

TREE DETECTION AND DIAMETER ESTIMATIONS BY ANALYSIS OF FOREST TERRESTRIAL LASERSCANNER POINT CLOUDS

A. Bienert^{a,*}, S. Scheller^a, E. Keane^b, F. Mohan^b, C. Nugent^c

^a Technical University of Dresden, Institute of Photogrammetry and Remote Sensing, D-01062 Dresden, Germany
anne.bienert@tu-dresden.de

^b TreeMetrics, The Rubicon Centre, CIT Campus, Bishopstown, Cork, Ireland

^c University College Cork, Cork Constraint Computation Centre, Cork, Ireland

Commission V, WG V/3

KEYWORDS: Terrestrial laserscanning, point cloud, segmentation, automation, forest inventory, diameters along the stem

ABSTRACT:

Forest inventory schemes collect, besides tree species and some area parameters, geometric tree parameters such as diameter at breast height (DBH), tree height, stem profiles, azimuth and distance. For some years the use of a terrestrial laserscanner for this forestry inventory task has been discussed. Dense 3D point clouds recorded in forest stands may form the basis for automatic determination of forest inventory parameters.

The paper presents an algorithm to detect trees in a horizontal cross section through a point cloud. This algorithm is divided in two segmentation steps to minimise the probability of false detections. The first segmentation step is a point cluster search in a cross section of the point cloud. In a second step all clusters are verified or discarded by analysing the point density in neighbouring cross sections. A study with 547 trees shows a detection rate of 97.4 % in single scan laserscanner data. Two other plots with heavy branching show a detection rate of 100 % and 94 %. Besides the tree detection, a new parameter is introduced to eliminate miss-fitted stem diameters. By using this parameter a least squares polynomial model is generated to smooth the diameters along the stem. Finally some results are demonstrated.

1. INTRODUCTION

Almost 73% of the European forest areas are used as production forest (Food and Agriculture Organisation of the United Nations, 2007). In order for efficient cultivation and planning to take place the current timber volume of standing trees inside a forest holding needs to be carefully monitored. Therefore forest inventories are carried out at regular intervals. Besides other inventory parameters, the most important geometric tree parameters are diameter at breast height and tree height, which are typically measured manually. Measurements on standing trees to determine the timber present is a time-consuming and costly process. The boles are typically measured after felling. Terrestrial laserscanning is an important technique which enables a non-destructive determination of standing timber. In the last 5 years terrestrial laserscanning has become an interesting tool for forest application. For instance, Aschoff et al. (2006) have researched the forest hunting habitats of bats with the use of terrestrial laserscanner data.

Laserscanner point clouds have a very high point density, which enables an extensive analysis and an automation of several utilisation processes. Different studies have been published with the aim of automatically determining forest inventory parameters from point clouds (Simonse et al., 2003; Hopkinson et al., 2004; Watt and Donoghue, 2005; Henning and Radtke, 2006b).

Segmentation and tree identification is done in a horizontal cut of the point cloud, to reduce processing time. Scan points on stem surfaces are mapped as an arc within the layer. Hopkinson

et al. (2004) demonstrated tree detection by overlaying a manually surveyed tree map on the point cloud layer. Fully automatic segmentation and tree identification is presented in Aschoff et al. (2006). A horizontal cut from the scanned point cloud is mapped to a constant raster and after a layer segmentation a Hough-transformation is performed. Based on this work Thies and Spiecker (2004) present a detection rate of 22% in single scan and 52% in multiple scan setups.

The approach presented here was first suggested in Bienert et al. (2006). The goal was to develop a program to determine forest inventory parameters by analysing laserscanner point clouds. In the following section, the plot acquisition and different data sets, provided in several study areas, will be outlined. Previous work had shown that the tree identification in single scans still produce false detections (classified as type I and type II errors). Heavy branching or dense undergrowth influence the tree identification. Sometimes a detection rate was obtained of less than 15%. Thus, the tree detection method was enhanced to minimise the probability of false detections. Section 3 will present the automatic identification of forest trees scanned in single scan setups by using two different approaches of segmentations. Beyond the use of segmentation, section 4 will show the profile fitting at several heights along the stem. A parameter which eliminates miss-fitted diameters will be introduced. In section 5 a diameter smoothing technique using a least squares polynomial model along the stem is given. Finally, section 6 will present the segmentation results of study areas containing a total of 547 trees of different species. Additionally results of the profile fitting and smoothing will be presented.

2. DATA SETS

2.1 Data recording

All data sets presented in this paper were acquired with the terrestrial laserscanner FARO LS 800 HE80. This full spherical laserscanner with a field of view of 360° horizontal and 320° vertical has a range up to 80 m. With a distance accuracy of ± 3 mm and a data rate of 120 000 points per second, point clouds with more than a million accurate measured surface points can be obtained in a short time (a typical 30 million point scan can be performed in 8 minutes). The LS 800 HE80 uses the phase-shift principle for range determination where a mirror rotates and directs the laser pulses (Faro, 2005).

To ensure a levelled instrument a built-in spirit level was used. The start direction was aligned to north to calculate azimuths without an offset. All plots were scanned in the single scan mode, so the scanner was placed at the plot centre.

2.2 Study areas

In May 2006 (leaf-on), 21 plots were acquired in several stands in the Vienna Woods 25 km west of Vienna. 52% of all plots were mixed plots with coniferous and deciduous trees, 33% were beech (*Fagus sylvatica*) and 5% spruce (*Picea abies*) plots. The stands ages are between 65 and 140 years. Manual reference data like DBH, tree height, azimuth and distance were measured subsequently. The radius of each plot was 15 m and the used scan resolution was 0.045°.

Another test site was acquired in March 2007, close to Aberfoyle in Scotland. These plots (plot A and B) contained 34 year old spruce (*Picea abies*) and were located on hilly terrain with less undergrowth. Both were scanned with a scan resolution of 0.036° and had a plot radius of 10 m. Figure 1 shows an intensity image of plot A. In contradiction to the study area of the Austrian forest, heavy branching on the lower region of the stems is present.



Figure 1. Intensity image of plot A

Yet another plot (plot C) was scanned 2006 in a Sitka spruce (*Picea sitchensis*) plantation in North Tipperary near Roscrea, Ireland. This plot was located on flat terrain and had a stand age of 35 years. A scan resolution of 0.036° was used. Inside a plot radius of 12 m harvester data was available to compare the diameter along the stem obtained from the program. Harvesters equipped with the appropriate sensors are able to measure diameter and length of boles.

3. TREE DETECTION

After a terrain model reduction, as described in Bienert et al. (2006), tree detection was undertaken. The segmentation was subdivided into two different algorithms to minimise classification errors, which is helpful on data sets from plots with heavy branching. Initially a segmentation based on a point cluster search was done. Then all clusters were analysed in a second step, to determine their point density inside a raster. The outcome of the tree detection was the number of trees in the plot and the approximate position of each tree.

3.1 Segmentation based on point cluster search

As outlined in Bienert et al. (2006) and Scheller and Schneider (2006) the segmentation was done on a slice with a thickness of 10 cm cut through the point cloud 1.3 m above the terrain model. The whole slice was analysed by a 2D-quadratic structure element with a size s . This element moves over the slice in X/Y projection and searches for point clusters with more than $n > n_{min}$ points. One object will be separated into two objects if the distance between one point and the nearest one is bigger than $s/2$. To classify the objects as trees, a circle fitting with all points of an object was done. As exclusion criteria the error of unit weight and the fitted diameter of the circle fitting are used to classify objects.

This algorithm produces type I and type II errors. Points which belong to a branch produce an error of unit weight bigger than the present threshold value and so the object will be rejected. As a result of this segmentation an object list was generated with all classified trees (containing false objects – type II errors) and a rejected object list with point clusters which fail the classification (containing tree cluster points – type I errors).

3.2 Segmentation based on point density raster analysis

Scanning techniques are characterised by a regular point grid area on a surface. The number of points inside this defined raster area depends on the scan resolution used and the alignment of the object surface to the scanner. A cut with a thickness greater than (minimum two times) the vertical scan resolution through a point cloud produces different numbers of points inside this defined raster in X/Y projection.

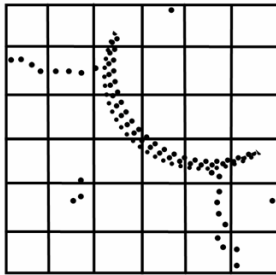
Subsequently the point density raster analysis of the object and rejected object list from the first segmentation step could be performed. The rejected object list was used to minimise type I errors, given that it contained stems which are not detected in the first segmentation phase. The 2D-bounding box for each cluster with x_{min} , x_{max} , y_{min} and y_{max} defined the object size and a raster with a cell size of 4 cm was overlaid (Figure 2a). Because of the distance from scanner to object and scan resolution are known, a maximum number of target points, which exist inside one raster (Figure 2c), can be calculated (equation 1 and 2). To get the number of redundant points a diagonal oriented vertical object surface (inside a raster element) was used. Because of the perpendicular stem direction, raster elements with a very high point number were produced (Figure 2b). Minimising the predicted target number by about 30% the minimum threshold for the raster analysis was defined. All scan points of a raster element above the threshold (Figure 2d) were copied to a new co-ordinate list, which includes only points of the stem surface. After analysing all rejected objects the new co-ordinate list was checked with the cluster search method mentioned in section 3.1.

$$n = \frac{\Delta z}{\Delta d} \cdot \frac{cell_{diag}}{\Delta d} \quad (1)$$

$$\Delta d = \sqrt{2 \cdot S^2 - 2 \cdot S^2 \cdot \cos \alpha} \quad (2)$$

where: n = number of points
 Δz = cut thickness
 $cell_{diag}$ = diagonal cell size
 Δd = average distance between object point
 S = distance from scanner to object
 α = scan resolution

* assumed that the scanner is on the same level like the object points



a)

0	0	4	1	0	0
4	3	15	0	0	0
0	0	16	1	0	0
0	1	2	14	16	9
0	2	0	0	4	1
0	0	0	0	2	2

b)

scan resolution: 0.1°
distance to scanner: 10 m

target number
on a raster element: 19
- 30%: 6
threshold: 13

c)

0	0	4	1	0	0
4	3	15	0	0	0
0	0	16	1	0	0
0	1	2	14	16	9
0	2	0	0	4	1
0	0	0	0	2	2

d)

Figure 2. a) Rejected object with overlaid raster; b) number of points inside the raster; c) threshold value calculation with simulated scan parameters; d) raster elements which pass the threshold (yellow elements)

The object list, which contained type II errors, was examined, repeating the technique of point density analysis. Type II errors caused by branching or undergrowth were characterized by a lower point density. Raster elements with a point number beneath the threshold were detected and therefore deleted from the object list.

4. DIAMETER PROFILE FITTING

Stem profiles at different height intervals (Figure 3) can be determined with the knowledge of the approximate position and diameter returned by the tree detection process. Starting from the reference ground point (Bienert et al., 2006) the profile fitting was done, using a least square circle fitting algorithm (Bienert et al., 2006). This procedure minimises the mean square distance from the fitted circle to the data points, which is similar to the algorithm shown in Henning and Radtke (2006a).

Henning and Radtke (2006a) show that surface points caused by branching appear outside of the stem cross-section. An overestimated diameter will result. If only a short section of the stem profile is visible, it is not possible to determine reliable diameters.

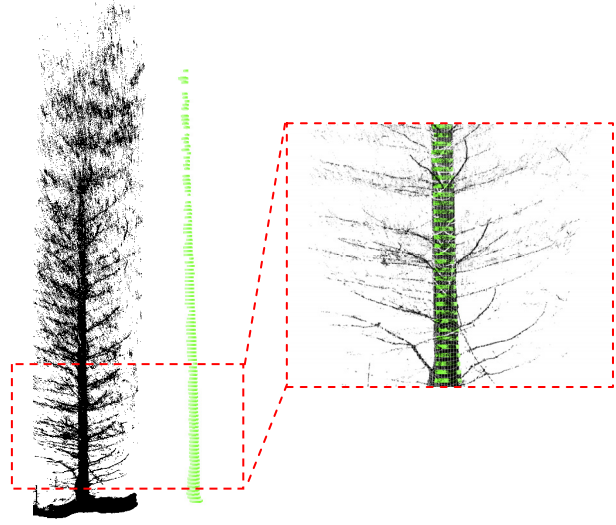


Figure 3. Point cloud of a spruce (left); generated profiles (middle); combined side view of a cut-out (right)

A new parameter, the “Reliability Factor”, was introduced to detect these over- and underestimated diameters as outlined in Bienert et al. (2007). Attributes of the circle fittings are used to evaluate the quality of the fit. The belonging to the universe is defined by the probability P_x of each parameter x . The following information are used:

- the error of unit weight of the circle fitting ($P_{\sigma 0}$),
- the standard deviation of the fitted diameter ($P_{\sigma \text{Diameter}}$),
- the visible stem section (angle of the visible circle section) (P_α),
- the ratio of the scanned point number and the calculated number of points (determined by scan resolution, distance to stem, cut thickness) ($P_\%$),
- a median filtering with 11 neighbour elements of the fitted diameters along the stem (P_{Median}).

The probability $P_{\sigma 0}$, $P_{\sigma \text{Diameter}}$ and P_α are determined by calculating the ratio of each value to a threshold ($P_x = x / TH_x$). To consider the median filtering the diameter is compared with the median of its five neighbour up- and downward diameters. If the difference between the diameter and the median diameter is smaller than 5% of the median diameter, P_{Median} is 1 otherwise 0. All five summands consist of values between 0 and 1. The reliability factor was calculated as an arithmetic mean (equation 3). In general, diameters with a factor greater than 0.7 are deemed as reliable.

$$ZP = (P_{\sigma 0} + P_{\sigma \text{Diameter}} + P_\alpha + P_\% + P_{\text{Median}}) / 5 \quad (3)$$

To predict the taper of standing trees taper equations and basic taper models are extant (Nieuwenhuis, 2002). A modified Kozak taper equation was used. Therefore a prediction of taper in non-observable stem heights can be done based on observed measurements.

5. UTILISING A POLYNOMIAL MODEL TO SMOOTH THE STEM

To improve the value recovery in forest stands, accurate stem shape and position of trees is essential. The efficiency of the harvesting operations can be greatly increased by determining the crop “profile”. Three important aims are to consider: maximising of (timber) volume, maximising of (timber) value and minimising of costs (Nieuwenhuis, 2006). Even just a few noisy values in the predicted stem profile can greatly impact on the processes used to determine the expected products extracted from a stem and its overall value. This in turn means that there are inaccuracies in the inventory data. In order to ensure the stem profile is as accurate as possible the effects of noisy values must be minimised as much as possible. In pursuit of this goal a weighted polynomial function is an invaluable technique to smoothen stem diameters.

In fitting a smoothing function, such as spline or polynomial model to the stem diameters, the important point to realise is that one can obtain a fit as close to the data as one wants, simply by adding more and more breakpoints. However, what one really wants is a smooth curve, flexible enough to capture the (unknown) functional relationship underlying the data, yet smooth enough not to follow the noise component in the data due to measurement errors. The problem of separating the noise from the underlying trend becomes more manageable if an indication of reliability can be determined. During the circle fitting process the reliability factor was determined as outlined in section 4.

This Reliability Factor was used to produce a set of weights for each point on which the polynomial model was to be built. This helps to ensure that unreliable data does not impact greatly on the fitted model. The inputted data to the model is the set of diameters from the circle fitting process and their corresponding reliability values as weighting factors. The polynomial model was fitted to the weighted set of inputs, using least squares regression. This minimised the residual sum of errors over the data (Hastie et al., 2001). A polynomial model was used instead of an interpolating spline, as the goal in fitting the model is to approximate the trend of the data and minimise the effects of noise. Approximating spline models exist and a smoothing spline was originally used for this purpose (De Boor, 1978). However, given the levels of noise and relatively simple shape of tree stems the least squared polynomial model was favoured. This simpler model was found to track the underlying trend adequately, given the noise and did not require the specification of a smoothing parameter.

6. RESULTS

6.1 Segmentation

The results shown here are derived from the 21 plots acquired in May 2006. Overall 533 of 547 trees could be detected correctly and therefore the detection rate was 97.4%. There were 14 type I errors, mainly caused by occlusions of trees and vegetation standing in the foreground. Therefore not enough laser points were landing on the stem surface within the horizontal data cut. 60 type II errors are produced. Figure 4 shows the number of errors caused in the different plots.

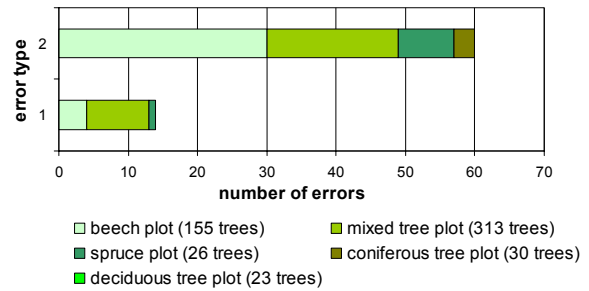


Figure 4. Detection errors

To assess the efficiency of both segmentations two plots (plot A and B) with heavy branching (Figure 5) were processed. Plot A and plot B consist of each 16 trees inside a plot radius of 10 m. By using the segmentation based on point cluster search only 2 (plot A) and 6 (plot B) stems could be found correctly. After the second segmentation using point density raster analysis all trees of plot A could be identified correctly and only one tree of plot B could not be found. The results of the segmentation of both plots are outlined in Table 1.

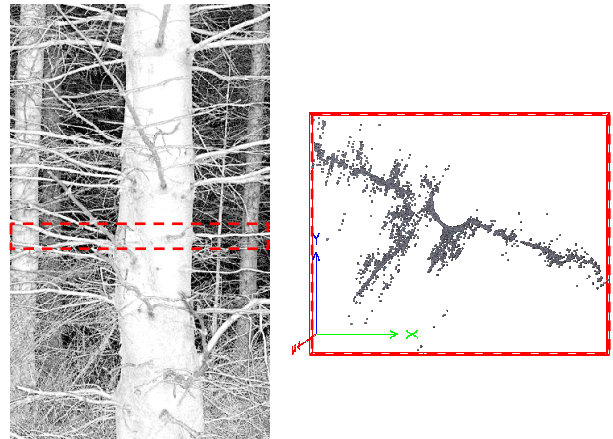


Figure 5. Spruce with heavy branching (left); top-view of a horizontal 10 cm cut

	Segmentation based on			
	point cluster search		point density raster analysis	
Data	Plot A	Plot B	Plot A	Plot B
Trees	16	16	16	16
Classified objects	23	21	16	21
Correct found stems	2	6	16	15
Detection rate	12.5%	37.5%	100%	93.75%
Type I error	14	10	0	1
Type II error	21	15	0	6

Table 1. Results of the segmentation

6.2 Diameter profile fitting

To assess the accuracy of profile fitting along the stem harvester stem file data was used. Figure 6 shows the comparison between the profile fitting and the diameters from the harvester along a Sitka spruce stem located 8.5 m from the laserscanner. The underestimation caused by too few survey points landing on a short stem segment, while the overestimation is caused by branching. The standard deviation of the profile differences of all fitted diameters (yellow rectangles in Figure 6) along the stem is 3.91 cm. Thus, all under- and overestimated diameters can be detected by using profiles with a “Reliability Factor” bigger than 0.69 (Figure 7). A standard deviation of 1.36 cm is obtained.

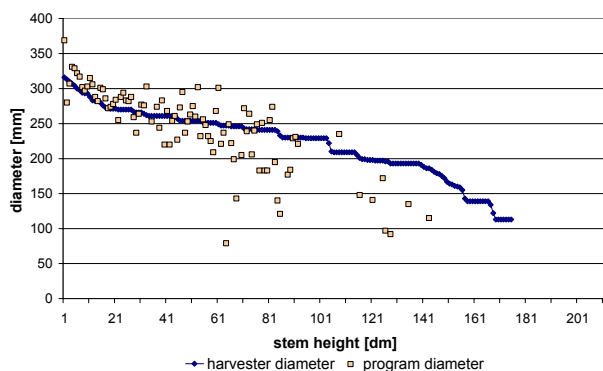


Figure 6. Derived diameters from the profile fitting of a Sitka spruce in comparison with the diameter obtained from the harvester

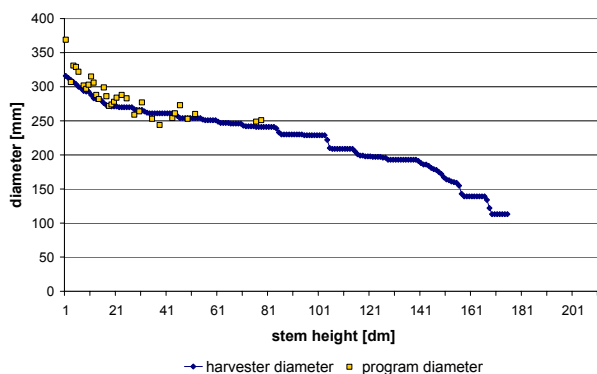


Figure 7. Derived diameters of a Sitka spruce with an “Reliability Factor” bigger than 0.69 in comparison with the diameter obtained from the harvester

The stem was smoothed up to a height of 7.8 m with a standard deviation of 0.64 cm by a polynomial model (Figure 8 - green triangles). As a height of 7.9 m a taper prediction based on the modified equation of Kozak, as outlined in section 4, was done (Figure 8 - red points). This equation was specific for spruce species with Ireland-Centric coefficients.

Table 2 summarises the accuracies of the profile fitting (all diameters, reliable diameters), the smoothing and the predicted taper for the tree from Figure 8. Finally the standard deviation of 22 trees inside a 12 m radius of the Sitka spruce plantation (plot C) are shown in Table 3.

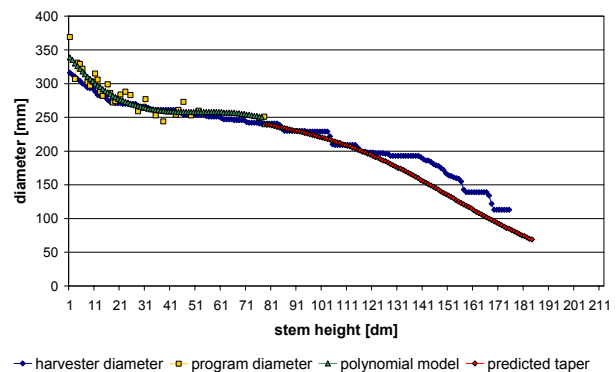


Figure 8. Reliable diameters, smoothed data and predicted taper compared with harvester data of one Sitka spruce

[cm]	Standard deviation	Arithmetic mean	Max. difference	Min. difference
All diameters	3.91	1.03	0.168	0
Reliable diameters	1.36	-0.94	-5.30	0
Polynomial model	0.64	-2.20	-2.30	0
Predicted taper	1.36	0.20	3.70	0

Table 2. Results of the profile fitting of one Sitka spruce located in plot C (compared with harvester data)

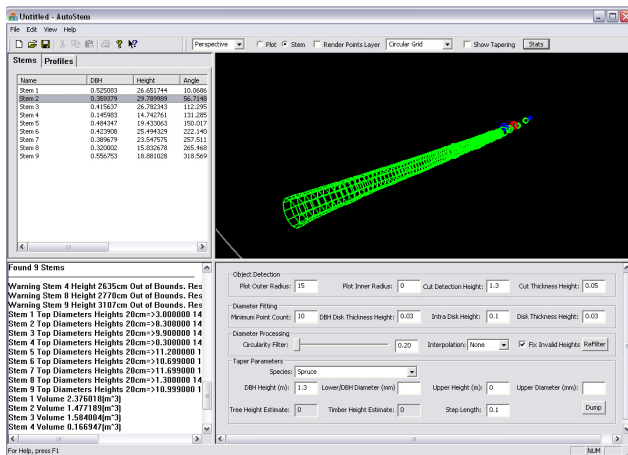
[cm]	Standard deviation	Arithmetic mean	Maximum difference	Minimum difference
Reliable diameters	2.48	-0.64	19.60	0

Table 3. Results of the profile fitting of plot C with 22 trees (compared with harvester data)

7. CONCLUSION

In collaboration with TreeMetrics (www.treemetrics.com), a computer program (*AutoStemTM*) was developed in the C++ programming language to automatically determine forest inventory parameters based on terrestrial laserscanner point clouds. Figure 9 shows the *AutoStemTM* user interface. By processing a point cloud (X, Y, Z), stem number ordered by azimuth, stem position, diameter at breast height (DBH - measured at a height of 1.3 m), tree height and profiles along the stem in user specific heights are displayed. Data sets recorded in natural or production forests from one position (single scan mode) or more positions (multiple scan mode) can be processed. Obstructions such as undergrowth, rocks and heavy branching do limit the effective range of a single scan. In these situations multiple scans are done to gather the necessary information.

By using two segmentation algorithms a slice of a point cloud was analysed. The rate of success was enhanced by using both algorithms for plots with heavy branching. Type I errors were minimised by the second segmentation process. Furthermore robust tree detection is possible, which the detection rate of 97.4% confirms. Nevertheless some type I and type II errors are present.

Figure 9. *AutoStem™* user interface

The newly introduced reliability factor enables the detection of badly fitted diameters (Figure 7 vs. Figure 8) allowing a better determination of the stem shape to be made. An average standard deviation of 2.48 cm for 22 Sitka spruces was presented for profile fitting with a height interval of 10 cm, compared with harvester data. Henning and Radtke (2006a) present a standard deviation of nine loblolly pine trees of 2.1 cm estimated at every 1-m bole section. However, this was done using co-registered datasets of three separate scanner positions around each tree. By processing point clouds obtained in multiple scan setups, occlusions can be reduced and accuracy increased.

REFERENCES

- Aschoff, T., Spiecker, H., 2004. Algorithms for the automatic detection of trees in laser scanner data. In: Thies, M.; Koch, B.; Spiecker, H.; Weinacker, H. (eds.): *Laser-Scanners for Forest and Landscape Assessment, Proceedings of NATSCAN-Conference on Laser-Scanners for Forest- and Landscape Assessment, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences Vol. XXXVI, Part 8/ W2: 66- 70*
- Aschoff, T., Holdenried, M. W., Spiecker, H., 2006. *Terrestrische Laserscanner zur Untersuchung von Wäldern als Jagdlebensräume für Fledermäuse*. In: Luhmann, T. und Müller, C. (Hrsg.): *Photogrammetrie - Laserscanning - Optische 3D-Messtechnik, Beiträge der Oldenburger 3D-Tage 2006*, Herbert Wichmann Verlag, Heidelberg, pp. 280-287
- Bienert, A., Scheller, S., Keane, E., Mullooly, G., Mohan, F., 2006. Application of terrestrial laser scanners for the determination of forest inventory parameters. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 36, Part 5*
- Bienert, A., Scheller, S., Keane, E., Mullooly, G., Mohan, F., 2007. Automatische Erfassung forstinventur-relevanter Geometrieparameter von Waldbäumen aus terrestrischen Laserscannerdaten. In: Luhmann, T. und Müller, C. (Hrsg.): *Photogrammetrie - Laserscanning - Optische 3D-Messtechnik, Beiträge der Oldenburger 3D-Tage 2007*, Verlag Herbert Wichmann Verlag, Heidelberg
- de Boor, C., 1978. *A practical guide to splines*. Springer, New York
- Faro Europe GmbH & Co. KG, 2005. *Aufbruch in neue Dimensionen: der FARO Laser Scanner LS. Laser Scanner Broschüre*
- Food and Agriculture Organisation of the United Nations, 2007. *State of the World's Forests 2007*. Rome, Italy
- Hastie, T., Tibshirani, R., Friedman, J. H., 2001. *The elements of statistical learning*. Springer
- Henning, J. G., Radtke, P. J., 2006a. Detailed stem measurements of standing trees from ground-based scanning lidar. *Forest Science* 52(1), pp. 67-80
- Henning, J. G., Radtke P. J., 2006b. Ground-based laser imaging for assessing three-dimensional forest canopy structure. *Photogrammetric Engineering & Remote Sensing*, 72(12), pp. 1349-1358
- Hopkinson, C., Chasmer, L., Young-Pow, C., Treitz, P., 2004. Assessing forest metrics with a ground-based scanning lidar. *Canadian Journal of Forest Research*, 34, pp. 573-583
- Nieuwenhuis, M., 2002. The development and validation of pre-harvest inventory methodologies for timber procurement in Ireland. *Silva Fennica*, 36(2), pp. 535-547
- Scheller, S., Schneider, D., 2006. Extraktion von Primitiven aus Laserscannerpunktwolken zur Rekonstruktion von Tragwerken. In: Luhmann, T. und Müller, C. (Hrsg.): *Photogrammetrie - Laserscanning - Optische 3D-Messtechnik, Beiträge der Oldenburger 3D-Tage 2006*, Herbert Wichmann Verlag, Heidelberg, pp. 156-163
- Simonse, M., Aschoff, T., Spiecker, H., Thies, M., 2003. Automatic determination of forest inventory parameters using terrestrial laserscanning. *Proceedings of the ScandLaser Scientific Workshop on Airborne Laser Scanning of Forests, Umeå*, pp. 251-257.
- Thies, M., Spiecker, H., 2004. Evaluation and future prospects of terrestrial laser scanning for standardized forest inventories. In: Thies, M.; Koch, B.; Spiecker, H.; Weinacker, H. (eds.): *Laser-Scanners for Forest and Landscape Assessment. Proceedings of NATSCAN-Conference on Laser-Scanners for Forest- and Landscape Assessment. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences Vol. XXXVI, Part 8/ W2: 192- 197*
- Watt, P. J., Donoghue, D. N. M., 2005. Measuring forest structure with terrestrial laser scanning. *International Journal of Remote Sensing*, 26(7), pp. 1437-1446

ACKNOWLEDGEMENTS

All data presented here were acquired by *TreeMetrics Ltd, Cork/ Ireland*. Their support is gratefully acknowledged. We also like to thank Bernhard Posch from *Österreichische Bundesforste AG* for providing the reference data for the Austrian plots.