SEGMENTATION OF TREE REGIONS USING DATA OF A FULL-WAVEFORM LASER

H. Gross, B. Jutzi, U. Thoennessen

FGAN-FOM Research Institute for Optronics and Pattern Recognition, 76275 Ettlingen, Germany
{gross,jutzi,thoe}@fom.fgan.de

Commission I, WG I/2

KEY WORDS: Laser data, full-waveform, point clouds, intensity, environment, covariance of points, eigenvalues, segmentation.

ABSTRACT:

The new sensor technique for receiving full-waveform laser data delivers not only height information but also additional features like intensity. The echoes generated by reflecting objects are approximated by amplitude and width of a Gaussian kernel. These features are used for tree region segmentation. The inclusion of the available information in the environment around each 3d point of the cloud can be used to stabilize the segmentation process. Two kinds of environments spherical and cylindrical are discussed. The calculation of the covariance inside these environments is discussed without and with weighting all points by their measured intensity. An application of the described method is demonstrated for a typical scene. A method for the correction of the number of points inside an environment produced by the movements of the sensor platform is presented. Two proposals are described for the segmentation of trees. One of them uses the eigenvalues of the covariance matrix in an environment. The other method considers only the features of the echoes for each laser pulse without considering an environment.

1. INTRODUCTION

The automatic generation of 3d models from laser scanning data to gain a description of man-made objects is of great interest to the photogrammetric research (Brenner *et al.*, 2001; Geibel & Stilla, 2000; Gross *et al.*, 2005). Spaceborne, airborne as well as terrestrial laser scanning systems allow a direct and illumination-independent measurement of laser scanning data from 3d objects in a fast, contact free and accurate way.

Recent developments of commercial airborne laser scanners led to small footprint laser systems that allow capturing the waveform of the backscattered laser pulse, namely the OPTECH ALTM 3100, TOPEYE MK II, and TOPOSYS HARRIER 56. The latter one is based on the RIEGL LMS-Q560. These systems mentioned above are specified to operate with a transmitted pulse width of 4-10 ns and allow digitization and acquisition of the waveform with approximately 0.5-1 GSample/s.

To interpret the received waveform of the backscattered laser pulse, a fundamental understanding of the physical background of pulse propagation and surface interaction is important. The waveform offers the possibility to study different features like the *range*, *elevation variations*, and *reflectivity* of the surface based on the inclination between the divergent laser beam and object plane. These specific features have an influence on the received waveform. The waveform of each pulse is described by a series of range values combined with amplitude values and can be approximated by one or more parameterized Gaussian curves (Hofton *et al.*, 2000). Due to this approximation, specific features like *temporal position*, *width* and *amplitude* (cf. Figure 1) caused by the object surfaces are estimated (Jutzi & Stilla, 2006).

Nowadays, the analysis of full-waveform data is more and more of interest especially for forestry applications, because it provides the opportunity for a detailed investigation of vertical distributed surface structures (Hug *et al.*, 2004). With full-waveform analysis additional points can be extracted from the measured data compared to the conventional first pulse and last pulse technique. These additional points and their corresponding surface can lead to a better description of vertical structures like vegetation (Persson *et al.*, 2005). Furthermore

these features can be used for segmentation and classification beside the geometric information (Wagner *et al.*, 2006). For instance, Reitberger *et al.* (2006) demonstrated that the features and the additional points, derived by full-waveform decomposition, are useful to classify tree species.

In this paper we focus on the features derived by full-waveform decomposition to segment vegetation, namely trees. Beside the raw features, like signal amplitude, signal width, and total number of echoes for each emitted laser beam, a spherical and cylindrical environment in the close neighborhood is used to improve the segmentation. Both methods are compared to each other by a ROC (Receiver Operating Characteristic) curve to evaluate the detection and the false alarm rate.

In Section 2 a brief description of the used full-waveform data is given. The used raw data and volumetric approaches are presented in Section 3. The methodology to calculate additional point features based on the covariance matrix is explained in Section 4. We visualize in Section 5 the scene and the corresponding features derived from the measured laser scanning data. A required correction of features is proposed in Section 6. The segmentation by using the volumetric and the raw data approach is described in Section 7. The results are presented in Section 8 including a comparison between ground truth data and ROC curves. Finally the used methods and derived results are discussed in Section 9.

2. FULL-WAVEFORM DATA

We operate on data measured by a RIEGL LMS-Q560 sensor with a field of view of 60° and a point density of about 3.7 points per m². The flying height was about 400m above ground. For each beam the total number of detected backscattered pulses is known and assigned to the corresponding echoes. Each echo is described by a point with its 3d coordinate, signal amplitude, and signal width derived from the Gaussian approximation.

Figure 1 shows the received signal produced by three different objects along the beam corridor. The shape of the received waveform depends on the illuminated surface area, especially on the material, reflectivity of the surface and the inclination angle between the surface normal and the laser beam direction.

The intensity (energy) is estimated by the width multiplied with the amplitude of the Gaussian approximation and corrected by the range between sensor and object. It describes the reflectivity influenced by geometry and material of the object at this point. For each particular echo caused by partially illuminated object surfaces an own intensity value is received.
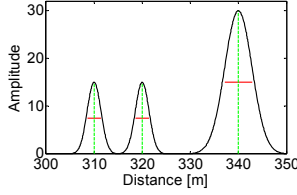

Figure 1. Example for three received echoes derived by a single emitted laser pulse with several signal widths (solid lines) and amplitudes (dashed lines).

3. RAW DATA AND VOLUMETRIC APPROACH

Due to the wide field of view the sensor delivers measurements between the nadir and an oblique angle. This varying angle causes problems discriminating objects. The angle influences the echoes even for the same kind of objects at different positions with respect to the flight path. Particularly for oblique emitted laser beams the received echoes are reflected by objects with essential different horizontal positions. Several approaches to overcome these problems are discussed in this paper.

The Gaussian decomposition of the full-waveform data allows the interpretation of the received echoes with signal widths and amplitudes. We call this the raw data approach.

Another possibility is a volumetric approach. For each point we mark all points inside a sphere (Figure 2b) or a vertical cylinder (Figure 2a) centred at the trigger point. The sphere includes all points inside a restricted 3d environment. The vertical cylinder includes all points inside a 3d environment without restriction on the vertical position.

All measured points calculated by the approximation of the waveforms of the same or different laser beams, which are located within the sphere or the cylinder, are considered as neighbors in the further processing.
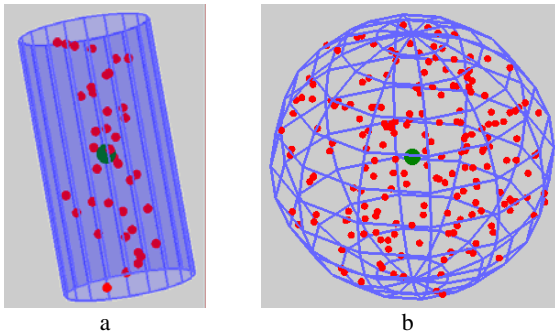

Figure 2. Environment with point cloud and centre point:
a) cylindrical environment,
b) spherical environment.

4. POINT FEATURES BASED ON THE COVARIANCE

For the defined volumes the 3d moments, as described by Maas & Vosselman (1999), are discussed and improved by considering the vertical dimension in the same manner as the horizontal ones (Gross & Thoennessen, 2006). After calculation of the covariance an interpretation of the eigenvalues is possible to define discriminating features for planes, edges and corners.

4.1 Moments of the covariance matrix

In a continuous domain, the moments are defined by

$$m_{ijk} = \int_V x^i y^j z^k f(x,y,z)\,dv , \qquad (1)$$

where $i,j,k \in \mathbb{N}$, and $i+j+k$ is the order of the moment integrated over a predefined volume weighted by $f(x,y,z)$. We use as weighting function the intensity of the reflected laser pulse. Considering only the moments with $i+j+k \leq 2$ we get the weight, the center of gravity and the matrix of covariance. Invariance against translation is achieved by subtraction of the center of gravity

$$\bar{x} = \frac{m_{100}}{m_{000}}, \quad \bar{y} = \frac{m_{010}}{m_{000}}, \quad \bar{z} = \frac{m_{001}}{m_{000}}, \qquad (2)$$

and invariance against the units of the coordinates is achieved by normalization with the radius R of the volume cell. The approximation of the integral by the sum over all points inside the cell yields the components of the covariance matrix

$$\widetilde{m}_{ijk} = \frac{\sum_{l=1}^{N}\left(x_l - \bar{x}\right)^i \left(y_l - \bar{y}\right)^j \left(z_l - \bar{z}\right)^k f(x_l,y_l,z_l)}{R^{i+j+k}\sum_{l=1}^{N} f(x_l,y_l,z_l)} . \qquad (3)$$

Finally we calculate for each point of the whole data set the covariance matrix

$$M = \begin{pmatrix} \widetilde{m}_{200} & \widetilde{m}_{110} & \widetilde{m}_{101} \\ \widetilde{m}_{110} & \widetilde{m}_{020} & \widetilde{m}_{011} \\ \widetilde{m}_{101} & \widetilde{m}_{011} & \widetilde{m}_{002} \end{pmatrix} . \qquad (4)$$

The eigenvalues λ_i and eigenvectors $\vec{e}_i$ with $i=1,2,3$ of the symmetrical matrix deliver additional features for each point. The eigenvalues are invariant concerning rotation of the coordinate system. If the weighting function $f(x,y,z)$ depends on the points we calculate the normalized weight by

$$\widetilde{m}_{000} = \frac{\sum_{l=1}^{N} f(x_l,y_l,z_l)}{N} . \qquad (5)$$

For object classification, West et al. (2004) uses the following features which depend on the eigenvalues:

$$\text{structure tensor omnivariance} = \sqrt[3]{\prod_{i=1}^{3} \lambda_i} , \qquad (6)$$

$$\text{structure tensor planarity} = \frac{\lambda_2 - \lambda_3}{\lambda_1} . \qquad (7)$$

4.2 Weighting of points

For comparison we selected two weighting functions $f(x,y,z)$ given by the equations (1) and (3). Therefore each point is weighted by a constant or weighted by its own intensity value.

5. SCENE DESCRIPTION

Our investigations are focused on a scene of a rural village as shown in Figure 3a. The scene includes streets, buildings, lawn and trees. The corresponding height image colored by the elevation is depicted in Figure 3b. The left part and the right upper corner show mainly grassland and trees. In the middle part several kinds of buildings with different shape and height can be recognized.

A more precise impression of the data as point cloud shows Figure 4. It demonstrates the influence of the intensity due to the material and the angle between sensor and object plane. We achieve a low intensity for trees and many streets. Higher values arise for grassland. Roof planes yield the highest intensities but vary depending on the angle between sensor and surface orientation (Figure 4c).
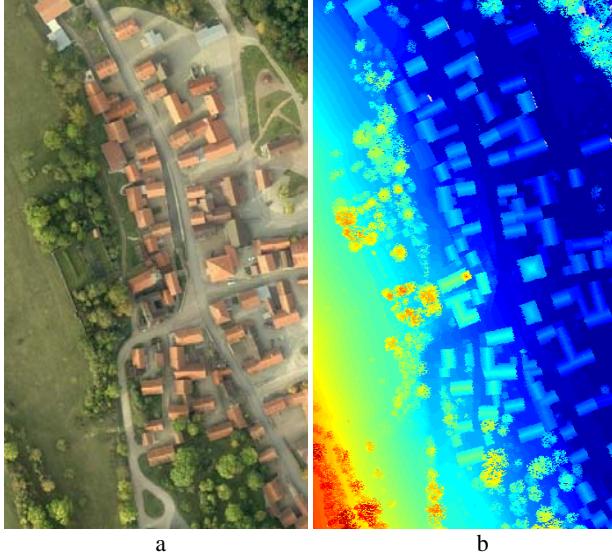

Figure 3.   a) RGB orthophoto,
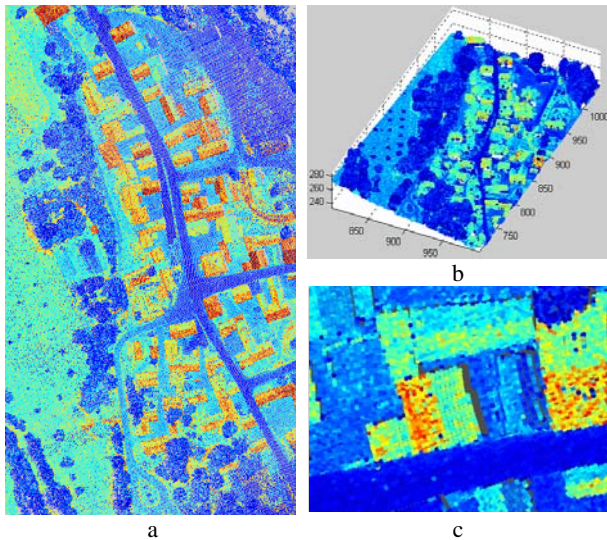b) laser elevation data colored by height.

Figure 4.   Point cloud of the laser elevation data colored by intensity:
a) vertical view,
b) oblique view,
c) dependency of the intensity signal by the surface orientation.

6.   DATA PREPROCESSING

The received number of echoes depends on the measurement situation. As an example we expect more echoes for tree regions than for grassland or roofs. Figure 5a shows the number of points inside the spherical environment at a radius of $R = 2m$ and demonstrates the variation of this number. This depends not only on the measured objects but also on the movements of the sensor platform. These movements produce not equal spaced scan lines. Therefore it is essential to define a normalization measure for the number of echoes in the sphere.

An image based method is proposed to compensate the point number variation, where the image is a rasterized representation of the point cloud. We filtered the image by a 2d Gaussian kernel with a support region of 17x17 pixels and a standard deviation of 8 pixels. Finally, we subtracted the result from the original image. This calculation steps yield to Figure 5b. The number of echoes has only a small variation at grassland and at roofs, but a high variation at walls and trees. The high values for the walls are caused by the wide field of view of up to 60 degrees. For cylindrical environment we get the same behavior. For high walls more measured points fall into the cylindrical environment especially if the viewing angle differs from nadir angle.
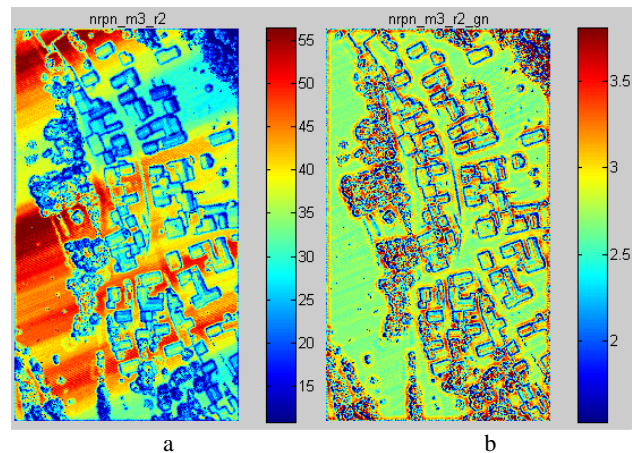

Figure 5.   Spherical environment:
a) number of points inside the sphere,
b) normalized number of echoes inside the sphere.

7.   SEGMENTATION OF TREE REGIONS

In the following sections the different methods presented above, will be applied to the segmentation of tree regions for example.

7.1   Influence of different environments and weighting

For the selected rural area the omnivariance (equation (6)) is calculated both with a cylindrical and with a spherical environment. Results are depicted in Figure 6a&b. The walls of buildings and tree regions are represented by a high omnivariance, whereas objects which appear like a plane inside the environment are showing small values.

Figure 6a includes only those points lying inside the spherical environment. We get a little more details due to the limited extension of the sphere in the vertical dimension.

Figure 6c&d demonstrate the influence of the weighting on the omnivariance feature. This is shown by the difference of the omnivariance with and without using the intensity of the laser beam as weighting factor during the calculation of the covariance. The dynamic of the difference lies within 10% of the dynamic of the omnivariance and is to the further consideration of minor importance.

By considering the different volumes it is observable, that for a cylindrical environment the trees and walls have greater differences than planes (Figure 6c). Whereas using the spherical environment the difference is not as significant to the walls of the buildings as in the former case (Figure 6d).

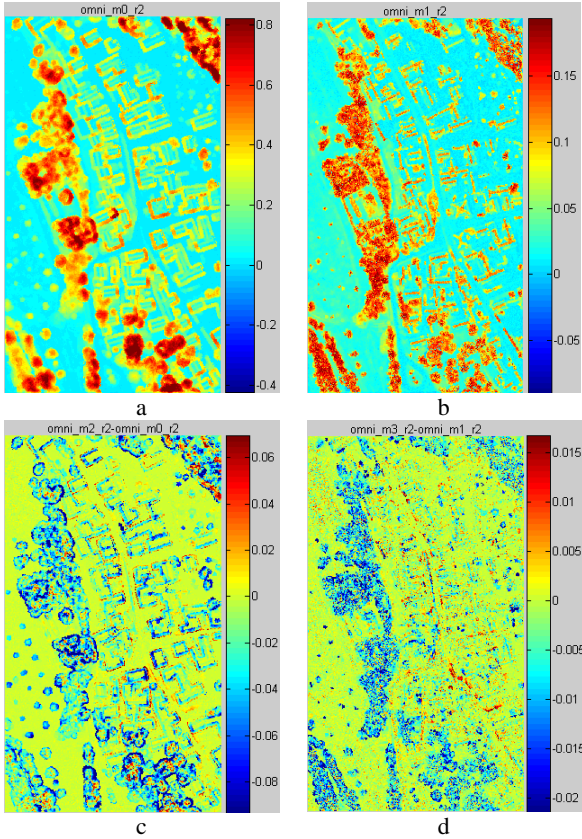## 7.2 Segmentation of tree regions

Figure 6. Omnivariance for different environments and weighting of the points: cylindrical environment (a&c), spherical environment (b&d), without weighting by the intensity (a&b), difference between the weighted and not weighted omnivariance (c&d).
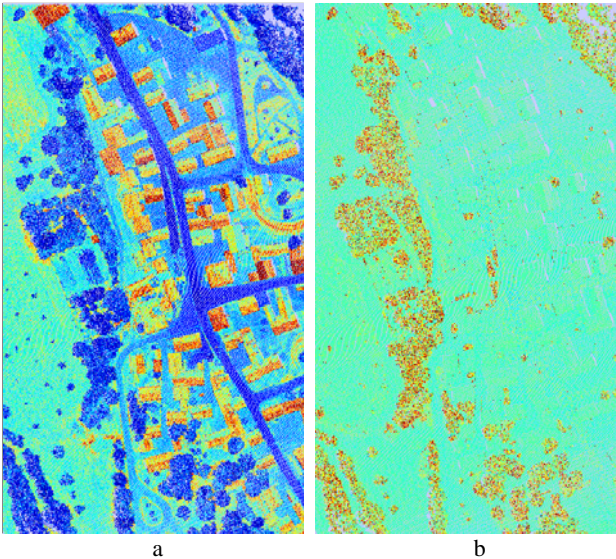

Figure 7. Features useable for tree segmentation:
a) signal amplitude,
b) signal width.

The results of the previous section have shown that the cylindrical volumetric approach is more appropriate for the segmentation of tree regions than the spherical solution.

Therefore in the following steps we will only discuss the cylindrical approach vs. the raw data feature exploitation.

The signal amplitude shown in Figure 7a and the signal width presented in Figure 7b are significant features for the detection of tree regions, because the geometry of trees deliver in most cases signal returns with very small amplitudes but high values for the width.

7.2.1 Volumetric approach

Investigating the different features the following behaviour can be observed: Trees have a low reflectivity resulting in a low value for averaged intensity W (equation (5)). But the low reflectivity of streets will have the same effect. Therefore we need an additional discriminating feature. The feature planarity P (equation (7)) shows high values especially for plane objects. In contrast to this the omnivariance O (equation (6)) delivers high values for volumetric objects.

Due to the different ranges of values these features are normalized by the sigmoid function defined by equation:

$$\sigma(x, x_0, s) = 1 / \left(1 + e^{s(x - x_0)}\right). \tag{8}$$

For trees we adapt the centre point and scaling factor.
The three features are transformed into

$$\sigma_W = \sigma(W, 1000, 0.3) \qquad \text{for the averaged intensity,}$$
$$\sigma_P = \sigma(P, 0.1, 5) \qquad \text{for the planarity,} \tag{9}$$
$$\sigma_O = \sigma(O, 0.4, -1) \qquad \text{for the omnivariance.}$$

Positive scaling values indicate a decreasing of the sigmoid function. Figure 8a shows the values for the planarity after normalization. Plane objects are marked by blue and volumetric objects by red regions. Finally, the three indicators are combined into a single measure

$$\sigma_T := \sigma_W \sigma_P \sigma_O > 0.25 \tag{10}$$

and restricted by a threshold. The result is given in Figure 8b where tree regions are marked. An optimization process with respect to the parameters of the sigmoid function has the aim to gain the best detection and false alarm rate. The results have to be compared to results derived from data of other regions.

7.2.2 Raw data approach

In this section we consider for each point of the cloud its own feature value delivered by the sensor. For segmentation of tree regions the features signal width and amplitude are used, derived by Gaussian decomposition of the full-waveform. Further the feature total number of echoes is considered. They are normalized by

$$\sigma_w = \sigma(w, 45, 0.7) \qquad \text{for the width,}$$
$$\sigma_a = \sigma(a, 50, -0.4) \qquad \text{for the amplitude,} \tag{11}$$
$$\sigma_t = \sigma(t, 2, -5) \qquad \text{for the total number of echoes.}$$

The result is shown in Figure 9a for the width of each echo. We define the tree regions for all points with

$$\sigma_{Tp} := \sigma_w \sigma_a \sigma_t > 0.1, \tag{12}$$

which delivers Figure 9b.

8. RESULTS

The different approaches discussed in the previous sections are applied to the data. A comparison with Figure 8b the result including the cylindrical environment indicates that tree regions may be segmented more precise by calculation of the

covariance matrix, their eigenvalues and derived features, than by considering only the raw data as shown in Figure 9b. For a final decision it would be necessary to optimize both results separately with respect to a target function by variation of all parameters of the used sigmoid function. The target function has to compare the calculated results to a ground truth definition of the tree regions.
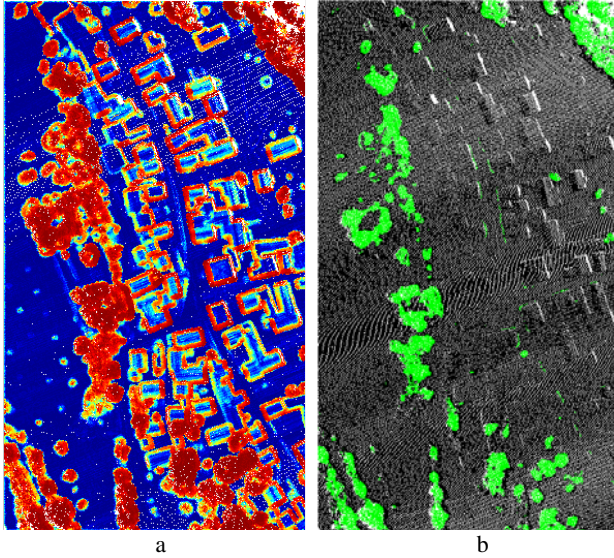

Figure 8.   Indicator for tree regions including the environment:
a) based on the planarity,
b) based on width, planarity and omnivariance with threshold.
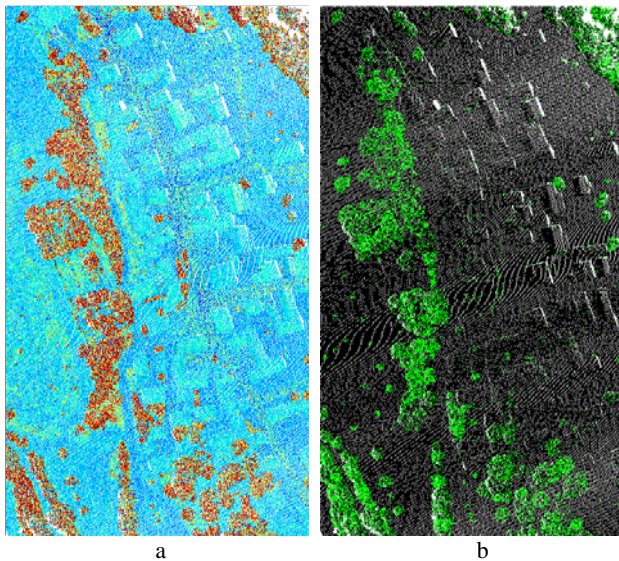

Figure 9.   Indicator for tree regions without consideration of the environment: a) based on the width of the echo, b) based on width, amplitude and total number of echoes with threshold.

After definition of real tree regions (Figure 10) by an interactive determination, we are able to calculate false alarm and detection rate by modifying the thresholds in equation (10) and (12). The tree region information is back propagated into the point cloud to assign the class information to each point.

The ROC curves for both methods are shown in Figure 11. Using eigenvalues by consideration of a cylindrical environment delivers a smaller false alarm rate than the

segmentation by the features of the raw data only. On the other hand the detection rate for the raw data approach method is higher than for the volumetric approach.
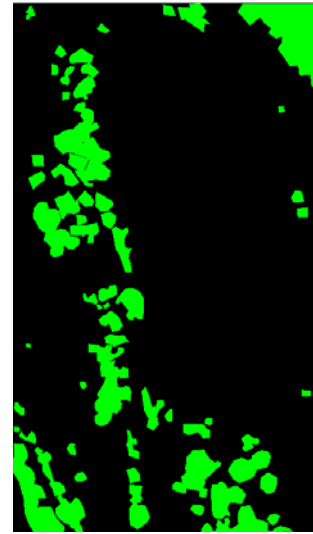

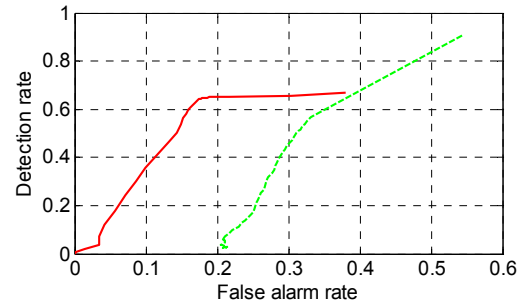Figure 10. Interactive defined tree regions.

Figure 11. Detection rate vs. false alarm rate for different thresholds of equation (10) including environment (solid line) and equation (12) for the features of the raw data (dash line).

Finally, a 3d visualization of the tree regions based on width, planarity and omnivariance is shown in Figure 12. The tree crowns are represented by ellipsoids. As Straub (2003) mentioned, this assumption is true, if the fine structure of the crown is ignored and the coarse structure is revealed in an appropriate level of the multi-scale representation of the surface model. For visualization purposes the scales of ellipsoids correspond to the eigenvalues and eigenvectors given by the relevant points inside a spherical environment with a $radius = 6m$. On the other side we know by analytical determination the eigenvalues of an ellipsoid as

$$\lambda_i = \frac{a_i^{\,2}}{5} \ \text{ with } \ i = 1, 2, 3, \tag{13}$$

where a_i are the half axes of an ellipsoid. After calculation of the eigenvalues of the covariance matrix we get the half axes by

$$a_i = \sqrt{5\lambda_i} \ . \tag{14}$$

For each point falling inside the tree regions the ellipsoids are determined. If its centre point lies near the considered point the ellipsoid is accepted. Already processed points within a tree crown are ignored for further processing. The tree trunks with a fix diameter are additionally plotted as junction between bottom and ellipsoid centre to give an impression of the tree position in the 3d visualization.

61

Figure 12. Virtual trees visualized by ellipsoids.

9. DISCUSSION AND CONCLUSION

New sensor technology of laser scanners delivers more detailed information. The new features include additional information about the measured objects.

The used data does not contain a homogeneous point distribution due to the movements of the sensor. This discrepancy is compensated by an adapted filter.

We consider features derived by calculation of the eigenvalues of the covariance matrix for cylindrical and spherical environments with respect to the intensity of the echoes. For our investigations only data with low point density was available. Therefore the spherical environment should have a large radius to get a representative number of points for a valuable distribution. This distribution of the points can be inhomogeneous (Gross & Thoennessen, 2006). For a cylindrical environment the trees and walls have greater significance than by calculation with a spherical environment. Therefore the segmentation process is executed only for cylindrical environments. The usage of normalized weight, omnivariance and planarity for tree segmentation results in acceptable detection and false alarm rate.

Without including the neighbored point information, only considering the features of the raw data like signal amplitude, signal width and total number of targets, we get a good detection rate but an unacceptable false alarm rate.

Modifications of the parameters of the sigmoid function may influence the values of detection and false alarm rate but not their principle behavior.

The usage of point clouds instead of images requires modified methods for data representation, processing and segmentation.

The analysis of full-waveform data delivers additional features beside the range value. These features are included in the data processing steps for visualizing relevant objects. The full-waveform analysis yields more details and allows a more precise segmentation of the different kind of objects. This will be supported by considering a spherical or cylindrical environment. Former investigations show that a spherical environment is suitable for edge, corner, and plane detection (Gross & Thoennessen, 2006). For the segmentation process of horizontal regions without considering the vertical dimension a cylindrical environment delivers more suitable results.

The decision of cylindrical or spherical environment depends on the kind of object class and should be investigated in more detail in the future.

REFERENCES

Brenner, C., Haala, N. Fritsch, D., 2001. Towards fully automated 3D city model generation. In: Baltsavias, E., Grün, A., van Gool, L. (Eds) Proceedings of the 3rd International Workshop on Automatic Extraction of Man-Made Objects from Aerial and Space Images, pp. 47-56.

Geibel R., Stilla, U., 2000. Segmentation of Laser-altimeter data for building reconstruction: Comparison of different procedures. International Archives of Photogrammetry and Remote Sensing 33 (Part B3), pp. 326-334.

Gross, H., Thoennessen, U., v. Hansen, W., 2005. 3D Modeling of Urban Structures. In: Stilla, U., Rottensteiner, F., Hinz, S. (Eds) Joint Workshop of ISPRS/DAGM Object Extraction for 3D City Models, Road Databases, and Traffic Monitoring CMRT05, International Archives of Photogrammetry and Remote Sensing 36 (Part 3/W24), pp. 137-142.

Gross, H., Thoennessen, U., 2006. Extraction of Lines from Laser Point Clouds. In: Förstner, W., Steffen, R. (Eds) Symposium of ISPRS Commission III: Photogrammetric Computer Vision PCV06. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 3), pp. 86-91.

Hofton, M.A., Minster, J.B., Blair, J.B., 2000. Decomposition of laser altimeter waveforms. IEEE Transactions on Geoscience and Remote Sensing 38 (4), pp. 1989–1996.

Hug, C., Ullrich, A., Grimm, A., 2004. LITEMAPPER-5600 - a waveform digitising lidar terrain and vegetation mapping system. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 8/W2), pp. 24–29.

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), pp. 95-107.

Maas, H., Vosselman, G., 1999. Two algorithms for extracting building models from raw Laser altimetry data. ISPRS Journal of Photogrammetry and Remote Sensing 54 (2-3), pp. 153-163.

Persson, Å., Söderman, U., Töpel, J., Ahlberg, S., 2005. Visualization and analysis of full-waveform airborne laser scanner data. In: Vosselman, G., Brenner, C. (Eds) Laser scanning 2005. International Archives of Photogrammetry and Remote Sensing 36 (3/W19), pp. 109-114.

Reitberger, J., Krzystek, P., Stilla, U., 2006. Analysis of Full Waveform LIDAR Data for Tree Species Classification. In: Förstner, W., Steffen, R. (Eds) Symposium of ISPRS Commission III: Photogrammetric Computer Vision PCV06. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 3), pp. 228-233.

Straub, B., 2003. Automatic Extraction of Trees from Aerial Images and Surface Models. In: Ebner, H., Heipke, C., Mayer, H., Pakzad, K. (Eds) Photogrammetric Image Analysis PIA'03. International Archives of Photogrammetry and Remote Sensing 34 (Part 3/W8), pp. 157-164.

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), pp. 100-112.

West, K.F., Webb, B.N., Lersch, J.R., Pothier, S., Triscari, J.M., Iverson, A.E., 2004. Context-driven automated target detection in 3d data. In: Sadjadi, F.A. (Ed) Automatic Target Recognition XIV. Proceedings of SPIE Vol. 5426, pp. 133-143.