatic detection of single street trees using airborne LiDAR data can be used effectively by local governments and city planners to plan urban horticulture, manage and maintenance land use and land covers studies.

Acknowledgement

We would like to thank the Metropolitan Municipality of Istanbul for supplying the LiDAR dataset covering the Davutpasa campus of Yildiz Technical University, Istanbul, Türkiye.

Author contributions

Zehra Cetin: Conceptualization, Methodology, Software, Data curation, Writing-Original draft preparation, Software, Validation, Visualization, Investigation, Writing-Reviewing and Editing. Naci Yastikli: Conceptualization, Methodology, Software, Data curation, Writing-Original draft preparation, Software, Validation, Visualization, Investigation, Writing-Reviewing and Editing.
Conflicts of interest

The authors declare no conflicts of interest.

References


