Automatic Stem Mapping Using Single-Scan Terrestrial Laser Scanning

Xinlian Liang, Paula Litkey, Juha Hyyppä, Harri Kaartinen, Mikko Vastaranta, and Markus Holopainen

***Abstract*—The demand for detailed ground reference data in quantitative forest inventories is growing rapidly, e.g., to improve the calibration of the developed models of airborne-laser-scanning-based inventories. The application of terrestrial laser scanning (TLS) in the forest has shown great potential for improving the accuracy and efficiency of field data collection. This paper presents a fully automatic stem-mapping algorithm using single-scan TLS data for collecting individual tree information from forest plots. In this method, the stem points are identified by the spatial distribution properties of the laser points, the stem model is built up of a series of cylinders, and the location of the stem is estimated by the model. The experiment was performed on nine plots with 10-m radius. The stem-location maps measured in the field by traditional methods were used as the ground truth. The overall stem-mapping accuracy was 73%. The result shows that, in a relatively dense managed forest, the majority of stems can be located by the automatic algorithm. The proposed method is a general solution for stem locating where particular plot knowledge and data format are not required.**

***Index Terms*—Forestry, laser measurement application, remote sensing.**

I. Introduction

AIRBORNE laser scanning (ALS), also known as airborne Light Detection And Ranging (LiDAR), has already been used in practical forest inventories in the Nordic countries. The two main estimation approaches for deriving forest information from ALS data have been those based on statistical canopy height distribution, also known as area-based method (e.g., [1]), and individual tree detection (e.g., [2]). These categories relate to the scale and accuracy requirements of the forestry information, the scale of the processing method, and the available point density [3]. In both approaches, the estimation models are developed from ALS and field reference measurements. It has been suggested that accurate ground reference data are important to estimate and calibrate stand characteristics in area-based and individual-tree-based inventories relying on ALS data (e.g., [4], [5]). It is also expected that the ALS-individual-tree-based inventory requires several hundreds of calibration trees in a similar way as the ALS-area-based inventory is using several hundreds of calibration plots.

The ground reference data are conventionally collected by means of manual measurements and interpretation in the field. It has widely been recognized as time consuming and labor extensive. Automatic means to obtain detailed properties of stems and trees are needed.

Terrestrial laser scanning (TLS), also known as ground-based LiDAR, provides a new solution for collecting the reference data in the forest environment. This technique acquires inventory parameters in an unequivocal, objective, and reproducible manner [6], and it makes automated, noninvasive, and expedient field mensuration possible [7]. The main advantages lie in its potential to improve the accuracy and efficiency of field inventories and to provide additional features for forestry applications. The data collected by TLS are high-precision three-dimensional (3-D) measurements of the targets in the form of a point cloud (x, y, z, intensity of the backscattered power). It has been shown to be promising in the estimation of various attributes for a single tree or a small number of individual trees, such as diameter at breast height (DBH) (e.g., [8]), leaf area index (e.g., [9] and [10]), crown shape (e.g., [11] and [12]), stem curve (e.g., [13]), volume (e.g., [14]), and detailed tree models (e.g., [15] and [16]).

The stem location within the plots is the fundamental parameter for the calibration of the ALS-individual-tree-based inventory. It is the main matching criterion between reference and ALS sets. If there exists good relative accuracy of tree positions within the plot, it is possible to use matching techniques [17] to compensate the relatively poor absolute accuracy of plot centers, which is typically about several meters. However, if the relative accuracy of tree positions inside the plot is poor, it is difficult to find the matching pairs, even though the absolute accuracy might be high.

Research on collecting stem locations from TLS data has been conducted since TLS has been used in forest-related studies. Nevertheless, the solutions available are not yet adequate. A more detailed description of the existing algorithms for stem locating is given in Section II. To improve the efficiency of reference collecting, methods for stem mapping from TLS data need to be improved.

When applying TLS data in forest inventories, both multi- and single-scan measurements can be used. In multiscan measurement, several scans are made from different positions in- and outside of the plot. Scans are accurately coregistered

Manuscript received September 28, 2010; revised March 4, 2011 and May 3, 2011; accepted June 11, 2011. Date of publication August 18, 2011; date of current version January 20, 2012. This work was supported in part by the Academy of Finland under the Projects "Improving Forest Supply Chain by Means of Advanced Laser Measurements" and "Science and Technology Towards Precision Forestry" and in part by the Finnish Funding Agency for Technology and Innovation under the Project "Development of Automatic, Detailed 3-D Model Algorithms for Forests and Built Environment."

X. Liang, P. Litkey, J. Hyyppä, and H. Kaartinen are with the Department of Remote Sensing and Photogrammetry, Finnish Geodetic Institute, 02431 Masala, Finland (e-mail: xinlian.liang@fgi.fi; paula.litkey@fgi.fi; juha.hyyppa@fgi.fi; harri.kaartinen@fgi.fi).

M. Vastaranta and M. Holopainen are with the Department of Forest Sciences, University of Helsinki, 00014 Helsinki, Finland (e-mail: mikko.vastaranta@helsinki.fi; markus.holopainen@helsinki.fi).

Digital Object Identifier 10.1109/TGRS.2011.2161613

typically using reference targets to obtain the complete data set. The merged point cloud normally covers the whole object and potentially improves the accuracy of the estimation. Yet, the automation of coregistration is not available but preferably needed in forest inventories. In single-scan measurement, the scanner is placed in the center of the plot, and one full-field-of-view scan is performed. As laser points are reflected from the nearest surface in the direction of laser beams, the opposite side of the object and the objects behind it are missed in the data set. However, the data volume is smaller compared to the merged point cloud, and point-cloud-level data matching is avoided. It is also possible to use several single scans, without the coregistration of multiple point clouds, to improve the accuracy. In this case, each scan is processed separately, and the registration of scans is done at the object level. For example, in the application of stem mapping, the registration of several scans can be performed by the found trees of each scan. To date, most studies have been based on multiscan TLS data.

This paper presents a fully automatic algorithm based on single-scan TLS data for stem mapping. The emphases are on the following: first, to develop a new and efficient point-processing method, and second, to explore the applicability of single-scan TLS data for stem locating in dense forests and to bring new knowledge in this context. The plot densities tested ranged from 500 to 1500 stems/ha (three of them were around 600 stems/ha and six were above 1000 stems/ha). In previous studies, the plot density has been between 300 and 600 stems/ha (e.g., [18] and [19]). In addition, the test plots were mostly mixed forest, composed of trees with different species and at different stages of growth.

This paper is organized as follows. Stem-mapping techniques are summarized in Section II. Section III describes the study material and the new mapping algorithm. Section IV provides the experimental results for nine field plots, the discussion concerning the method, the applicability of the single-scan data in dense forests, and future work. Section V is the conclusion. The Appendix presents the parameterization of a right circular cylinder.

II. Techniques in Stem Mapping

Conventionally, two measuring methods are applied when using circular and rectangular plots. With circular plots, the stem positions are measured relative to the plot center (throughout this paper, the term "stem" is used interchangeably with the term "trunk"). The plot center coordinate is collected by means of a Global Positioning System (GPS) device, and the relative stem location within the plot is measured by means of, for example, a rangefinder and a bearing compass. With rectangular plots, the corner points of the plot are determined by means of GPS or tacheometric measurements and ground control points, and the distances between the stems and the plot edges are determined using, for instance, a tape measurement. Absolute stem locations are then calculated in both cases.

Alternative solutions to improve position measurement have been proposed. In [20], the laser relascope is used to measure the trunk locations relative to the plot center. Distance is measured by means of a laser rangefinder, and angle is determined by means of an electronic compass. In [21], the tree locations are calculated from known tree positions by photogrammetric measurements, angle and distance measurements by field observations, and least squares adjustment. Other instruments, like Haglöf PosTex (Haglöf Sweden AB, Långsele, Sweden), also give the tree coordinates in the plot. In general, these solutions improve the data-collection efficiency but still rely heavily on manual measurements in the field. The GPS-based method provides an accurate method for achieving absolute position measurements. The main drawbacks are, however, that the dense canopy in the forest environment often disturbs GPS measurements and that the GPS antenna is usually somewhere apart from the measured object. A combination of GPS and inertial observations to collect terrestrial coordinates under the forest canopy has been reported in [22]. Inertial observations are used to collect coordinate data when the GPS signal is unavailable. It is possible to use this method to collect stem positions in the plot, but the individual stem measurement is required and the instrument used is quite expensive and heavy.

TLS data can be used to determine the relative stem location within a plot. Absolute mapping requires matching of the result to known coordinates, measured by, for instance, GPS. The application of TLS data significantly reduces the amount of manual measurement. So far, three main method types have been reported. They are two-dimensional (2-D)-layer searching, range-image clustering, and point-cloud processing.

In the 2-D-layer searching technique, a slice with a certain thickness is cut from the original point cloud. Points inside the slice are projected onto the layer. Tree trunks are identified by means of point clustering or circle finding [19], [23]–[25] or by waveform analysis using the Echidna[1] validation instrument [26]. Multilayer estimation has also been proposed (e.g., [23]). Since the searching takes place within the 2-D space, the amount of computation needed is fairly modest. However, trees within the plot need to be uniform, e.g., of the same age and/or species, as all trees are assumed to present a clear trunk at the same height where the slice goes through the point cloud. In the case where branches are present or nearby branches are overlapped in the layer, the detection of stems becomes a problem. It has been reported that, in a mixed deciduous stand, manual trunk detection in a TLS data layer meets difficulties [7]. In addition, knowledge of the terrain is necessary in order to enable constructing the slice at a particular height. This is achieved by assuming that the terrain is flat, which is not always the case, or by producing a digital terrain model (DTM), which is still a challenge in the forest area. As a result, the applicability of this method depends on the forest type and the type of terrain or the method used to generate DTM.

When applying the range-image clustering method, points, or pixels, in the range image are grouped according to local properties, e.g., the distance or surface curvature [27]–[29]. The main advantage of this technique lies in that the amount of the computation required is modest, as searching for the neighboring points in the 3-D space can be performed in an approximate manner in the 2-D space. Another potential benefit is that it is also possible to detect trunks from far away [30].

[1] http://pubs.casi.ca/doi/abs/10.5589/m08-046.

However, the applicability of this method is also partly limited. It is not always possible to get or reconstruct the 2-D image structure. The range image is not always maintained during the processing pipeline. It may be deleted before data delivery or destroyed at some stage to reduce the data volume for example.

In the point-cloud processing method, the attributes of an individual point are estimated in a local neighborhood, and the trunk is identified by means of semantic interpretations. In [31], the points are supposed to belong to one of the following three classes: surface, linear structure, and porous volume. Stems can be found in the linear structure class. Points are labeled with Bayesian classification, and stems are identified by means of semantic interpretation. In general, this technique needs more computation than the other two, but it does not require particular plot knowledge or data structure. The method processes the point cloud itself and can be used in different types of forest. So far, this technique requires the hand-labeled training data set in order to learn the parametric distribution model. In addition, it is mainly designed for robot perception applications, where the tasks, targets, and emphases are clearly different from those in forest inventories.

In previous studies on stem locating in TLS data, the proportions of located stems from single- and multiscan data have been 22% and 52% within a plot with 556 stems/ha [18], respectively; 97% within plots with an average density of 321 stems/ha and 100% within a plot with 310 stems/ha [19]. In [30], 85% of the trunks, which can be manually identified from TLS data, can be automatically detected in single-scan data. In [25], it is reported that all trunks, which can be located manually, can be automatically detected in multiscan data. In other studies, accuracy is not reported or the stems are located by means of manual detection [7], [32]. As a summary, there is a need to further develop techniques for stem mapping with TLS. In the future, benchmarking tests with various methods should be performed in various types of forest.

III. Material and Methods

A. Study Area and Data

The study area is a managed forest, located in the vicinity of Evo, Finland (61.19° N, 25.11° E). The research material is composed of nine circular plots each with a fixed radius of 10 m, and it includes both treewise field references and TLS measurements.

The main tree species growing on the plots are pine, spruce, and birch at various stages of growth. The pine and spruce account for 37% and 31% of all the trees, respectively. The deciduous accounts for 25%. The descriptive statistics of the plots are summarized in Table I.

The field measurement data were collected in 2008. The plots were located using a Trimble GEOXM 2005 GPS device and postprocessed using local base station data. The locations of the trees were determined using a Suunto bearing compass (Suunto Oy, Vantaa, Finland), and the distance from the plot center was measured using a Haglöf Vertex laser rangefinder (Haglöf Sweden AB, Långsele, Sweden). All trunks with DBH greater than 5 cm are included in the reference measurements.

TABLE I
Statistical Summary of Plot Attributes

Plot	Plot density (stems/ha)	DBH (cm)				Tree height (m)			
		Min	Max	Mean	std	Min	Max	Mean	std
1	509	7.5	29.0	19.4	7.3	8.4	22.0	17.1	4.8
2	541	8.0	31.5	20.7	6.4	11.9	25.0	18.6	4.1
3	668	7.3	29.4	22.1	5.9	6.5	25.8	20.5	4.3
4	1019	6.7	27.5	18.3	5.4	5.9	21.7	16.6	3.4
5	1082	7.5	22.7	16.5	4.2	6.5	19.2	14.6	3.1
6	1273	7.9	26.5	15.9	4.7	8.2	22.2	15.2	3.4
7	1305	7.8	30.5	15.4	5.5	7.7	25.2	17.5	4.2
8	1369	7.0	51.3	16.1	8.1	6.3	26.0	15.3	4.6
9	1432	7.1	21.8	13.4	4.0	7.0	16.1	12.8	2.0

The TLS data were collected in 2008 using a Leica HDS6000 terrestrial laser scanner (Leica Geosystems AG, Heerbrugg, Switzerland). The scanner uses phase-shift measurements of continuous waves to measure the distances, and its data acquisition rate is 500 000 points/s. In each plot, the scanner was positioned at the center of the plot. The point spacing is 6.3 mm at a distance of 10 m. For this study, the point cloud was sampled every 10 points, resulting in a rough point spacing of 20 mm at 10 m.

Fig. 1 shows the box plot of the laser point distribution along a planar distance for nine plots. The X-axis is the planar distance to the plot center measured at 1-m intervals, while the Y-axis is the percentage of total point numbers. In the first interval, the point number is smaller than that in the second one. This is because the ground points close to the scanner are automatically filtered out during the data exportation, which is the default setting in the Leica software. The amount of data in this interval is, however, not expected to be clearly smaller than that in the second one.

B. Method

The proposed approach is a point-cloud processing technique where a specific forest type or 2-D image data structures are not required. A two-step procedure is designed for stem mapping. First, the spatial distribution properties of the laser points are studied. Trunk points are recognized by means of features such as flatness, direction, and shape. The trunks are identified from the point cloud as individual groups. Second, the trunk model is constructed by applying robust estimation. In a dense forest environment, it is difficult and expensive to identify the exact trunk points by means of geometric attributes. By employing other features, e.g., the spectrum and waveform, the problem is very likely the same, as the trunk and branch share many similar attributes. Therefore, a robust modeling method is needed to estimate trunk parameters when branch or foliage points are present. This becomes more important when single-scan TLS data are applied. In that data set, laser points cover, at best and very often less than, 50% of the object surface. Consequently, the influence of outliers becomes significant. In this step, trunk points are further identified from other points, and the stem is located.

1) Trunk Detection: Trunk detection is a process of identifying the laser points in the original point cloud belonging to an individual trunk. In this context, the main objects of concern are the ground, the canopy, and the trunk. The properties of point

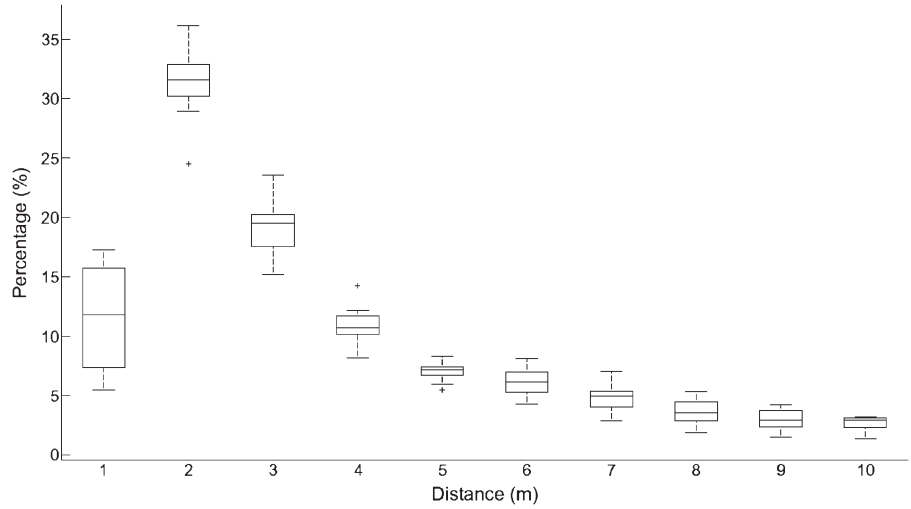

Fig. 1. Distribution of laser points along a planar distance.

distribution, such as flatness and normal vector direction, are used to distinguish the trunk points from the other points.

For a point $p_i = (x_i, y_i, z_i)^{\mathrm{T}}$ in the point cloud P, $P \subset R^3$, its surrounding finite space is defined by N nearest points $p_j = (x_j, y_j, z_j)^{\mathrm{T}}$ with mean $u = (1/N)\sum_{j=1}^{N} p_j$. The covariance matrix C_p of points p_j is defined by

$$C_p = \frac{1}{N}\sum_{j=1}^{N}(p_j - u)(p_j - u)^{\mathrm{T}}. \tag{1}$$

The point set p_j itself does not have context information. The eigenvalue/eigenvector approach is usually employed to study point distribution, for example, to find building surface from ALS data (e.g., [33]). By utilizing eigenvalue decomposition, a new coordinate system can be defined, where axis directions are given by eigenvectors and point variances along axes are given by the corresponding eigenvalues. Let e_i be the eigenvector and λ_i be the corresponding eigenvalue, where $i = 0,\ 1,\ 2$, and $\lambda_0 \leq \lambda_1 \leq \lambda_2$. In the new coordinate system, e_0 gives the direction in which the points exhibit the least variance and e_2 gives the one in which the points exhibit the most variance. Where a surface structure is spanned in the neighborhood, e_0 approximates the normal vector at point p_i. For convenience, e_0 is denoted as the normal vector in the following.

The eigenvalue λ_i quantitatively shows the data variance along the axis e_i, or the compactness of the point distribution along the axis. Therefore, flatness FL can be defined as

$$FL = 1 - \frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2} \tag{2}$$

which shows the significance of the point distribution in two main directions. A high value of FL means that the point is approximately on a planar surface.

The normal vector's Z component is measured by

$$Z_n = |e_0 \cdot n_z| \tag{3}$$

where $n_z = [0, 0, 1]$. For a point on a planar surface, a low value of Z_n means that the point is on a vertical structure.

Fig. 2 shows typical canopy-, ground-, and trunk-point distributions in the forest environment in subfigures (a)/(d), (b)/(e), and (c)/(f), respectively. The points are plotted in the XY (a)–(c) and XYZ spaces (d)–(f), respectively. The point under study is marked by an asterisk, its neighborhood points are dots, the arrow shows the normal vector direction, and the smaller points give a larger neighborhood for illustration. In general, the trunk points exhibit flat and vertical shapes in the local space.

The identified individual points are further studied. The points are grouped such that one point in a group is within a certain distance to at least one point in the same group. The distance is set as 10 cm in the experiment. The trunk groups are recognized by their spatial distribution, which usually present a vertical shape. The trunk groups are further studied and aligned. The group is represented by a vector through the center of gravity, and the vector is projected onto the XY plane. Two groups are merged if the corresponding vectors are close to each other. This connects groups belonging to the same trunk but separated in the data set by an occlusion in the dense forest.

2) Trunk Modeling: A trunk is a tree structure between the root and the branches. Its shape shows great diversity, and it is characterized by its changing direction along the height, its irregular perimeter in the cross section, and the different forms exhibited by the different tree species. In addition, external defects, such as knots and bulges, cause many irregularities. In practice, however, the shape of a section of a trunk can be generalized by geometrical primitives.

The horizontal circle and vertical cylinder are most commonly used to model the tree trunk. The assumption about these primitives is that the axis of the trunk section is strictly vertical. As this is rarely the case, a systematic bias prevails. The circle or free-form curve can be used in the cross section that is orthogonal to the trunk axis. This, however, requires that the axis direction be a known parameter, which can be accurately derived in the modeling procedure with additional computation. Alternatively, a finite right circular cylinder can model the trunk section. The advantage is that the radius and axis direction are estimated at the same time. This balances the requirements of accuracy and the amount of computation needed. Therefore, trunk modeling can be defined by finding a series of cylinders along the trunk to describe the radius and orientation of the stem.

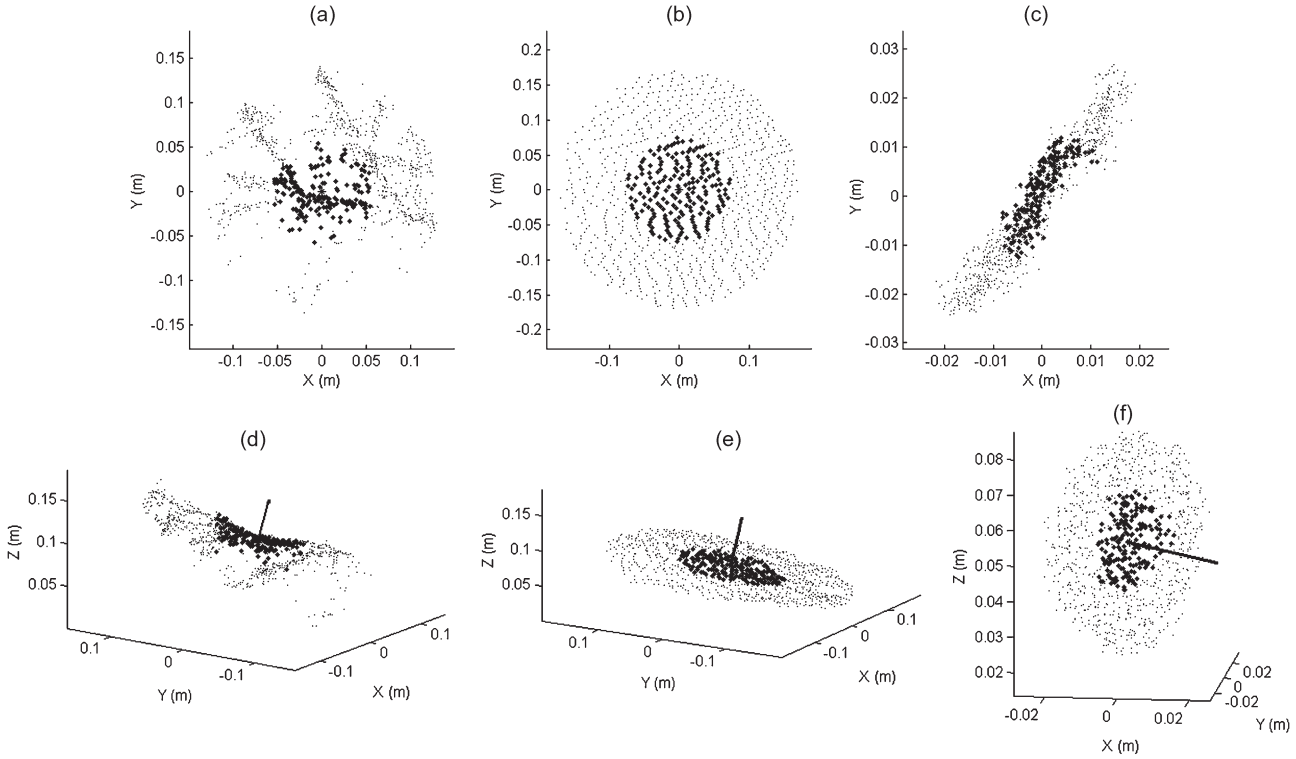

Fig. 2. Typical forest scenes and point distributions.

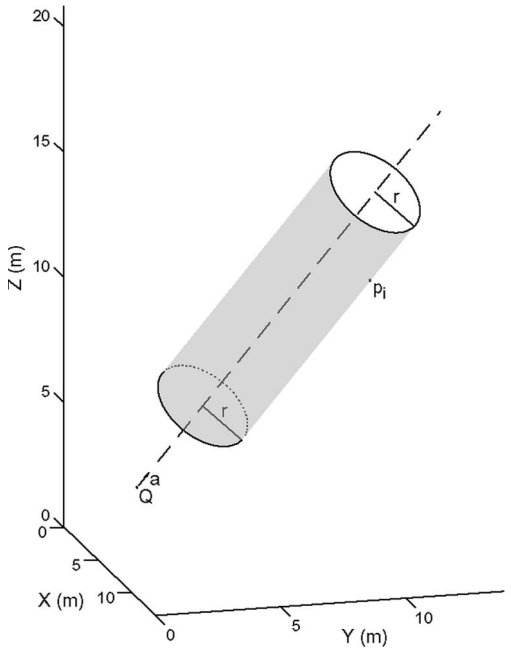

Fig. 3. Part of the infinite right circular cylinder.

a) Robust cylinder fitting: The infinite right circular cylinder in 3-D space is defined by

$$\|(P-Q)\times a\| - r = 0 \tag{4}$$

where $P = \{p_i, i = 1, 2, \ldots, n\}$ is the point set on the cylinder surface, $Q = (x_q, y_q, z_q)^{\mathrm{T}}$ is a point on the axis, $a = (a_x, a_y, a_z)^{\mathrm{T}}$ is the direction of the axis of the cylinder with a unit length, r is the radius, $\|\ldots\|$ is the Euclidian norm of the vector, and $\times$ is the outer product. A part of the infinite right circular cylinder is shown in Fig. 3.

In (4), there are five independent parameters, $C = \{c_j, j = 1, 2, \ldots, 5\}$, where one is for the radius and the rest are for the axis, a line in 3-D space. The details of the parameterization are given in the Appendix. Suppose that points in P are not perfectly located on the surface. The distance, or the residual $V = (v_1, v_2, \ldots, v_n)^{\mathrm{T}}$, between the point and the cylinder surface is given by $V = \|(P-Q)\times a\| - r$.

When all measurements come from the cylinder surface, the standard least squares fitting can be used to determine the parameters in the sense that the square sum of residual, $\sum_{i=1}^{n} v_i^2$, is minimized. However, in the problem addressed in this paper, outliers such as branch points are present in the measurements. Therefore, the measurement errors are supposed to be distributed according to the mixed model, $H = (1-\varepsilon)F + \varepsilon G$, where part of the data ε contain gross errors with an unknown distribution G and the rest of the data have a standard normal distribution F.

Tukey's estimator [34] is employed to reduce the effect of large errors, where the estimator's biweight function suppresses the outliers. Tukey's estimator is given by

$$\rho(Y) = \begin{cases} b^2/6 * \left\{1 - \left[1 - (Y/b)^2\right]^3\right\}, & |Y| \le b \\ b^2/6, & |Y| > b. \end{cases} \tag{5}$$

In (5), b is a factor and picked to be 5, which corresponds to approximately 95% efficiency on the standard normal distribution. $Y = V/s$, where s is an estimation of the unknown standard deviation σ of V. $s = MAD(V)/0.6745$, where MAD is the median absolute deviation defined by $Median\{|V - Median(V)|\}$, and the constant 0.6745 is chosen so that s is asymptotically unbiased for σ if V is with the normal distribution $N(0, \sigma^2)$.

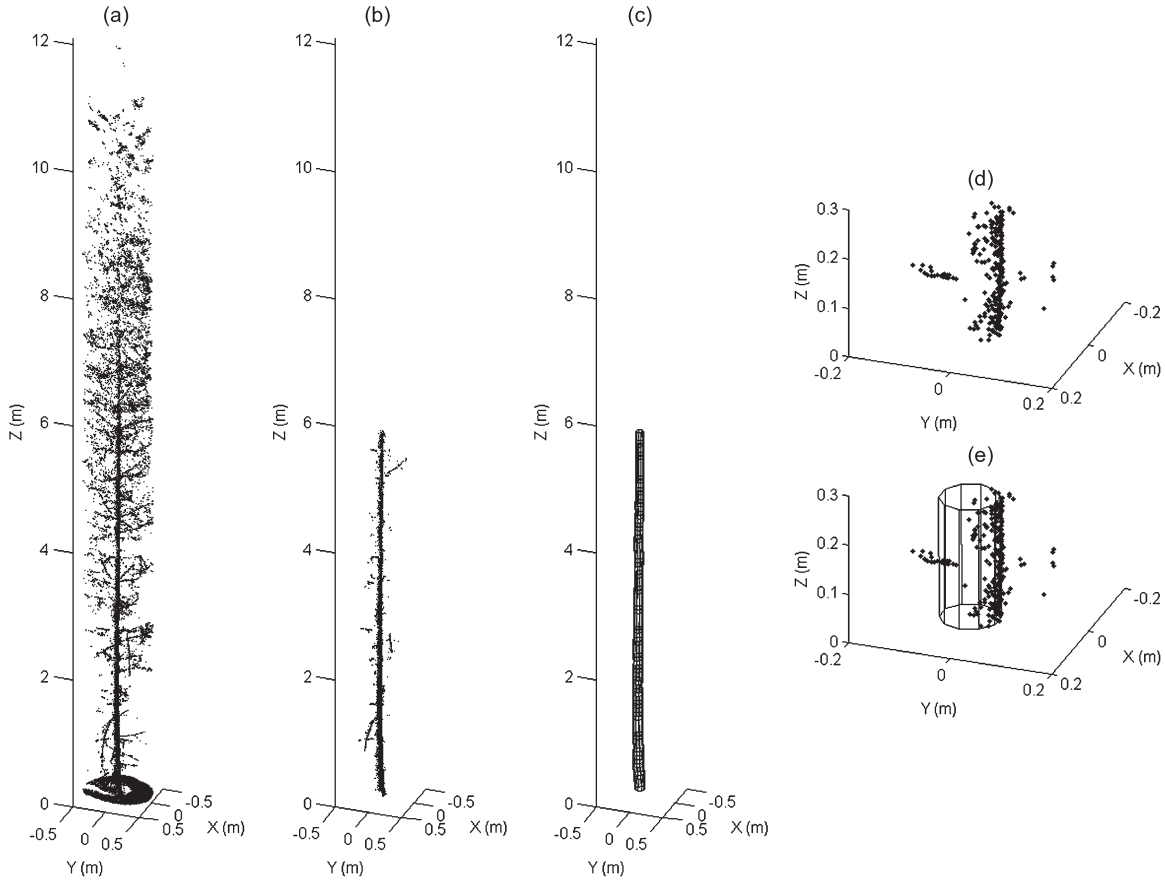

Fig. 4. Stages in the presented trunk modeling procedure.

The estimate of the parameters is the solution to

$$\frac{\partial \rho}{\partial C} = \sum_{i=1}^{n} \frac{\partial \rho(y_i)}{\partial y_i} \frac{\partial y_i}{\partial c_j} = \sum_{i=1}^{n} \psi(y_i) \frac{\partial y_i}{\partial c_j} = \sum_{i=1}^{n} \omega(y_i) y_i \frac{\partial y_i}{\partial c_j} = 0 \tag{6}$$

where $\psi(y_i)$ is the influence function and $\omega(y_i)$ is the weight function. Let $X = \partial y_i / \partial c_j$ be the n-by-5 matrix, W be the diagonal n-by-n weight matrix derived from $\omega(y_i)$, and $Y = (y_1, y_2, \ldots, y_n)^{\mathrm{T}}$. Function (6) is in matrix format, $X^T W Y = 0$.

As the residual equation is not linear, linearization of v_i is required. The solution is obtained by the iteratively reweighted least squares algorithm. The initial adjustment is carried out using W^0 estimated from C^0. In the following iterations, the weight matrix and the parameters are recalculated until convergence. At the $(k+1)$th iteration, the increment Δ is given by

$$\Delta^{(k+1)} = \left(X^T W^{(k)} X \right)^{-1} X^T W^{(k)} Y. \tag{7}$$

The initial approximation values for the axis are estimated from the distribution and the center of gravity of the points P. As the radius enters only linearly into the cylinder equation, no initial approximation is needed.

b) Trunk modeling: In stem mapping, one fitted cylinder at a certain height provides the needed XY position. However, a full trunk model, composed of a series of cylinders along the stem, is constructed to improve the overall identification accuracy. The detected groups from the trunk detection step are mainly from trunks, but some may belong to other objects like the crown. The false groups show vertical distribution similar to those of trunks. However, they do not have cylinder shapes. In general, it is not possible to build the whole stem model for them. By building the full stem model, the false trunk-point groups are detected and deleted. In addition, the full stem model brings additional flexibility by providing a position anywhere along the trunk.

Trunk modeling begins with a point selection procedure. The selection is designed to ensure that the majority of points in the group come from the trunk. If the main distribution direction of the selected points is away from the Z-axis, the region will be extended to find more points. This procedure is important where branch points exist. A cylinder is fitted to the point group, and the uppermost and lowest points in the axis direction are selected to give the position of the cylinder. The modeling process is performed in such a way that the next group of points, along the axis and with a certain overlap with the current one, is selected for cylinder fitting, and the parameters of the current cylinder are used as the initial estimation. The procedure stops when there are not enough points to be found.

As the field reference position is measured at the base of the trunk, the lowest cylinder is selected, and the lower endpoint of the axis is used as the trunk location estimation.

Fig. 4 shows the stages in trunk modeling. The original point cloud, the detected trunk points, and the trunk model are shown in Fig. 4(a)–(c), respectively. The points around the second upmost cylinder and its model are shown in Fig. 4(d) and (e).

TABLE II
SUMMARY OF MAPS PRODUCED BY FIELD AND TLS MEASUREMENTS

Plot	Reference trees	Mapped trees	Type I error	Type II error
1	16	11	5	2
2	17	14	3	9
3	21	17	4	2
4	32	26	6	3
5	34	24	10	0
6	40	21	19	4
7	41	32	9	7
8	43	25	18	11
9	45	40	5	4
Total	289	210	79	42

IV. RESULTS AND DISCUSSION

The location map calculated using the methods depicted in Section III was manually matched with the reference data. The mapping was evaluated based on all reference trees in the plots, where the smallest measured DBH was 5 cm. Table II summarizes the mapping results. Type I error, or omission, is the number of stems not modeled. Type II error, or commission, is the number of stems falsely modeled. The overall accuracy was 73%.

Fig. 5 shows the mapping accuracy as a function of range. The X-axis is the distance between the scanner and the reference stem when applying an interval of 1 m. The Y-axis is the accumulated mapping accuracy.

The algorithm presented in this paper is a point-processing technique developed for plotwise stem-location mapping. Among the three main method types (summarized in Section II), this technique is the most general solution for stem mapping. It does not require *a priori* plot knowledge or data format. So far, with regard to its application to TLS data over large area, the point-processing technique is mainly used for robot-based perception. In the perception problem, the objects in the surrounding area need to be classified based on their obstacle properties, by the estimated distribution model. In stem mapping, however, the context of the application is clearly different. The tree trunk is the main object concerned, and it can be defined by certain features.

In the proposed algorithm, a two-step procedure is designed to identify the stem location within the plot. The two steps are independent, but they are closely integrated. As shown in the test, it is an efficient procedure. Moreover, focusing on the further method development, it is a valuable framework. On the one hand, as robust estimation is not unduly affected by gross errors, it is possible to identify trunk points in the first step by some simplified method, which may commit more errors but require less computation. For example, the point property may be estimated at a group level. Compared with the point-level study, this can significantly reduce the amount of computation in the point identification phase. Further research works need to study this possibility. On the other hand, the two-step procedure is an extendable framework. New techniques can be easily integrated into this procedure. For instance, in the future, when waveform or multispectrum TLS is more commonly used, the first step, i.e., trunk-point recognition, can be implemented by means of waveform or spectral analysis. After that, the estimation method presented in this paper can be used to further locate trunk points. In the meantime, it is worth noting that the spatial data, or point coordinates, are the only data employed in this method. It shows that the point coordinates provide adequate data to locate stems in the plots. The employment of additional data, e.g., waveform or spectrum, is not compulsory in this application. They should be used when the overall performance can be clearly improved. It also shows the efficiency of the proposed method, where only the most essential data provided by TLS are needed.

Fig. 5 shows that single-scan TLS data can be employed in stem locating, and the applied algorithm works with an accuracy of 73% when the average stem number per hectare is 1022 stems/ha. When applying a range of 5 m, the accuracy is 85%. Comparing the stem-locating accuracy of this study to the result of previous TLS studies, it can be stated that the presented method is an efficient algorithm that achieves high mapping accuracy in dense forest stands.

For plots characterized by high stem density, the multiscan TLS approach or the combination of several single-scan TLSs within the same plot should be adopted. Interactive editing is probably the only practical solution in the near future to find the missing stems that are still visible from the point-cloud data. In this procedure, the overall accuracy and the efficiency of manual work can be clearly enhanced. The hybrid mechanism will be of high practical value for the forest inventory system.

The accuracy of stem mapping using single-scan TLS data is attributed to several factors. They are the shadow effect in single-scan TLS data, the complexity of the forest scene, and the characteristic of the laser point distribution.

The shadow area is the place where the laser beams are blocked by a foreground object. It appears in both single and multiscans but poses more problems for applications using single-scan data. The effects of shadows introduced by impermeable or permeable structures, such as the stem and canopy here, are different. Stems block all laser beams in that direction. The stem standing in the shadow is completely missed in the data set. The canopy structure partially blocks the laser beams. The stem behind appears as several disconnected parts in the data set. This effect makes the objects in the shadow either impossible to identify or difficult to extract.

The plots employed in this paper were originally used for field reference collection when applying traditional forest survey techniques. The need for single-scan TLS measurement was not considered in this case. For example, the plot center, also being the scanner position, was not selected with the purpose of reducing the shadow effect. In some cases, there was a tree that was very close to the plot center, and this placed a large area behind it in its shadow.

The numbers of reference stems within each plot, manually measured in the field and single-scan TLS data, are shown in Table III. The reference stems in TLS data are those that are at least partly available in the point cloud. The percentage of identifiable stems varies among the plots because of the geometry between the stems and the scanner. Approximately 10% of the reference stems are not available in single-scan TLS data.

The second aspect is the complexity of the forest scene. In reality, the spatial pattern of nearby trees can be complicated in a dense forest stand. When two or more trees are close to each

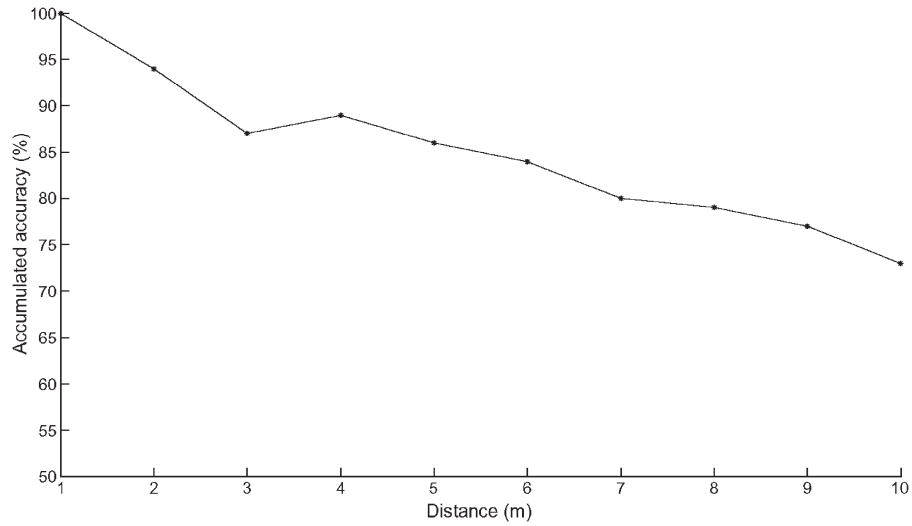

Fig. 5. Stem-locating accuracy as a function of range.

TABLE III
TREES MEASURED IN THE FIELD AND TLS DATA

Plot	1	2	3	4	5	6	7	8	9
Trees in field	16	17	21	32	34	40	41	43	45
Trees in TLS	15	16	18	29	31	34	40	35	45
Trees from TLS data(%)	94	94	86	91	91	85	98	81	100

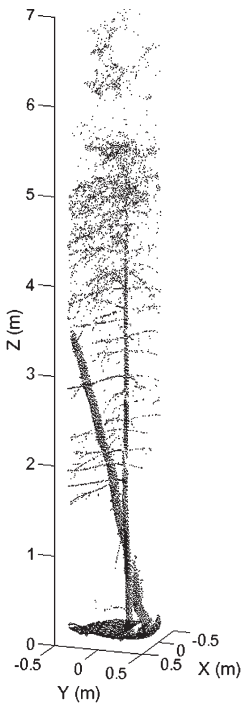

Fig. 6. Two trees close to each other.

other, the method designed for single-stem modeling tends to give false results. Fig. 6 shows an example where two trees are standing next to each other. As the two trunks are recognized as one group, the model construction is not successful. Roughly 5% of the missing reference stems is attributed to this factor.

The third issue is the characteristic of the laser point distribution in single-scan TLS data. The point density decreases with increasing planar distance. This means that the farther away an object is, the smaller the amount of laser beams is reflected by the object. Fig. 1 shows that, in a dense forest environment, the proportion of points within 5 m decreases rapidly as a function of range; between 6 and 10 m, the proportions are rather small and show little variation. For stem mapping, the accuracy also shows the decrease as a function of distance to the scanner position. This correspondence indicates that a high point density is a contributing factor in the trunk detection.

Small trees and bushes in the lower canopy layer, between 0.5 and 3 m in height, also affect the trunk-locating rate. On the one hand, laser points are blocked by them, and fewer points are reflected from trunks standing farther away. On the other hand, when lower level vegetation is close to tree trunks, the vertical planar distribution property of trunk points becomes less significant. As a result, it becomes harder to distinguish the trunk from other objects. Tree species and growth stages, such as young spruces with an abundance of lower branches, affect the results in the same way as lower level vegetation. In addition, the uncertainty of small trees introduces some commission error. A small tree with DBH close to 5 cm may or may not be recorded in the field.

In this paper, the stem density did not significantly impact on the mapping accuracy. However, its influence is an indirect one. When the plot density is high, there are greater possibilities for some trees to have complicated geometric features, standing partly in the shadow or reflecting limited points due to the occlusion effect. Consequently, it is more difficult to detect and model such trunks than to work in a low-density plot.

To improve the performance of the algorithm, some issues need further discussion. For example, the rough estimation for the approximation values for the cylinder axis was employed in trunk modeling. The method is capable of giving good parameters when the points are evenly distributed on the trunk. In single-scan TLS data, however, the estimation is biased. A better routine can be expected to contribute to the modeling procedure. Another topic is to improve the trunk-point detection method. The applicability of feature estimation for a group of points needs to be studied.

So far, TLS is supposed to be used mainly as a technique to collect references for large-area forest inventories. In this circumstance, certain features, such as tree position and basal area, are the ones needed to be collected from the point cloud. In addition, it is possible to collect more features automatically using TLS measurements. In practice, some of them are seldom recorded, like stem curve, before the harvesting stage, though widely recognized as useful data sources. It is possible to improve the understanding of the stand properties by additional features. The application of those features in forest inventories needs further studies.

V. Conclusion

This paper has presented a fully automatic algorithm of single-scan TLS data for stem mapping in a typical boreal forest environment. The algorithm is a general solution for stem mapping in the sense that it is independent of *a priori* plot knowledge and data format. The method was evaluated with regard to stem mapping using nine field plots each with a radius of 10 m. The average stem count was 1022 stems/ha. The overall detection accuracy was 73%. This shows that, in a dense managed forest stand, the majority of trunks can be located using single-scan TLS data, but the combination of several single scans, the application of multiscan, and the employment of interactive editing from point-cloud data need to be further studied to provide a practical system feasible in various forest environments.

Appendix

A parameterization of cylinder is given in [35]. Here, the radius r is employed instead of the parameter $1/k$. The infinite right circular cylinder in 3-D space is

$$\|(P-Q)\times a\| - r = 0$$

where $P = \{p_i, i = 1, 2, \ldots, n\}$ is the point set on the cylinder surface, a is the direction of the axis of the cylinder with a unit length, and r is the radius. Suppose that the point Q is the orthogonal projection of the origin O on the cylinder axis, $n = (n_x, n_y, n_z)^{\mathrm{T}}$ is the unit vector in the direction OQ, and ρ is the distance from the origin to the closest point of the cylinder. Therefore

$$Q = \langle \rho + r, n \rangle$$

where $\langle \ldots \rangle$ is the inner product. Parameterize n in its polar coordinates, $n = (\cos\varphi \sin\theta, \sin\varphi \sin\theta, \cos\theta)$, where θ is the angle between n and the Z-axis, and φ is the angle between the projection of n onto the XY plane and the X-axis. Considering $\langle n, a \rangle = 0$, a can be parameterized in the form of

$$a = n^{\theta} \cos\alpha + n^{\varphi} \sin\alpha / \sin\theta$$

where n^{θ} and n^{φ} are partial derivative vectors with respect to θ and φ, respectively, and α is the angle between a and n^{θ}. The cylinder is parameterized by $C = (\rho, \theta, \varphi, \alpha, r)$.

Acknowledgment

The authors would like to thank the anonymous reviewers for their helpful comments. Author X. Liang would also like to thank Dr. E. Honkavaara and Dr. J. Liu at the Finnish Geodetic Institute for the helpful discussion.

References

[1] E. Næsset, "Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data," *Remote Sens. Environ.*, vol. 80, no. 1, pp. 88–99, Apr. 2002.
[2] J. Hyyppä and M. Inkinen, "Detecting and estimating attributes for single trees using laser scanner," *Photogramm. J. Finland*, vol. 16, no. 2, pp. 27–42, 1999.
[3] J. Hyyppä, H. Hyyppä, D. Leckie, F. Gougeon, X. Yu, and M. Maltamo, "Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests," *Int. J. Remote Sens.*, vol. 29, no. 5, pp. 1339–1366, Mar. 2008.
[4] X. Yu, J. Hyyppä, H. Holopainen, and M. Vastaranta, "Comparison of area-based and individual tree-based methods for predicting plot-level forest attributes," *Remote Sens.*, vol. 2, no. 6, pp. 1481–1495, Jun. 2010.
[5] K. Zhao and S. Popescu, "Lidar-based mapping of leaf area index and its use for validating GLOBCARBON satellite LAI product in a temperate forest of the southern USA," *Remote Sens. Environ.*, vol. 113, no. 8, pp. 1628–1645, Aug. 2009.
[6] M. Leeuwen and M. Nieuwenhuis, "Retrieval of forest structural parameters using LiDAR remote sensing," *Eur. J. Forest Res.*, vol. 129, no. 4, pp. 749–770, Jul. 2010.
[7] C. Hopkinson, L. Chasmer, C. Young-Pow, and P. Treitz, "Assessing forest metrics with a ground-based scanning lidar," *Can. J. Forest Res.*, vol. 34, no. 3, pp. 573–583, Mar. 2004.
[8] J. G. Henning and P. J. Radtke, "Detailed stem measurements of standing trees from ground-based scanning lidar," *Forest Sci.*, vol. 52, no. 1, pp. 67–80, Jan. 2006.
[9] F. Hosoi and K. Omasa, "Voxel-based 3-D modeling of individual trees for estimating leaf area density using high-resolution portable scanning lidar," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 12, pp. 3610–3618, Dec. 2006.
[10] I. Moorthy, J. R. Miller, B. Hu, J. Chen, and Q. Li, "Retrieving crown leaf area index from an individual tree using ground-based lidar data," *Can. J. Remote Sens.*, vol. 34, no. 3, pp. 320–332, Jun. 2008.
[11] N. Pfeifer and D. Winterhalder, "Modelling of tree cross sections from terrestrial laser-scanning data with free-form curves," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 36(8/W2), pp. 76–81, 2004. [Online]. Available: http://www.isprs.org/publications/archives.aspx.
[12] S. Fleck, N. Obertreiber, I. Schmidt, M. Brauns, H. F. Jungkunst, and C. Leuschner, "Terrestrial lidar measurements for analysing canopy structure in an old-growth forest," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. 36(3/W52), pp. 125–129, 2007.
[13] M. Thies, N. Pfeifer, D. Winterhalder, and B. G. H. Gorte, "Three-dimensional reconstruction of stems for assessment of taper, sweep and lean based on laser scanning of standing trees," *Scand. J. Forest Res.*, vol. 6, no. 19, pp. 571–581, 2004.
[14] M. Lefsky and M. R. McHale, "Volume estimates of trees with complex architecture from terrestrial laser scanning," *J. Appl. Remote Sens.*, vol. 2, no. 1, p. 023 521, May 2008.
[15] H. Xu, N. Gosset, and B. Chen, "Knowledge and heuristic-based modeling of laser scanned trees," *ACM Trans. Graph.*, vol. 26, no. 4, p. 19, Oct. 2007.
[16] Z. Cheng, X. Zhang, and B. Chen, "Simple reconstruction of tree branches from a single range image," *J. Comput. Sci. Technol.*, vol. 22, no. 6, pp. 846–858, Nov. 2007.
[17] X. Yu, J. Hyyppä, A. Kukko, M. Maltamo, and H. Kaartinen, "Change detection techniques for canopy height growth measurements using airborne laser scanning data," *Photogramm. Eng. Remote Sens.*, vol. 72, no. 12, pp. 1339–1348, Dec. 2006.
[18] M. Thies and H. Spiecker, "Evaluation and future prospects of terrestrial laser scanning for standardized forest inventories," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. 36, pp. 192–197, 2004.
[19] H.-G. Maas, A. Bienert, S. Scheller, and E. Keane, "Automatic forest inventory parameter determination from terrestrial laser scanner data," *Int. J. Remote Sens.*, vol. 29, no. 5, pp. 1579–1593, Mar. 2008.
[20] J. Kalliovirta, J. Laasasenaho, and A. Kangas, "Evaluation of the laser-relascope," *Forest Ecology Manage.*, vol. 204, no. 2/3, pp. 181–194, Jan. 2005.

[21] I. Korpela, T. Tuomola, and E. Välimäki, "Mapping forest plots: An efficient method combining photogrammetry and field triangulation," *Silva Fennica*, vol. 41, no. 3, pp. 457–469, 2007.
[22] S. E. Reutebuch, W. W. Carson, and K. M. Ahmed, "A test of the Applanix POS LS inertial positioning system for the collection of terrestrial coordinates under a heavy forest canopy," in *Proc. 2nd Int. Precision Forestry Symp.*, Seattle, WA, 2003, pp. 1–11.
[23] T. Aschoff, M. Thies, and H. Spiecker, "Describing forest stands using terrestrial laser-scanning," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. 35, pp. 237–241, 2004.
[24] G. Király and G. Brolly, "Tree height estimation methods for terrestrial laser scanning in a forest reserve," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. 36(3/W52), pp. 211–215, 2007.
[25] K. Tansey, N. Selmes, A. Anstee, N. J. Tate, and A. Denniss, "Estimating tree and stand variables in a Corsican Pine woodland from terrestrial laser scanner data," *Int. J. Remote Sens.*, vol. 30, no. 19, pp. 5195–5209, Oct. 2009.
[26] A. H. Strahler, D. L. B. Jupp, C. E. Woodcock, C. B. Schaaf, T. Yao, F. Zhao, X. Yang, J. Lovell, D. Culvenor, G. Newnham, W. Ni-Miester, and W. Boykin-Morris, "Retrieval of forest structural parameters using a ground-based lidar instrument (Echidna®)," *Can. J. Remote Sens.*, vol. 34, no. S2, pp. 426–440, 2008.
[27] N. Haala, R. Reulke, M. Thies, and T. Aschoff, "Combination of terrestrial laser scanning with high resolution panoramic images for investigations in forest applications and tree species," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 34(5/W16), 2004, 4p.
[28] P. Forsman and A. Halme, "3-D mapping of natural environments with trees by means of mobile perception," *IEEE Trans. Robot*, vol. 21, no. 3, pp. 482–490, Jun. 2005.
[29] X. Liang, P. Litkey, J. Hyyppä, A. Kukko, H. Kaartinen, and M. Holopainen, "Plot-level trunk detection and reconstruction using one-scan-mode terrestrial laser scanning data," in *Proc. Int. Workshop Earth Observation Remote Sens. Appl.*, Beijing, China, 5p. 2008.
[30] P. Litkey, X. Liang, J. Hyyppä, A. Kukko, H. Kaartinen, and M. Holopainen, "Single-scan TLS methods for forest parameter retrieval," in *Proc. SilviLaser*, Edinburgh, U.K., pp. 295–304, 2008.
[31] J. F. Lalonde, N. Vandapel, D. F. Huber, and M. Hebert, "Natural terrain classification using three-dimensional ladar data for ground robot mobility," *J. Field Robot.*, vol. 23, no. 10, pp. 839–861, Oct. 2006.
[32] P. J. Watt and D. N. M. Donoghue, "Measuring forest structure with terrestrial laser scanning," *Int. J. Remote Sens.*, vol. 26, no. 7, pp. 1437–1446, Apr. 2005.
[33] V. Verma, R. Kumar, and S. Hsu, "3D building detection and modeling from aerial LiDAR data," in *Proc. IEEE Comput. Soc. Conf. CVPR*, 2006, vol. 2, pp. 2213–2220.
[34] Z. Zhang, "Parameter estimation techniques: A tutorial with application to conic fitting," *Image Vis. Comput.*, vol. 15, no. 1, pp. 59–76, Jan. 1997.
[35] G. Lukács, R. Martin, D. Marshall, and R. R. Martin, "Faithful least-squares fitting of spheres, cylinders, cones and tori for reliable segmentation," in *Proc. 5th Eur. Conf. Comput. Vision*, Feiburg, Germany, 1998, pp. 671–686.

Xinlian Liang was born in Beijing, China, in 1980. He received the B.S. degree in geomatics engineering from Wuhan University, Wuhan, China, in 2002 and the M.S. degree in photogrammetry and remote sensing from the Chinese Academy of Surveying and Mapping, Beijing, China, in 2005.

In 2006, he joined the Finnish Geodetic Institute, Masala, Finland, where he is currently a Senior Research Scientist with the Department of Remote Sensing and Photogrammetry. His research interests include the development of algorithms of laser scanning data and their applications in environmental studies, particularly forest inventories.

Paula Litkey received the M.Sc.Tech. and Lic.Tech. degrees in electrical engineering from the Helsinki University of Technology, Espoo, Finland, in 1998 and 2004, respectively.

She is currently a Specialist Research Scientist with the Department of Remote Sensing and Photogrammetry, Finnish Geodetic Institute, Masala, Finland. Her research interests are in computational methods and algorithms for laser scanner data and their applications for full-waveform data.

Juha Hyyppä received the M.Sc., Dr.Ing., and Dr.Sc. degrees from the Helsinki University of Technology, Espoo, Finland, in 1987, 1990, and 1994, respectively.

He is currently an Adjunct Professor and a Head of Department with the Finnish Geodetic Institute, Masala, Finland. He was/has been the President of European Spatial Data Research Network Com II (information extraction) from 2004 to 2010, the Vice-President of International Society for Photogrammetry and Remote Sensing (ISPRS) Com VII from 2008 to 2012, the Cochair of ISPRS WG III/3 from 2004 to 2008, the Principal Investigator in European Space Agency (ESA)/National Aeronautics and Space Administration Announcement of Opportunity studies, and a Finnish delegate at the ESA Programme Board on Earth Observation. He has coordinated ten international projects. His references have been represented by over 100 refereed journal papers. His research interests include all kinds of new laser scanning algorithms and applications.

Harri Kaartinen received the M.Sc.(Tech.) degree in geodesy from the Helsinki University of Technology, Espoo, Finland, in 1993 and the Lic.Sc.(Tech.) degree in remote sensing from the Aalto University School of Science and Technology, Espoo, in 2010.

He is currently a Specialist Research Scientist with the Department of Remote Sensing and Photogrammetry, Finnish Geodetic Institute, Masala, Finland. His research interests include terrestrial and mobile laser scanning and reference measurements.

Mikko Vastaranta received the M.Sc. degree from the Department of Forest Resource Management, University of Helsinki (UH), Helsinki, Finland, in 2007, where he is currently working toward the Ph.D. degree.

Since then, he has been working with UH in the projects "Improving the Forest Supply Chain by Means of Advanced Laser Measurements" and "Science and Technology Towards Precision Forestry." He has authored or coauthored some 40 publications, more than 20 of which have been published in peer-reviewed scientific journals or books. His current research interests include forest inventory using airborne laser scanning, terrestrial laser scanning, mobile laser scanning, and synthetic aperture radar data.

Markus Holopainen received the M.Sc. and Ph.D. degrees in forestry from the University of Helsinki (UH), Helsinki, Finland, in 1993 and 1998, respectively, and the M.Sc. and Dr.Tech. degree from the Department of Surveying, Helsinki University of Technology (HUT, Aalto University), Espoo, Finland, in 1995 and 2011, respectively.

He was with the Graduate School of Forest Ecology, UH, in 1995–1998. In 1998–2001, he coordinated the graduate school's "Forests in GIS." He has been with UH since 2001, where he was/has been a University Lecturer (2002–2005 and 2009–present) and an Acting Professor of Forest Mensuration and Management (2001–2002 and 2006–2007), and a Professor of Geoinformatics (2007–2008). He was also an Adjunct Professor in Forest Mensuration and Management as well as Remote Sensing through docentships at UH in 2007 and HUT in 2009. He has authored or coauthored some 120 publications, more 60 of which have been published in peer-reviewed scientific journals or books. His current research interests include forest inventory using airborne laser scanning, terrestrial laser scanning, and synthetic aperture radar data; forest valuation; and uncertainty in forest management planning and valuation.