ISPRS Journal of Photogrammetry and Remote Sensing 67 (2012) 134–147

Contents lists available at SciVerse ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Urban vegetation detection using radiometrically calibrated small-footprint full-waveform airborne LiDAR data

Bernhard Höfle [a,*], Markus Hollaus [b], Julian Hagenauer [a]

[a] University of Heidelberg, Institute of Geography, Chair of GIScience, Berliner Strasse 48, 69120 Heidelberg, Germany
[b] Vienna University of Technology, Institute of Photogrammetry and Remote Sensing, Gußhausstraße 27-29, 1040 Vienna, Austria

ARTICLE INFO

Article history:
Received 1 May 2011
Received in revised form 23 November 2011
Accepted 14 December 2011
Available online 8 January 2012

Keywords:
Laser scanning
LiDAR
Calibration
Vegetation
Object based image analysis
Full-waveform

ABSTRACT

This paper introduces a new GIS workflow for urban vegetation mapping from high-density (50 pts./m^2) full-waveform airborne LiDAR data, combining the advantages of both raster and point cloud based analysis. Polygon segments derived by edge-based segmentation of the normalized digital surface model are used for classification. A rich set of segment features based on the point cloud and derived from full-waveform attributes is built, serving as input for a decision tree and artificial neural network (ANN) classifier. Exploratory data analysis and detailed investigation of the discriminative power of selected point cloud and full-waveform LiDAR observables indicate a high value of the occurrence of multiple distinct targets in a laser beam (i.e. 'echo ratio') for vegetation classification (98% correctness). The radiometric full-waveform observables (e.g. backscattering coefficient) do not suffice as single discriminators with low correctness values using a decision tree classifier (⩽72% correctness) but higher values with ANN (⩽95% correctness). Tests using reduced point densities indicate that the derived segment features and classification accuracies remain relatively stable even up to a reduction factor of 10 (5 pts./m^2). In a representative study area in the City of Vienna/Austria the applicability of the developed object-based GIS workflow is demonstrated. The unique high density full-waveform LiDAR data open a new scale in 3D object characterization but demands for novel joint strategies in object-based raster and 3D point cloud analysis.

1. Introduction

Airborne LiDAR, also referred to as Airborne Laser Scanning, has already proven to be a state-of-the-art technology for high-resolution and highly accurate topographic data acquisition with active and direct determination of the Earth surface elevation (Vosselman and Maas, 2010). Most commercial sensors are multi-sensor systems, which utilize the laser pulse time-of-flight principle for distance measurement combined with a Global Navigation Satellite System (GNSS) assessing the sensor position and an Inertial Measurement Unit (IMU) that is recording the sensor attitudes with high frequency (Lemmens, 2009; Wehr and Lohr, 1999). The laser scanner emits nanosecond long single infrared laser pulses with a certain divergence of the beam, resulting in an illuminated area (i.e. footprint) at the object with typical diameters between 0.2 m and 1 m depending on the distance between sensor and target and beam divergence (Wehr and Lohr, 1999). For example, a beam divergence of 0.5 mrad leads to a footprint diameter of 0.5 m at a distance of 1000 m. Illuminated objects in different elevations can cause more than one echo that can be recorded at the receiver. Generally, two different generations of receiver units exist: (1) discrete echo recording systems, which are able to record multiple echoes on-line and typically store up to four echoes per laser shot (Lemmens, 2009) and (2) full-waveform (FWF) recording systems capturing the entire time-dependent variation of the received signal power with a defined sampling interval such as 1 ns (Mallet and Bretar, 2009; Wagner et al., 2006). While traditional systems determine and record single echoes in real-time, full-waveform recording allows the detection of single echoes (i.e. the decomposition of the waveform) in a post-processing step with higher range determination accuracy for complex geometries, e.g. as shown for stepped surfaces by Jutzi and Stilla (2006). Additionally, full-waveform systems may enable the derivation of a higher number of intermediate echoes compared to discrete systems with first/last echo capability (Chauve et al., 2007). Furthermore, the derivation of physical quantities for each echo is possible with full-waveform data, such as the signal amplitude, echo width and backscatter cross-section (Wagner et al., 2006; Section 2.1).

* Corresponding author. Tel.: +49 6221 54 5594; fax: +49 6221 54 4529.
E-mail address: hoefle@uni-heidelberg.de (B. Höfle).

0924-2716/$ - see front matter © 2011 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS) Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.isprsjprs.2011.12.003

Airborne LiDAR data have been used in various applications in urban environments, particularly aiming at mapping and modeling the city landscape in 3D with its artificial land cover types such as buildings (Oude Elberink and Vosselman, 2009), power lines (Melzer, 2007), bridges (Sithole and Vosselman, 2006) and also non-elevated objects such as roads (Clode and Rottensteiner, 2007). Furthermore, as urban environments are active regions with respect to alteration in land cover, change detection plays important role for the assessment and quantification changes (e.g. construction of buildings) and following map updating (cf. Matikainen et al., 2010). The high potential of airborne LiDAR data is witnessed even in urban traffic analysis where single-pass airborne LiDAR data can be used to detect moving objects (e.g. cars) by applying 3D segmentation such as adaptive mean shift (Yao et al., 2011). Also detection of natural surfaces within urban areas in LiDAR data such as vegetation (Höfle and Hollaus, 2010; Rutzinger et al., 2008) and water surfaces (Höfle et al., 2009a; Yuan and Sarma, 2011) have gained increasing interest. In particular vegetation monitoring and tree inventory play an important role in modern urban spatial data management, as many applications make use of detailed up-to-date data sources, such as mitigation of noise (cf. Stoter et al., 2008) and pollution, the creation of 3D city models (Haala and Kada, 2010) and urban GIS database compilation. For example, 3D city models and detected 3D objects such as trees and buildings enable further analysis such as solar potential assessment of urban roof planes. Jochem et al. (2009) show that solar potential assessment can be improved, if roof planes and 3D tree segments are detected in 3D LiDAR point clouds and are included in solar radiation modeling. Assigning vegetation "transparency values" (i.e. penetrability of laser shots) to the trees allows improving the modeling of solar energy rates on roof planes nearby high vegetation.

To date, high point density ($\sim$50 echoes/m^2) LiDAR data are increasingly available, in particular for urban areas. This high point density data constrains the operational usage of existing computationally intensive 3D point cloud analysis strategies for large areas due to the projected long processing times. This study presents a novel GIS framework for full-waveform LiDAR data land cover classification, making use of the strengths and advantages of both image and point cloud analysis (cf. Höfle et al., 2009b). The main objective is to combine object-based image (Blaschke, 2010) and point cloud analysis to derive a GIS-ready vegetation mask (i.e. binary classification) in a complex urban area including high vegetation (i.e. trees). Particular focus is laid on assessing the potential of full-waveform LiDAR echo attributes for segment-based vegetation classification. Echo attributes that can be derived directly from the recorded waveforms (e.g. echo width) as well as attributes that have been derived by radiometric calibration based on physical concepts known from radar remote sensing (e.g. backscattering coefficient) are considered. This paper is based on initial investigations presented by Höfle and Hollaus (2010).

Background on full-waveform observables and related work on urban scene classification are given in Section 2. Thereafter, Section 3 introduces the study area and the available LiDAR and reference datasets. The methodology and the developed workflow are presented in Section 4. In Section 5 the results are described and discussed, and the last section concludes with the major findings and future work considerations.

2. Background and related work

2.1. Full-waveform LiDAR observables

Processing of the recorded waveforms includes the localization and extraction of echoes representing single distinct targets in the path of laser beam. Several methods have been developed so far, such as decomposing the waveform into Gaussian components (Wagner et al., 2006) or modeling the waveform using a stochastic framework with a set of parametric functions (Mallet et al., 2009). Typically, for each detected and modeled echo the echo width and amplitude are derived. From a target perspective the echo width represents the range distribution of individual scatterers contributing to an echo. Generally, the echo width is given in the temporal domain (mostly nanoseconds) as Full-Width-at-Half-Maximum (FWHM) or standard deviation in the case of Gaussian decomposition (Alexander et al., 2010). The signal amplitude is a measure for the strength of the received echo comprising the target area and reflectance in the laser's wavelength. In order to use the amplitude for surface classification, correcting for e.g. range dependency, system settings and atmospheric attenuation is required (cf. Höfle and Pfeifer, 2007). Further details on full-waveform LiDAR principles, systems and waveform processing are described in the extensive review of Mallet and Bretar (2009) and references therein.

With full-waveform LiDAR systems recording the complete return signal radiometric calibration as common in radar remote sensing is feasible. Wagner et al. (2006) adapted the radar equation (Jelalian, 1992) to ALS and presented the theoretical considerations for backscatter cross section (σ [m^2]) estimation and radiometric calibration. In Wagner (2010) these considerations have been extended by promoting the backscattering coefficient γ [m^{2}m^{-2}] as additional physical quantity for Earth surface characterization and particularly to be used for radiometric calibration. The coefficient γ is defined as σ normalized to the cross section of the incoming beam (i.e. footprint area; Wagner, 2010). The advantage of using γ compared to σ is the increased comparability between sensors and flight campaigns with e.g. different beam divergence and range, and thus footprint area (Alexander et al., 2010; Wagner, 2010). Additionally with γ there is no need for local incidence angle estimation, which is often not feasible for e.g. complex targets and vegetation.

Geometric and radiometric information are often seen to be complementary, which is true to a certain degree, but they are de facto not fully independent of each other. For example, range accuracy and separability of echoes are dependent on the target's backscatter properties and geometry (e.g. vertical distribution), the used system settings (e.g. emitted pulse energy and waveform sampling interval; Lin and Mills, 2010) and the echo detection method (Stilla and Jutzi, 2008). Hence, the recorded waveforms result from mixed effects of geometric (e.g. slope) and radiometric object properties (e.g. reflectance), which has to be considered when working with FWF data (Steinvall, 2000).

2.2. Detection and classification of vegetation

Compared to predominant studies on vegetation detection and characterization mainly in purely forested areas (e.g. Hyyppä et al., 2001), this study concentrates on urban environments, which have a high structural complexity with a multitude of different objects (e.g. temporary objects, vegetation on top of buildings, road signs, power lines and cables). Previous studies using airborne LiDAR for vegetation and single tree detection, respectively, used image-based methods, e.g. including orthophotos (Bretar and Chehata, 2007; Hirschmugl et al., 2007; Iovan et al., 2008; Rottensteiner et al., 2007) and LiDAR derived raster layers, e.g. first-last-pulse difference (Liang et al., 2007). Tóvári and Vögtle (2004) use a region growing segmentation on the normalized digital surface model in order to detect buildings and vegetation. The classification is based on segment features such as gradients on segment borders, height texture, first-last-pulse-difference, segment size and shape as well as signal intensity. They classify vegetation with classification rates between 86% and 96%. Secord and Zakhor (2007) merge aerial image and LiDAR data to detect trees in a workflow consisting of a prior image-based segmentation using region growing followed

by an object-based classification step using weighted support vector machines (SVM). Segment features for classification are derived from the aerial image (e.g. hue, saturation and value) as well as from the point cloud such as height, local height variation defined as difference between maximum and minimum height in a 1.5×1.5 m^2 area, and the x- and y-components of the estimated surface normal vector. They found that (i) the segment-based classification outperforms a direct pointwise classification without prior segmentation and (ii) larger segments improve classification results compared to small segments with only few input values for segment feature calculation.

Due to the three-dimensional nature of urban landscapes (e.g. tree on top of building, roof above tree), several authors propose to use the original 3D laser point cloud for urban land cover classification. Carlberg et al. (2009) classify four urban target classes – water, ground, roof and trees – where trees can be classified with >89% precision. They use a cascade of binary classifiers based on unsupervised region growing followed by supervised, segment-based classification. Trees are detected by identifying scattered regions in the 3D point cloud in the sense of 3D spatial distribution of the laser points (e.g. taking into account the eigenvalues derived from principle component analysis over a segment). Also working directly in the 3D point cloud, Rutzinger et al. (2007) derived a roughness parameter as indicator for urban vegetation for each laser point by using the ratio between 3D and 2D point density in a fixed distance neighborhood. Their object-based point cloud analysis (OBPA) applies a 3D point cloud region growing segmentation with subsequent segment-based classification into vegetation and non-vegetation and is successfully tested on a dataset with moderate laser point density of ca. 4 pts./m^2.

Most recent approaches include FWF information (e.g. echo width and amplitude) for urban object detection in the 3D point cloud directly and for 3D segmentation of single trees (Gross et al., 2007; Reitberger et al., 2009). Rutzinger et al. (2008) developed a workflow for point cloud based classification of single laser echoes from high vegetation (i.e. trees and shrubs) based on FWF echo attributes in an urban study area with high point density (>20 pts./m^2). Seeded region growing segmentation solely using the echo width attached to the laser points is used to derive homogeneous 3D point cloud segments. Then, segment attributes are derived by aggregation of the segment's laser echo attributes such as echo width, amplitude and roughness parameters. Finally, different decision trees are tested for deriving a binary classification (vegetation/non-vegetation) of the 3D point cloud segments achieving completeness and correctness rates above 90%. Mallet et al. (2008) additionally include further urban land cover classes (e.g. building and ground) in their single laser point based classification approach. They train a SVM with full-waveform laser point attributes such as echo amplitude, width and shape, as well as e.g. the number of distinct targets per laser shot. They achieve a discrimination of buildings, vegetation and ground with 92% accuracy. Furthermore, clear trends in the full-waveform attributes for the class vegetation are characterized: medium signal amplitude, high echo width and variable shape of the echoes as well as an increasing number of multiple targets indicate vegetation. Additionally Guo et al. (2011) include aerial multispectral image data and the echo backscattering cross-section (cf. Wagner, 2010) in a point cloud based classification using the Random Forests classifier. The results demonstrate the advantages of combining aerial imagery and FWF LiDAR data as well as the benefits of FWF LiDAR features for dense urban scenes. They characterize vegetation by large height difference within one laser shot and increased number of echoes. Furthermore, echo amplitude and cross-section are found to be important for vegetation detection, where the cross-section exhibits low values for vegetation.

Alexander et al. (2010) apply the radiometric FWF features (i) echo amplitude, (ii) backscatter cross section [m^2] and (iii) backscatter coefficient [m^{2}m^{-2}] (Wagner, 2010) for point cloud based urban area classification using decision tree classifiers with different input datasets. They defined six target classes comprising grass (<0.5 m height), shrubs (0.5–2.5 m), trees (>2.5 m), buildings either with flat or pitched roofs, and roads. They conclude that classification accuracy of the results using the backscatter coefficient as additional input is higher than with the results achieved with amplitude or backscatter cross section. Particularly the transferability to different test sites with different acquisition settings (e.g. flying height and scan angle) strengthen the usage of the radiometrically calibrated backscatter coefficient for urban land cover classification.

Besides the task of separating vegetation from non-vegetation objects, recent studies focus on assessing the potential of FWF LiDAR for tree species classification (cf. Heinzel and Koch, 2011), assuming a forested study area or a prior removal of non-vegetation objects. Höfle et al. (2008) showed the suitability of aggregated FWF attributes attached to segments derived from LiDAR raster analysis for tree species discrimination. Segment-based statistical values (e.g. average, coefficient of variation) of echo width and backscatter cross section are found to be a suitable discriminator between tree species (e.g. beech, larch and oak) in forested areas.

To sum up, the majority of presented studies does not calibrate the radiometric attributes and thus does not retrieve physical observables for classification. As consequence, the derived and published classification thresholds and object descriptions remain specific for the study site, used sensor and settings, and hence lack of transferability. Furthermore, object-based approaches are increasingly presented, which exhibit more significant features and robust classification compared to single echo classification approaches. Object-based image and object-based point cloud analysis (Rutzinger et al., 2008) make use of such larger units (i.e. segments) and thereby aggregated segment attributes. For example, a single backscatter cross section value does not allow for tree species determination, whereas the average cross section per tree segment can give a good estimate (cf. Höfle et al., 2008).

3. Materials

3.1. Study area

The area of investigation is located in the city centre of Vienna (Austria) and comprises the three city parks: Rathauspark, Volksgarten and Burggarten (Fig. 1). The study area is characterized by a great variety of planting and tree species e.g. alley of trees, short-cut trees, hedgerows and shrubs. Deciduous trees are predominant such as beech (*Fagus sylvatica*), Norway maple (*Acer platanoides*), plane (*Platanus acerifolia*), linden (*Tilia cordata*, *platyphyllos*) and chestnut (*Aesculus hippocastanum*) and sparsely coniferous species. The study area mainly contains large building complexes (e.g. city hall, Burgtheater and parliament) as well as artificial objects such as fences, cars, power lines, park benches. Furthermore, a high amount of people are usually present in this central part of the city and are therefore also visible in the airborne LiDAR elevation datasets.

3.2. Airborne LiDAR and reference datasets

The full-waveform airborne LiDAR data were obtained in the framework of the city-wide LiDAR acquisition campaign in the winter season 2006/2007 under leaf-off conditions. The RIEGL LMS-Q560 system using near-infrared (1550 nm) laser pulses with

Fig. 1. Study area located in the city centre of Vienna, Austria. Diversity in tree species and structure are shown in detailed photographs (B1–B3).

a pulse width of 4 ns was employed. The chosen LiDAR system settings comprise a pulse repetition frequency of 200 kHz, scan angle range of ±22.5° and a beam divergence of 0.5 mrad. The average flying height above ground was 500 m resulting in a footprint diameter of about 0.25 m on ground. Full-waveform recording was done with 1 ns temporal resolution and Gaussian decomposition of the waveforms was performed (Wagner et al., 2006) using RIEGL's software RiANALYZE by the company *AREA Vermessung ZT-GmbH* in the course of the city-wide LiDAR project. Theoretically, the number of echoes that can be detected for a laser shot is not limited by full-waveform recording sensor systems, and therefore, the number of detected echoes is generally higher as with traditional discrete echo recording, particularly in high vegetation (cf. Mallet et al., 2008). The average echo density of the LiDAR dataset covering the study area exhibits 50 echoes per m^2 with about 6.8% first echoes, 1% intermediate echoes (e.g. 2nd, 3rd echo), 6.8% last and 85.4% single echoes (i.e. only one reflection per shot). Furthermore, a Digital Terrain Model (DTM) with 0.5 m resolution is made available by terrain point filtering and interpolation using the iterative robust interpolation approach implemented in the software SCOP++ (SCOP, 2010). For classification (training data) and evaluation, a reference dataset of single tree positions of alley trees and the official land cover map of Vienna are available (MA41, 2011): Mehrzweckkarte (MZK), Flächenmehrzweckkarte (FMZK).

4. Methodology

The high point density of the LiDAR dataset of about 50 pts./m^2 makes a pure point cloud based vegetation detection procedure unfeasible for large areas (400 km^2 for entire Vienna) due to the projected long computation times. Therefore, a combined object-based image and point cloud approach is used, which takes advantage of both, fast raster data processing compared to processing of the unstructured 3D point cloud and detailed, 'interpolation-free' 3D point cloud based information extraction including full-waveform laser point attributes. Additionally, the high point density of the point cloud enables the derivation of high-resolution raster layers (e.g. DTM with 0.5 m cell size), providing a high and sufficient planimetric accuracy of the final results. The developed workflow is implemented in GRASS GIS (GRASS Development Team, 2010) and the OPALS software (Mandlburger et al., 2009; IPF, 2011). Fig. 2 shows the single processing steps that are described in detail below.

4.1. Radiometric calibration procedure

The radiometric calibration in this study is performed by following the procedure of Briese et al. (2008) and Lehner and Briese (2010) where a portable reflectometer and Spectralon targets are used to estimate the calibration constant of in situ radiometric reference targets (e.g. asphalt and roof areas). Meteorological data for atmospheric attenuation estimation was obtained from three meteorological stations located in the city of Vienna. The large asphalt areas of the Maria Theresia casern south of Schönbrunn are used as reference surfaces. The surface reflectance of the reference area is measured using a RIEGL reflectometer in the wavelength of the airborne LiDAR system (i.e. 1550 nm), where a mean reflectance of 23.5% could be measured for the large parade yard of the casern. Next, a mean calibration constant value is assessed, which is then applied to all flight strips of the study area. The processing procedure of (i) calibration constant estimation and (ii) the calculation of calibrated values are documented and implemented in the OPALS module *opalsRadioCal* (IPF, 2011). Furthermore, the signal amplitude is normalized by range, i.e. corrected for the loss of signal of one over range squared (cf. Höfle and Pfeifer, 2007), and is attached to each echo using a normalization range of 500 m. Note that from a theoretical point of view the normalization using range squared is only valid for area-extended targets (Jelalian, 1992). This *normalized amplitude* is derived and included in this study because it can be easily derived without any assumptions and references data. Further it has already proven a high suitability for classification of natural surfaces.

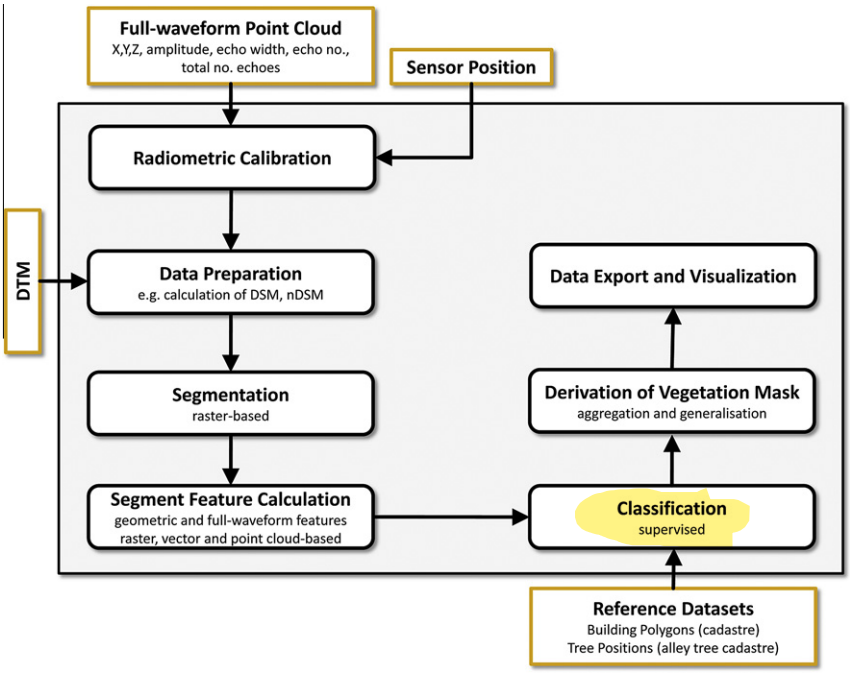

Fig. 2. Workflow of urban vegetation detection combining raster and full-waveform point cloud input data.

4.2. Data preparation

The input laser point cloud is classified into first, last, single and intermediate echoes. The available laser point attributes echo number (*ENo*) and total number (*TNo*) of echoes of the corresponding laser shot are used for assigning each echo to the classes defined as follows: single echoes: $ENo = TNo = 1$; first echoes: $ENo = 1$ && $TNo > 1$; intermediate echoes: $ENo > 1$ && $ENo != TNo$. In a next step, a digital surface model (DSM) is derived by using the maximum height per 0.5 m cell. Once having the DSM, a normalized DSM (nDSM) can be calculated by subtracting the DTM from the DSM in order to derive relative object heights. Empty cells are assigned a normalized height of 0 m. In urban areas generally complete and high quality cadastral data layers are available, which could be used as a priori mask for the subsequent steps. The buildings can be extracted from the official land cover map and could be used optionally to exclude these areas from further investigation. In our study the building polygons are used for evaluation and thus are not considered as input (i.e. mask).

4.3. Raster-based segmentation

The segmentation is based on an object-based raster analysis that makes use of an edge-based segment boundary detection applied to the high-resolution nDSM. Höfle et al. (2008) already tested the procedure in densely forested areas where no buildings are present. In this study this procedure is transferred to the densely built-up urban area of Vienna. The main idea of the segmentation is to delineate convex objects in the nDSM by finding concave edges in between single objects (e.g. trees). Additionally, constraints on normalized height (nDSM > 1.0 m) as well as the number of distinct targets are taken into account. For this purpose, multiple reflections are parameterized by a multi-echo based echo ratio (ER_{ME}) defined as (Eq. (1)):

$$ER_{ME}[\%] = (n_{first} + n_{intermediate})/(n_{last} + n_{single}) \times 100 \quad (1)$$

where n is the respective number of echoes per cell. If no laser echo lies within a raster cell, ER_{ME} is set to zero and if no last and single echoes are found in a cell, ER_{ME} is set to 100 [%], exhibiting a high vertical height extension as well as transparency of the object in the sense of penetrability of parts of the laser beam through small gaps. High values of ER_{ME} indicate vegetation, which is characterized by a high number of first and intermediate echoes compared to other elevated objects in urban areas (e.g. buildings). Other definitions of echo ratios such as a purely geometric computation (e.g. Höfle et al., 2009b; Rutzinger et al., 2007) where number of local neighbors are counted in a defined 2D and 3D neighborhood are computationally far more expensive than the presented ER_{ME} approach. Additionally, ER_{ME} is more robust towards multi-temporal effects caused by temporary or moving objects. For example, a car scanned in one flight strip but not present in a second strip covering the same area will cause an artificial object with a certain height above terrain and a very high penetrability (Fig. 3c). However, the echoes still have low echo widths and a low number of multiple echoes per shot and thus ER_{ME} is not affected by temporary objects such as the geometry-based version (Fig. 3).

The edge detector searches for edges defined as the skeleton of concave areas in a curvature map. First, a curvature map of the minimum curvature in direction perpendicular to the direction of maximum curvature in a certain neighborhood of each cell (e.g. window of 7×7 cells) is derived. Second, a binary classification of this curvature map is performed and potential edge areas are identified, where concave areas have a curvature <0.0. Thereafter, the potential edge areas are skeletonized in order to derive a final edge map. These derived edges correspond to the segment boundaries between neighboring objects. The final segment raster is achieved by combining the edge map with areas fulfilling height (e.g. >1.0 m) and ER_{ME} threshold (e.g. >5%) (Fig. 4). The conversion from raster to vector in order to derive segment polygons is done by connected component labeling followed by a vectorization step of the connected region boundaries.

4.4. Segment feature calculation

The segment feature calculation provides the input for the subsequent classification making use of the available segment features

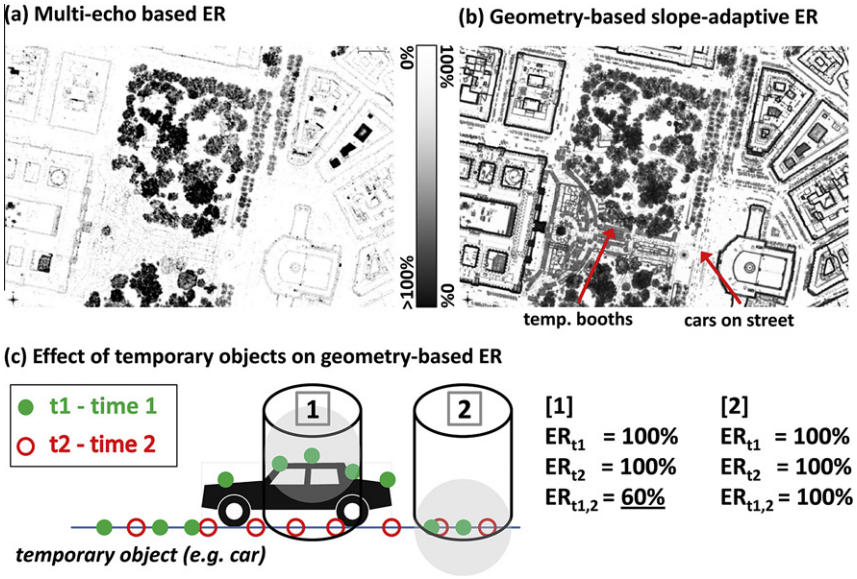

Fig. 3. (a) Multi-echo based echo ratio (ER_{ME}), (b) geometry-based echo ratio (ER_{geom}) (Höfle et al., 2009b). Temporary objects can clearly be identified in ER_{geom}. (c) Temporary objects not present in all scans cause a decrease in the values of ER_{geom}.

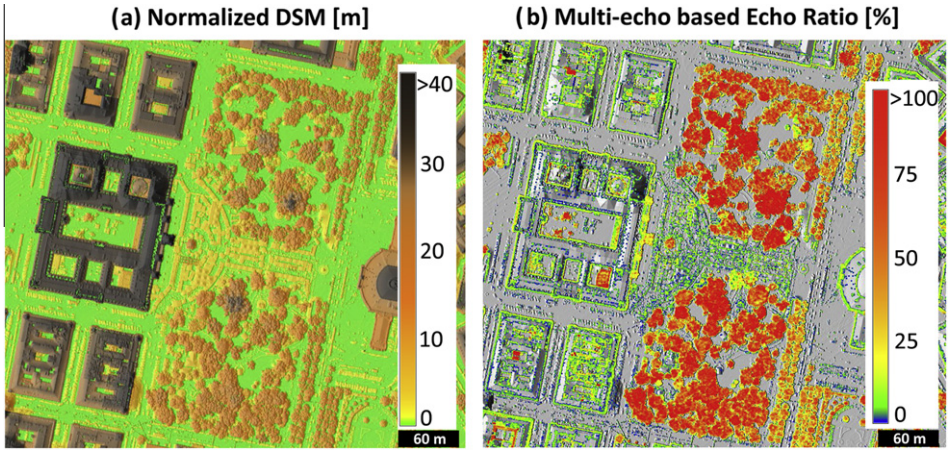

Fig. 4. Input layers for segmentation of convex regions in the (a) nDSM having (b) high penetrability of parts of the laser beam through small gaps parameterized by ER_{ME}.

Table 1
List of derived segment features serially numbered (46 features in total). SD (=Standard Deviation).

	Normalized height mean/SD	Normalized amplitude mean/SD	Echo width mean/SD	Backscatter cross section σ mean/SD	Backscatter coefficient γ mean/SD	No. of echoes
All echoes	1/2	3/4	5/6	7/8	9/10	11
First echoes	12/13	14/15	16/17	18/19	20/21	22
Multi echoes	23/24	25/26	27/28	29/30	31/32	33
Last echoes	34/35	36/37	38/39	40/41	42/42	44

45: Segment based ER_{ME}.
46: Percentage of laser echoes >1 m height above DTM ($Perc_{above}$).

(i.e. attributes attached to the segment polygons). An extensive segment feature database is generated, comprising segment features based on the point cloud and derived from full-waveform attributes (Table 1). The computed segment features are attached to the attribute table of the GIS polygon layer. First, the laser points are assigned to the respective segments by simple point-in-polygon test. Second, for each segment the laser point attributes (1) normalized height, (2) normalized amplitude, (3) echo width, (4) backscatter cross section σ, and (5) backscattering coefficient γ are aggregated per laser echo classes (1) all, (2) first, (3) multi echoes = first and intermediate, and (4) last and single echoes using the descriptive statistics (1) mean and (2) standard deviation (SD). Additionally the number of echoes per echo class is attached (i.e. four features in total). In this aggregation step the laser echoes are filtered by the minimum vegetation height value of 1 m above ground, in order to exclude the terrain signature from the segment statistics, except the descriptive statistics for the 'all' echo class for which all echoes spatially lying within a segment are used. This

means that the derived segment features based on all echoes may also contain a large number of terrain echoes contributing to the derived values. Additionally the ER_{ME} on segment basis is computed using only echoes above the minimum vegetation height (i.e. 1.0 m). Thus, terrain echoes, primarily single and last echoes, are not considered in the computation of the segment based ER_{ME}, assuming that building segments have a far higher number of single and last echoes above 1 m than trees. Furthermore, the percentage of points above the minimum vegetation height is calculated and attached to the segment feature database. Altogether 46 point cloud based features could be obtained for each segment.

4.5. Classification of segments

Compared to previous studies using e.g. a decision tree (Rutzinger et al., 2008) or Support Vector Machines (Mallet et al., 2008) for full-waveform LiDAR point cloud classification of urban areas, this approach uses the (3D) segment features as input to classify 2D polygons. Following the review of Blaschke (2010) and e.g. the study of Secord and Zakhor (2007) it can be assumed that a segment-based approach may lead to more stable features for classification than directly classifying single pixels and laser points, respectively.

For classification two different machine learning techniques are applied, (1) decision trees and (2) artificial neural networks (ANNs). ANNs imitate the brain's model of an interconnected system of neurons, enabling computers to detect patterns and to learn complex relationships within data (Anderson, 1995). However, because of the complexity of the interactions between neurons of an ANN, it is difficult to gain deep insights into the interior working of an ANN. Hence, ANNs basically provide a 'black box' model. Decision trees are tree structures where each branch of an internal node represents an outcome of a test for an attribute. The leaf nodes of the tree represent the classes or class distributions (Han and Kamber, 2001). In contrast to ANNs, decision trees are easy to interpret and understand, but they do not inherently include interactions between the input variables as well as the relationship between such interactions and any output pattern as ANNs do (Pyle, 1999).

The ANN used in this study consists of a single hidden layer with $(n + 1)/2$ hidden neurons, where n is the number of attributes. The ANN is trained for 500 cycles by backpropagation (Rumelhart et al., 1986) with a learning rate of 0.3. For decision tree induction basically the algorithm of Breiman et al. (1984) is used. For splitting nodes, at minimum 4 samples must be assigned to a node and an information gain ratio of 0.1 must be achieved. The algorithm stops if no more nodes can be split or the maximal depth of 20 is reached. Finally, the tree is pruned from leaves that do not improve classification accuracy.

Eight different input feature sets are chosen as input for both classifiers in order to assess the effect on the final classification result and thus their significance for vegetation detection (Table 2). Training and reference datasets are generated by assigning the class 'vegetation' to all segments containing at least one alley tree reference point. 'Non-vegetation' segments are limited to segments that overlap at least 50% with the building polygons of the digital cadastral map. A stratified sample of 70% is selected for training the classifiers, while the remaining data are used for validation purposes.

4.6. Derivation of vegetation mask GIS layer

The GIS vector topology allows a straightforward generation of a final vegetation mask product from a classified segment GIS polygon layer. Common boundaries between segments classified as vegetation are dissolved, i.e. neighboring polygons that 'meet' are merged. This procedure is followed by removal of small isolated polygons and closing of small holes (i.e. islands) falling below the defined minimum mapping unit of 20 m^2.

Table 2
Eight different segment feature sets used for segment classification. For each point cloud attribute (e.g. height, echo width) the aggregated segment mean (MEAN) and standard deviation (SD) values for (i) all, (ii) first, (iii) first and intermediate (i.e. multi) and (iv) last echoes are taken.

Input for features set	# Features	Derived features
Geometric point cloud features		
(1) Segment based ER_{ME}	1	Same as input[a]
(2) Percentage of laser echoes >1 m height above DTM ($Perc_{above}$)	1	Same as input[a]
(3) Normalized laser point height	8	MEAN and SD of all, first, multi and last echoes, resp.
(4) Echo width	8	MEAN and SD of all, first, multi and last echoes, resp.
Full-waveform derived features		
(5) Normalized amplitude	8	MEAN and SD of all, first, multi and last echoes, resp.
(6) Backscattering coefficient γ	8	MEAN and SD of all, first, multi and last echoes, resp.
(7) Backscattering coefficient σ	8	MEAN and SD of all, first, multi and last echoes, resp.
(8) Normalized Amplitude, Backscattering coefficient γ and σ	24	for all three input features MEAN and SD of all, first, multi and last echoes, resp.

[a] No aggregation per segment required for input.

5. Results and discussion

This section gives detailed insight into the results of the developed workflow of urban vegetation detection and explores the value of single segment-based full-waveform LiDAR features for binary classification of vegetation. The results comprise (i) ER_{ME} computation and raster layer derivation, (ii) raster-based segmentation, (iii) the derivation of radiometrically calibrated full-waveform LiDAR input data and segment feature derivation, exploratory data analysis as well as (iv) the presentation and accuracy assessment of different classification results with respect to varying input data and classifiers.

5.1. Multi-echo based echo ratio raster layer

The derived ER_{ME} raster layer clearly indicates a high agreement with potential urban vegetation areas (Figs. 3 and 4). Both the high point density and the increased echo detection sensitivity (cf. Chauve et al., 2007) provided by full-waveform LiDAR allows faster ER_{ME} derivation without the need for computationally expensive 3D point cloud neighborhood analysis. As ER_{ME} uses the number of echoes per shot as indicator for vegetation, there is a dependence on certain sensor settings and waveform processing determining the occurrence of multiple echoes. For example, larger beam widths increase the chance of multiple echoes. Higher waveform sampling intervals (e.g. <1 ns) and more sensitive waveform decomposition algorithms may lead to a higher number of echoes per shot. Thus, this has to be considered when deriving and using ER_{ME} data layers from different sensor systems and settings as well as waveform processing algorithms. On the sensor side the occurrence of multiple echoes can be increased by using larger beam divergence and hence footprint sizes. Although the accuracy of the position determination of the actually contributing target(s) decreases with increasing footprint size, for vegetation mapping

and characterization larger footprints (e.g. >1 m diameter) may be preferable. In particular, a LiDAR system composed of a high density small-footprint scanner together with a lower frequency large-footprint system could provide both a detailed geometric description of the objects as well as representative backscatter information of larger units (i.e. footprints).

5.2. Raster-based segmentation

Due to the integration of the thresholds (i.e. masks) on normalized height and ER_{ME} in the segmentation step, the area for further detailed analysis can be clearly reduced (to 24.4% of the full extent), which generally depends on the percentage of vegetated area in the study area. By assuming that all derived segments are of class 'vegetation', 92% of the segments do not intersect with buildings, indicating already a high agreement after segmentation. For deriving the segment boundaries not determined either by normalized height or ER_{ME}, the curvature based edge detection defines the segment boundary. The selected window size for curvature computation affects the degree of detail that will be expressed by the segment boundaries and results in over- or under-segmentation (cf. Höfle et al., 2008). The larger the size of the moving window the fewer segments are found while the size of the segments increases. Additionally, the threshold on the curvature raster controls the sensitivity for edge detection. The more moderate the threshold value is set, the more potential edges can be detected at cost of wider potential edge zones (Fig. 5c), which consequently decreases the planimetric accuracy of the "thinned" edge line. As a further improvement of the segmentation, deriving the depression line between two segments could be performed by weighted skeletonization (weighted by curvature and height) or by determination of the point with highest concavity in the boundary zone of two segments.

The result of the segmentation are ca. 31,000 segments with an average size of 11 m^2 compared to 67×10^6 laser echoes within the study site. The edge detection was performed with a window size of 7×7 pixels (i.e. 3.5x3.5 m) for curvature calculation with a curvature constraint of <-0.2. The segment boundaries indicate that trees with a compact and convex crown shape result in one segment (Fig. 5). Large deciduous trees with multiple tree tops lead to more than one segments (one per convex tree top). Minor over-segmentation can be seen for the alley trees (181 selected trees) where the average number of segments per tree is 2.6, indicating trees with multiple tops. Particularly in the park areas with larger deciduous trees, the number of segments per tree is clearly increased. Due to missing reference tree positions, the over-segmentation factor could not be assessed in the park areas. To avoid over-segmentation, prior filtering (e.g. Gauss filter) of the DSM could be performed in order to reduce canopy roughness and suppress small structures (cf. Hirschmugl et al., 2007).

5.3. Full-waveform segment features

Complementary to the multi-echo based parameterization using ER_{ME}, the radiometric information obtained by radiometric calibration provides further sub-footprint description of the scanned objects. These full-waveform features are available for each laser echo and are then aggregated per segment to derive the segment statistics. In order to get a better impression of the information content of the FWF observables in urban areas and how different objects are represented, Fig. 6 compares images derived from the point cloud using the recorded echo width, normalized amplitude and backscattering coefficient γ. The vegetation areas are characterized by high values of echo width and low values of signal amplitude, which finally results in low mean values of backscattering coefficient γ. For multiple targets in one waveform

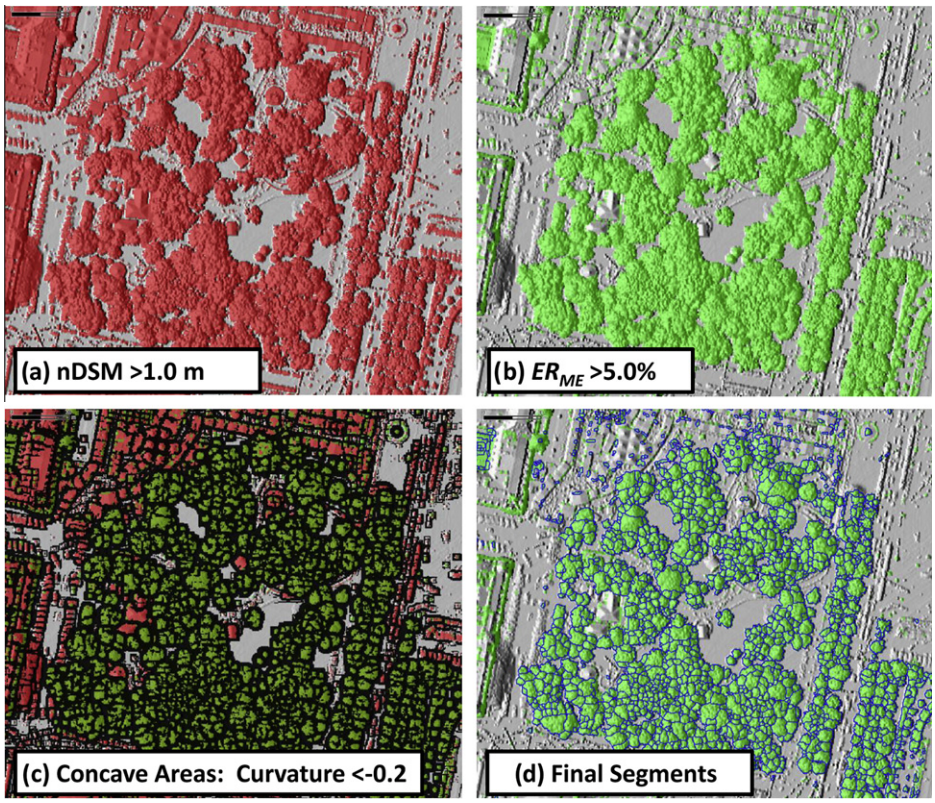

Fig. 5. Input layers for segmentation of convex regions in the nDSM having high multi-echo based echo ratio (ER_{ME}).

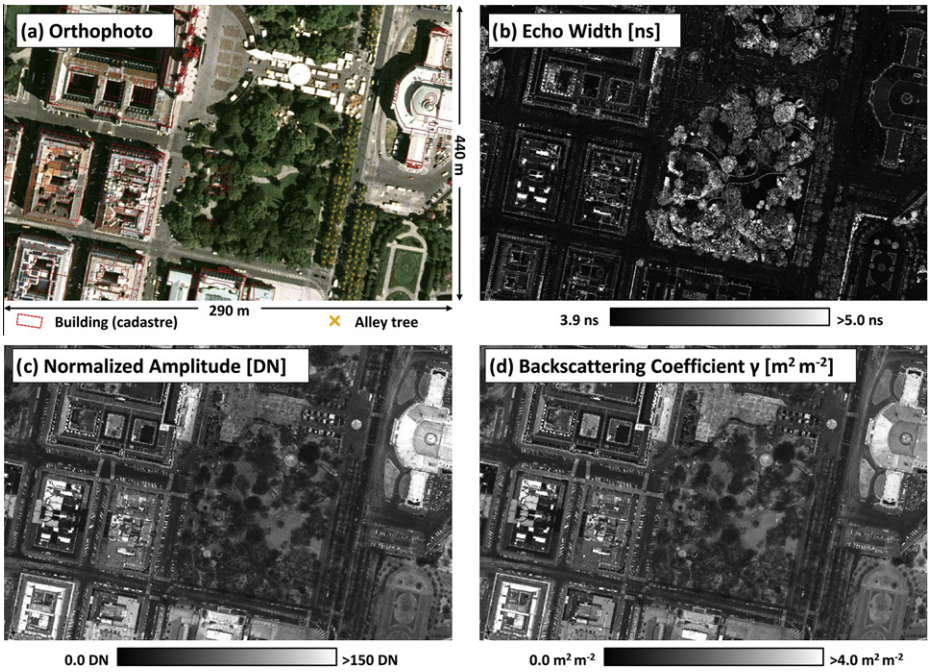

Fig. 6. (a) Orthophoto and images (0.5 m cell size; mean value of all echoes per cell) of full-waveform LiDAR data layers: (b) recorded echo width, (c) normalized amplitude and (d) backscattering coefficient γ derived by radiometric calibration.

the individual target areas are smaller than the size of the laser footprint and hence may be independent of the footprint size (Wagner, 2010). Thus, γ values of the same small multi-echo targets differ when changing the footprint size due to the normalization by footprint area. This has to be taken into account when using LiDAR data with different footprint areas, particularly for γ values of high vegetation where target scattering areas are generally small (Alexander et al., 2010). Furthermore, directional backscattering effects can be seen in the amplitude and γ images (cf. Lehner and Briese, 2010; Wagner, 2010), particularly at building roofs where higher values are generally apparent at low incidence angles (i.e. roof facing towards the sensor). A detailed evaluation of the quality of the applied radiometric calibration procedure is given by Lehner and Briese (2010).

The derived segment features (i.e. attributes) are used as input for classification. Thus, the selection and derivation of the attached features is of importance to achieve high classification accuracy. Fig. 7 shows the histograms of the segment based mean values of echo widths for three different surface classes. It can be seen that the segment mean echo widths of first echoes have larger variation than the segment mean values of all echoes. Using all echoes to derive mean segment echo width values leads to a large averaging effect, which does not allow separating the three shown classes. It can be seen that buildings tend to have more homogeneous echo widths in last echoes than trees. These segment based signatures correspond to the general findings identified by Mallet et al. (2008) and Alexander et al. (2010) based on single laser point attributes. However, the overlap of the class distributions of the segment based features is very large using all echoes. Certain separability can be denoted in the histograms of the first echo mean values, where vegetation exhibits higher echo widths. The alley trees have an average segment echo width using first echoes of 4.46 ns with 0.16 ns standard deviation (SD), the park trees 4.60 ns (SD = 0.24 ns) and the buildings have lower values in average of 4.22 ns (SD = 0.19 ns).

A major advantage of the combined object-based image and point cloud approach is the full access to the 3D information inherent in the point cloud. The vertical profiles of boxplots in Fig. 8 explore the vertical distribution of laser echo attributes within a representative single tree and building segment, clearly demonstrating the benefit of accessing the point cloud for segment feature derivation. The selected building and alley tree segment have similar average echo width but show different vertical profiles. The first echoes in the tree top have larger echo widths than the laser echoes from below the canopy (e.g. tree stem) whereas the building has a relatively constant echo width through the vertical profile except the few outliers with very high values. In contrast, the alley tree has lower γ values throughout the vertical profile than the building segment exhibiting large variation in values and higher average values.

Fig. 8 points out that the vertical distribution within segments contains further potential for deriving significant features. In this study, the features are derived per echo class, which can be seen as parameterization of the vertical profile as first echoes tend to originate from the top layer. Further investigations should be performed on the vertical distribution of full-waveform attributes such as Höfle et al. (2008) used a dynamic minimum height threshold (e.g. 50% nDSM height) to select laser echoes for feature calculation in order to discriminate tree species.

5.4. Classification of segments

As evaluation dataset 4171 non-vegetation segments were identified by spatial overlay with the buildings of the digital cadastral map as well as 642 vegetation segments by using the alley tree positions of the city's tree cadastre. The segmentation step already excludes most of non-vegetation areas but the resulting segments still contain non-vegetation objects such as mainly building walls, roof overhangs and power lines. Thus, the main task is to identify and remove non-vegetation segments in order to derive an accurate vegetation mask.

Table 3 lists the achieved correctness (i.e. precision) and completeness (i.e. true positive rate) for the 16 different results. The most prominent results are the high correctness and completeness using either (i) the segment based ER_{ME} (0.98 correctness), (ii) percentage of points above >1 m normalized height (0.97 correctness)

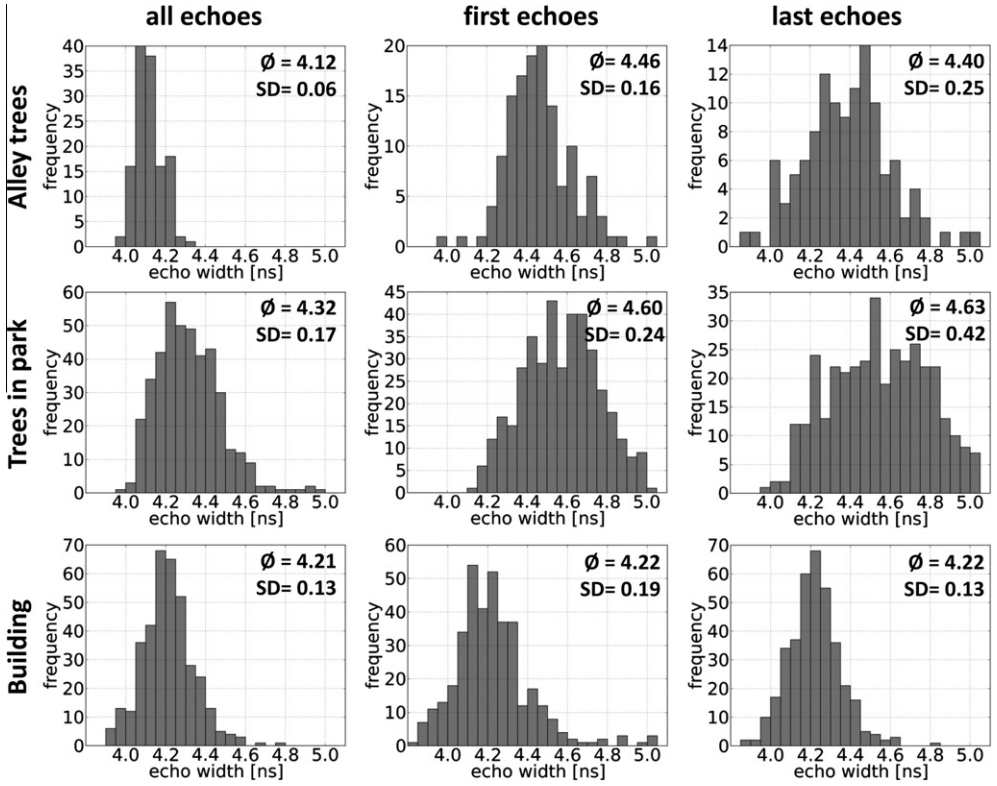

Fig. 7. Histograms of mean echo width per segment (in the range of 3.8 ns to 5.0 ns) for selected alley trees, trees in park area and buildings. The mean echo width values are derived for all echoes, first echoes and last echoes only. Mean values (*Ø*) and standard deviations (*SD*) of the distributions shown in the histograms are attached to each subplot.

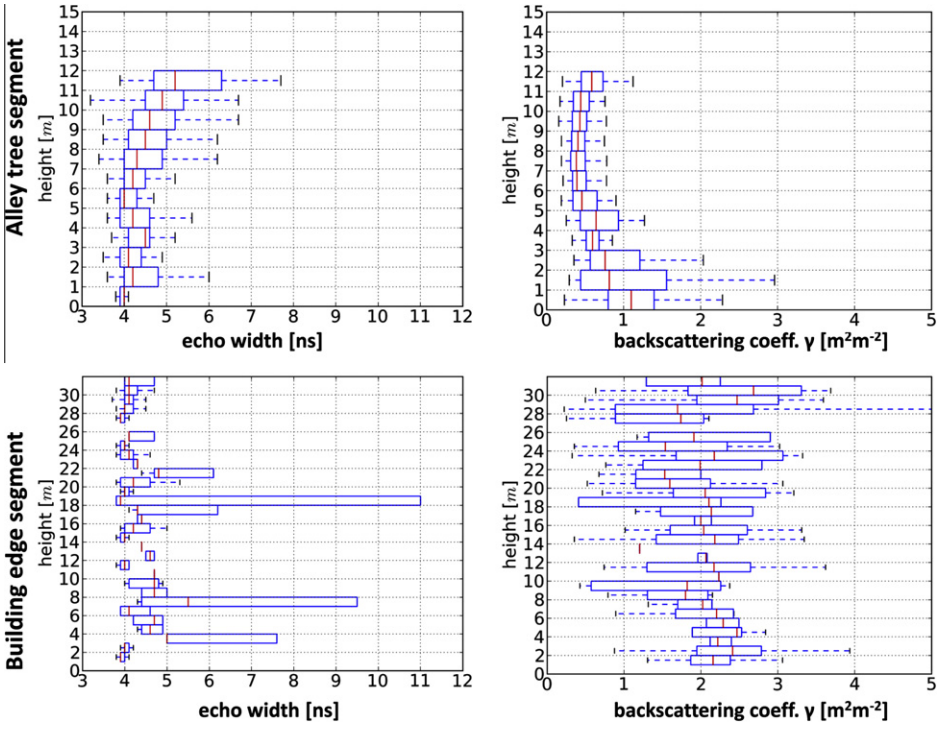

Fig. 8. Boxplots in vertical profiles (1 m intervals) of echo width and backscatter coefficient γ using all echoes of a selected alley tree and building segment.

and (iii) echo width ($\geqslant$0.95 correctness). The differences between the two chosen classifiers are marginal for the geometric point cloud features compared to the differences in correctness for the full-waveform features, where mostly the ANN achieves higher correctness but no clear trend can be deduced. Looking for example at the normalized amplitude feature set, the ANN achieves a higher

Table 3
Accuracy assessment of binary classification (vegetation/non-vegetation) using different feature sets based on different input data (cf. Table 2). Abbreviations: True Positive (TP), False Positive (FP), False Negative (FN) and True Negative (TN).

Input for feature set	Classifier	TP	FP	FN	TN	TN-Rate	Completeness (TP-Rate)	Correctness (Precision)
Geometric point cloud features								
Segment based ER_{ME}	Dec.Tree	182	4	11	1247	1.00	0.94	0.98
	ANN	180	4	13	1247	1.00	0.93	0.98
Perc. of points >1 m height [%]	Dec.Tree	184	5	9	1246	1.00	0.95	0.97
	ANN	182	5	11	1246	1.00	0.94	0.97
Normalized point height	Dec.Tree	181	6	12	1245	1.00	0.94	0.97
	ANN	190	2	3	1249	1.00	0.98	0.99
Full-waveform derived features								
Echo width	Dec.Tree	118	6	75	1245	1.00	0.61	0.95
	ANN	169	2	24	1249	1.00	0.88	0.99
Normalized amplitude	Dec.Tree	179	72	14	1179	0.94	0.93	0.71
	ANN	106	6	87	1245	1.00	0.55	0.95
Backscattering coefficient γ	Dec.Tree	171	67	22	1184	0.95	0.89	0.72
	ANN	154	18	39	1233	0.99	0.80	0.90
Backscattering coefficient σ	Dec.Tree	180	106	13	1145	0.92	0.93	0.63
	ANN	38	2	155	1249	1.00	0.20	0.95
Normalized amplitude, γ, σ	Dec.Tree	174	26	19	1225	0.98	0.90	0.87
	ANN	184	73	9	1178	0.94	0.95	0.72

correctness of 0.95 but much lower completeness of 0.55 compared to the decision tree classifier with a completeness of 0.93 at correctness of 0.71. Different results of the chosen classifiers may be due to that decision trees – compared to ANNs - do not inherently include interactions between the input variables and relationships between such interactions. For further general details on the two machine learning algorithms it shall be referred to e.g. Pyle (1999).

The radiometrically calibrated full-waveform features (normalized amplitude, γ and σ) do not allow a precise classification into vegetation and non-vegetation, if they are used without any other attributes and a decision tree classifier (correctness <0.72). The geometric point cloud features have very high true negative rates of 1.0 as well as relatively high completeness of >0.93. This indicates that both vegetation and non-vegetation segments can be classified at high rates. The full-waveform features show (i) lower completeness than the geometric features as many vegetation segments are classified as buildings but (ii) have low rates of buildings classified as vegetation (high true negative rate), indicating that non-vegetation is more significantly characterized by full-waveform features. The results suggest that the LiDAR signatures of natural objects (i.e. trees) are far more heterogeneous (and partly similar to buildings) than of artificial objects (i.e. buildings) which can be clearly separated from trees.

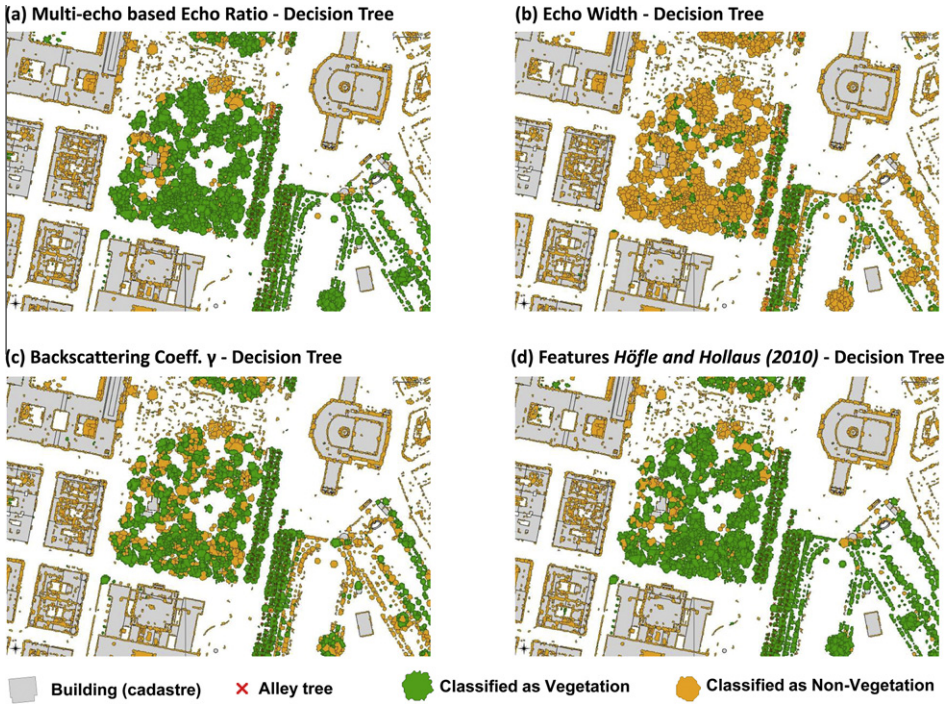

Fig. 9. Classification results using a decision tree classifier and four different input data: (a) ER_{ME} per segment, (b) echo width, (c) backscattering coefficient γ and (d) features used in Höfle and Hollaus (2010).

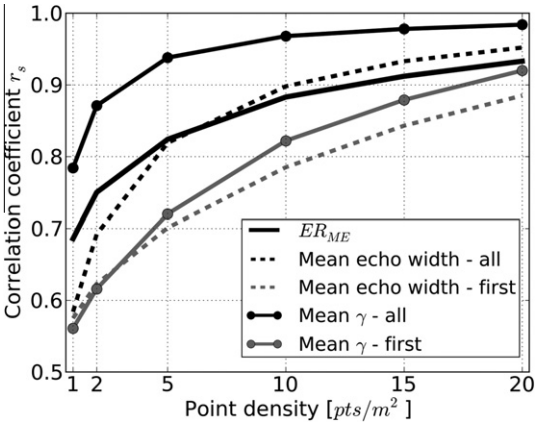

Fig. 10. Effect of point density on segment features: Correlation between segment features derived from the entire point cloud (50 pts./m^2) and the corresponding features derived from artificially reduced point clouds by randomly selecting a maximum of pts./m^2. The selected point densities are 1, 2, 5, 10, 15 and 20 pts./m^2.

The segment based ER_{ME} seems to be the most effective discriminator with hardly any requirements for calculation (Fig. 9a). A simple rule-base is derived by the decision tree classifier: $ER_{ME} > 108.4$: vegetation; $ER_{ME} \leqslant 108.4$: non-vegetation. For ER_{ME} there is no need for prior DTM generation, e.g. compared to the feature where points above 1 m of the DTM are counted. However, ER_{ME} can be affected by the waveform decomposition strategy as well as the sensor settings (e.g. beam width). Furthermore, it can be seen that except the echo width the full-waveform features are quite 'instable' as they achieve quite different results w.r.t. the used classifier. In Fig. 9 these large differences in the results are displayed by the classified segment maps. Fig. 9b points out that the alley trees (i.e. reference data) are detected quite well using the echo width, but the large park area in the middle of the image has been classified as non-vegetation. Without independent training data for large park trees and shrubs, the accuracy assessment of the vegetation mapping remains limited to alley trees. For this study only alley tree positions were available because the tree positions in park areas are not included in the city's tree cadastre. However, ER_{ME} (Fig. 9a) delivers comparable results as the features used in Höfle and Hollaus (2010) who used a manually derived rule-base covering park and alley trees. The feature set of Höfle and Hollaus (2010) – e.g. mean amplitude of first echoes, mean echo width of first echoes, segment based ER_{ME} etc. – has been tested using a decision tree classifier resulting in a correctness of 0.96 (Fig. 9d).

5.5. Effect of point density on segment features and classification accuracy

To address the contribution of the relatively high point density (50 pts./m^2) of the study area to the accuracy of the results, artificially reduced point clouds were generated by randomly selecting a maximum number of points per square meter. In order to sustain comparability, the segment polygons derived from the original point cloud are used as spatial unit for feature calculation and classification with varying point densities.

Fig. 10 shows that segment features derived from a point cloud with 20 pts./m^2 have a high correlation ($r_s > 0.89$) in all selected segment features expressed by the Spearman's rank correlation coefficient r_s. With even lower point densities the correlation decreases and thus the segment feature values significantly change compared to the values derived from entire point cloud. The highest stability w.r.t. point density can be attested for the mean values of the backscatter coefficient γ. which still has a $r_s = 0.94$ at 5 pts./m^2. The ER_{ME} is also relatively stable but the correlation drastically decreases below 5 pts./m^2. The decrease of correlation with point density can be explained by the low number of echoes falling in one segment, i.e. sample size, resulting in noisier features (cf. Secord and Zakhor, 2007). The low value for ER_{ME} may also be explained by the fact that not laser shots with all corresponding echoes have been selected in the reduction process but simply all echoes have been randomly reduced, which may alter the ER_{ME} value for small segments.

The effect of reduced point densities on the final classification result is investigated in Fig. 11 using a solely a geometric and full-waveform derived feature, respectively. As already indicated in Fig. 10 the segment features of γ are relatively insensitive to reduced point densities compared to the geometric feature ER_{ME}, which is also expressed in the high classification accuracies using features of γ derived from point clouds with <5 pts./m^2. The classification using the segment feature ER_{ME} still exhibits a high correctness of 0.92 with densities lower 5 pts./m^2 but the completeness drops below 0.42.

6. Conclusions

This paper presents a novel workflow of GIS-based urban vegetation mapping using high density full-waveform LiDAR data. The joint utilization of object-based image and point cloud analysis for segment feature derivation and classification combines the strengths of both domains: (i) fast raster data processing compared to processing of the original unstructured 3D point cloud and (ii) 3D geometric and radiometric point cloud LiDAR object

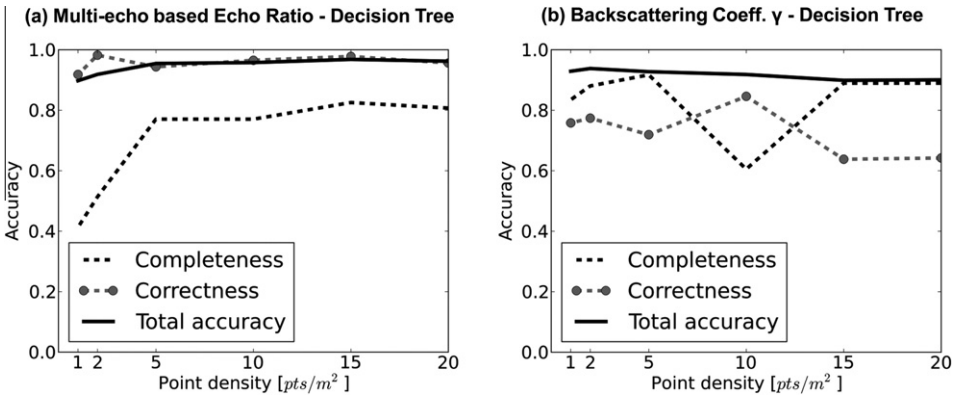

Fig. 11. Effect of point density on classification accuracies using the decision tree classifier with different input datasets: (a) segment based ER_{ME} and (b) segment features derived from the backscatter coefficient γ. The point cloud is artificially reduced to point densities of 1, 2, 5, 10, 15 and 20 pts./m^2.

characterization. Multi-level LiDAR analysis, such as prior image based detection of candidate regions with following point cloud based object detection increasing delineation and classification accuracy offers the possibility to process large areas even with very high point densities in an operational manner without major loss in classification accuracy if the pre-selection has high completeness.

The investigation of the binary classification performance of the derived segment features (i.e. geometric and radiometric) points out higher correctness values and stability with different classifiers for geometric point cloud features than for full-waveform derived features. Particularly the occurrence of multiple distinct targets parameterized by an echo ratio already leads to high correctness and detection rates, achieving comparable results as with complex classifications with many input features. Furthermore, by using echo ratios (either multiple echoes or vertical echo distribution) there is no need for prior DTM generation and height normalization. Thus, fast image-based detection of potential vegetation areas for improved terrain point filtering could be a promising option. The echo width achieves high correctness and fits very well with the reference data. However, the detected vegetation areas are limited to the training areas (i.e. alley trees) and do not cover trees in parks, indicating a difference in echo width between different tree species. The radiometric LiDAR echo attributes (e.g. normalized amplitude and backscattering coefficient) do not suffice as single discriminator between vegetation and non-vegetation in densely built-up urban areas. However, as already shown by several authors these additional full-waveform LiDAR signatures are a valuable source of information for further vegetation characterization (e.g. tree species discrimination) as well as the derivation of physical quantities of the scanned surface (e.g. surface roughness; Hollaus et al., 2011). Tests using datasets with different point density for segment feature calculation and classification clearly point out that the segment features as well as the classification accuracy remains relatively stable for point densities higher than 5 pts./m^2 (i.e. reduction of factor 10). Even lower point densities lead to noisier features due to the lower number of echoes lying within a segment.

A more diverse reference and training dataset is required to be able to improve the detection of trees in park areas as well as shrubs. The great amount and variety in potential segment features could be treated in a more stable way by combining robust optimization methods with classification algorithms, e.g. using genetic algorithms for selecting significant features prior to an ANN or decision tree classification. Furthermore, in order to make use of the distribution of the point cloud attributes within a segment, approaches such as self-organizing maps should be tested.

To be able to use LiDAR for tree inventory, the precise planimetric position of the single trees (i.e. stem position) is required. Even with high density airborne LiDAR data urban areas remain a challenging and complex environment for single tree detection as a strong confusion with many other artificial objects exists. Thus, it should be considered to combine airborne and mobile laser scanning (from cars; cf. Rutzinger et al., 2011) to gather complete 3D geometric models, tree positions and further tree attributes (e.g. crown width). Furthermore, future work should investigate the effect of scan settings and scan geometry (e.g. beam divergence and sampling interval) and waveform processing algorithms and settings on the quality of the radiometric LiDAR observables.

Acknowledgments

We would like to thank the MA41-Stadtvermessung, City of Vienna, for their support and providing the airborne LiDAR data and reference datasets. Furthermore, we thank Hubert Lehner for the radiometric calibration of the LiDAR data and Marco Helbich for fruitful input to the manuscript. This work was partly funded by the Federal Ministry of Economics and Technology (BMWi), Germany, in the framework of the project "HyLand" (FKZ 50EE1014).

References

Alexander, C., Tansey, K., Kaduk, J., Holland, D., Tate, N.J., 2010. Backscatter coefficient as an attribute for the classification of full-waveform airborne laser scanning data in urban areas. ISPRS Journal of Photogrammetry and Remote Sensing 65 (5), 423–432.

Anderson, J.A., 1995. An Introduction to Neural Networks. MIT Press, Cambridge, MA.

Blaschke, T., 2010. Object based image analysis for remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing 65 (1), 2–16.

Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. Classification and Regression Trees. Wadsworth & Brooks/Cole Advanced Books & Software, Monterey, CA.

Bretar, F., Chehata, N., 2007. Digital terrain model on vegetated areas: joint use of airborne lidar data and optical images. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 3/W49A), 19–24.

Briese, C., Höfle, B., Lehner, H., Wagner, W., Pfennigbauer, M., Ullrich, A., 2008. Calibration of full-waveform airborne laser scanning data for object classification. In: Turner, M.D., Kamerman, G.W. (Eds.), Proceedings of SPIE Laser Radar Technology and Applications XIII, vol. 6950, 19–20 March, Orlando, FL, pp. 0H1–0H8.

Carlberg, M., Gao, P., Chen, G., Zakhor, A., 2009. Classifying urban landscape in aerial LiDAR using 3D shape analysis. In: Proceedings of the 16th IEEE International Conference on Image Processing (ICIP), 7–10 November, Cairo, Egypt, pp. 1701–1704.

Chauve, A., Mallet, C., Bretar, F., Durrieu, S., Pierrot-Deseilligny, M., Puech, W., 2007. Processing full-waveform lidar data: modelling raw signals. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 3/W52), 195–200.

Clode, S., Rottensteiner, F., 2007. Detection and vectorization of roads from lidar data. Photogrammetric Engineering and Remote Sensing 73 (5), 517–536.

GRASS Development Team, 2010. Geographic Resources Analysis Support System (GRASS) Software, Version 6.4.0. Open Source Geospatial Foundation, <http://grass.osgeo.org> (accessed 1.5.2011).

Gross, H., Jutzi, B., Thoennessen, U., 2007. Segmentation of tree regions using data of a full-waveform laser. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 3), 57–62.

Guo, L., Chehata, N., Mallet, C., Boukir, S., 2011. Relevance of airborne lidar and multispectral image data for urban scene classification using random forests. ISPRS Journal of Photogrammetry and Remote Sensing 66 (1), 56–66.

Haala, N., Kada, M., 2010. An update on automatic 3D building reconstruction. ISPRS Journal of Photogrammetry and Remote Sensing 65 (6), 570–580.

Han, J., Kamber, M., 2001. Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers, San Francisco, CA.

Heinzel, J., Koch, B., 2011. Exploring full-waveform LiDAR parameters for tree species classification. International Journal of Applied Earth Observation and Geoinformation 13 (1), 152–160.

Hirschmugl, M., Ofner, M., Raggam, J., Schardt, M., 2007. Single tree detection in very high resolution remote sensing data. Remote Sensing of Environment 110 (4), 533–544.

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.

Höfle, B., Hollaus, M., Lehner, H., Pfeifer, N., Wagner, W., 2008. Area-based parameterization of forest structure using full-waveform airborne laser scanning data. In: Hill, R.A., Rosette, J., Suárez, J. (Eds.), Proceedings of SilviLaser 2008: 8th international conference on LiDAR applications in forest assessment and inventory, 18–19 September, Edinburgh, UK, pp. 227–235.

Höfle, B., Vetter, M., Pfeifer, N., Mandlburger, G., Stötter, J., 2009a. Water surface mapping from airborne laser scanning using signal intensity and elevation data. Earth Surface Processes and Landforms 34 (12), 1635–1649.

Höfle, B., Mücke, W., Dutter, M., Rutzinger, M., Dorninger, P., 2009b. Detection of building regions using airborne LiDAR - A new combination of raster and point cloud based GIS methods. In: Car, A., Griesebner, G., Strobl, J. (Eds.), Geospatial Crossroads @ GI_Forum' 09: Proceedings of the Geoinformatics Forum Salzburg, 7–10 July, Salzburg, pp. 66–75.

Höfle, B., Hollaus, M., 2010. Urban vegetation detection using high density full-waveform airborne LiDAR data – Combination of object-based image and point cloud analysis. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (Part 7B), 281–286.

Hollaus, M., Aubrecht, C., Höfle, B., Steinnocher, K., Wagner, W., 2011. Roughness mapping on various vertical scales based on full-waveform airborne laser scanning data. Remote Sensing 3 (3), 503–523.

Hyyppä, J. et al., 2001. HIGH-SCAN: The first European-wide attempt to derive single-tree information from laserscanner data. The Photogrammetric Journal of Finland 17, 58–68.

Iovan, C., Boldo, D., Cord, M., 2008. Detection, characterization, and modeling vegetation in urban areas from high-resolution aerial imagery. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 1 (3), 206–213.

IPF, 2011. OPALS – Orientation and Processing of Airborne Laser Scanning data, <http://www.ipf.tuwien.ac.at/opals/> (accessed 1.5.2011).
Jelalian, R., 1992. Laser Radar Systems. Artech House, Boston, London.
Jochem, A., Höfle, B., Hollaus, M., Rutzinger, M., 2009. Object detection in airborne LIDAR data for improved solar radiation modeling in urban areas. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (Part 3/W8), 1–6.
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.
Lehner, H., Briese, C., 2010. Radiometric calibration of full-waveform airborne laser scanning data based on natural surfaces. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (Part 7B), 360–365.
Lemmens, M., 2009. Airborne Lidar Sensors. GIM International 23 (2), 16–19.
Liang, X., Hyyppä, J., Matikainen, L., 2007. Deciduous-coniferous tree classification using difference between first and last pulse laser signatures. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 3/W52), 253–257.
Lin, Y.C., Mills, J.P., 2010. Factors influencing pulse width of small footprint, full waveform airborne laser Scanning data. Photogrammetric Engineering and Remote Sensing 76 (1), 49–59.
MA41, 2011. Stadtvermessung Wien - MA41, http://www.wien.gv.at/stadtentwicklung/stadtvermessung/ (Accessed 1 May, 2011).
Mallet, C., Bretar, F., Soergel, U., 2008. Analysis of full-waveform LiDAR data for classification of urban areas. Photogrammetrie Fernerkundung Geoinformation (PFG) 2008 (5), 337–349.
Mallet, C., Bretar, F., 2009. Full-waveform topographic LiDAR: state-of-the-art. ISPRS Journal of Photogrammetry and Remote Sensing 64 (1), 1–16.
Mallet, C., Lafarge, F., Bretar, F., Roux, M., Soergel, U., Heipke, C., 2009. A stochastic approach for modelling airborne LIDAR waveforms. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (Part3/W8), 201–206.
Mandlburger, G., Otepka, J., Karel, W., Wagner, W., Pfeifer, N., 2009. Orientation And Processing Of Airborne Laser Scanning Data (OPALS) – Concept and first results of a comprehensive ALS software. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (Part 3/W8), 55–60.
Matikainen, L., Hyyppä, J., Ahokas, E., Markelin, L., Kaartinen, H., 2010. Automatic detection of buildings and changes in buildings for updating of maps. Remote Sensing 2 (5), 1217–1248.
Melzer, T., 2007. Non-parametric segmentation of ALS point clouds using mean shift. Journal of Applied Geodesy 1 (3), 159–170.
Oude Elberink, S., Vosselman, G., 2009. Building reconstruction by target based graph matching on incomplete laser data: analysis and limitations. Sensors 9 (8), 6101–6118.
Pyle, D., 1999. Data preparation for Data Mining. Morgan Kaufmann Publishers, San Francisco, CA.
Reitberger, J., Schnörr, C., Krzystek, P., Stilla, U., 2009. 3D segmentation of single trees exploiting full waveform LIDAR data. ISPRS Journal of Photogrammetry and Remote Sensing 64 (6), 561–574.
Rottensteiner, F., Trinder, J., Clode, S., Kubik, K., 2007. Building detection by fusion of airborne laser scanner data and multi-spectral images: performance evaluation and sensitivity analysis. ISPRS Journal of Photogrammetry and Remote Sensing 62 (2), 135–149.
Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning representations by back-propagating errors. Nature 323, 533–536.
Rutzinger, M., Höfle, B., Pfeifer, N., 2007. Detection of high urban vegetation with airborne laser scanning data. In: Proceedings of ForestSat 2007, 5–7 November, Montpellier, France, 5 p. (on CDROM).
Rutzinger, M., Höfle, B., Hollaus, M., Pfeifer, N., 2008. Object-based point cloud analysis of full-waveform airborne laser scanning data for urban vegetation classification. Sensors 8 (8), 4505–4528.
Rutzinger, M., Pratihast, A.K., Oude Elberink, S., Vosselman, G., 2011. Tree modelling from mobile laser scanning datasets. The Photogrammetric Record 26 (135), 361–372.
Secord, J., Zakhor, A., 2007. Tree detection in urban regions using aerial Lidar and image data. IEEE Geoscience and Remote Sensing Letters 4 (2), 196–200.
SCOP, 2010. SCOP++ software, Institute of Photogrammetry and Remote Sensing (I.P.F.), <http://www.ipf.tuwien.ac.at/products/products.html> (accessed 1.5.2011).
Sithole, G., Vosselman, G., 2006. Bridge detection in airborne laser scanner data. ISPRS Journal of Photogrammetry and Remote Sensing 61 (1), 33–46.
Steinvall, O., 2000. Effects of target shape and reflection on laser radar cross sections. Applied Optics 39 (24), 4381–4391.
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. CRC Press, pp. 215–234.
Stoter, J., De Kluijver, H., Kurakula, V., 2008. 3D noise mapping in urban areas. International Journal of Geographical Information Science 22 (8), 907–924.
Tóvári, D., Vögtle, T., 2004. Classification methods for 3D objects in laserscanning data. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 35 (Part B3), 408–413.
Vosselman, G., Maas, H.G., 2010. Airborne and Terrestrial Laser Scanning. Whittles Publishing, 318 p.
Wagner, W., 2010. Radiometric calibration of small-footprint airborne laser scanner measurements: basic physical concepts. ISPRS Journal of Photogrammetry and Remote Sensing 65 (10), 505–513.
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.
Wehr, A., Lohr, U., 1999. Airborne laser scanning – an introduction and overview. ISRPS Journal of Photogrammetry and Remote Sensing 54 (2–3), 68–82.
Yao, W., Hinz, S., Stilla, U., 2011. Extraction and motion estimation of vehicles in single-pass airborne LiDAR data towards urban traffic analysis. ISPRS Journal of Photogrammetry and Remote Sensing 66 (3), 260–271.
Yuan, X.H., Sarma, V., 2011. Automatic urban water-body detection and segmentation from sparse ALSM data via spatially constrained model-driven clustering. IEEE Geoscience and Remote Sensing Letters 8 (1), 73–77.