ISPRS Journal of Photogrammetry and Remote Sensing 64 (2009) 561–574

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

3D segmentation of single trees exploiting full waveform LIDAR data

J. Reitberger [a,*], Cl. Schnörr [b], P. Krzystek [a], U. Stilla [c]

[a] Department of Geoinformatics, University of Applied Sciences Muenchen, 80333 Munich, Germany

[b] Department of Computer Science and Mathematics, University of Applied Sciences Muenchen, 80335 Munich, Germany

[c] Photogrammetry and Remote Sensing, Technische Universitaet Muenchen, 80290 Munich, Germany

ARTICLE INFO

Article history:
Received 26 June 2008
Received in revised form
9 April 2009
Accepted 12 April 2009
Available online 17 May 2009

Keywords:
LIDAR
Segmentation
Aerial survey
Clustering
Forestry

ABSTRACT

This paper highlights a novel segmentation approach for single trees from LIDAR data and compares the results acquired both from first/last pulse and full waveform data. In a first step, a conventional watershed-based segmentation procedure is set up, which robustly interpolates the canopy height model from the LIDAR data and identifies possible stem positions of the tallest trees in the segments calculated from the local maxima of the canopy height model. Secondly, this segmentation approach is combined with a special stem detection method. Stem positions in the segments of the watershed segmentation are detected by hierarchically clustering points below the crown base height and reconstructing the stems with a robust RANSAC-based estimation of the stem points. Finally, a new three-dimensional (3D) segmentation of single trees is implemented using normalized cut segmentation. This tackles the problem of segmenting small trees below the canopy height model. The key idea is to subdivide the tree area in a voxel space and to set up a bipartite graph which is formed by the voxels and similarity measures between the voxels. Normalized cut segmentation divides the graph hierarchically into segments which have a minimum similarity with each other and whose members (= voxels) have a maximum similarity. The solution is found by solving a corresponding generalized eigenvalue problem and an appropriate binarization of the solution vector. Experiments were conducted in the Bavarian Forest National Park with conventional first/last pulse data and full waveform LIDAR data. The first/last pulse data were collected in a flight with the Falcon II system from TopoSys in a leaf-on situation at a point density of 10 points/m^2. Full waveform data were captured with the Riegl LMS-Q560 scanner at a point density of 25 points/m^2 (leaf-off and leaf-on) and at a point density of 10 points/m^2 (leaf-on). The study results prove that the new 3D segmentation approach is capable of detecting small trees in the lower forest layer. So far, this has been practically impossible if tree segmentation techniques based on the canopy height model were applied to LIDAR data. Compared to a standard watershed segmentation procedure, the combination of the stem detection method and normalized cut segmentation leads to the best segmentation results and is superior in the best case by 12%. Moreover, the experiments show clearly that using full waveform data is superior to using first/last pulse data.

1. Introduction

The development of new approaches to forest inventory utilizing LIDAR data has been an important research issue in the past.Beside area-based methods (e.g. Naesset (2004)), techniques for single tree extraction from LIDAR data have been investigated for mapping forests at the tree level and for identifying important parameters, such as tree height, crown size, crown base height, and tree species. Most of the methods for single tree detection have in common to find the trees from local maxima in the canopy height model (CHM). The stem position usually corresponds with a peak in the CHM, and the crown diameter is basically found from the segment polygon which is delineated from the CHM surface. For instance, Hyyppä et al. (2001) interpolate a local CHM from the highest laser reflections, Persson et al. (2002) apply the active contour algorithm, and Solberg et al. (2006) interpolate the CHM with a special gridding method and subsequently smooth the CHM with an appropriate Gaussian filter. Stem positions are derived from the interpolated CHM at the highest positions (Solberg et al., 2006) or from a special local tree shape reconstruction (Brandtberg, 2007). Tree crowns are typically derived with the watershed algorithm (Pyysalo and Hyyppä, 2002) or by a slope-based segmentation (Persson et al., 2002; Hyyppä et al., 2001).

* Corresponding author.
E-mail addresses: josef.reitberger@hm.edu (J. Reitberger), schnoerr@cs.hm.edu (C. Schnörr), krzystek@hm.edu (P. Krzystek), stilla@tum.de (U. Stilla).

0924-2716/$ – see front matter © 2009 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.isprsjprs.2009.04.002

Recently, Solberg et al. (2006) proposed a region growing method that starts from local surface maxima and finds crown polygons optimized in shape. The application of this method to a structurally heterogeneous spruce forest leads to an overall detection rate of 66% and a commission rate (= false detections) of 26% if the CHM is smoothed three times with a Gaussian filter of size 30 cm. The study of Persson et al. (2002) reports a detection rate of 71% for a Scandinavian forest dominated by spruce and pine. Heurich (2006) demonstrates that the segmentation method of Persson et al. (2002) leads to a detection rate of 51% for coniferous trees, 40% for deciduous trees and on average 45% in the Bavarian Forest National Park. Thus, the results of Persson et al. (2002) are by far better than the results of Heurich (2006), who used the same algorithm. This shows the strong dependency of segmentation results on the forest type. Timber volume is usually estimated from the tree height and crown diameter. In the study of Persson et al. (2002), 91% of the timber volume could be determined with 22% root mean square error (RMSE). Heurich (2006) estimates 85% of the timber volume with 31% RMSE.

The drawback of these methods is that they are solely oriented on the CHM, which is reconstructed from the LIDAR points in an interpolation process that smoothes the data to some extent. The degree of smoothing directly affects the success rate in terms of false positives and negatives. Moreover, in some cases neighbouring trees do not appear as clear separated local maxima. Thus, approaches that solely use the CHM will be restricted in the success rate anyway, especially in heterogeneous forest types where groups of trees grow close together. Since smaller trees below the canopy do not appear in the CHM the detection rate of trees with small diameter at breast height (DBH) is considerably low. Maltamo et al. (2004) developed a procedure to predict small occluded trees in lower forest levels by theoretical distribution functions. Timber volume and stem density can be estimated with 16% RMSE and 49% RMSE, respectively. The distribution functions must be calibrated with field data. Mehtätalo (2006) dealt with the same subject and estimated the number of stems using a function that describes the dependence of the probability of observing a tree on the crown radius. The study refers to simulated data and reports in the best case an RMSE of 11% for the number of detected stems. So far, little attention has been paid to reconstruct single trees in three dimensions using LIDAR data, mainly because of the low spatial point density of conventional first/last pulse systems. Morsdorf et al. (2003) segments trees in a 3D voxel space with the *k*-means clustering method using the local maxima of the CHM as seed points. Recently, Wang et al. (2008) presented a new method for 3D reconstruction of trees and tackled the problem with a hierarchical morphological approach. The 3D segments of trees are found by combining hierarchically the individual 2D tree crown regions detected in each layer of a voxel space. If using standard parameters, the overall detection rate in mature stands (tree age $>$ 90 years) is 84%. In more complex younger stands (tree age $>$ 60 years) the mean detection rate degrades to 75% if the process parameters are individually optimized. The method remains dependent on the correct findings of tree tops. Moreover, the procedure has limitations in younger stands (tree height $<$ 20 m) in the case of first/last pulse data.

In general, the single tree approach has some clear advantages compared to the area-based approach. If successfully applied, it has the potential to provide tree species information and to estimate more precisely the timber volume. However, Maltamo et al. (2007) pointed out that single tree detection methods need to be more accurate than standard forest inventory methods assuming that reduced cost must not be the only driving force to replace standard methods. The study has shown that timber volume can be estimated with about 30% RMSE accuracy if derived as a function of the tree height. Clearly, the parameter DBH is not a deterministic parameter but is also dependent on many other parameters which are not of allometric nature. Thus, in view of the apparent advantages of single tree methods, the detection rate, the 3D reconstruction of single trees and the accuracy of tree species classification must be improved. New technologies like small footprint full waveform systems will be one new driving force to push the methods to a new level since they detect significantly more reflections in the tree crown and provide the intensity and the pulse width as reflecting parameters. Moreover, new reconstruction methods for single trees must truly work in three dimensions and must be flexible to use the various types of information that modern LIDAR techniques provide.

Recent advances in LIDAR technology have rendered possible new full waveform scanners that provide a higher spatial point density as well as additional information on the reflecting characteristics of the forest structure like branches and stems. Wagner et al. (2006) focused on the calibration issue and the decomposition of full waveform data with a series of Gaussians. They also showed that different types of vegetation, such as trees and bushes, can be separated using the cross section calculated from the waveforms. Jutzi and Stilla (2006) fit Gaussians to the surface response that is obtained by the measurement of the individual emitted waveform and a corresponding deconvolution of the received waveform. Recently, Kirchhof et al. (2008) presented a method to improve the reconstruction of buildings partly occluded by vegetation by pre-segmenting reflections from the vegetation using the surface response. Reitberger et al. (2008) also showed the decomposition of waveforms by fitting Gaussians to the waveform. Compared to conventional first/last pulse data, an increase in the point density by a factor of 2 could be verified. The classification of coniferous and deciduous trees is successfully demonstrated using salient features calculated from the pulse width and the intensity of the decomposed waveform. Finally, Stilla and Jutzi (2008) illustrated the general principles of designing small-footprint pulsed laser systems and analysing the received waveform.

Obviously, full waveform data render possible new approaches to reconstruct and classify objects, such as trees. However, an important question that still needs to be resolved is: how much do full waveform data really improve forest inventory methods compared to conventional first/last pulse data? More precisely, it will certainly be interesting to find out whether new extraction methods for single trees can be developed that take advantage of full waveform data and improve the detection rate significantly.

The objectives of this paper are (i) to highlight a new segmentation method based on normalized cuts that extracts single trees using full waveform LIDAR data, (ii) to present the results of the method when applied to first/last pulse data and full waveform data acquired in the same area in leaf-on and leaf-off situations at different point densities, and (iii) to evaluate the new segmentation method with respect to the success rate.

2. Preprocessing of LIDAR data

2.1. Decomposition

We assume a point cloud in a region of interest (ROI) represented by $X_i^{\mathrm{T}} = (x_i, y_i, z_i, W_i, I_i)$ $(i = 1, \ldots, N_{ROI})$ LIDAR points that result from reflections of the laser beam at the positions (x_i, y_i, z_i) (Fig. 1). In the case of a full waveform scanner, a waveform decomposition provides the pulse width W_i and the intensity I_i as attributes (Jutzi and Stilla, 2005; Reitberger et al., 2008). Note that I_i is the integral of the return signal and, hence, this intensity value is physically the pulse energy. In comparison, conventional discrete return LIDAR systems provide only the coordinates of the reflections (x_i, y_i, z_i) and – in many cases

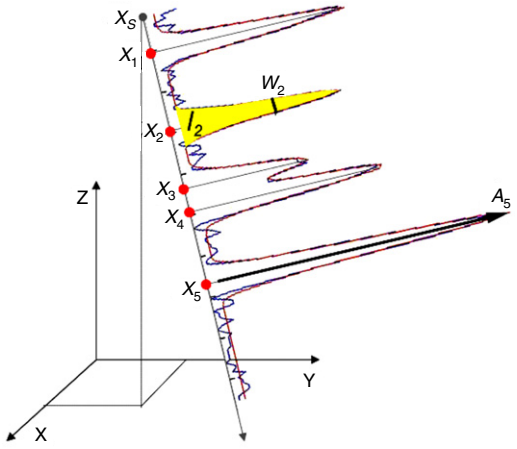

Fig. 1. 3D points and attributes derived from a waveform.

– also the intensity. However, in general, the users have only little knowledge how this intensity value is obtained in these systems. Typically, it is the maximum amplitude of the returned pulse, which is incorrectly called intensity. The advantage of full waveform LIDAR data is that, basically, each reflection can be detected by the waveform decomposition. Hence, neighbouring targets with a minimum distance of $v_g\tau/2$ in the direction of the laser beam can be separated, where v_g is the group velocity of the laser pulse in the atmosphere and τ is the pulse duration (full duration at half maximum) of the Airborne Laser Scanning (ALS) system (Wagner et al., 2006). For example, the pulse duration of 4 ns of the Riegl LMS-Q560 scanner leads to a minimum distance of 0.6 m. This is remarkable since conventional discrete return LIDAR systems have a dead zone of at least twice the above-mentioned minimum distance due to internal electronic characteristics, which makes these systems effectively blind after a single reflection. The length of this zone varies from system to system and even amounts to about 3 m for older systems.

2.2. Calibration

The calibration of the values W_i and I_i is achieved by using the pulse width W^e and the intensity I^e of the emitted Gaussian pulse. Additionally, the intensity is corrected with respect to the varying distances s_i between the laser scanner and the reflecting object with

$$W_i^c = W_i/W^e \tag{1}$$

$$I_i^c = (I_i \cdot s_i^k)/(I^e \cdot s_0^k) \tag{2}$$

assuming a target size larger than or equal to the footprint (Wagner et al., 2006). The calibrated values I_i^c refer to the same nominal distance s_0, which can be chosen arbitrary. However, for numerical reasons a value around the mean flying height of the campaign is advisable. The exponent k in Eq. (2) amounts theoretically to the value 2. In practice, however, it might deviate depending on the LIDAR system. In Section 4.3 a procedure for determining an optimal value of k for a given data set is demonstrated. A detailed treatment of the calibration issue can be found in Höfle and Pfeifer (2007). They compare several data driven calibration models with a theoretical model.

3. Segmentation

3.1. Watershed segmentation

A first crude segmentation of the tree crowns is calculated from a CHM by the watershed algorithm according to Vincent and Soille (1991), which is a non-parametric segmentation method. The CHM is derived by subdividing the ROI into a grid having a cell spacing of c_p and N_R cells and filtering the highest 3D points within each cell of size c_p^2. The height z_j of these points is corrected with respect to the ground level z_j^{ground} which is estimated from a given DTM by bilinear interpolation:

$$z_j^{\text{CHM}} = z_j - z_j^{\text{ground}} \quad (j = 1, \ldots, N_R). \tag{3}$$

In the next step, all the highest 3D points $X_j^{\text{T}} = (x_j, y_j, z_j^{\text{CHM}})$ $(j = 1, \ldots, N_R)$ of all N_R cells are robustly interpolated into a grid that has $N_{\text{CHM}} = N_X \times N_Y$ grid points

$$X_{\text{Int}\,k,l}^{\text{CHM}\,\text{T}} = (x_{\text{Int}\,k,l}^{\text{CHM}}, y_{\text{Int}\,k,l}^{\text{CHM}}, z_{\text{Int}\,k,l}^{\text{CHM}})$$
$$(k = 1, \ldots, N_X; l = 1, \ldots, N_Y) \tag{4}$$

and a grid width g_w, where N_X and N_Y are the numbers of grid lines. For this purpose a method called 'gridfit' (D'Errico, 2006) is adopted which smoothens the surface by keeping the surface gradients as small as possible. The balance between interpolation and regularization is determined by the adjustable smoothing factor λ. Both steps are carried out simultaneously in a least squares adjustment using the observation equations

$$\begin{bmatrix} A \\ \lambda B \end{bmatrix} (z_{\text{Int}}^{\text{CHM}})^{\text{T}} = \begin{pmatrix} (z_1^{\text{CHM}}, \ldots, z_{N_R}^{\text{CHM}})^{\text{T}} \\ 0 \end{pmatrix}. \tag{5}$$

The matrix A $(N_R \times N_{\text{CHM}})$ contains the coefficients of the bilinear interpolation and the matrix B $(N_{\text{CHM}} \times N_{\text{CHM}})$ comprises the coefficients of the first partial derivatives of neighbouring cells. The vector $z_{\text{Int}}^{\text{CHM}}$, representing the interpolated posts, has the length N_{CHM}. The result is a smoothed CHM having N_{CHM} equally spaced posts to which the watershed segmentation is applied. Note that the factor λ controls both the smoothing and the regularization in the case of ill posed situations. The local maxima of the watershed segments define possible tree positions $(X_{\text{stem}\,i}^{\text{CHM}}, Y_{\text{stem}\,i}^{\text{CHM}})$ $(i = 1, \ldots, N_{seg})$ (Fig. 2). The heights $Z_{\text{stem}\,i}^{\text{CHM}}$ are derived from the highest reflections in the segments.

This segmentation procedure is mainly controlled by the parameters c_p, g_w and λ for the CHM construction, whose values have been optimized in a sensitivity analysis (Section 4.4.2). In this process, the parameter settings $c_p = 0.5$ m, $g_w = 0.5$ m and $\lambda = 4$ turned out to be optimal. The watershed algorithm itself does not take any parameters. Note that a certain smoothing of the CHM is necessary in order to avoid an oversegmentation. On the other hand, neighbouring trees often merge thereby and form a tree group instead of single trees. Furthermore, smaller trees in the intermediate and lower height level cannot be recognized since they are not visible in the CHM. Fig. 3 shows the CHM of a dense spruce stand, where the black lines represent the positions and heights of the reference trees. Apparently, most of the reference trees are not represented in the CHM.

3.2. Detection of tree stems

The key idea of the stem detection is to separate neighbouring trees forming a tree group and to improve the accuracy of the tree positions calculated from the local maxima of the CHM. The approach is motivated by the observation that in case of full waveform data the stems of trees are very often clearly visible. For example, the centre segment of Fig. 4a contains three trees, which cannot be separated by the watershed segmentation, whereas the stems of at least two of the trees can be clearly seen in Fig. 4b.

The method works in a three-step algorithm.

First, the N_S points X_j^{Seg} $(j = 1, \ldots, N_S)$ within a watershed segment are separated into ground points, stem points and crown

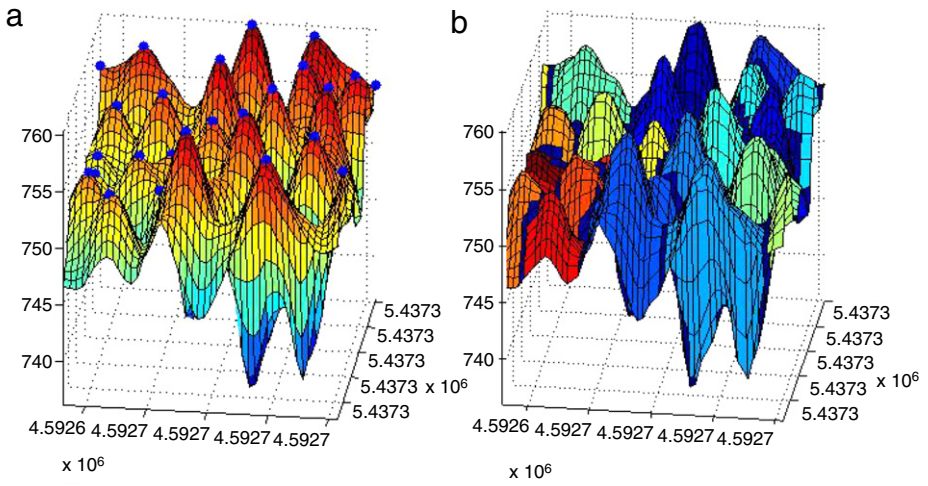

Fig. 2. CHM with (a) local maxima and (b) watershed segments.

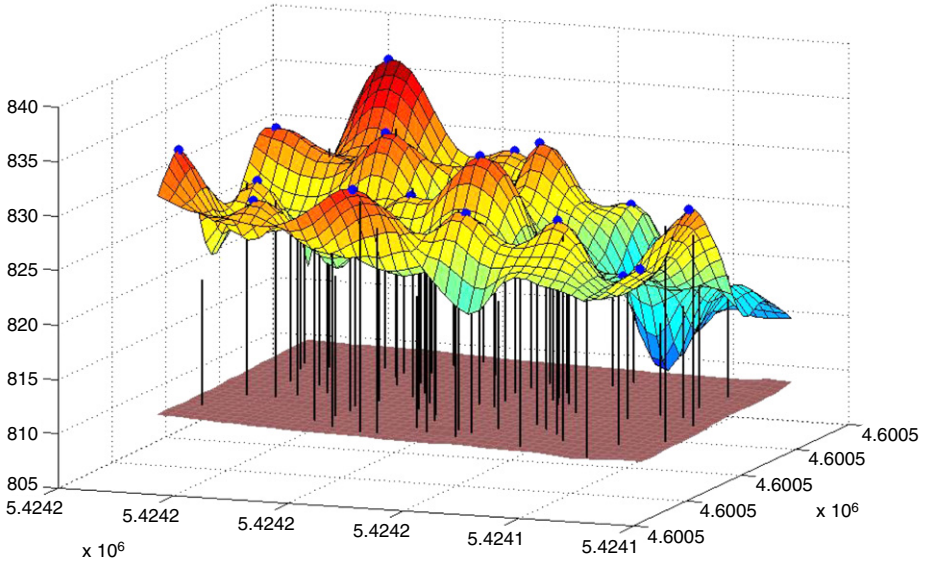

Fig. 3. CHM with local maxima and reference trees (black lines).

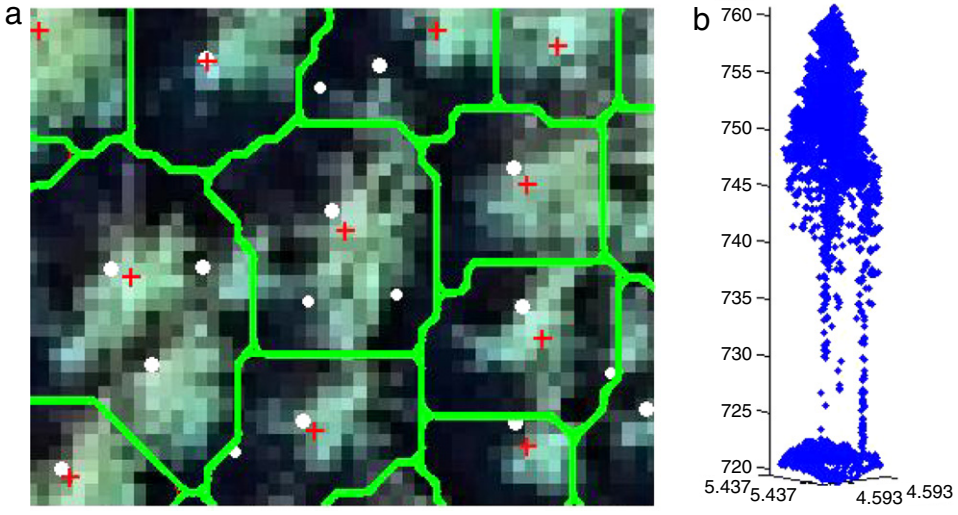

Fig. 4. (a) Orthophoto with watershed segments (green), reference trees (white) and local maxima (red) and (b) laser reflections for the centre segment. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

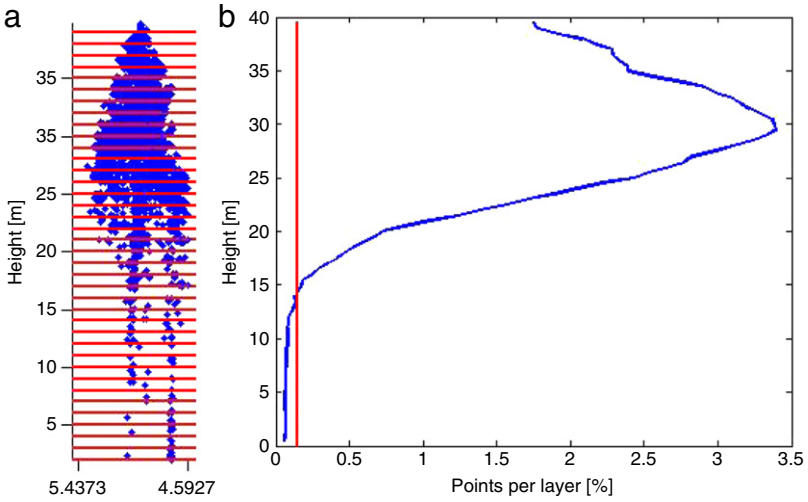

Fig. 5. (a) Tree height layers and (b) determination of the crown base height.

points. Ground points are points within a small height bound (e.g. 1 m) to the DTM. The remaining $N_{S,\text{tree}}$ tree points are divided into stem points and crown points by finding an appropriate crown base height h_{base}. This is achieved by (i) splitting the tree points into l layers with height of 0.5 m (Fig. 5a), (ii) calculating the number of points n_i per layer, (iii) forming the vector $N_p = \{n_i/N_{S,\text{tree}}\}$ $(i = 1, \ldots, l)$, (iv) smoothing N_p with a 3×1 Gaussian filter and, finally, (v) defining h_{base} as the height that corresponds to 0.15% of the total number of tree points (Fig. 5b). The points below h_{base} are the N_{stem} potential stem points and can result from one or several stems. Of course, points from the understorey or from isolated branches may be contained, too.

In the second step these points are clustered according to their horizontal distances, in order to get a first estimation of the number of stems and an assignment of the points to the stems. The following hierarchical clustering scheme is applied after calculating the Euclidian distance matrix $D_{\text{stem}} = \{d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}; i = 1, \ldots, N_{\text{stem}}; j = 1, \ldots, N_{\text{stem}}; i\#j\}$ (Heijden et al., 2004):

1. Assign each point to its own cluster, resulting in N_{stem} clusters.
2. Find the closest pair of clusters and merge them in to one cluster. The number of clusters reduces by one.
3. Compute the distance d between the new clusters and each of the old clusters.
4. Repeat steps 2 and 3 until all items are clustered into a single cluster of size N_{stem} or a predefined number of clusters is achieved.

In this clustering process the distance between two clusters C_i and C_j is defined as the shortest distance from any point in one cluster to any point in the other cluster. The clustering yields a dendrogram which shows at which distance the clusters are grouped together. By defining a minimum distance $d_{\min} = 1.2$ m between the cluster centres the most suitable number of clusters N_{cluster} is selected.

In the third step, the stems are finally found by applying a RANSAC-based 3D line adjustment to all the N_{cluster} clusters and labelling all 3D lines with an incident angle smaller than $\alpha = 7^\circ$ and a minimum number of 3 points as stems g_{stem}. This procedure eliminates clusters that result from the understorey and do not show a vertical main direction. Also, it cleans the cluster point cloud from erroneous points, e.g. from branches. The stem positions $(X^{StDet}_{\text{stem}\,i}, Y^{StDet}_{\text{stem}\,i})$ $(i = 1, \ldots, N_{StDet})$ are calculated as the intersection of the 3D line g_{stem} with the DTM. Finally, the tree

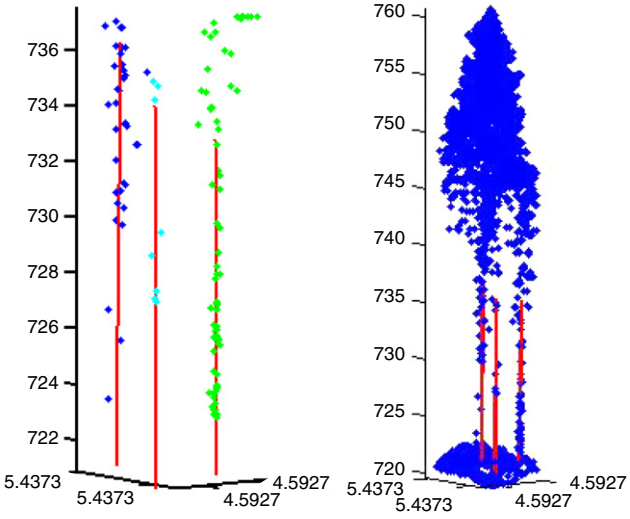

Fig. 6. Stem point clusters and stems reconstructed with RANSAC.

height Z^{StDet}_{stem} is derived from the highest laser point that lies within the cylinder V_{stem}. The cylinder V_{stem} is defined by g_{stem} as the centre line of the cylinder and the radius $R = 1$ m. If there is a gap of more than 3 m above h_{base}, the highest point below the gap is used as tree height, because the points above the gap most probably belong to another tree.

Fig. 6 shows the stem point clusters of the hierarchical clustering and the reconstructed stems for the sample segment in Fig. 4. Three stems can be clearly isolated. A comparison of the stem positions with reference trees (Fig. 7) evidences that all trees within the watershed segment in the centre could be found with a high accuracy.

In general, the approach improves the detection rate of single trees in the intermediate and upper tree layers and the accuracy of the tree positions originally provided by the watershed segmentation, even for segments with only one tree. The stem detection works successfully, if there are enough stem reflections and if the stem area can be reliably separated from the crown points. It fails, of course, when many small trees are located below the taller trees. Another drawback is that the crown points belonging to the original segment are not separated with respect to the detected stems.

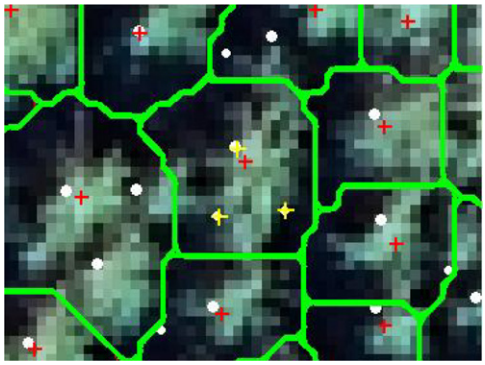

Fig. 7. Orthophoto with watershed segments (green), reference trees (white dots), local maxima (red crosses) and detected stems (yellow crosses). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

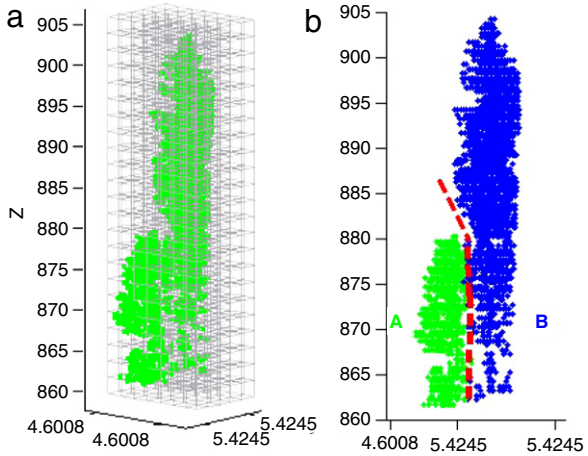

Fig. 9. (a) Subdivision of the ROI into a voxel structure and (b) division into two tree segments.

3.3. Normalized cut segmentation

3.3.1. Approach

In order to tackle the drawbacks of the 2D watershed segmentation and the stem detection we have set up a true 3D segmentation of single trees using the normalized cut method known from image segmentation (Shi and Malik, 2000). Before giving a detailed description of this method, some results of complex situations (Fig. 8) are shown, to demonstrate the potential of the approach. Here normalized cut segmentation works excellently and the watershed segmentation and stem detection approaches fail.

In the first step, the ROI is subdivided into a voxel structure with a voxel spacing of vp and $N_v = N_v^x \times N_v^y \times N_v^z$ voxels (Fig. 9). Within each voxel (l, m, n) $(l = 1, \ldots, N_v^x; m = 1, \ldots, N_v^y; n = 1, \ldots, N_v^z)$ of size vp^3 we collect N_{lmn} reflections $X_i^{\mathrm{T}} = (x_i, y_i, z_i, W_i, I_i)$ $(i = 1, \ldots, N_{lmn})$, where only voxels comprising at least one reflection are used in the segmentation. The voxel space is represented as a graph $G = \{V, E\}$ with V as the voxels representing the nodes and E as the edges formed between every pair of nodes. The similarity between two nodes $\{i, j\} \in V$ is described by the weights w_{ij} which are computed from features associated with the voxels. Basically, the similarity between voxels decreases with increasing distance between two voxels and drops down to zero beyond the threshold r_{XY} in order to keep the graph G at a reasonable size for computational reasons. The goal of normalized cut segmentation is to divide the graph G into disjoint voxel segments A and B (Fig. 9) by maximizing the similarity of the segment members and minimizing the similarity between the segments A and B.

The corresponding cost function is

$$N\mathrm{Cut}(A, B) = \frac{\mathrm{Cut}(A, B)}{\mathrm{Assoc}(A, V)} + \frac{\mathrm{Cut}(A, B)}{\mathrm{Assoc}(B, V)} \tag{6}$$

with $\mathrm{Cut}(A, B) = \sum_{i \in A, j \in B} w_{ij}$ as the total sum of weights between the segments A and B and $\mathrm{Assoc}(A, V) = \sum_{i \in A, j \in V} w_{ij}$ representing the sum of the weights of all edges ending in the segment A. The minimization of $N\mathrm{Cut}(A, B)$ is solved by the corresponding generalized eigenvalue problem

$$(D - W)y = \lambda D y. \tag{7}$$

The $n \times n$ weighting matrix W is representing the weights w_{ij} between all n nodes of the graph G and is inherently symmetric and positive semi-definite. The $n \times n$ degree matrix D is directly derived from W and holds the degree of a node i at the diagonal element $D(i, i) = \sum_j w_{ij}$.

The minimum solution y_1 for (7) corresponds to the second smallest eigenvalue λ_1. Since y_1 is real-valued, but may only have two distinct indicator values (+1,-1) we need to binarize it by introducing a threshold $thres_{\mathrm{bin}}$ into the histogram of y_1. An appropriate value for $thres_{\mathrm{bin}}$ can be found in different ways. Using simply the values 0 or $median(y_1)$ leads already to proper results. The results can be improved by testing several possible values with respect to the resulting value of $NCut$ and choosing the value causing the smallest $NCut$ value. We find the optimal $thres_{\mathrm{bin}}$ value

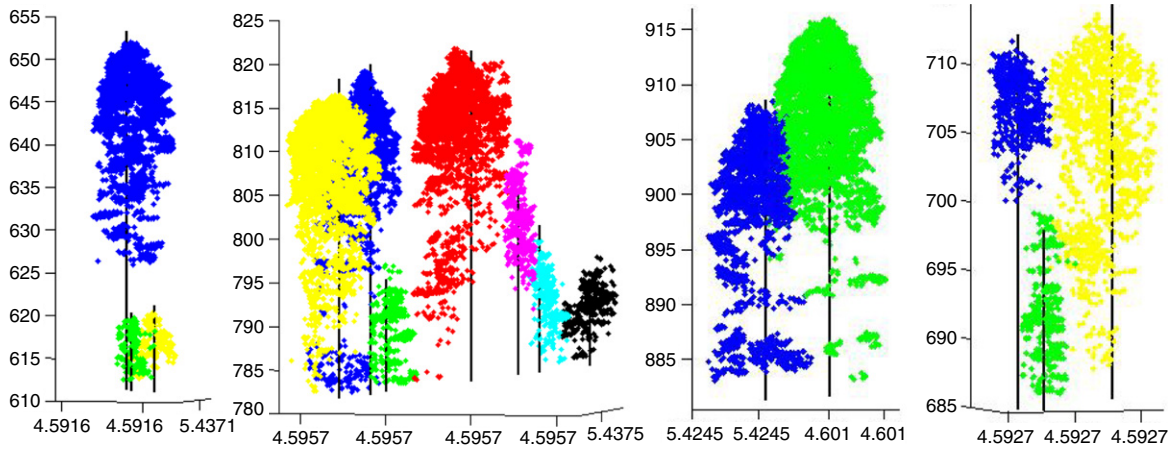

Fig. 8. Excellent examples of normalized cut segmentation.

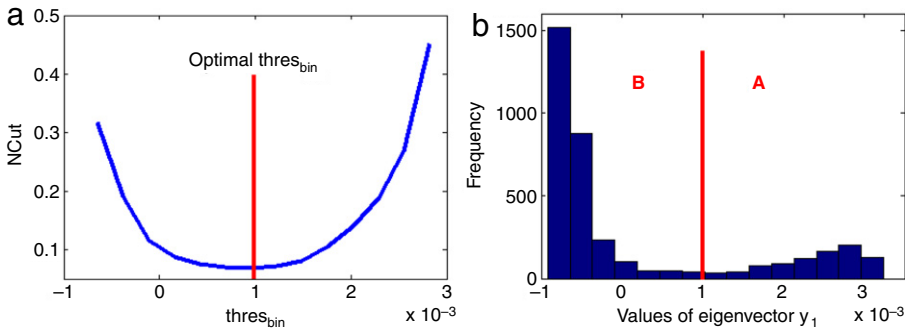

Fig. 10. (a) Finding the optimal $thres_{bin}$ and (b) dividing the histogram of y_1 for the segments in Fig. 9.

by stepwise testing all values within the range of y_1 in a small interval. Thus, the graph G is subdivided into two disjoint segments A and B. For example, Fig. 10 shows the finding of the optimal value for $thres_{bin}$ and the division of the histogram of y_1 for the segments in Fig. 9.

The subdivision of the G into several segments is found in the following hierarchical procedure.

Step 1: Create the graph G by computing W and D for all nodes.
Step 2: Find the solution of the eigenvalue problem (7).
Step 3: Binarize the solution vector y_1 with $thres_{bin}$ and cut the graph G into two new graphs G_1 and G_2.
Step 4: Apply steps 2 and 3 to the graphs G_1 and G_2. Stop if the value for $NCut$ reaches or exceeds the threshold $NCut_{thres}$ or if the number of voxels is smaller than a certain minimal number of voxels.

3.3.2. Similarity function and combined segmentation

One crucial point is the finding of similarity measures reflecting the likelihood that two voxels belong to one tree. In order to set up the weighting matrix W we introduce the function

$$w(i,j) = \begin{cases} e^{-X(i,j)} \times e^{-Z(i,j)} \times e^{-F(i,j)} \times e^{-G(i,j)} & \text{if } (D_{ij}^{XY} < r_{XY}) \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

with

$$X(i,j) = \left(\frac{D_{ij}^{XY}}{\sigma_{xy}}\right)^2, \qquad Z(i,j) = \left(\frac{D_{ij}^{Z}}{\sigma_{z}}\right)^2,$$
$$F(i,j) = \left(\frac{|f_i - f_j|}{\sigma_f}\right)^2, \qquad G(i,j) = \left(\frac{G_{ij}^{\max}}{\sigma_G}\right)^2 \quad (9)$$

that computes the similarities w_{ij} between two voxels i and j within a cylinder of radius r_{XY} around the voxel i. The key idea of function (8) is to multiplicatively combine several factors impacting the similarity.

The components $X(i,j)$ and $Z(i,j)$ are the basic modules and weight the quadratic Euclidian distances between the voxels, where D_{ij}^{XY} is the horizontal and D_{ij}^{Z} the vertical distance. The horizontal and vertical distances are weighted separately to take into consideration the prior knowledge of a typical 3D tree shape.

The component $F(i,j)$ describes the quadratic Euclidian distance between two arbitrary feature vectors f_i and f_j derived from the data points (= reflections) in the voxels or in certain surroundings of the voxels. We use the features $f_i = \{I_{mean}, W_{mean}\}$ for each voxel i which are calculated from the N reflections in a given space segment S of size 2 m × 2 m × 6 m as

$$I_{mean} = \frac{1}{N}\sum_{j\in S} I_j \qquad W_{mean} = \frac{1}{N}\sum_{j\in S} W_j. \quad (10)$$

The features in (10) represent the mean intensity and the mean pulse width for each voxel. The computation in (10) is equivalent to a low pass filtering of the raw reflections with a box filter.

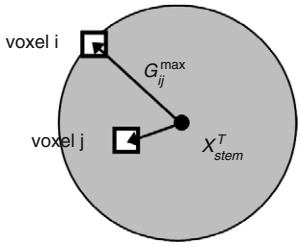

Fig. 11. The dependence of weighting of voxels i and j on their distance to X_{stem}^{T}.

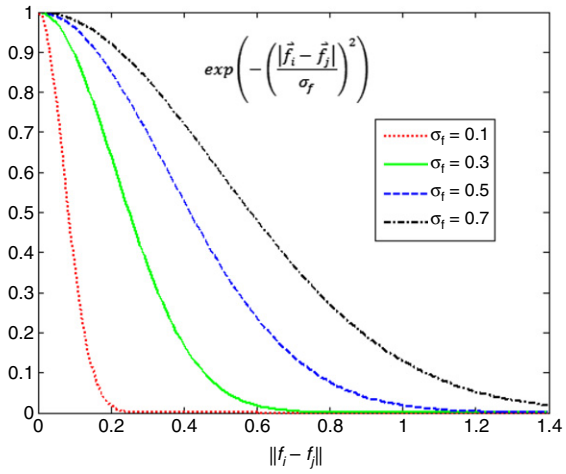

Fig. 12. Influence of σ_f on the values of $e^{-F(i,j)}$.

Finally, the fraction $G(i,j)$ renders it possible to introduce prior knowledge about the stem position, e.g. from the watershed segmentation or the stem detection. Two voxels are weighted depending on their maximum horizontal distance G_{ij}^{max} to a known stem position X_{stem}^{T} (Fig. 11). Hence, it is modelled that voxels nearby X_{stem}^{T} belong most probably to one tree.

The parameters σ_{xy}, σ_z, σ_f and σ_G control the sensitivity of the impact factors $X(i,j)$, $Z(i,j)$, $F(i,j)$ and $G(i,j)$ in the numerator of Eq. (9). For example, Fig. 12 illustrates the change of $e^{-F(i,j)}$ for different values of σ_f. The higher the values of σ_f are, the slower is the loss of similarity for an increasing feature difference.

It is also noteworthy that the two latter terms in (8) are optional and the approach works just having the coordinates of the reflections. If additional features or results from other segmentation methods are available, this can be used advantageously.

3.3.3. Postprocessing of results

The results of normalized cut segmentation are cleaned in the following steps. Firstly, segments with a number of voxels lower

Table 1
Characteristics of sample plots.

Plot name	21	22	55	56	57	58	59	60	64	65	74	81	91	92	93	94	95	96
Age (a)	160	160	240	170	100	85	40	110	100	100	85	70	110	110	110	110	110	110
Size (ha)	0.20	0.20	0.15	0.23	0.10	0.10	0.10	0.10	0.12	0.12	0.30	0.30	0.36	0.25	0.28	0.29	0.25	0.30
Altitude (m)	860	885	610	640	765	710	810	890	835	875	720	690	764	767	766	768	750	781
N/ha	500	540	830	340	450	440	2150	380	430	810	700	610	260	170	240	250	240	200
N lower layer	37	19	77	31	0	10	76	8	13	53	11	29	31	13	7	15	6	30
N interm. layer	14	60	21	19	4	4	85	22	4	26	33	59	11	3	2	4	0	3
N upper layer	48	29	20	27	41	30	54	27	35	35	165	96	54	27	59	54	53	26
Deciduous (%)	66	79	5	10	0	14	1	100	87	96	29	100	75	100	66	97	10	86

than a certain threshold are cancelled. The threshold depends on the height of the segment and is 30 for segments below 12 m and 60 for segments higher than 12 m. Secondly, we subdivide each segment into layers of height 2.0 m and search for layers containing only empty voxels (= no reflections!) at an absolute segment height of more then 10 m. All the voxels above the empty layer are ignored.

3.3.4. Control parameters

Normalized cut segmentation is controlled by several parameters whose values have been optimized in a sensitivity analysis (Section 4.4.2). Firstly, the size of the voxels *vp* was set equal to 0.5 m. The threshold r_{XY} defining the maximum horizontal distance for calculating $w(i,j)$ was set to 4.5 m. The most important parameter $NCut_{\text{thres}}$, which controls the subdivision (= cut) of a graph G, was set equal to 0.16. Moreover, a graph G is not subdivided any more if the number of voxels of the graph undershoots the limit of 40 voxels. We used the empirical values $\sigma_{xy} = 1.35$ m, $\sigma_z = 11.0$ m, $\sigma_f = 0.5$ m and $\sigma_G = 3.5$ m to control the influence of the impact factors $X(i,j)$, $Z(i,j)$, $F(i,j)$ and $G(i,j)$. The value for σ_z is larger than σ_{xy} assuming that the tree height is larger than the tree crown diameter.

4. Experiments

4.1. Field data

Experiments were conducted in the Bavarian Forest National Park which is located in South-Eastern Germany along the border with the Czech Republic (49° 3′ 19″ N, 13° 12′ 9″ E). There are four major test sites of size between 591 ha and 954 ha containing alpine spruce forests, mixed mountain forests and spruce forests as the three major forest types. 18 sample plots with an area size between 1000 m^2 and 3600 m^2 were selected in the mixed mountain forests dominated by Norway spruce (*Picea abies*) and European beech (*Fagus sylvatica*). Some fir trees (*Abies alba*), sycamore maples (*Acer pseudoplatanus*), Norway maples (*Acer platanoides*) and lime trees (*Tilia europaea*) also occur in the sample plots. The height above sea level varies between 610 m and 890 m. Reference data for all trees with diameter at breast height (DBH) larger than 10 cm were collected for 688 Norway spruce, 812 European beech, 70 fir, 71 sycamore maple, 21 Norway maple and 2 lime trees. Several tree parameters like the DBH, total tree height, stem position and tree species were measured and determined with the help of GPS, tacheometry and the 'Vertex' III system. Furthermore, the trees are subdivided into three layers with respect to the top height h_{top} of the plot, where h_{top} is defined as the average height of the 100 highest trees per ha (Heurich, 2006). The lower layer contains all trees below 50% of h_{top}, the intermediate layer refers to all trees between 50% and 80% of h_{top}, and finally, the upper layer contains the rest of the trees. Naturally, the reference data were updated for the individual flying dates of the LIDAR data. Table 1 summarizes the characteristics of the individual sample plots. The reference trees are plotted in Figs. 3, 8, 16, 17 and 20–22 as black vertical lines.

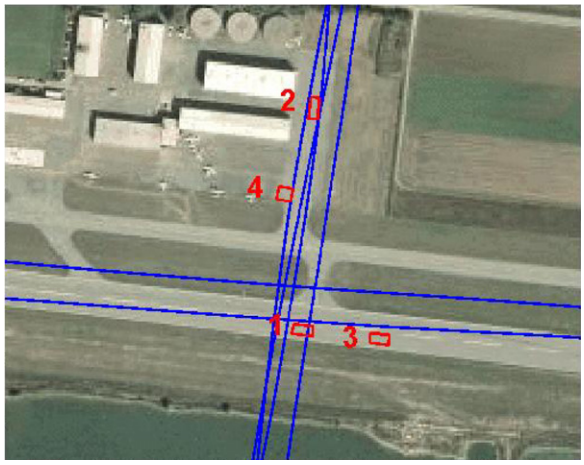

Fig. 13. Calibration airfield with tracks (blue) and calibration areas (red). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.2. LIDAR data

LIDAR data of several ALS campaigns are available for the test sites. First/last pulse data were recorded by TopoSys with the Falcon II system. Full waveform data were collected by Milan Flug GmbH using the Riegl LMS-Q560 scanner. Table 2 contains details on the point density, leaf-on and leaf-off conditions during the flights, the footprint size and the flying height above average terrain (HAAT). The term point density refers to the nominal value influenced by the pulse repetition frequency, flying height, flying speed and strip overlap. These data sets allow the comparison of conventional and full waveform systems, which were flown in the same area. However, data set IV is only available for 12 reference plots, referred to as 'Area E'. This has to be considered when comparing results of other data sets with this data set.

Furthermore, a DTM with a grid size of 1 m and an absolute accuracy of 25 cm was available for all the test sites. It was generated from LIDAR data which had been acquired in 2003 (Heurich, 2006).

4.3. Calibration

The calibration of the Riegl full waveform system was determined from special calibration flights performed over an airfield. Several tracks were flown at different flying heights (200 m and 400 m) along and across the airfield (Fig. 13). The mean intensity I_i, corrected with respect to the emitted intensity I^e, and the mean run length s_i were calculated in four homogeneous areas (122 m^2–133 m^2) for each track i. In order to reduce the impact of the scan angle we only used measurements whose scan angles were smaller than 10°.

According to Eq. (2), the best coefficient k was estimated from all possible observation equations

$$I_i \cdot s_i^k = I_j \cdot s_j^k \tag{11}$$

Table 2
Different ALS campaigns.

Time of flight	September 2002	May 2006	May 2007	May 2007
Data set	I	II	III	IV
Foliage	Leaf-on	Leaf-off	Leaf-on	Leaf-on
Scanner	TopoSys Falcon II	Riegl LMS-Q560	Riegl LMS-Q560	Riegl LMS-Q560
Pts/m^2	10	25	25	10
HAAT (m)	850	400	400	500
Footprint (cm)	85	20	20	25
Reference plots	All	All	All	Area E

Table 3
Estimation of calibration parameter k.

	Flight 2006 (data set II)	Flight 2007 (data sets III and IV)
Calibrated parameter k	1.902	1.736

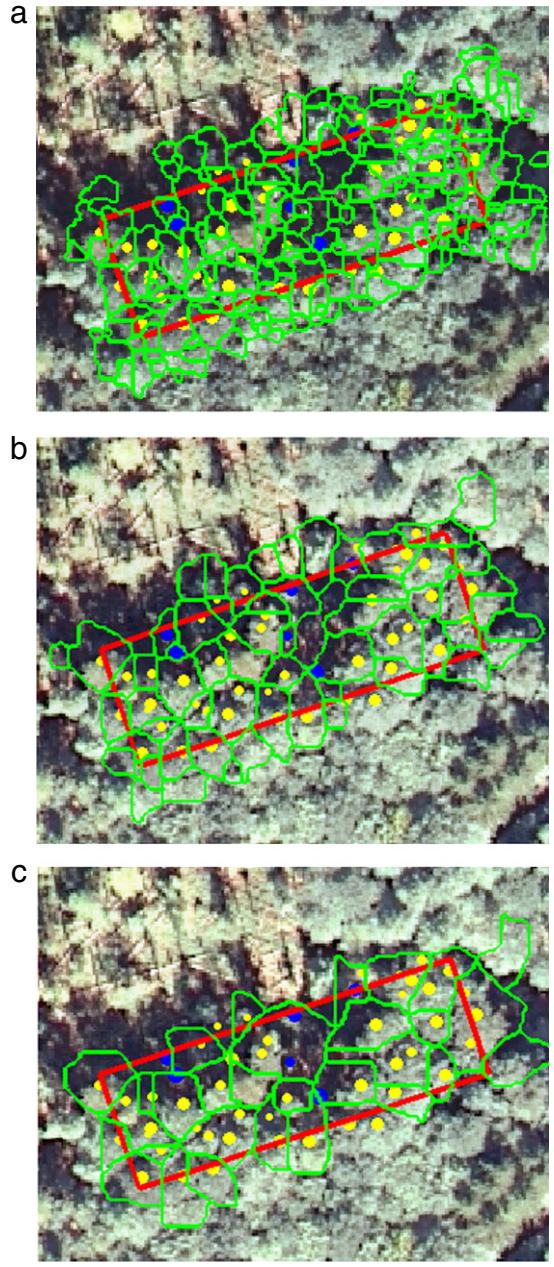

Fig. 14. Watershed segments for (a) low, (b) intermediate and (c) high smoothing of the CHM.

Table 4
Options of normalized cut segmentation.

Option	Parameterization of weighting function
$NCut_C$	Coordinates
$NCut_{CF}$	Coordinates, Features
$NCut_{CPm}$	Coordinates, Position of maxima from CHM
$NCut_{CPms}$	Coordinates, Positions of maxima from CHM and stem detection
$NCut_{CFPms}$	Coordinates, Features, Positions of maxima from CHM and stem detection

which can be formulated for two tracks i and j flown at different heights. Table 3 shows the results obtained for the two flights of data sets II, III and IV. It is remarkable that the mean values for k are smaller than the value of 2. Thus, the attenuation of the signal is smaller, as theoretically expected. One possible reason is a non-linearity in the sensor system which amplifies low intensities more than high intensities. However, no detailed information from the hardware supplier is available on that subject. The calibrated values for k were used in the following analysis.

4.4. Segmentation

4.4.1. Procedure and evaluation

First, watershed segmentation was applied to all the reference plots in a batch procedure without any manual interaction. The results are the stem positions and heights $(X^{CHM}_{stem\,i}, Y^{CHM}_{stem\,i}, Z^{CHM}_{stem\,i})$ $(i = 1, \ldots, N_{seg})$ from the local maxima and the segment borders from the watershed algorithm.

Second, stem detection was conducted for each of the watershed segments. Depending on the number of detected stems three cases are distinguished:

(1) No stem is found. The position and height $(X^{CHM}_{stem}, Y^{CHM}_{stem}, Z^{CHM}_{stem})$ from the watershed segmentation remains unchanged.
(2) One stem is found. The position and height $(X^{CHM}_{stem}, Y^{CHM}_{stem}, Z^{CHM}_{stem})$ are replaced by the more exact values $(X^{StDet}_{stem}, Y^{StDet}_{stem}, Z^{StDet}_{stem})$.
(3) Several stems are found. The position and height $(X^{CHM}_{stem}, Y^{CHM}_{stem}, Z^{CHM}_{stem})$ are replaced by $(X^{StDet}_{stem\,i}, Y^{StDet}_{stem\,i}, Z^{StDet}_{stem\,i})$ $(i = 1, \ldots, N_{StDet})$. Thus, additional tree positions are detected, whereas the crown points belonging to the original segment are not separated with respect to the detected stems.

Third, normalized cut segmentation was applied to all the reference plots. As described, this segmentation method does not only use voxel coordinates but also optionally features derived from the intensity and pulse width. Furthermore, the results from watershed segmentation and stem detection can be utilized. Thus there are multiple possibilities to parameterize the weighting function (8), which describes the similarity between the voxels. Table 4 summarizes the options that were tested.

The accuracy and reliability of the presented methods are evaluated in the following way. The tree positions from the segmentation are compared with reference trees if (i) the distance to the reference tree is smaller than 60% of the mean tree distance within the plot and (ii) the height difference between h_{tree} and the height of the reference tree is smaller than 15% of the top

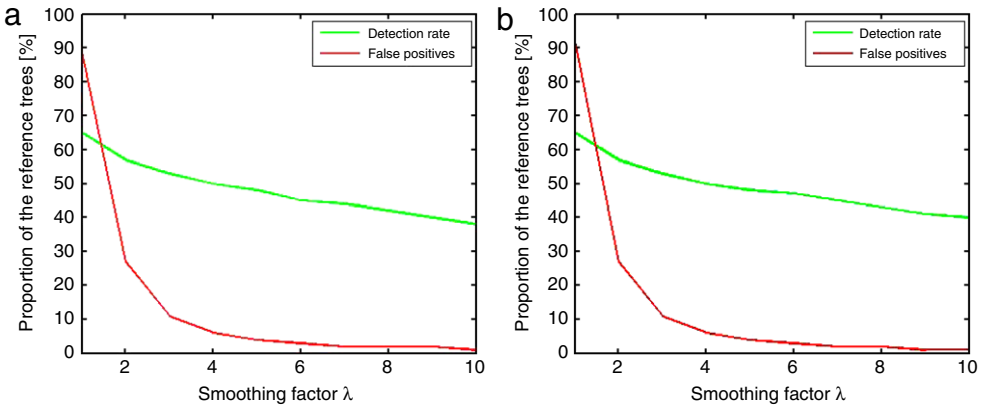

Fig. 15. Influence of smoothing factor λ on detection rate and false positives for (a) data set II and (b) data set III.

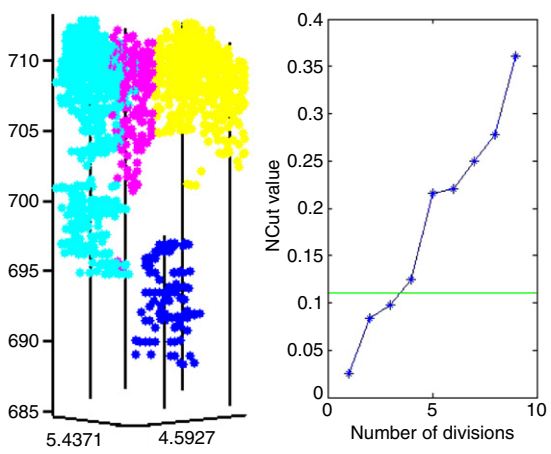

Fig. 16. Tree segments resulting from an *NCut* value of 0.11. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

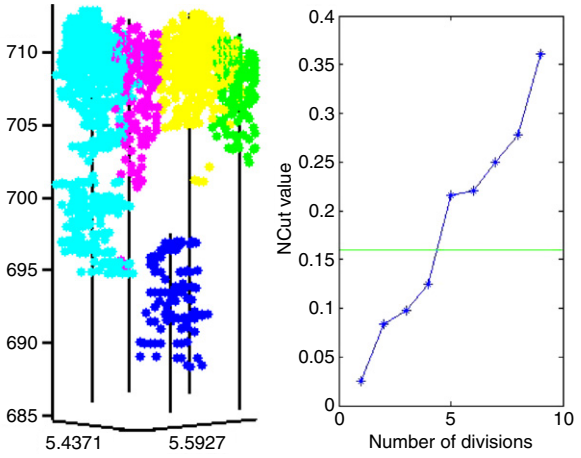

Fig. 17. Tree segments resulting from an *NCut* value of 0.16. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

height h_{top} of the plot. If a reference tree is assigned to more than one tree position, the tree position with the shortest distance to the reference tree is selected. Reference trees that are linked to one tree position are so-called 'detected trees' and reference trees without any link to a tree position are treated as 'non-detected' trees. Finally, a tree position without a link to a reference tree is referred to as a 'false positive' tree.

4.4.2. Sensitivity analysis

Obviously, the proposed segmentation methods require proper values for many control parameters. Tests were conducted to find these values and to test their sensitivity. Concerning the 2D segmentation this is exemplarily demonstrated for the smoothing factor λ. Fig. 14 shows the segmentation results for one reference plot using different values of λ.

The green polygons represent the watershed segments and the yellow and blue dots show the reference positions of beeches and spruces. Apparently, low smoothing leads to an oversegmentation whereas high smoothing merges neighbouring trees. The influence of λ on the detection rate and the number of false positives for all reference plots of data sets II and III is presented in Fig. 15.

The number of false positives significantly increases for values of λ smaller than 4. Thus, the value 4 was chosen for the control parameter λ.

As a further example we demonstrate the influence of the $NCut_{thres}$ value that controls the subdivision of the segments in step 4 of the normalized cut segmentation procedure. The larger

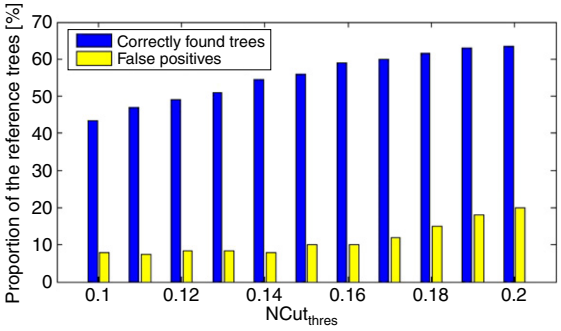

Fig. 18. Influence of control parameter $NCut_{thres}$ on detection rate and false positives.

the value is set the higher is the similarity between the resulting segments. Figs. 16 and 17 show in an example how the value 0.16 splits the upper right (yellow) segment into two correct tree segments.

Note that the higher the value of $NCut_{thres}$ is set the better is the detection rate of single trees. However, this also leads normally to a larger number of false positives. This is nicely depicted in Fig. 18, where the sensitivity of the segmentation process in terms of correctly found trees and false positives is shown by using the LIDAR data set II. The value of 0.16 was found as the best compromise and – more importantly – was robust for all the reference areas. Tests with the other data sets also showed that a value around 0.16 is optimal.

Table 5
Results of segmentation methods with data set II (high density full waveform data, leaf-off).

Method	Detected trees per height layer (%)				False positives (%)
	Lower	Intermediate	Upper	Total	
W	5	21	77	48	4
$W+S$	7	27	82	52	5
$NCut_C$	18	32	77	53	6
$NCut_{CF}$	19	35	77	54	6
$NCut_{CPm}$	20	36	83	57	9
$NCut_{CPms}$	20	37	86	59	9
$NCut_{CFPms}$	21	38	87	60	9

Table 6
Accuracy of the tree position without and with stem detection for data set II (high density full waveform data, leaf-off).

	W (m)	$W+S$ (m)
Mean positioning error, coniferous	0.83	0.68
Mean positioning error, deciduous	1.49	1.12
Total mean positioning error	1.17	0.91

4.4.3. Results of segmentation methods

4.4.3.1. Benefit of stem detection and normalized cut segmentation. In the first instance, we compare the results of the different segmentation methods and options achieved with data set II (Table 5).

The watershed-based segmentation (W) leads to an overall detection rate of about 48%. As expected, the detection rate is rather poor in the lower forest layer since most of these trees are covered by taller trees. A combination of the watershed segmentation and the stem detection ($W+S$) works successfully in the intermediate and upper layers and improves the overall detection rate by 4%. However, the improvement achieved in the lower layer is low since (i) laser hits at the stems of small trees happen rarely, (ii) the crown base height h_{base} is inaccurate for trees beneath taller trees, and (iii) smaller trees in particular often have an h_{base} which is close to the ground. The improved overall detection rate is only one advantage of the stem detection method. A further advantage is the higher positional accuracy of the tree locations (X_{stem}^{StDet}, Y_{stem}^{StDet}), which is in particular useful when the stem positions are used as prior knowledge for normalized cut segmentation. Table 6 shows the mean positioning errors with and without stem detection for all sample plots.

The mean positioning error of coniferous trees gets better by 18%, which corresponds to 15 cm. The improvement for the deciduous trees is higher: 25% (=37 cm). All in all, the overall improvement of the tree positions amounts to 22% (=26 cm).

3D segmentation ($NCut$) generally increases the detection rate considerably in the lower and intermediate layers, in the range 11%–17% for all options. The detection rate in the upper layer does not improve if only voxel coordinates and features are used as similarity measures. If the tree positions – calculated from the local maxima of the CHM and by stem detection – are used, the detection rate in the upper and intermediate layers is significantly improved, by up to 10%. If features are used, the gain is small (by 1%) but always significant. In summary, 3D segmentation and stem detection increase the overall detection rate by 12%. Most interestingly, the improvement is more evident in the lower and intermediate levels: it is about 16%. This is remarkable, and shows that the new 3D segmentation technique can successfully detect smaller trees below the CHM. The high spatial point density of the full waveform data, which practically contain all relevant reflections of the laser beam, turns out to be the key factor for segmenting in three dimensions not only the dominant trees but also the dominated smaller trees in the lower and intermediate layers. Fig. 19 illustrates the improvement of the detection rate graphically, but also shows that there are still many undetected smaller trees.

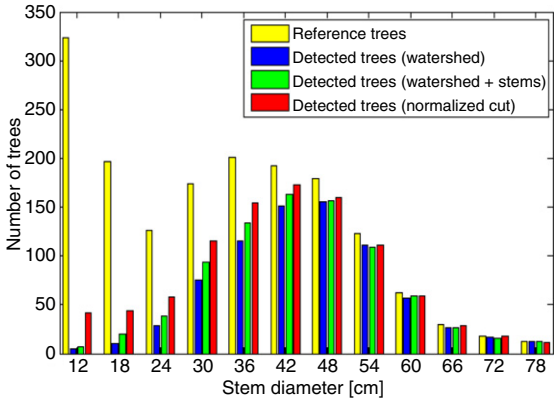

Fig. 19. Comparison of detected trees and reference trees for different stem diameters using data set II.

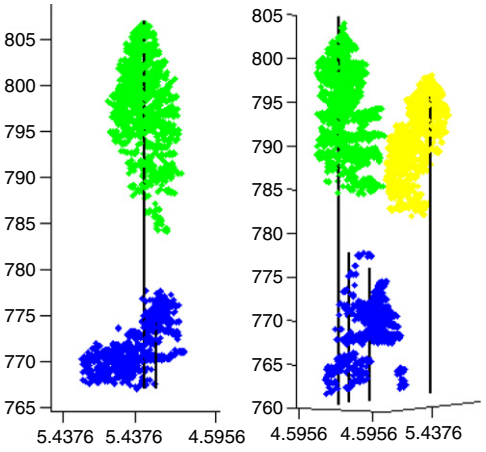

Fig. 20. Two examples where several small trees are merged into one segment (blue). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Possible reasons, therefore, are shown in Fig. 20, where several smaller trees are merged into one segment. These trees remain statistically undetected since the positions and heights of the segmented trees do not correspond with the reference trees within the assumed error tolerances. However, as the Fig. 20 clearly shows, the 3D structure of the forest area is fully captured and separated into dominant trees and the group of small trees of the understorey. Thus, the segments of dominant trees are mostly not affected by trees from the understorey. Subsequent analyses of timber volume will be more precise than they would be with 2D segments calculated from the watershed segmentation method.

Another important aspect is that the increased detection rate of normalized cut segmentation reduces the reliability of the segmentation process by a factor of 2 in terms of false positives. The reasons for false tree segments are manifold. The left plot in Fig. 21 shows that the segmentation wrongly assigns branches of the tall tree to the smaller tree beneath this tree. Thus, the smaller segment is completely inaccurate and the segment of the tall tree is falsified. The right plot shows two broad trees whose crowns are partly merged into a third segment for which a reference tree does not exist. An increase of the control parameter σ_{xy} solves this problem. However, the drawback of this parameter change is that smaller trees are merged into one tree, hence increasing the number of non-detected trees considerably for all reference plots. Thus, a change of σ_{xy} is commonly not reasonable.

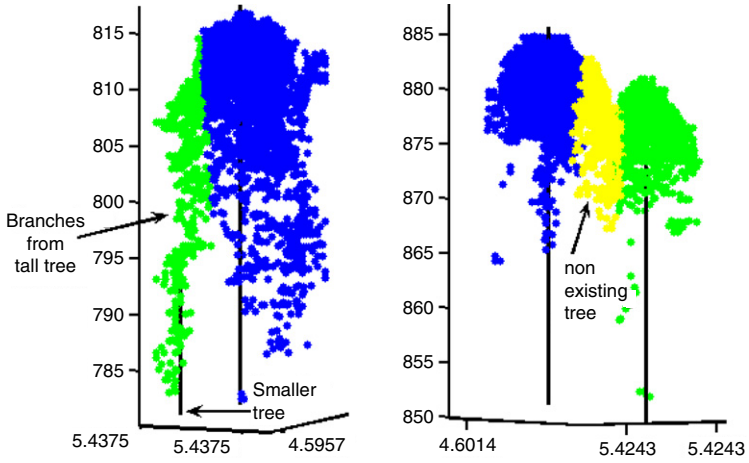

Fig. 21. Examples for false segments.

Table 7
Results of segmentation methods with data set III (high density full waveform data, leaf-on).

Method	Detected trees per height layer (%)				False positives (%)
	Lower	Intermediate	Upper	Total	
W	5	20	79	48	4
$W+S$	6	27	83	52	6
$NCut_{C}$	13	29	73	49	7
$NCut_{CF}$	16	29	73	50	7
$NCut_{CPm}$	15	29	82	54	9
$NCut_{CPms}$	15	31	86	57	10
$NCut_{CFPms}$	17	32	86	58	10

Table 8
Results of segmentation methods with data set III (high density full waveform data, leaf-on) by evaluating only area E.

Method	Detected trees per height layer (%)				False positives (%)
	Lower	Intermediate	Upper	Total	
W	5	20	82	55	5
$W+S$	6	29	87	60	7
$NCut_{CFPms}$	24	35	88	66	11

Table 9
Results of segmentation methods with data set IV (low density full waveform data, leaf-on, only area E).

Method	Detected trees per height layer (%)				False positives (%)
	Lower	Intermediate	Upper	Total	
W	6	21	84	57	6
$W+S$	7	22	86	58	7
$NCut_{CFPms}$	26	33	87	65	11

4.4.3.2. Leaf-on versus leaf-off. The results given in Table 5 for leaf-off conditions can also be compared with full waveform data captured in the same area and with the same point density in leaf-on condition (data set III). Table 7 shows that watershed segmentation and stem detection do not change in all layers, both in detection rate and in reliability. Apparently, the reconstructed surface of the CHM is identical in both foliage conditions, and the crown shape of deciduous trees is well represented even in a leaf-off situation. As expected, the detection rate deteriorates in the case of normalized cut segmentation in the lower and intermediate layer by roughly 5% due to the reduced penetration rate of the laser beam, which in turn causes a worse spatial distribution of the reflections. Finally, the number of false positives does not change significantly for normalized cut segmentation.

4.4.3.3. Impact of point density. If we restrict data set III to area E and compare it with data set IV, the impact of the nominal point density on the segmentation methods can be shown. Tables 8 and 9 demonstrate that the detection rate and false positives are practically the same for both point densities. Obviously, although the number of penetrating laser beams is significantly reduced, the most relevant tree structures are still detected by reflections.

Table 10
Results of segmentation methods with data set I (first/last pulse data, leaf-on) by evaluating only area E.

Method	Detected trees per height layer (%)				False positives (%)
	Lower	Intermediate	Upper	Total	
W	2	12	80	52	5
$W+S$	3	13	80	52	6
$NCut_{CPms}$	15	27	77	55	13

4.4.3.4. First/last pulse versus full waveform. Finally, we compare the segmentation methods with respect to conventional first/last pulse data (data set I; Table 10) and full waveform data that have the same nominal point density (data set IV; Table 9). The foliage condition is leaf-on in both cases.

The total detection rate of the 2D watershed-based segmentation is better by 5% for the full waveform data. The number of false positives is basically the same. The main reason for this is that the full waveform data represent the tree shape more precisely since the waveform decomposition basically detects each reflection, even weak reflections and reflections resulting from adjacent targets. This leads to an effective point density, which is higher by a factor of 2–3 compared to first/last pulse data.

In contrast, first/last pulse systems ignore most of these reflections due to the inherent detection method and the dead zone after the first reflection, which amounts at least to one complete pulse duration (e.g. 1.5 m for the TopoSys Falcon II system). For example, Fig. 22a and b show two neighbouring spruce trees captured both with first/last pulse data and with full waveform data.

In Fig. 23b, a typical waveform is depicted that results from two adjacent targets (black points in Fig. 23a) with a distance of 1 m. Obviously, the tree shape is better detected by the full waveform data.

If we focus on normalized cut segmentation in Tables 9 and 10, the benefit of full waveform data becomes clearer. The total detection rate amounts to 65%, which is 10% better than with first/last pulse data. What is remarkable is the fact that normalized cut segmentation increases the detection rate in the lower and intermediate layers even for the first/last pulse data. Most importantly, the combination of full waveform with normalized cut segmentation is by more than 20% better than a conventional 2D segmentation using first/last pulse data.

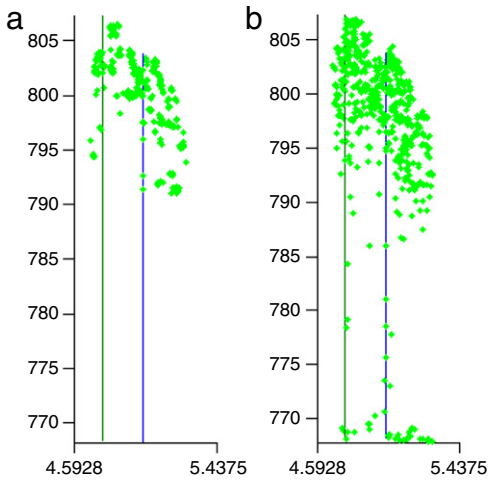

Fig. 22. Neighbouring spruces captured (a) with first/last pulse data and (b) with full waveform data.

5. Discussion

This paper addresses three methods for single tree detection from airborne LIDAR data. The results of the watershed segmentation are similar to the results of Heurich (2006) for the first/last pulse data of data set I. This comparison proves that our 2D segmentation approach is on a par with the approach of Persson et al. (2002). The presented approach of detecting tree stems goes one step further than existing methods by using additional information inside the tree. It leads to an improvement of the detection rate of single trees in the intermediate and upper forest layer. This refinement of the detection rate could be expected since (i) in many cases neighbouring trees do not appear as two clear maxima in the raw data and (ii) the smoothing of the CHM blurs the maxima. The second advantage of this method is that the position of detected trees is significantly improved. This is also not very surprising since the intersection of the detected tree stem with the DTM must be more precise than the tree position derived from the CHM maximum. Thirdly, the stem detection checks the hypothesis of stem positions which have been derived from the CHM. The restrictions of the approach are that only trees in the upper and intermediate forest layer can be additionally detected. It fails in the lower layer where stem hits are rare and stems points cannot be clearly clustered. Also, the tree height still depends on the highest point found in the raw data contained in the stem cylinder V_{stem}. Thus, in cases where the tree belonging to the detected stem is covered by a taller tree and the gap between the trees is small, the derived tree height h_{tree} can be erroneous. Stem detection is not a real segmentation method because the crown points are not assigned to the stems, but the results can be used advantageously by normalized cut segmentation.

Normalized cut segmentation represents substantial progress in the single tree approach and constitutes a natural hierarchical segmentation based on an energy minimization. Thereby, a global minimum is approximated. The 3D segments allow a direct computation of the crown volume and, hence, the timber volume is expected to be estimated more precisely. In comparison to other methods that tackle tree segmentation in three dimensions it is important to mention that normalized cut segmentation needs no seed points for initialization and works fully in three dimensions. For instance, Morsdorf et al. (2003) use the local maxima of the CHM to initialize a *k*-means 3D clustering of LIDAR data points. The approach from Wang et al. (2008) firstly splits the tree area into layers and finally connects the 2D tree crowns of each layer to a 3D tree model. Apparently, the tree extraction depends on the correct finding of proper tree tops. The mathematical formulation makes the approach presented very flexible in using various information derived from LIDAR data. At the moment, the information used is split up into a geometrical part, a feature-based part and a part which uses prior knowledge about probable stem positions. Note that only the geometrical part is mandatory. Thus, our approach also works basically with conventional first/last pulse data and is not dependent on full waveform data or the stem positions. Also, the list of features can be extended arbitrarily.

In general, the experiments with the different data sets show that a combination of normalized cut segmentation with watershed segmentation and stem detection methods always provides the best results. A comparison of the different foliage conditions demonstrates a higher detection rate for the leaf-off data set mainly in the lower and intermediate layers because of the higher penetration of the deciduous trees in the leaf-off situation. Thus, the leaf-off situation seems to be the more appropriate flying time to segment trees in three dimensions, at least for mixed mountain forests that are scanned with a high point density. Note that the leaf-off condition is also advantageous for deriving a DTM and for classifying coniferous and deciduous trees (Reitberger et al., 2008).

The experiments show that a nominal point density higher than 10 pts/m^2 does not improve the detection rate considerably. However, it remains to be seen whether a higher density would be advantageous for estimating other parameters, such as timber volume. In general, the use of full waveform data is clearly superior to the use of first/last pulse data. All segmentation approaches shown in this paper work better with full waveform data. Most notably, only normalized cut segmentation can take

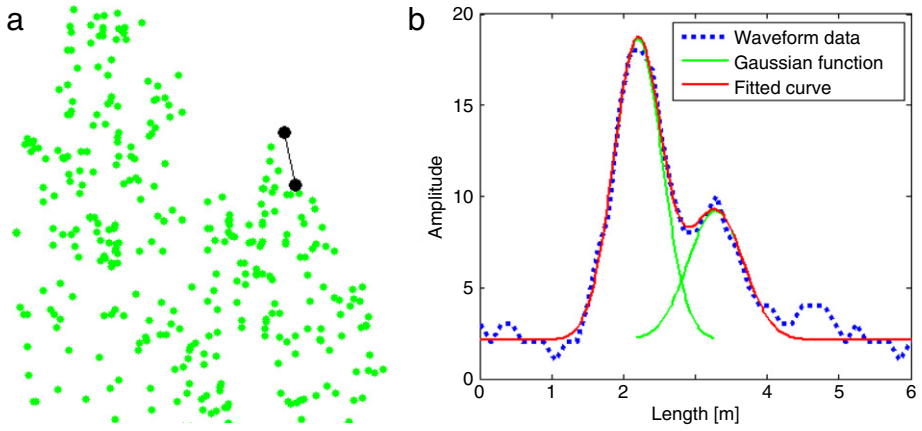

Fig. 23. (a) Detailed view on crown points of Fig. 22b and (b) waveform of the two marked points.

advantage of the high spatial point density that the full waveform technique provides. In summary, the significant improvement of the detection rate – especially apparent in the lower and intermediate layers – is influenced both by the full waveform data and the new normalized cut segmentation.

Finally, further improvements are possible by a sophisticated post-processing of the segmentation results using a priori knowledge about trees. For example, segments could not be accepted if their size exceeds tolerances about the tree shape. Also, segments could be tested with respect to closeness. Although the new 3D segmentation is not capable of separating all single trees in the lower layer, it provides at least important information about such forest structures that can be used for further statistical analysis and, hence, can reduce costly calibration procedures. Moreover, the entire segmentation approach could be organized in a hierarchical way. For example, if the segmented tree geometry does not appear to be feasible the *NCut* parameter can be changed appropriately in order to adapt to a reasonable tree geometry.

6. Conclusions

We have shown in this paper how several segmentation methods work with conventional first/last pulse data and full waveform data. The combination of a new normalized cut segmentation with the tree positions derived from a watershed segmentation method or a stem detection method leads to a significant improvement in the detection rate, especially in the case of full waveform data. In order to make further progress, the behaviour of the laser beam in the tree structure must be analysed in more detail. A better understanding of the correlation between the tree structure and its representation in the LIDAR data might help to identify salient features, which are useful for segmentation and tree species classification purposes. A comparison of airborne and densely captured terrestrial LIDAR data may be helpful in this context.

Acknowledgements

We thank Dr. Marco Heurich and the Administration of the Bavarian Forest National Park for their productive contributions and for giving us the opportunity to use their remote sensing test sites and data set I. This research has been funded by the German Department of Education and Research (BMBF) under contract number 17311B04.

References

Brandtberg, T., 2007. Classifying individual tree species under leaf-off and leaf-on conditions using airborne lidar. ISPRS Journal of Photogrammetry and Remote Sensing 61 (5), 325–340.

D'Errico, J., 2006. Surface fitting using gridfit. http://www.mathworks.com/matlabcentral/fileexchange (accessed on 10.01.09).

Heijden, F. van der, Duin, R.P.W., de Ridder, D., Tax, D.M.J., 2004. Classification, Parameter Estimation and State Estimation — An Engineering Approach using MATLAB. John Wiley & Sons Ltd., The Attrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England.

Heurich, M., 2006. Evaluierung und Entwicklung von Methoden zur automatisierten Erfassung von Waldstrukturen aus Daten flugzeuggetragener Fernerkundungssensoren. Forstlicher Forschungsbericht München, Nr. 202. ISBN: 3-933506-33-6, (accessed on 18.02.07) http://mediatum2.ub.tum.de/.

Höfle, B., Pfeifer, N., 2007. Correction of laser scanning intensity data: Data and model-driven approaches. ISPRS Journal of Photogrammetry and Remote Sensing 62 (6), 415–433.

Hyyppä, J., Kelle, O., Lehikoinen, M., Inkinen, M., 2001. A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. IEEE Transactions on Geoscience and Remote Sensing 39 (5), 969–975.

Jutzi, B., Stilla, U., 2005. Waveform processing of laser pulses for reconstruction of surfaces in urban areas. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 8/W27), on CD-ROM.

Jutzi, B., Stilla, U., 2006. Range determination with waveform recording laser systems using a Wiener Filter. ISPRS Journal of Photogrammetry and Remote Sensing 61 (2), 95–107.

Kirchhof, M., Jutzi, B., Stilla, U., 2008. Iterative processing of laserscanning data by full waveform analysis. ISPRS Journal of Photogrammetry and Remote Sensing 63 (1), 99–114.

Maltamo, M., Eerikäinen, K., Pitkänen, J., Hyyppä, J., Vehmas, M., 2004. Estimation of timber volume and stem density based on scanning laser altimetry and expected tree size distribution functions. Remote Sensing of Environment 90 (3), 319–330.

Maltamo, M., Packalén, P., Peuhkurinen, J., Suvanto, A., Pesonen, A., Hyyppä, J., 2007. Experiences and possibilities of ALS based forest inventory in Finland. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 3/W52), on CD-ROM.

Mehtätalo, L., 2006. Eliminating the effect of overlapping crowns from aerial inventory estimates. Canadian Journal of Forest Research 36 (7), 1649–1660.

Morsdorf, F., Meier, E., Allgöwer, B., Nuesch, D., 2003. Clustering in airborne laser scanning raw data for segmentation of single trees. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 34 (Part 3/W13), 27–33.

Naesset, E., 2004. Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. Scandinavian Journal of Forest Research 19 (2), 164–179.

Persson, A., Holmgren, J., Söderman, U., 2002. Detecting and measuring individual trees using an airborne laserscanner. Photogrammetric Engineering & Remote Sensing 68 (9), 925–932.

Pyysalo, U., Hyyppä, H., 2002. Reconstructing tree crowns from laser scanner data for feature extraction. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 34 (Part 3B), 218–221.

Reitberger, J., Krzystek, P., Stilla, U., 2008. Analysis of full waveform LIDAR data for the classification of deciduous and coniferous trees. International Journal of Remote Sensing 29 (5), 1407–1431.

Shi, J., Malik, J., 2000. Normalized cuts and image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (8), 888–905.

Solberg, S., Naesset, E., Bollandsas, O.M., 2006. Single tree segmentation using airborne laser scanner data in a structurally heterogeneous spruce forest. Photogrammetric Engineering & Remote Sensing 72 (12), 1369–1378.

Stilla, U., Jutzi, B., 2008. Waveform analysis for small-footprint pulsed laser systems. In: Shan, J., Toth, C.K. (Eds.), Topographic Laser Ranging and Scanning: Principles and Processing. Taylor & Francis, Boca Raton, FL, pp. 215–234.

Vincent, L., Soille, P., 1991. Watersheds in digital spaces: An efficient algorithm based on immersion simulations. IEEE Transactions of Pattern Analysis and Machine Intelligence 13 (6), 583–598.

Wagner, W., Ullrich, A., Ducic, V., Melzer, T., Studnicka, N., 2006. Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner. ISPRS Journal of Photogrammetry and Remote Sensing 60 (2), 100–112.

Wang, Y., Weinacker, H., Koch, B., Sterenczak, K., 2008. LIDAR point cloud based fully automatic 3D single tree modelling in forest and evaluations of the procedure. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 37 (Part B6b), 45–51.