Automated detection of 3D individual trees along urban road corridors by mobile laser scanning systems

Wei Yao[a] and Hongchao Fan[b]

[a]Department of Geoinformatics, Munich University of Applied Sciences, 80333 Munich, Germany, yao@hm.edu
[b]Chair of GIScience, University of Heidelberg, Heidelberg, 69120 Germany hongchao.fan@geog.uni-heidelberg.de

KEY WORDS: mobile laser scanning, urban areas, tree detection, 3D segmentation

ABSTRACT(300 words): Recently, mobile laser scanning (MLS) has emerged as an efficient means to acquire massive 3D point clouds along urban road corridors for the several applications such as building footprint reconstruction, facade modeling and road traffic inventories. In this paper, we propose an automated strategy to address the issue of detecting 3D individual trees in urban traffic corridors using MLS data. Firstly, an approach based on analyzing the features in the spatial accumulation map of MLS point cloud is developed to identify man-made objects from natural objects by sequentially extracting man-made structures. The remaining points of natural objects should mainly contain vegetation and few vertical disturbing objects. This point class is further delivered to the next step to undergo a 3D segmentation process to obtain individual object instances by using a spectral clustering method. Once the point cloud is partitioned into various point segments corresponding to object instances such as single trees, poles or traffic signs, different fine classes can be further distinguished by using local shape descriptor of point segments. After the refinement of the segmentation, vegetation can be separated from other disturbing objects and detected as individual trees. An experimental study shows and evaluates the feasibility and capability of the proposed strategy towards detecting 3D individual trees from MLS data of complex urban corridors.

1. INTRODUCTION

Vegetation analysis has recently become a very important topic in the community of urban remote sensing. The results of such research can be potentially used in a wide range of data modeling across diverse applications, such as 3D city modeling, road inventory, natural resources mapping and renewable energy use (Secord and Zakhor, 2007; Jaakkola et al., 2010). So far imaging sensors are mostly used to characterize the urban vegetation from the airborne viewpoint. However, the detection is limited to describing and measuring the tree canopy surface and its 2D outline. On the other hand, in addition to extracting man-made objects airborne laser scanner (ALS) is increasingly used to acquire the 3D information about urban green areas such as vegetation in parks (Mallet et al., 2008; Hecht et al., 2008).. The methods for extracting urban objects from ALS data usually consist of three steps: geometric filtering, feature calculation and classification (Mallet et al., 2008). Other methods included an object-based point cloud analysis strategy for object extraction (Yao et al., 2011). Former studies showed that vegetation in city canyons still cannot be extracted very well from such data sources and imposed a great challenge to the data acquisition from an airborne platform. Therefore, a potential integration with mobile laser scanning system (MLS) is needed to accomplish the task due to synergistic benefits. Unlike ALS data, vehicle-borne MLS acquires object information from a ground perspective and is capable of providing high-fidelity and comprehensive data for the 3D modeling of objects in the vicinity of city transportation networks (Yang et al., 2012; Rutzinger et al., 2011a). It leads to a great potential for object extraction at street level and high-fidelity reconstruction of object models at single instance level. Although MLS data cannot be extended to cover the complete city area, it can be used as reference data for the model calibration by transferring the specific knowledge (e.g. fine tree species, stem diameter, etc.) to airborne observations.

However, most of recent works related to MLS data analysis focused on extracting man-made objects by designing class-specific features and adopting model-based approaches. To classify and extract various objects from MLS data, basic geometric structures such as plane and vertical pole are first extracted to represent the primitives for urban objects. Former works (e.g. Yang et al., 2012; Rutzinger et al., 2011a) presented approaches for the automatic extraction of vertical walls, road markings and poles, respectively. So far only little attention is paid to the detection and characterization of trees along urban road corridors from MLS data, although the analysis and estimation of vegetation parameters using ALS and TLS data have been intensively investigated in the context of forestry science (Maas et al., 2008). Since MLS is installed on a moving vehicle platform, it can not only acquire the tree object with high-detailed geometry but also cover lager areas. Tree features such as stem diameter and branch structure can be expected to be resolved with a much better accuracy. Additionally, Rutzinger et al. (2011b) have developed a strategy for the extraction and modeling of trees from MLS data. The tree detection algorithm is based on applying a segmentation followed by a classification. The detected point clouds of high vegetation are delivered to the model parameterization process where crown length, stem diameter and crown shape are determined. Most

recently, Monnier et al. (2012) have presented a tree detection method from heterogeneous and complex 3D point clouds acquired for the occlusion mapping in dense urban environments. This method couples 2D and 3D methods which represented the local geometrical shape of point clouds by 3D descriptors and retrieved individual objects by clustering similar points on the descriptor accumulation map.

The goal of this work is to develop an automated workflow for individual tree detection in city road corridors from MLS point clouds, which could retain the important geometric information for 3D tree modeling and characterization. The constructed tree models not only contribute to building a complete road tree cadastral but also to providing an accurate data basis for estimating biophysical properties at the single tree level. Firstly, an approach of analyzing the features in the horizontal accumulation space of MLS point cloud is developed to identify man-made objects from natural objects. The remaining points which mainly contain vegetation and some disturbing vertical objects are furthered delivered to a 3D segmentation process to obtain individual object instances. After the refinement of the segmentation step, the vegetation can be separated from other disturbing objects and detected as individual trees.

2. IDENTIFYING MAN-MADE OBJECTS IN MLS DATA BY CLASSIFICATION

In general, off-ground urban objects can be divided into man-made and nature objects, whereas man-made objects are referred to a variety of buildings, traffic facilities, power lines, poles, and cars, while nature objects are referred to vegetation. In this section, the approach for identifying man-made objects from natural objects is presented. The process consists of three steps: the preprocessing; the extraction of seed points for man-made objects; and the extraction of man-made objects.

2.1 Preprocessing

In this stage, ground points are detected by analyzing the height histogram (Yao et al., 2010) and are further delivered to a surface fitting algorithm (Gridfit) to obtain a ground level raster. The vertical distance of each point to the ground level is calculated as the height above the ground level (AGL). Man-made objects can be separated into three classes according to the AGL height: (i) fences or cars whose AGL heights are less than two meters; (ii) buildings whose AGL heights are larger than two meters; and (iii) power lines hanging in the air and having the lowest AGL height of five meters. In the vertical dimension, different objects could co-exist. For instance, a vehicle might be parked under a tree, while branches of a tree grow over the fence. This leads to error detection when using the method of projecting 3D points on a horizontal accumulator used in Hammoudi, (2009). To overcome such deficiency, the off-ground points of MLS data are divided into three layers according to the AGL height.

2.2 Extracting seed points for man-made objects

The second step aims to extracting seed points of man-made objects in each height layer. The ideal is rooted in the fact that man-made objects feature geometric regularity like vertical planes, while vegetation reveals a huge diversity of shapes and point distribution. If 3D points are projected onto 2D horizontal plane, the position where man-made objects are located will indicate a large accumulation density.

For each height layer, an horizontal accumulation space of a 2D $M \times N$ grid is created respectively, where $M = ceil((X_{max} - X_{min})/S_x)$, and $N = ceil((Y_{max} - Y_{min})/S_y)$. The cell size (S_x, S_y) depends on the density of MLS data. In this work, the grid cell size is set to 0.25m in square. It counts how many points fall within each cell. Vertical structures such as walls, fences, poles, and trucks of trees are scanned with high density and accumulated with large number of points when projected onto horizontal cells. Consequently, for a cell in which a large number of points fall, the centroid of these points is calculated and recoded for the cell. If there are only few or even no points falling within a cell, there might be leaves or horizontal objects hanging over the cell. These points are omitted as noises and the cell is kept as empty. The process uses a global threshold of ten points.

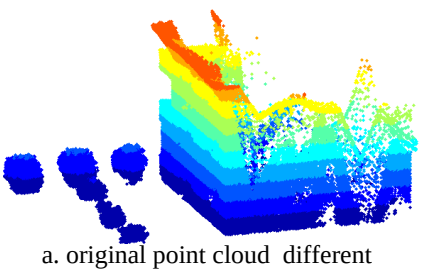
$$G_{footprint} = \begin{cases} \emptyset, & N_p < 10 \\ X_{foot} = \frac{1}{N_p}\sum_{i=1}^{N_p} X_i \,,\; Y_{foot} = \frac{1}{N_p}\sum_{i=1}^{N_p} Y_i \,, & N_p \gg 10 \end{cases} \quad (1)$$

The result of this process is called footprint map, where not only the footprint of buildings and vehicles, but also roots of poles and trees are recorded in cells, as well as shrubs with high dense leaves. In order to get rid of these influences, linear segments are indicated, since other than that of buildings or vehicles the footprint of vegetation always forms as a cluster. The footprint map is first converted into a binary image by taking grid cells as pixels, where the pixel value is set as 1 for a non-empty cell and 0 for an empty cell. Then edges are detected using Canny operator (Canny, 1986) with sensitivity thresholds of [0.1 0.2] and the sigma of the Gaussian filter is set to 1. The extracted line segments (Figure 1d) are then aggregated when two line segments are collinear. In the next step, the detected edges are filtered by applying a length threshold which is set to two meters, as the shortest edge of a building footprint should be longer. On the other hand, the long side of a car is normally longer than two meters. As the result, only line segments of building footprints, fences and vehicles can remain after filtering. When transferring extracted line segments back to the footprint-map, a cell should contain a set of seed points of man-

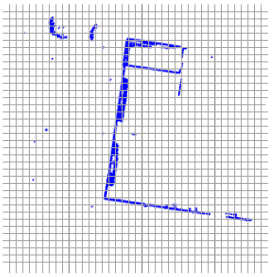
made objects, if the cell is located on or closed to a line segment. Therefore, a rectangular buffer zone is generated around a line segment by taking the long side parallel to the line segment and a width of one meter ($d = 0.5m$) with the consideration of the accuracy of line extraction and noisy measurement. The rectangle is generated as a buffer zone for the line segment according to Eq.2.

$$P_{b,i} = P_m + P_i \cdot \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}, i = 1,2,3,4 \quad (2)$$

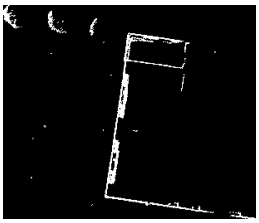
where $P_{i=1,2,3,4} = \left[\left(\frac{s}{2}+d,-d\right)\left(\frac{s}{2}+d,d\right)\left(-\frac{s}{2}-d,d\right)\left(-\frac{s}{2}-d,-d\right)\right]$, P_m is the middle point of the line segment $\overline{P_a,P_e}$, s is the length of the line segment, and θ is the orientation. After buffer zones are generated for extracted line segments, the spatial relationship between the buffer rectangle and the laser points are examined regardless of the vertical dimension. If a laser point falls into a buffer zone, it is regarded as the seed point of a man-made object.



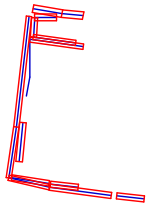
a. original point cloud different heights are coded with different colors)



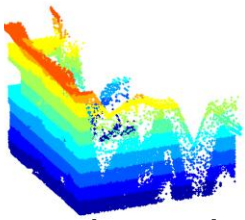
b. Projection onto 2D accumulator



c. Binary image of the accumulation space



d. Line segments and their buffer rectangles



e. The extracted facade structure

Figure 1: Extracting 3D points on vertical structure of man-made objects

2.3 Identifying man-made objects by distinguishing seed points

The last step provided seed points of man-made objects in the three layers. Every cluster formed by seed points corresponds to an individual urban object. They will be distinguished to different types of objects at first. As indicated in step of pre-processing, cars and fences exist only in the lower layer ($h \leq 2m$), while power lines can only be found in the upper layer ($h \geq 5m$), and building points could be found in every layer. To this end, the seed points are distinguished by cross-comparison through the three layers.

- For a cluster G_{lower} in the lower layer, if there is a cluster G_{middle} at the same location in the middle layer, then the points of the cluster G_{lower} are seed points of a building. Otherwise, the points of cluster G_{lower} are seed points of a car or a piece of fence.
- For a cluster G_{upper} in the upper layer, if there is no cluster in the same location in the middle layer, then the points in the cluster G_{upper} are seed points of power lines.

Point pairs are found out between two clusters of seed points (e.g. $G_{lower,a}(P_{i=1,\dots,M})$, $G_{middle,b}(P_{j=1,\dots,N})$) on different layers. Two points (P_i, P_j) are assumed to be paired if they are almost located at identical position (the distance is smaller than 5cm,$d(P_i, P_j) \leq 0.05m$) on the horizontal plane. If most of points of a cluster in a layer can be paired to points of a cluster in another layer with a condition: $\frac{K}{\min(M,N)} \geq 95\%$, the two clusters are deemed to be at the same location, where K stands for the number of the paired points, min(M,N) indicates the minimal number of points within the two clusters.

After seed points are extracted for different man-made objects, 3D points of which can be identified according to topological relations to the seed points. This is achieved by a two-step method. Firstly, the convex hull of each cluster of seed points is calculated in order to reduce the search area. The spatial relation of points falling within the convex hull of the seed points is calculated. The convex hull is calculated for each cluster of seed points:

- the seed points of an object class are checked regardless of their heights. If a point is located inside of (and on) the convex hull of the cluster of seed points, this point is regarded as candidate point of object class

In the next step, a circular neighborhood is generated for every seed point in the cluster, whose radius is set to 0.5m. Then, candidate points of each object class are examined with respect to the spatial relationship to the circle. Those candidate points are regarded to ought to be projected to the same footprint as the seed point, if they fall within the circle of this seed point. And they are the points belonging to the object which is represented by the cluster of seed points. Finally, all the points belonging to the certain object class are identified. The points which do not fall in any circle of seed points will be released from the candidate points. So far, man-made objects are extracted and distinguished. The remaining points represent natural objects consisting of vegetation (trees), poles of power lines, and street lights.

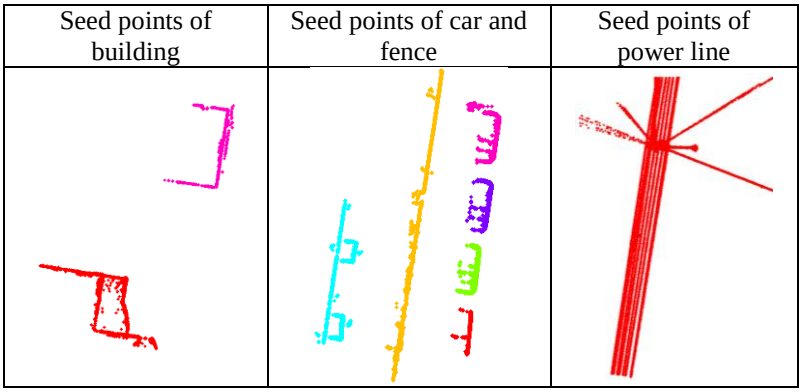


Figure 2: Seed points of different types of man-made objects

3. SEGMENTATION OF 3D INDIVIDUAL TREES

After natural objects are separated from man-made objects, 3D individual objects are to be segmented by using a spectral clustering method performed to analyze a graph structure of point clouds. The similar algorithm has been successfully applied to airborne LiDAR data (Reitberger, et al., 2009; Yao and Wei, 2013). However, MLS data feature a much higher point density and model fidelity which posed a new challenge for 3D tree segmentation. It will be of great interests to examine whether the spectral graph clustering method can also be successfully applied to MLS data and how it will perform.

Likewise, point clouds of natural objects can be converted into the Object Height Model (OHM) in which the object height is recorded in a 2D grid cell. Then, single objects are segmented by delineating contours in the OHM. However, all the OHM based segmentation methods suffer from the drawback that sub-dominated object in the lower height layer cannot be recognized since they are invisible in the OHM. The key idea of the 3D segmentation technique is to apply the normalized cut segmentation to a voxel structure (e.g. voxel size = 0.5m) of point clouds representing natural objects. Here, the voxel structure especially makes sense to the MLS data, since the point density of MLS data is very high and the required computational effort can be considerably reduced by using the voxel as the segmentation unit. Voxel size is adjustable according to the point density and average size of trees to be segmented. The segmentation uses the positions (x_i, y_i, z_i) of the laser reflections and optionally the pulse intensity I_i. The normalized cut segmentation applied to the voxel structure of point clouds is based on constructing and analyzing a region adjacency graph G. The two disjoint segments A and B of the graph are found by maximizing the similarity within each segment and minimizing the similarity between the two segments via solving the cost function:

$$NCut(A,B) = \frac{Cut(A,B)}{Assoc(A,V)} + \frac{Cut(A,B)}{Assoc(B,V)} \tag{3}$$

with $Cut(A,B) = \sum_{i \in A, J \in B} w_{ij}$ as the total sum of weights between the segments A and B and $Assoc(A,V) = \sum_{i \in A, j \in V} w_{ij}$ as the sum of the weights of all edges ending in the segment A. The weight w_{ij} specifies the similarity between each two graph node voxels and are a function of the LiDAR point distribution and features calculated from I_i or other spectral domain. A minimum solution for Eq. (3) is found by means of a corresponding generalized eigenvalue problem. It turned out that the spatial distribution of the LiDAR points mainly influences the weighting function. Consequently, the weighting matrix consists of following similarity functions between voxel i and voxel j which is located within a cylinder of radius r_{XY} around the voxel i:

$$w(i,j) = \begin{cases} e^{-X(i,j)} \times e^{-Z(i,j)} \times e^{-F(i,j)} & \text{if } (D_{ij}^{XY} < r_{xy}) \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

The components $X(i,j)$ and $Z(i,j)$ weight the quadratic Euclidian distances between the voxels, where $X(i,j)$ is the horizontal and $Z(i,j)$ is vertical distance. The component $F(i,j)$ describes the quadratic Euclidian distance between

mean LiDAR intensities The position and height of a segmented object is defined as the centroid and maximum height of the tree segment from the point cloud.

Since the classified point set of natural objects represent not only vegetation (trees) object but also some other vertical objects including poles, street lights and traffic signs, 3D segmentation results need to be refined in order to exclude disturbing objects. This refinement process can be done based on the shape analysis of each point segment, where a shape descriptor is created to measure and discriminate point segments, since almost all the non-tree objects within the class are characterized by the pole-like form. The shape descriptor for measuring the maximal diameter and area of the horizontal distribution of point segments are the most useful features, which can be derived by analyzing the convex hull of 2D projection of point segments onto the x-y plane.

4. EXPERIMENT AND ANALYSIS

4.1 Dataset

The point clouds of urban road corridor used in this study were captured by an Optech Lynx Mobile Mapper system, which consists of two Optech laser scanners, one GPS receiver, and an Applanix IMU, which collects LiDAR data at more than 100,000 measurements per second with a 360° field of view. The point clouds of two sites were selected for assessing the performance of the proposed method. The two test sites are located in an urban residential area with dimensions of 190 m by 130 m. It contains vegetation (e.g., trees, bushes), grass field, fences, buildings, power lines, multi-lanes roads, and cars. Reference trees were selected by human inspection and their position and height were obtained by manually locating the tree-top position directly in LiDAR point clouds and measuring its height above the ground level.

4.2 Results and discussion

The detected results are evaluated by linking detected trees to the reference ones. Segmented trees that are linked to one tree position in the reference are so-called 'detected trees', if (i) the distance to the reference tree position is smaller than 1.5 m and (ii) the height difference between the reference and detected tree is smaller than 15% of top tree height. If a reference tree is assigned to more than one tree position in segmentation results, the tree position with the minimum distance to the reference tree is selected. The evaluation measures are defined as follows: Completeness = $Number_{\text{true positives}}/Number_{\text{reference trees}}$; Correctness = $Number_{\text{true positives}}/Number_{\text{detected trees}}$; Tree height deviation = $| Height_{\text{detected tree}} - Height_{\text{reference tree}} |$; Tree position deviation = $\text{sqrt}[(X_{\text{reference tree}} - X_{\text{detected tree}})^2 + (Y_{\text{reference tree}} - Y_{\text{detected tree}})^2]$.

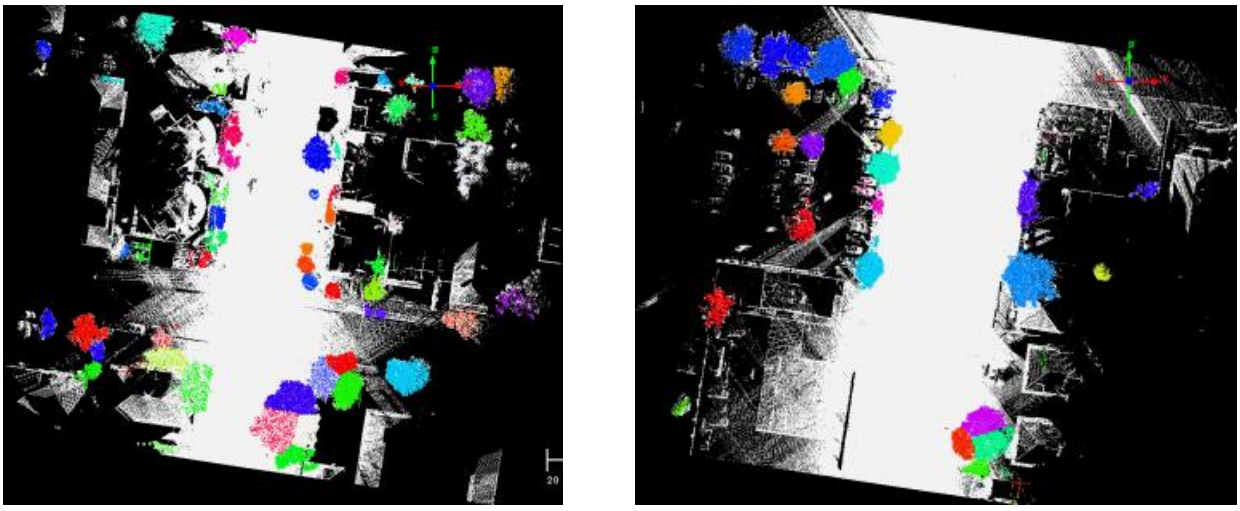


Figure 3: Extraction and representation of 3D individual trees in two test sites

The detection results of 3D individual trees for the two datasets are illustrated in the original point clouds using different colors in Figure 3. The quantitative results of the evaluation can be found in Table 1. A promising result has been obtained for the first dataset while the result of second one is slightly better. Although the point density and sensor system of the two datasets are identical, the results of the two datasets are not directly comparable, since the scene complexity, structure and the composition of tree species could still lead to different impacts on algorithm performance. It can be seen that properties of the first dataset, such as more dense trees, building occlusion and complex residential courtyard, could make the detection of individual trees in this area more difficult. Most of miss-detected trees in this case are traced back to the misclassification caused by confusion with adjacent man-made objects, building occlusion and under-segmentation. The position and height of extracted individual trees have been

determined and compared to that in reference data. The mean and standard deviation of residuals for the tree position and height reached 1.15m/0.86m and 0.9m/ 0.82m, respectively.

Table 1: Quality for individual tree detection

Dataset I		Dataset II	
Completeness	Correctness	Completeness	Correctness
80.7%	70.2%	81.2%	75.5%

5. CONCLUSION

This paper has presented an automated approach to detect 3D individual trees along road corridors of complex urban areas based on employing MLS data. It is a rigorous attempt made to characterize the individual trees along urban road corridors in full three dimensions from a vehicle platform. The experiments upon two urban datasets showed that promising results can be achieved for both residence areas. Up to 81% of dominate trees located along urban road corridors can be detected and characterized in the point cloud via 3D segmentation, while the promising accuracies for both position and height of detected trees are achieved. Since it is still impossible to acquire reference trees in true 3D, the scheme to evaluate detection results with respect to the tree position and height is believed to be a plausible method for performance assessment at this stage. Future works will be put on deriving extended tree properties based on 3D segments such as stem diameter and crown base height, and studying the feasibility of distinguishing tree species and vitality conditions.

6. REFERENCES

Canny, J., 1986. A Computational Approach to Edge Detection. In: IEEE Trans. Pattern Analysis and Machine Intelligence, 8(6):679–698.

Hammoudi, K., Dornaika F. and Paparoditis N. 2009. Extracting building footprints from 3D point clouds using terrestrial laser scanning at street level. In: Stilla U., Rottensteiner F., Paparoditis N. (Eds) CMRT09. IAPRS, Vol. XXXVIII, Part3/W4, Paris, France, 3-4 September, 2009.

Hecht, R., Meinel, G. and Buchroithner, M. F., 2008. Estimation of urban green volume based on singlepulse LiDAR data. IEEE Transactions on Geoscience and Remote Sensing, 44(11): 3832–3840.

Jaakkola A., Hyyppä J., Kukko A., Yu X., Kaartinen H., Lehtomäki M., Lin, Y., 2010. A low-cost multi-sensoral mobile mapping system and its feasibility for tree measurements, ISPRS Journal of Photogrammetry and Remote Sensing, Volume 65, Issue 6, Pages 514-522.

Mallet, C., Soergel, U., Bretar, F., 2008. Analysis of fullwaveform LiDAR data for an accurate classification of urban areas. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. 37 (Part 3A). Beijing, China.

Maas, H.G., Bienert, A., Scheller, A. and Keane, E., 2008. Automatic forest inventory parameter determination from terrestrial laser scanner data. International Journal of Remote Sensing, 29(5): 1579–1593.

Monnier, F., Vallet, B., Soheilian, B., 2012. Trees Detection from LASER Point Clouds Acquired in Dense Urban Areas by a Mobile Mapping System. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume I-3, pp. 245-250, 2012, XXII ISPRS Congress, Melbourne.

Reitberger, J., Schnörr, Cl., Krzystek, P. and Stilla, U., 2009. 3D segmentation of single trees exploiting full waveform LIDAR data, ISPRS Journal of Photogrammetry and Remote Sensing, vol.64, no.6, pp.561-574.

Rutzinger, M.; Hoefle, B.; Oude Elbernik, S.; Vosselman, G., 2011a.Feasibility of facade extraction from mobile laser scanning data. Photogrammetrie, Fernerkundung und Geoinformation. 2011, Heft 3 (2011), p. 97 – 107.

Rutzinger, M., Pratihast, A. K., Oude Elberink, S. J. and Vosselman, G., 2011b. Tree modelling from mobile laser scanning data-sets, The Photogrammetric Record, Papers 26, 361-372.

Secord, J., Zakhor, A., 2007. Tree Detection in Urban Regions Using Aerial Lidar and Image Data, Geoscience and Remote Sensing Letters, IEEE, vol.4, no.2, pp.196-200.

Yang, B., Fang, L., Li, Q., Li, J., 2012. Automated Extraction of Road Markings from Mobile Lidar Point Clouds. Photogrammetric engineering and remote sensing.vol. 78, no4, pp. 331-338.

Yao, W., Hinz, S., Stilla, U., 2011. Extraction and motion estimation of vehicles in single-pass airborne LiDAR data towards urban traffic analysis. ISPRS Journal of Photogrammetry and Remote Sensing 66(3), 260-271.

Yao, W., Hinz, S., Stilla, U., 2010. Automatic Vehicle Extraction from Airborne LiDAR Data of Urban Areas aided by Geodesic Morphology. Pattern Recognition Letters, Vol.31, no.10, pp.1100-1108.

Yao, W., Wei, Y., 2013. Detection of 3D Individual Trees in Urban Areas by Combining Airborne LiDAR Data and Imagery. Geoscience and Remote Sensing Letters IEEE, Volume 10, Issue 6, DOI: 10.1109/LGRS.2013.2241390.