AUTOMATIC RECONSTRUCTION OF SINGLE TREES FROM TERRESTRIAL LASER SCANNER DATA

Norbert Pfeifer¹, Ben Gorte¹ and Daniel Winterhalder²

¹: Section of Photogrammetry and Remote Sensing, TU Delft, Kluyverweg 1, 2629HS, The Netherlands
²: Institute for Forest Growth, Univ. Freiburg, Germany
{n.pfeifer,b.g.h.gorte}@lr.tudelft.nl, daniel.winterhalder@iww.uni-freiburg.de

KEY WORDS: Laser scanning, Close Range, Modelling, Forestry, Automation, Measurement, Algorithms

ABSTRACT:

The investigation of single trees in a forest is of ecological and economical interest. One aim is to capture the geometric aspects of a tree: the length and diameter of the trunk and individual branches, the change of the radius along the branch and similar measures. These measures can be determined automatically from terrestrial laser scanner data. The conditions for scanning in the forest, but also the irregular structure and surface of trees aggravate the reconstruction process. The branches of the trees are locally modelled by circular cylinders. With the radius, the axis direction and the axis position the main parameters of interest are captured. We describe a set of algorithms for automatically fitting and tracking cylinders along branches and reconstructing the entire tree. Especially for coniferous trees the computation of an outer hull, giving the extent in different directions and at different heights is an alternative, as the dense foliage coverage renders a distinction between branches and needles impossible. Examples for the different reconstructions of trees are presented.

1 INTRODUCTION

The world forest area covers roughly one quarter to the total land area of the world. Considered this, it is obvious that forests play an important role in our lives for ecological and economical reasons. Assessing various forest parameters is performed with photogrammetric techniques (optical and radar satellite remote sensing, aerial photography, and airborne laser scanning), but also with (terrestrial) field surveys. In this paper a new measurement method is added to the existing ones, offering the possibility for objectively determining parameters of single trees.

The tree parameters considered in this study are of geometric type, including the diameter of the trunk and the branches, but also the angles between different branches and their location, as well as the crown diameter are to be determined. These parameters are of interest due to ecological reasons (habitat investigations, studying growth reactions to wind and other environmental influences, etc.) and economical reasons (timber volume estimation for wood production, detection and quantization of failures during the growth process, etc.) and can be described in terms of an ‘as-grown’ analysis. Current measurement methods are either based on human estimation and experience (e.g. for crown diameters) or performed with very simple tools (e.g. tape measurements). Generally, it can be said that there is a lack of automation in the current methodology, making it expensive and subjective (i.e. dependent on the operator).

Airborne laser scanning with its ability to penetrate the tree crown cover is investigated and applied in forestry (e.g. (Pyysalo and Hyyppä, 2002)) for measuring tree height and crown diameter, or forest height respectively, depending on the data density. Satellite laser scanning, combined with full waveform capturing (e.g. NASA’s ICESat mission) offers the possibility to measure the biomass in forests. Naturally, these methods provide not much information on a single tree, but their strength lies in providing overview information on logs or complete forests. In aerial imaging only the upper crown surface can be seen, but automation (i.e., detecting and analyzing single trees) is low. With aerial imaging only the parameters visible from above (e.g. crown diameter, health state) can be determined. Ground based imaging methods, on the other hand, are not suitable due to the irregular structure of the tree surface (considering image matching), which would require further manual processing, and due to the often poor lighting conditions in the forests. With the advent of terrestrial laser scanning an active measurement technology – independent of the sun or an additional artificial light source –, capable of providing millions of points on highly irregular surfaces is now available for measuring inside forests.

This paper presents a method for automatic reconstruction of branches, and therefore trees, from terrestrial laser scanning data. Special consideration has to be given to their irregular structure and the problems of data recording in forests. In Section 2.1 the requirements for the reconstruction are stated and in Section 2.2 the reconstruction methodology is presented. A collection of algorithms for the reconstruction is presented in Section 3. In Section 4 examples are presented and discussed.

2 SINGLE TREE MODELLING

2.1 Modelling Requirements and Scanning Environment

The tree model we want to reconstruct has to provide information on i) the start point and end point of each branch, and ii) the radii at these points. This captures in a straightforward manner the essential measures of the tree components (i.e. the branches), giving sufficient input to the above described tasks in forestry.

As pointed out in the Introduction, an automated reconstruction of the tree model is only suitable with terrestrial laser scanning data. Laser scanner data can be acquired in short time (a few minutes per scan) and provides a dense point cover on the surface with an accuracy of a few centimeters or even better. In Fig. 1 a part of one scan (thinned out) is shown. However, point density decreases with distance from the device, providing less and less points in the higher parts of the tree. Likewise, shadow effects from lower branches will generate gaps in the coverage of higher branches (cf. Fig. 1 and Fig. 2). Additionally, the chances of hitting a branch with the laser beam drop with its radius. Therefore, the outer branches and the higher branches will be covered.
The point cloud has been thinned out by averaging points in a 10x10x10cm³ raster. The scanner was positioned on the right side of the image in the background (not shown), yielding a view on the inner side of the trunks.

Figure 1: Point cloud of one scan. To improve visual appearance, the point cloud has been thinned out by averaging points in a 10x10x10cm³ raster. The scanner was positioned on the right side of the image in the background (not shown), yielding a view on the inner side of the trunks.

less dense with “laser points”, making reconstruction in the above described manner difficult or even impossible. The evergreen foliage of coniferous trees can be very dense and presents another difficulty. A view to the tree top becomes almost impossible. Another difficulty in scanning trees is caused by the wind. It makes branches move, again depending on their size, but small branches can move a few centimeters (e.g. 5cm) given only a light breeze.

Finally, one last problem is the registration (relative orientation) of scans from different positions to each other. The ICP algorithms (Besl and McKay, 1992) or variants for orienting two scans to each other require that the same object is covered in both scans. Due to shadow effects this overlapping coverage is naturally given only at the trunk. This allows for horizontal alignment in the trunk region but rotation around the vertical axis and positioning along it remain problematic. As there are no natural tie-points (no natural sharp corners), artificial targets have to be used as tie-points. In planimetry it is possible to spread the tie-points over the complete scene, but due to practical limitations it is very difficult to place tie-points higher than e.g. 3m. This leads to an extrapolation of orientation information in the zenith direction.

Summing up, the environmental conditions under which the measurements are performed, cannot be controlled well, rendering the reconstruction of a tree with all its branches from laser scanning data a difficult, if not impossible task. Thus, an alternative reconstruction method will be presented as well, not giving information on individual branches, but aiming at reconstructing the outer hull of a tree.

2.2 Reconstruction methodology

The principle idea exploited in the reconstruction is that branches are described as right circular cylinders. This model deviates from reality in two aspects. First, the cross-section is in most cases not circular, but of a more general form. Branches growing sideways tend to have a more elliptical cross-section due to gravity, while other influences, e.g. wind, make the cross-sections even more complicated. Second, the axis of a branch is not a straight line, but curved. This already shows, that modelling of trees is a challenging task. The choice taken here was to – first – ignore the deviation of the cross-section from the circular form, and – second – to fit cylinders piecewise to branches, so that the cylinder axis can adopt locally to a curved branch. The first decision was also influenced by the low point coverage of branches, typically only 50% or less of a cross-section is covered with points, and by the expectation that a fewer number of parameters would be easier to estimate from the data. The second decision of piecewise approximation allows to fit multiple cylinders to one branch, each in a restricted area. Approximation values from one branch to the next can easily be generated, if these areas are made to overlap.

As mentioned above, especially for coniferous trees the reconstruction of an outer hull is needed. Choosing a convex hull is not a good choice, because it contradicts to the behavior of trees. Environmental forces, e.g. competition with neighboring trees for sunlight, lead to different horizontal extents of the foliage in different levels. The methodology applied here is that the outer hull is reconstructed slice-wise for horizontal levels. In these different height levels the recorded points on the foliage determine the circumference of the hull.

3 ALGORITHMS FOR TREE RECONSTRUCTION

In the following a collection of algorithms is presented which are used for the tree reconstruction.

3.1 Normal estimation

For each given point its surface normal vector provides additional information that can be used in the subsequent steps. Assuming that the point coverage on the measured surfaces is dense enough, the normal vector can be estimated from the point and its nearest neighbors. We used the k nearest neighbors (ANN, 2003) to each point and fit a plane to these points by minimizing the sum of squares of orthogonal distances between the points and the plane. The normal n is the eigenvector to the smallest eigenvalue λ, of:

\[ A^T A n = \lambda n, \quad A = \begin{pmatrix} \pi_1 & \pi_2 & \pi_3 \\ \vdots & \vdots & \vdots \\ \pi_k & \pi_k & \pi_k \end{pmatrix} \] (1)

The values \( \pi_1, \pi_2, \pi_3 \) are the coordinates reduced to the barycenter of the k points. The r.m.s.e. of this adjustment is \( \sqrt{\lambda_k/(k-3)} \) (for a proof see e.g. (Shakarji, 1998)).

The value k was set to 40, and given a minimal point spacing (in one scan) of 3mm, this corresponds to an area of 2cm diameter. Naturally, point coverage gets lower in higher parts of the tree, corresponding to bigger areas used for normal estimation. An average fitting accuracy below 2cm was achieved for points on the tree surface. This value is larger than the measurement accuracy of the laser scanner used, because of the roughness of the trees.

3.2 Removal of ‘wrong’ points

During the normal vector computation for the given points filtering of points can be performed simultaneously. Single measurement errors typically lead to points ‘hanging’ in mid-air without neighbors in close proximity. This can be caused, for example, by measuring ranges to surfaces outside of the uniqueness range.
of phase-difference laser rangers (Fröhlich et al., 2000). These points, but also points on leaves, which typically have no, or only a few neighbors on the same leaf, can be eliminated by removing all those points where the plane fitting accuracy is below a certain threshold. Choosing a threshold of 2 cm reduces the number of points by a few percent only, removing points that are not suited for tree modelling anyway.

3.3 Cylinder fitting

Cylinders have 5 parameters: the radius is one parameter and the axis (a line in 3-dimensional space) has 4 parameters. It is, however, easier to use more parameters for the cylinder axis and introduce some conditions between those. The following parameterization of a cylinder is adopted:

$$\| (P - Q_i) \times a \| - r = 0$$  \( (2) \)

The symbol \( \times \) denotes the outer product between two vectors and the symbol \( \| \ldots \| \) denotes the euclidian norm of the vector. In the above equation \( P = (x_P, y_P, z_P) \) is a point on the axis, \( a = (x_a, y_a, z_a) \) is a vector of unit length that points into the direction of the axis, and \( r \) is the radius. All points \( Q_i = (x_i, y_i, z_i) \) fulfilling this equation lie on the cylinder. Demanding that \( a \) has length 1 and that \( P \) lies in a fixed, predefined plane reduces the parameters from 7 to 5 again.

In order to fit a cylinder to a given number of points \( Q_i, i = 1, \ldots, n \), the unknowns (i.e. \( P, a, r \)) are determined by least squares adjustment. The residuals \( v_i \) are the the distances from the measured points to the ideal cylinder:

$$v_i = \| (P - Q_i) \times a \| - r$$

As the problem is not linear the observation equations of the least squares problem are linearized and the adjustment is iterated. Approximation values are required for the parameters \( P^0, a^0 \) of the cylinder. The radius enters only linearly into the cylinder equation, and therefore no approximation is needed. An initial guess for the axis direction can be obtained from the normals estimated in each point. The normal vectors lie approximately in one plane, which is perpendicular to the cylinder axis \( a^0 \). The point \( P^0 \) can be estimated from the center of gravity of the points \( Q_i \). A method for cylinder fitting without additional constraint equations is described in detail in (Lukacs et al., 1998).

Typical accuracies achieved in cylinder fitting are in the order of \( \pm 3 \) cm. As for the normal estimation this value is below the accuracy of the laser scanner, but the rough surface of the tree, as well as the deviation of a real branch from a circular cross-section explain this difference. Fig. 2 shows an example of a cylinder fit to a set of points. The r.m.s.e. of the fit is \( \pm 2.7 \) cm, the cylinder radius is \( 29 \) cm.

Applying cylinder fitting to a point cloud, does not only determine radius and axis, but by projecting the given points to the axis, two outermost points can be identified, marking the top and bottom end (or start and end point) of the cylinder.

3.4 Cylinder following

Assuming that the axis direction and radius of a branch are known at a certain position, this branch can be tracked (in both axis directions) in the given point cloud, using the cylinder fitting described above. The base cylinder is described by the known radius and an axis segment (i.e. the two axis end points). Shifting this cylinder forward (or backward) in the axis direction for a certain percentage of its length, an approximate position for the next cylinder is found. Points that are close (i.e. within a certain distance) to this new cylinder can be selected and used for cylinder fitting. If the parameters of the fitted cylinders are in accordance with quality criteria, the cylinder is accepted and tracking continues. If the cylinder does not pass the quality checks, another (e.g. a smaller) forward shift is applied, or a shorter cylinder can be used for selecting the points for the cylinder fitting.

The method of cylinder following has many parameters, but given the domain of tree reconstruction, these values can be set easily. The measurement accuracy, the tree surface roughness, the curvature of branches, the change in radius along the axis, and finally the deviation of the circular cross-section guide the selection of the necessary threshold values. Values have been determined empirically and checked for compatibility to knowledge of forest experts. To account for gross measurement errors in the forest environment, these values must not be set too tight.

For trunks, usually growing straight, the length of the cylinder can be set to 50 cm. Smaller branches showing stronger curvatures require shorter lengths. Forward shift has been set to 30% of the length, which gives enough flexibility to adopt to the curved branch. If cylinder fitting is not accepted, shorter forward shifts (16%, ...) are tried first, assuming that a point set similar to the previously successfully used points will give better results. If these shifts fail too, longer forward shifts are tried (100%, ...), assuming that a major irregularity in the branch disables the good fit of a cylinder. If this fails too, cylinder following is stopped in this direction.

Points being close to the shifted cylinder are selected for the next least squares fit. Closeness is determined in radial direction to the cylinder side surface, and good results have been achieved with a threshold value of 7 cm, both for the inner and the outer distance. This corresponds to a 3σ bound of the achievable fitting accuracy. Points lying inside this bound are counted, but not used for fitting of the cylinder. If the forward shifts are enlarged (e.g. 100%), the selection of points has to be made more generous, too, because of the unknown curving direction of the branch (e.g. 10 cm for 100%).
Deciding if the cylinder fit is successful, is based on the r.m.s.e.
of the fit (3cm), on minimum and maximum radius of the cylinder
(5cm–50cm), maximum allowed angle deviation between consecutive
cylinders (e.g. 3°), maximum allowed radius change in percent
(e.g. 2.5%) and maximum allowed percentage of points within the
cylinder, diminished in radius by 3σ (5%). Other possible measures
would include checking the angular distribution of points on the cylinder skin.

3.5 Segmentation

The algorithm of cylinder following requires a starting point in
the given point cloud which lies approximately on the branch
surface and approximate values for the cylinder axis direction, the
algorithm for cylinder fitting requires a set of points of which all
points are supposed to lie on one cylinder. Apart from segmenting
the point cloud manually into branches, and thereby choosing the
relevant branches, this can also be done automatically. (Gorte and
Pfeifer, 2004) describe a method for transforming the point cloud
into voxel space, filling the empty spaces inside the branches
with mathematical morphology (closing), and skeletonizing this
tree. Each skeleton voxel belongs to one 3D curve, representing
a branch in object space. The original point cloud is segmented
based on distance to these 3D curves in voxel space.

3.6 Clustering aligned cylinder axes

The cylinder following provides a number of cylinders, more
precisely a start and an end points and one radius for each pair
of axis points. This can be performed for many datasets, i.e.
different positionings, and for many trees within one dataset. As
mentioned above, precise registration of different scans can pose
a severe problem in forest environments, and in these cases it
makes sense to combine not the (poorly registered) point clouds,
but results on a later processing stage. An exemplary shift of 10cm
between two scans will make the cylinder fitting impossible for
the combined point cloud, but merging two stems which are 10cm
shifted apart, will still give reliable and improved results of the
radius. In order to combine the results from multiple runs of
the cylinder following within one plot, the computed axes are
combined to form individual ‘trees’. A tree is in this sense a
collection of branch elements that are aligned along each other.
For this, a measure of alignedness needs to be defined.

Given are two axes, i.e. two lines in 3D space, \( a \) and \( b \), and a
start and an end point forming a segment on either axis, \( S_a \) and
\( E_a \) for axis \( a \) and \( S_b \) and \( E_b \) for axis \( b \) (see Fig. 3). The distance
\( d(S_a, E_a, S_b, E_b) \) between those two segments is defined as
\[
d(S_a, E_a, S_b, E_b) = \frac{1}{4}(S_a b + E_a b + S_b a + E_b a)
\] (3)

with \( P \) denoting the (shortest) Euclidean distance between point
and line \( l \). For two segments on the same line the distance is zero.
For two segments on parallel lines the distance is equal to
the distance of the lines. For intersecting segments, the distance
grows with the length of the segments and their angle. For intersecting
or skew lines, the distance also grows with the distance of the
segments. This means that two short segments at – more or
less – the same location but with different axis direction are still
close to each other. It is noted that this measure is not a metric,
because the triangle inequality is not fulfilled. However, this mea-
sure is invariant to congruency transformations. Additionally, the
measure is given in length units (i.e. in meters).

The algorithm for finding sets of aligned segments is a global
algorithm and computes the distance between every pair of seg-
ments. All segments, which have a distance below a threshold
value (e.g. 30cm) are considered neighbors to each other. This
threshold has to be set with the expected registration error in
mind. The grouping of aligned segments begins with choosing
the segment which has the highest number of neighbors within
this distance. In case there are multiple segments with the same
highest number of neighbors, the one segment is chosen, were
the largest distance to a neighbor is smallest. All the neighbors of
this segment form a group and are removed from the set of segments.
Then the procedure is repeated, until no segments can be grouped
anymore.

3.7 Computing continuous branch models

From applying the above algorithms of cylinder following and
possibly aligning of axes, a set of start and end point pairs with
associated radius is determined to describe one branch. These
sets are used to determine a model of a branch that is geometri-
cally continuous of order one (continuous tangent planes, yielding
smoothness). Note that the collection of cylinders obtained so far
is not even a continuous model of the branch.

To compute these improved branch models, the two end points of
all cylinders are parameterized according to their position on a 3D
line. It is sufficient to take the average direction of the individual
axes as the direction of this parameterization axis, weighting each
with the length of the segment. Each cylinder end point provides
one observation of the radius along this line.

Different models can be fitted to these observations. Assuming a
linear decrease of the radius (e.g. with height) an adjusting line is
fitted to the radius observations, parameterized over the 3D line.
Alternatively moving least squares (Lancaster and Salkauskas,
1986) can be applied to get a smooth model of radius change with
branch extent.

3.8 Hull determination

Leaf trees can be scanned for crown reconstruction ideally in the
winter half year when they bear no foliage, whereas the crowns of
coniferous trees are opaque for the laser scanner the whole year.
Therefore very little points are measured on the upper trunk.
Additionally, the separation of needles and branches in the point
cloud is even visually very hard. In Fig. 4 a detail of a scan from
a pine (one branch) is shown. For these reasons, an alternative is
necessary to determine the extent of the tree, given in the form of
an outer hull. Given the irregular structure of trees, this hull is
naturally far from being convex.
Figure 4: Example of laser scanner point cloud of a branch on a coniferous tree. Scattering of the points in the top-to-bottom direction in the image can partly be caused by the wind, other possible sources are the mixing of ranges to targets within one laser beam.

In the following a simple algorithm for the reconstruction of the outer hull is described, requiring that the stem axis is known. The outer hull is defined as a collection of closed polygons. Each polygon describes the hull at a certain height, and the polygons are sorted with ascending height. The vertices of such a polygon represent the outermost point of the tree in a specific direction and height. The height interval (e.g., 20cm) and the number of directions (e.g., 12, corresponding to equally spaced angular sectors of 30°) are user-specified values. It is chosen based on point density and, for coniferous trees, the average distance of branches in vertical direction. Each point of the outer hull has 3 co-ordinates (x,y,z).

The outer hull is computed by first aggregating the points of the given point cloud in the height slices and the angular sectors. If the outer hull is to be determined for a plot of trees standing close together, points are first assigned to a tree by choosing the one where it has the minimum horizontal distance to the axis. Taking the outermost point in each sector is not sufficiently robust for computing the outer hull, because measurement errors, or points from neighboring trees, without reconstructed axis, may fall into that sector as well. A simple analysis considering the distances of the points to the axis is performed, starting from the points closest to the axis. If large gaps (i.e., large distance differences) are found between point in the sector, the outer points are discarded and considered to be measurement errors or originating from other trees. From the accepted points, the one with the largest distance to the axis is taken.

This algorithm provides the outer hull in the form of polygons in different heights.

4 EXAMPLES AND DISCUSSION

The set of presented algorithms can be used differently. A fully automatic reconstruction of leaf trees can be performed in the following way. After removing spurious points in the original point cloud the segmentation algorithm in the voxel domain is carried out, and for each found segment (i.e., a set of points) cylinder following is applied.

In Fig. 5 the result of the segmentation (45 segments) and the reconstruction of an oak tree are shown. The segmentation performed very well, but it can be seen that some smaller branches were grouped to larger segments. The tree was scanned from all sides, using 4 scans, summing up to 2.4 million points. Not all branches could be reconstructed, but the 36 found branches are correct – based on visual comparison to the point cloud. Apparently some failures in the segmentation process (mentioned above), but also the fact, that some branches are only covered poorly with points cause these results. In these cases it should – at least – be possible to estimate rough start and end points of the branch. During the cylinder following all together 90 cylinders were determined, with minimum and maximum r.m.s.e. of 1.1cm and 3.0cm, respectively. The average fitting accuracy (r.m.s.e.) is 1.8cm.

In Fig. 6 the fully automatic reconstruction of 3 coniferous trees with their outer hulls is shown. Left the point clouds are shown (5 scans, all together 7 million points), each scan is given a different color for each tree. Right, the reconstructed stems and the hulls are shown. For the stem reconstruction one point was selected manually on each of the three trees. As the registration was not accurate enough, leading to distances of the tree axis in higher parts of 10cm, each scan was processed individually. After running the cylinder following, the branches were aligned automatically and a model of linear decreasing radius was fit to the sets of cylinder start and end points and radii. This yielded the stems shown in the right part of Fig. 6. For computing the outer hulls, the slices were given a height of 1m, and split up into 16 sectors (22.5° each). The hulls are shown as closed polygons around the tree axes.

These results demonstrate that the method works well, but it can also be seen, that the upper part of the stem of the left tree was not reconstructed. A lower part of the stem is occluded by lower points on the side branches and needles, and the cylinder following could not bridge this gap. Tuning the parameters of the cylinder following can help here, but this makes the process less automatic.

Another application (not shown) is to apply cylinder following to manually selected point clouds. This can give results in short time, too, and a quality check can be performed during the work.
5 CONCLUSION

We presented a set of algorithms for the reconstruction of individual trees from dense laser scanning point clouds in forests. Terrestrial laser scanning is, as this example shows, not only competing with image based methods, but allows also to open new application fields. The technology of laser scanning and the immediate processing steps (i.e., registration) are not completely fit for difficult and rough environments as forests, yet.

The cylinder models of the individual tree stems are being used by forest managers and timber industry for determining important parameters describing timber quality like taper, sweep and lean which are closely related to the proportion of reaction wood. The reconstructed crowns or outer hulls of the crowns, respectively, allows the calculation of the crown projection area which one the one hand describes the vitality or the competitive status of forest trees and on the other hand is an important criteria for target-oriented management of forest stands (Biging and Dobbertin, 1995).

It has been demonstrated that fully automatic modelling is possible, but visual interpretation of the point cloud allows to extract more information, than the models currently deliver. As was to be expected, reconstruction works better for lower, thicker branches, where the point coverage is denser. The different threshold values should be set automatically from the data. Future research direction aims therefore at i) extracting more branches, and ii) doing away with the restriction of circular cross-sections. This will allow the extraction of more information (e.g. quality measures for wood production) from the laser scanner point cloud.

REFERENCES