Full Wave Analysis in 3D Laser Scans for Vegetation Detection in Urban Environments

Jan Elseberg, Dorit Borrmann, Andreas Nüchter
School of Engineering and Science, Jacobs University Bremen gGmbH, 28759 Bremen, Germany
Email: {j.elseberg|d.borrmann|a.nuechter}@jacobs-university.de

***Abstract*—This paper presents a novel technique for detecting vegetation of virtually all forms in terrestrial laser scanning data of urban environments. We make use of a modern laser range finder capability to measure multiple echoes per laser pulse via Full Wave Analysis. The algorithm is able to efficiently, i.e., less than acquisition time, identify vegetation to a high degree of accuracy (more than 99 percent). We present and evaluate three alternatives to classify candidate regions as either vegetation or non-vegetation.**

***Index Terms*—Semantic Mapping, Object Recognition, Full Wave Analysis**

I. Introduction

Recognizing and understanding sensor data of laser range scanners is beneficial for many applications. The robotic mapping problem has received a lot of attention in recent years and the current trend is to add meaningful labels to the maps. Annotating metric maps with semantic labels is a necessary prerequisite to achieve true autonomy for any mobile robot. This paper deals with identifying vegetation in unorganized point clouds. Unlike other approaches, the presented algorithm makes no implicit assumption about a single non-moving viewpoint. This is particularly important for urban modelling applications where data is usually acquired continuously on a moving vehicle instead of in a stop-and-go fashion. Therefore, the algorithm is also inherently able to deal with registered range images, allowing it to consider all available data. The points are classified into two categories, vegetation and non-vegetation. This topic is relevant for several reasons. First, it is an important first step in a comprehensive semantic robotic mapping system. Detecting vegetation means that subsequent interpretation algorithms can focus their attention on parts of the data that contain relevant structures. Second, for geometric modeling applications in urban environments with terrestrial laser scanner like 3D city modeling it is often of interest to only reconstruct houses and streets to a high level of detail. In this context, the possibility of removing vegetation from point clouds as in Fig. I is of obvious importance.

We intend to solve this problem by purposefully extracting vegetation from laser scans. Last but not least, the proposed method can be used for automated forest surveys and forestry assessment.

II. State of the Art

Fig. 2 shows the robot platform that acquired the data sets presented in this paper. The main sensor is a RIEGL VZ-400 laser scanner [11]. The device distinguishes itself from common range scanner used in surveying or robotics by the Full-Waveform-Analysis (FWA) [12]. Common pulsed or continuous-wave laser scanner are based on the time-of-flight principle to measure distances to the target. Reception is typically detected by an analog threshold. Most non FWA capable devices are therefore restricted to one range measurement per pulse. Very often a pulse hits multiple surfaces, so that multiple reflections are received. This usually occurs on

Fig. 1. A scan from the Campus data set with all echoes (left) and without points that were classified as vegetation (right). Points are colored in accordance to their calibrated reflectance values.

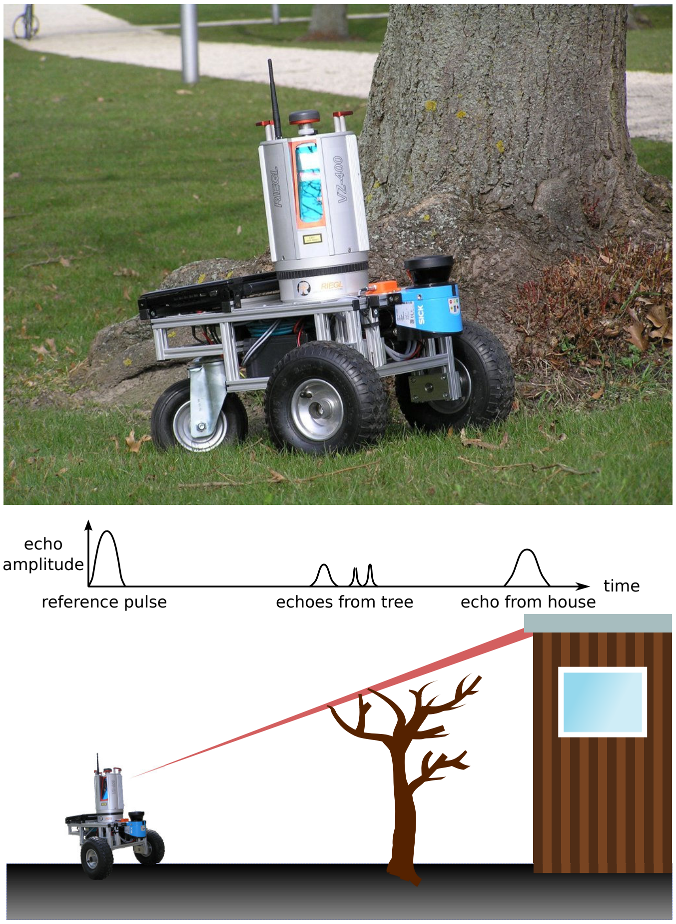


Fig. 2. The mobile robot Irma3D (Intelligent robot for mapping applications in 3D) with its main sensor, the RIEGL VZ-400, which is mounted roughly 60cm above ground. The scanner typically acquires $125,000$ points per second and requires 6 seconds per rotation. Bottom: The principle of Full Wave Analysis. The beam sent out by the laser scanner causes an echo to be received for each obstruction in proportion to the object's distance. The response is sampled with a very high temporal resolution so that multiple echoes and their properties (distance, reflectance, amplitude, deviation) can be recovered.

the edges of structures, predominantly on trees and vegetation. FWA is the technique of modeling the received signal with several Gaussian functions, one for each range measurements. The laser scanner VZ-400 is therefore capable of measuring multiple targets per pulse with additional features such as amplitude and standard deviation per range measurement (cf. Fig. 2. The device has a maximum range of up to 500m and an absolute minimum range of 1.5m. Within the entire range multiple echoes are received so that vegetation detection is possible (cf. Fig. 6).

The FWA property has been used in aerial laser scanning for identifying vegetation for some years [14]. Yu et al. segment range data into vegetation and non-vegetation to measure forest growth and individual tree-sizes [15]. Using such data, single trees are identified in dense forests by employing normalized cut segmentation [4]. Reitberger et al. detect the stem of individual trees by a robust RANSAC-based estimation of the stem points. With single trees identified, Holmgren and Persson classify trees by their species [3]. Using only an unorganized point cloud and no further feature aside from the type of echo (first, single, last, etc.) they were able to distinguish between spruce and pine trees with an accuracy of 95 percent. As we will be conducting terrestrial instead of aerial laser scanning, approaches to vegetation detection must change with the perspective. Trees seen from the side instead of from above will not have their trunk obscured as much, making multiple echoes less prominent. The distance between scanning device and tree will also vary in a terrestrial scenarios. Pfeifer et al. identify trees in terrestrial data taken from forests by cylindrical fitting [9]. They model the branch structure of a tree by cylinder following, while the foliage is captured by an outer hull model. Barnea et al. employ the k-nearest neighbors algorithm for classifying range images of urban environments [1]. As they work directly on the depth image taken by the scanner, the algorithm is viewpoint dependent and does not benefit from registering multiple scans to each other.

Recently conditional random fields (CRF) have become a popular approach for classifying point cloud data. Lim and Suter [5] have presented a classification algorithm based on CRFs that uses super-voxels to reduce the amount of data to be processed. Other machine learning approaches like Markov networks and support vector machines (SVM) have also been applied for point cloud classification tasks. Munoz et al. [6] have adapted the functional gradient technique to max-margin Markov networks for learning high-order classification models. To handle the usually large amount of data and the ensuing high computation time Triebel et al. [13] adaptively reduce the point cloud by utilizing k-D trees before using associative Markov networks. Posner et al. [10] combine camera images with 3D point cloud data taken on a continuously moving vehicle and classify it into multiple classes using a support vector machine.

III. THE ALGORITHM

Based on the assumption that trees and other vegetation contain a high amount of multiple echoes, the algorithm presented in this paper first segments the point cloud into regions of interest. Then for each region shape description features are computed. The regions are afterwards classified according to the features into vegetation and non-vegetation. All points from the entire data set falling into regions classified as vegetation are then labeled vegetation as well. As a large number of points from the floor will be included in this last step, a post-processing step identifies large planar surfaces and removes them from the regions. This is done by horizontally sweeping over the tree to compute the largest horizontal slice underneath the vegetation. This post-processing is very fast as it is linear in the number of input points.

The following two sections will lay out the efficient segmentation of the point cloud and three features we employ to classify the segments.

A. *Multi-Resolution Clustering*

To get an impression of the distribution of the received multiple echoes, refer to the third image in Fig. 7. Trees and vegetation will have a high incidence of echoes. In aerial laser scanning, vegetation, particularly trees are virtually the only source of such echoes. In terrestrial laser scanning additional structures such as fences, power lines, people and window or door frames are present. All of these will produce more than a single range measurement per pulse. Echos are of several types, according to how many there are and when they occur. There is the single echo (only one measurement), the first and last echoes (closest and farthest echo respectively) and the inner echoes that lie between the first and last echoes. To model regions without vegetation in aerial range data it usually suffices to use only single and last echoes. In that scenario the scanner is always several 100 meters away from any tree so that it is very unlikely echoes of that type will stem from trees. This is not the case at all in terrestrial data acquired in urban environments. However, we assume that vegetation will always contain inner or first echoes. We therefore segment regions of the scan with a high concentration of those types of echoes.

Our approach is based on the two-pass clustering algorithm [2]. Initially the point cloud is leveled by rotating it according to the inclinometer on the measuring platform. The scan is then projected on the horizontal ground plane and converted to a binary occupancy grid. A grid cell is set if the corresponding patch of the scan contains inner or first echoes, otherwise it is not set. Segments are then identified by the classical two-pass algorithm. One pass to label each pixel with its neighbor's label, or with a new unique label if no neighbors exist. The second pass ensures each pixel is labeled equivalent to the connected component the pixel belongs to. Ideally we wish to segment the scans with a resolution of 10cm, so that each pixel represents a patch of $100\mathrm{cm}^2$. As the employed laser range finder VZ-400 has a maximum range of up to 500m and a registered data set encompasses several times that area, a resolution of 10cm would be prohibitive. To cope with large areas, we employ a multi resolution two-pass algorithm. The scan is clustered on multiple resolutions from coarse to fine. An initial coarse image is segmented using the two-pass algorithm. Any identified cluster containing more than 100 echoes is converted to a finer resolution and segmented. The recursion stops when the finest resolution is reached. At any one point the memory requirements of the algorithm is kept to a minimum, while the runtime is close to linear in the number of points.

B. *Features*

Each of the segments identified have to be analyzed to distinguish vegetation from clutter. We have implemented 2 features that were used for SVM classification and a fast classification based on a threshold verification.

Inspired by the trunk detection employed by Reitberger et al. [4], we compute a histogram of point densities along the trunk of the hypothetical tree within the cluster. Point density is computed for 10cm wide horizontal slices along

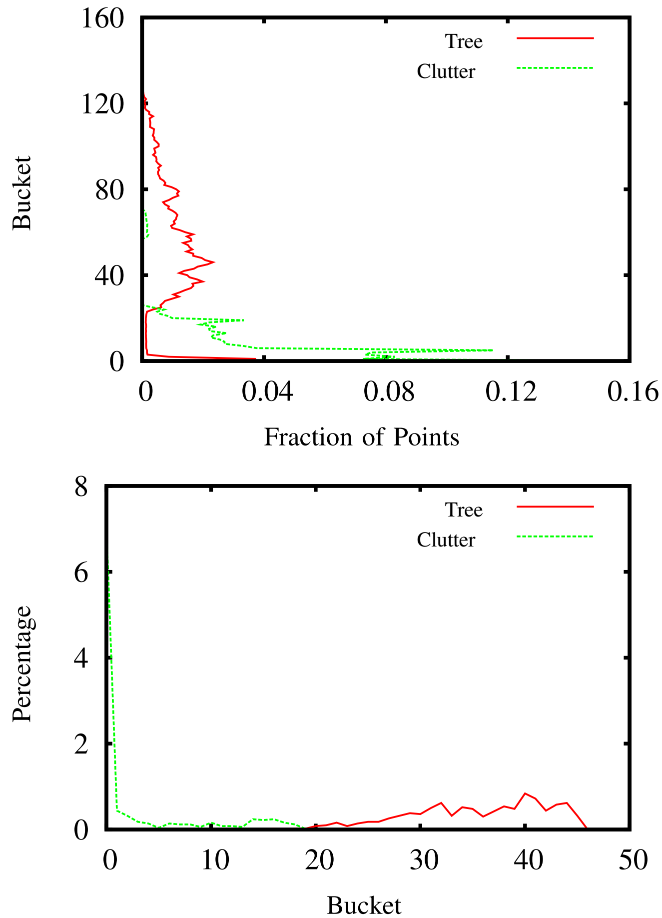


Fig. 3. Top: Histogram feature where each bucket contains the number of points in a horizontal slice of 10cm in proportion to the total number of points. Bottom: Eigenvalue histogram of the smallest eigenvalue of 500 randomly selected points and their neighborhood.

the supposed supposed trunk. Fig. 3 shows the histogram feature for two exemplary clusters. The vegetation is clearly distinguishable from the clutter. As a second feature describing the shape of vegetation in a more encompassing fashion we also implemented a histogram of eigenvalues. Vegetation of any type should generally appear volumetric in the point cloud, while clutter as discussed above stems primarily from close to planar structures. This is captured by a principal component analysis (PCA), i.e., the eigenvalues $\lambda_1, \lambda_2, \lambda_3$ of the mean centered covariance matrix. The ratio $\frac{\lambda_1}{\sum_i \lambda_i}$ of the smallest eigenvalue λ_1 to the sum of eigenvalues is a measure of the planarity of the measured points. We compute this ratio of 500 randomly selected points and their neighborhood and record it in a histogram of 50 evenly spaced buckets. The result of this computation for two exemplary clusters is presented in Fig. 3.

As an alternative to the two histogram features, we devised a fast classification based on the PCA of the entire cluster. As shown in Fig. 4 the ratios of the two smallest eigenvalues are a rough predictor of the class of the cluster. We therefore classify segments using two threshold. If the ratio of second

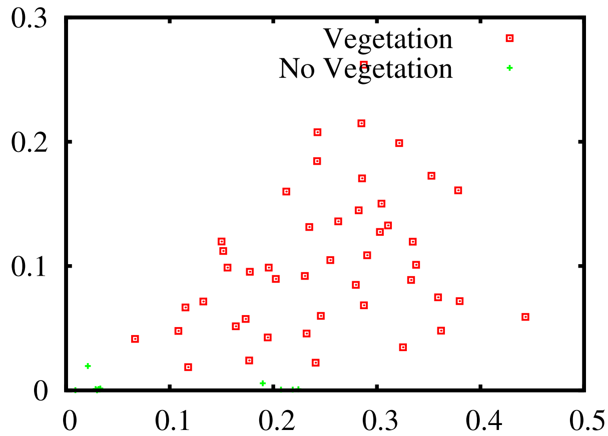
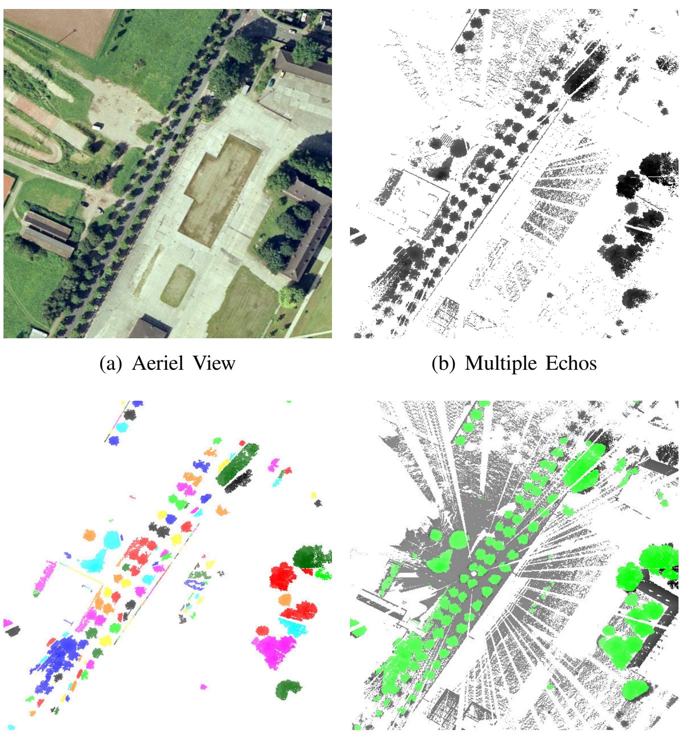


Fig. 4. Smallest eigenvalue plotted against the 2nd smallest eigenvalue of clusters in a scan of the city data set. Clusters that do not contain trees tend to be planar and have smaller eigenvalues.



(a) Aeriel View

(b) Multiple Echos

(c) Clusters

(d) Detected Trees

Fig. 5. Scans from the campus data set. a) Aerial view of the scanned regions. Image courtesy of Google Maps. b) Overview of a single scan containing only multiple echoes. Single as well as last echoes are not shown. c) The same scan clustered, where each cluster is colored distinctly. d) The entire scan from above. All identified trees are colored green.

smallest eigenvalue is below 0.05 or the ratio of the smallest is below 0.01 the cluster is labeled non-vegetation. These values have proven themselves to be reliable across many different environments.

TABLE I
RESULTS FOR EACH CLASSIFICATION STRATEGY.

(a) SVM with point density histogram feature

Clusters	true	false
positive	0.5537	0.0473
negative	0.0959	0.3031
$\sum$	0.6496	0.3504

Points	true	false
positive	0.1070	0.0025
negative	0.4689	0.4216
$\sum$	0.5759	0.4241

(b) SVM with eigenvalue histogram feature

Clusters	true	false
positive	0.7775	0.0115
negative	0.1317	0.0793
$\sum$	0.9092	0.0908

Points	true	false
positive	0.4788	0.0009
negative	0.4705	0.0499
$\sum$	0.9493	0.0508

(c) Eigenvalue threshold classification

Clusters	true	false
positive	0.8133	0.0166
negative	0.1266	0.0435
$\sum$	0.9399	0.0601

Points	true	false
positive	0.5252	0.0009
negative	0.4704	0.0034
$\sum$	0.9956	0.0043

IV. RESULTS

The algorithm is evaluated with all three classifications on a data set comprising 9 scans taken on our university campus. This data set contains a wide variety of trees and shrubs. Trees were of all sizes between 3 and 30 meters. The data set is very challenging due to the trees' proximity to buildings and the presence of clutter like cars and people. An overview of the data set is given in the top of Fig. 6.

For the histogram features, a training data set in a different urban environment was acquired. 13 scans were taken in the historic city center with enough overlap to allow for a complete registration. The result of scan matching is presented in Fig. 7. The city data set also contains a wide variety of trees and clutter. The amount of clutter is clearly much higher in the city data set as the scans were taken on a highly populated street, with many trams, cars, people and other. SVM training and classification was performed by the SVM implementation of the Bioinformatics Toolbox for Matlab with the Gaussian Radial Basis Function.

Ground truth for both data sets was manually acquired. We identified regions in the scan that represent trees, shrubs and parts thereof. Points falling into these regions should be identified positively. Note: Lawn is not considered vegetation for this classification task. All points that are not classified thusly are considered non-vegetation.

The effectiveness of the clustering algorithm can be observed in Fig. 5. With the exception of highly occluded and distant trees (more than 100 meters from the scan position) all of the vegetation is present within the clusters. A disadvantage of this clustering approach is manifested in the upper right corner of the scan. A large tree is split up into several segments due to occlusion caused by a trunk closer to the scanner. In this case, each segment was correctly identified as vegetation, however smaller segments stand a higher chance of being wrongly classified.

The results for the evaluation of the SVM based classifications are given in Table I. A single tree may be present in

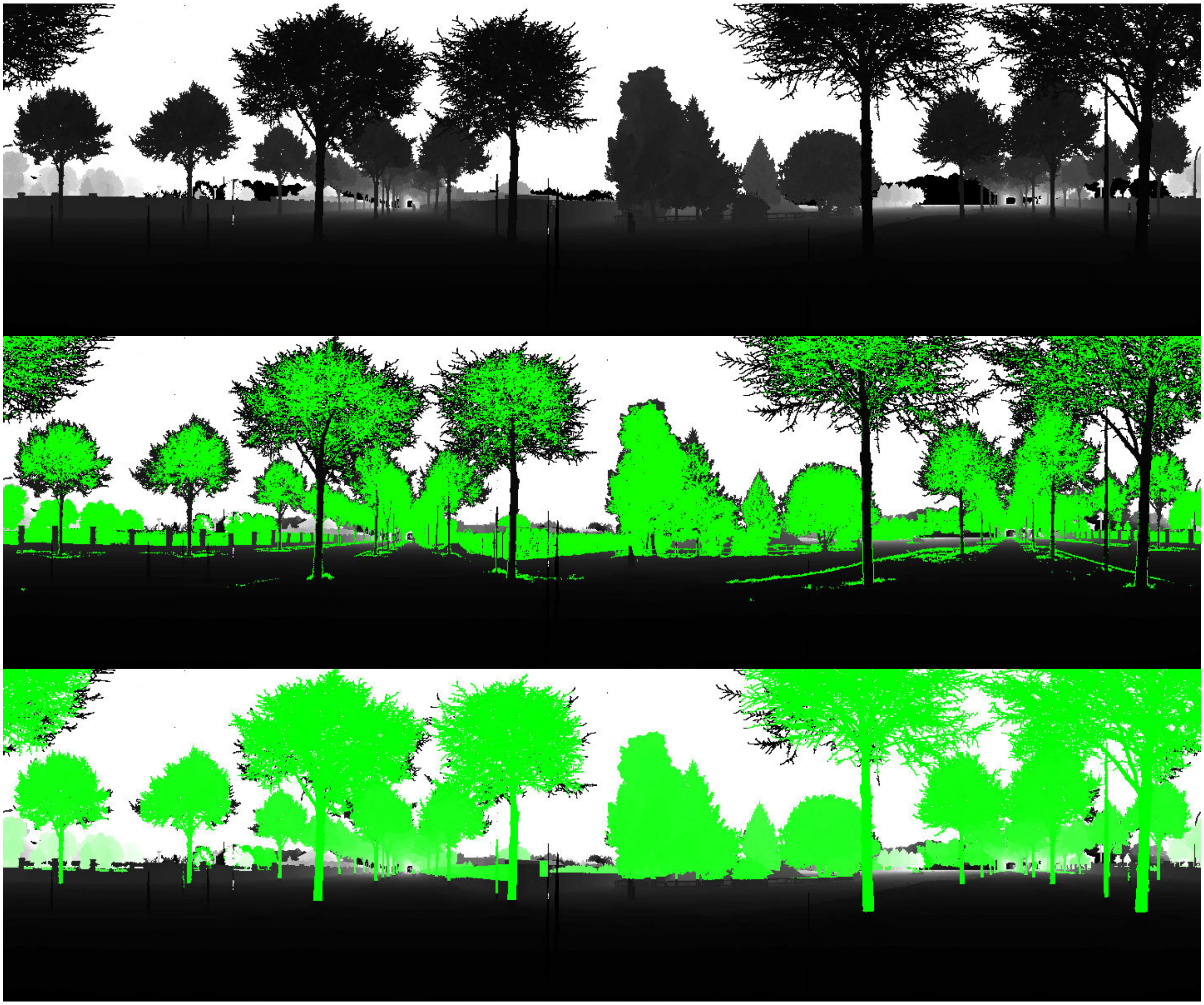


Fig. 6. The Campus data set consisting of 9 scans with approx. 22 million points each used for evaluation. In entirety the data set contains $186,172,085$ with approx. 53% vegetation. Top: An overview of 8 of the scans shown as a range image. Bottom: The last scan from the Campus data set in the perspective of the laser scanner. The first is a depth image where each pixel is shaded according to distance. In the second depth image pixels containing multiple echoes are colored green. In the third depth image points labeled as vegetation are represented as green.

several clusters, and clusters are of varying sizes. Therefore, accuracy is measured both in terms of correctly/incorrectly labeled clusters as well as points. Interestingly, the point density histogram fares badly on the campus data set. This is in contrast to earlier trials on the city data set, where the overall accuracy was high. The density histogram is evidently not adequate for general vegetation detection across different domains. The city data set's vegetation is generally smaller and contains less shrub-like vegetation. This is likely the cause of the lower performance.

The histogram of eigenvalue features fares far better in the evaluation. Points were classified correctly with an accuracy of about 95 percent, suggesting that this feature vector distinguishes vegetation very well from clutter. Interestingly, the simple eigenvalue threshold achieves even better results as shown in Table I bottom row. Since computing the eigenvalues of the covariance matrix of the points is linear in the number of points, the most efficient feature also proved to be the most accurate.

V. Conclusions and Future Work

This paper aimed at bringing the Full Wave Analysis for scene interpretation to the attention of researchers in the robotics community. We presented an approach to detect vegetation with a very high success rate that is based on the inner echoes returned by a FWA capable laser range finders. We evaluated three approaches to classify point clouds, two of which were based on SVM classification. However, the most

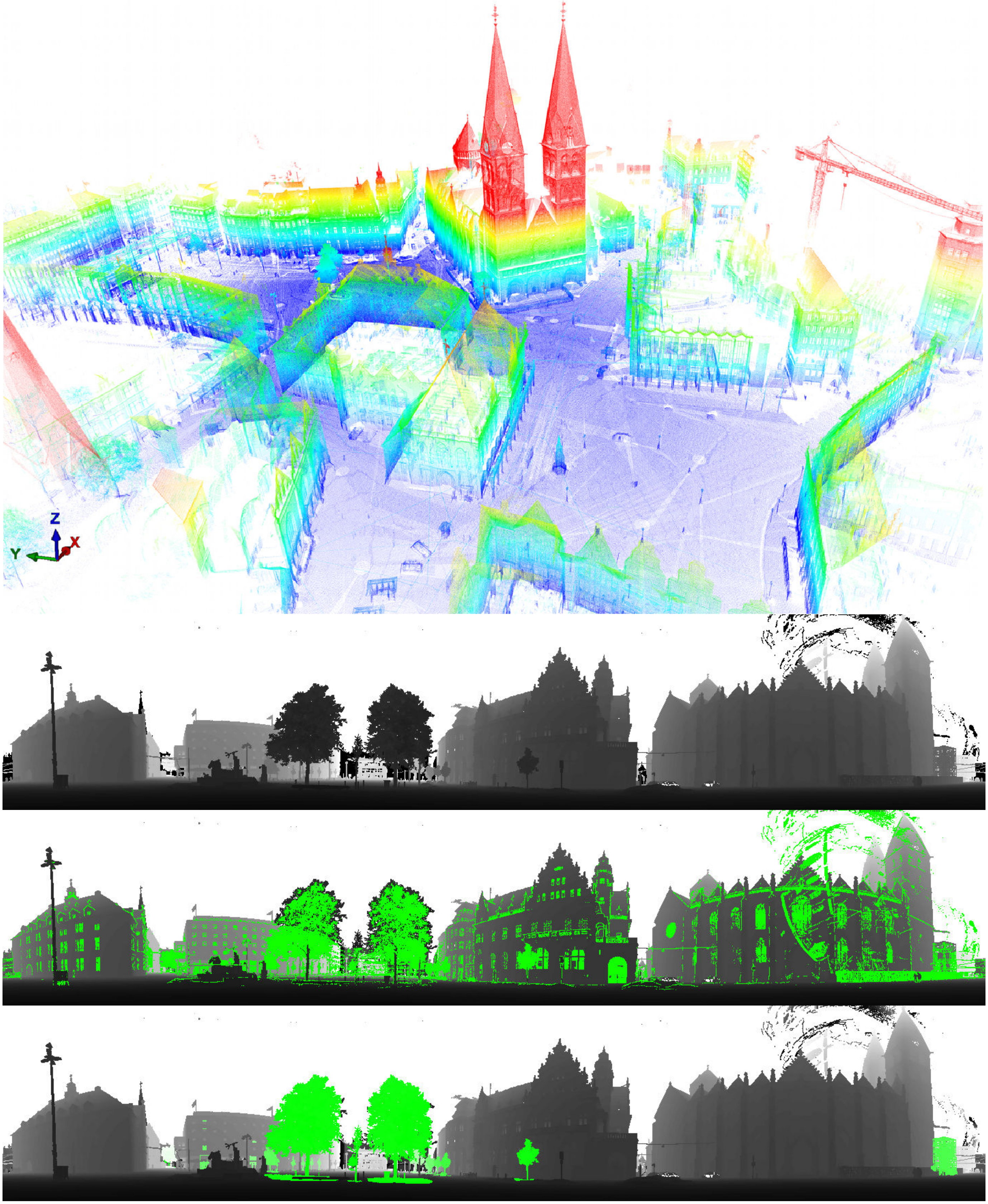


Fig. 7. City data set comprising 13 scans with approx. 22 million measurements each used for training. Top: The data set (available at [7]) was registered with the 6D SLAM algorithm of Nüchter et al. [8]. Points are colored according to their height. Second: Scan 2 from the city data set as a 1440×400 range image representation shaded according to distance. Third: Areas of the scan with inner echoes are indicated in green. Bottom: Areas in green were classified as vegetation by the eigenvalue threshold of the clusters. The incorrectly classified scaffold in the right is a notorious issue due to its structure. The window frame on the left being labeled vegetation is a consequence of the cluttered rooms behind the windows, i.e. window blinds, decoration and furniture.

efficient approach in terms of runtime, i.e., the threshold of eigenvalues was also the most accurate one. The histogram of eigenvalues as the most computationally expensive is a viable alternative for classification as it achieves a well enough accuracy.

In the future, the post processing will be expanded to include strategies to identify individual trees as in [4]. In addition to providing semantically richer labels, this should further increase the accuracy of the vegetation identification. Furthermore, regions should be merged to compensate for occlusions that occur frequently due to larger branches. This in turn will drive down the rate of false negatives, as small regions from only parts of a tree make identification in isolation difficult.

Further future work will also concentrate on life-long mapping. We will verify our algorithm on data sets taken in different seasons to analyze the green of leaves influence. This will result in a season independent mobile mapping system.

Acknowledgment

This work was partially supported by the SEE-ERA.NET project ThermalMapper under the project number ERA 14/01.

References

[1] S. Barnea, S. Filin, and V. Alchanatis. A Supervised Approach for Object Extraction from Terrestial Laser Point Clouds Demonstrated on Trees. *Photogrammetric Image Analysis*, 36(3):135 – 140, 2007.

[2] M. B. Dillencourt, H. Samet, and M. Tamminen. A General Approach to Connected-Component Labeling for Arbitrary Image Representations. *J. ACM*, 39:253–280, April 1992.

[3] J. Holmgren and Å. Persson. Identifying species of individual trees using airborne laser scanner. *Remote Sensing of Environment*, 90(4):415 – 423, 2004.

[4] J. Reitberger and C. Schnörr and P. Krzystek and U. Stilla. 3D Segmentation of Single Trees Exploiting Full Waveform LIDAR data. 64(6):561 – 574, 2009.

[5] E. H. Lim and D. Suter. 3D Terrestrial LIDAR Classifications with Super-Voxels and Multi-Scale Conditional Random Fields. *Computer-Aided Design*, 41(10):701 – 710, 2009.

[6] D. Munoz, J. A. Bagnell, N. Vandapel, and M. Hebert. Contextual Classification with Functional Max-Margin Markov Networks. In *Computer Vision and Pattern Recognition (CVPR)*, 2009.

[7] A. Nüchter and K. Lingemann. http://kos.informatik.uni-osnabrueck.de/3Dscans, 2010.

[8] A. Nüchter and K. Lingemann. http://slam6d.sourceforge.net, 2010.

[9] N. Pfeifer, B. Gorte, and D. Winterhalder. Automatic Reconstruction of Single Trees from Terrestrial Laser Scanner Data. In *ISPRS XXth World Congress*, 2004.

[10] I. Posner, D. Schröter, and P. Newman. Online Generation of Scene Descriptions in Urban Environments . *Robotics and Autonomous Systems*, 56(11):901 – 914, 2008.

[11] RIEGL Laser Measurement Systems. http://www.riegl.com, 2010.

[12] U. Stilla. Full Waveform Laser Scanning: Technologie, Datenprozessierung und Anwendungen. In *Geodätische Kolloquium*, 2009.

[13] R. Triebel, K. Kersting, and W. Burgard. Robust 3D Scan Point Classification using Associative Markov Networks. In *ICRA '06*, pages 2603–2608, 2006.

[14] W. Wagner, M. Hollaus, C. Briese, and V. Ducic. 3D Vegetation Mapping using Small-Footprint Full-Waveform Airborne Laser Scanners. *International Journal of Remote Sensing*, 29(5):1433 – 1452, 2008.

[15] X. Yu and J. Hyyppä and H. Kaartinen and M. Maltamo. Automatic Detection of Harvested Trees and Determination of Forest Growth using Airborne Laser Scanning. *Remote Sensing of Environment*, 90(4):451 – 462, 2004.